

C.P. Wong

Kyoung-Sik Moon

Yi Li

Editors

Nano-Bio-Electronic, Photonic and MEMS Packaging



Springer

Nano-Bio- Electronic, Photonic and MEMS Packaging

C.P. Wong · Kyoung-Sik (Jack) Moon · Yi Li
Editors

Nano-Bio- Electronic, Photonic and MEMS Packaging



Editors

C.P. Wong
School of Materials Science &
Engineering
Georgia Institute of Technology
771 Ferst Drive NW,
Atlanta GA 30332-0245, USA
cp.wong@mse.gatech.edu

Kyoung-Sik (Jack) Moon
School of Materials Science &
Engineering
Georgia Institute of Technology
771 Ferst Drive NW,
Atlanta GA 30332-0245, USA
jack.moon@gatech.edu

Yi Li
Intel Corporation
CH5-159
5000 W. Chandler Blvd.
Chandler AZ 85226, USA
yi.li@intel.com

ISBN 978-1-4419-0039-5 e-ISBN 978-1-4419-0040-1
DOI 10.1007/978-1-4419-0040-1
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2009941849

© Springer Science+Business Media, LLC 2010

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

For the past decade, we have witnessed tremendous advances in nano- and bio-related science, with enormous technical information and literature published worldwide. However, limited numbers of commercialized and industrialized items have appeared. Reproducibility, low yield, and difficulties in packaging, which are critically needed to manufacture usable devices and systems, are issues that researchers strive to resolve.

This book provides comprehensive reviews and overviews on the latest developments and cutting edges on nano- and bio-packaging technologies and their science, including nano- and biomaterials, devices and thermal issues for nanobio-packaging, and the molecular or atomistic scale modeling to predict those small world phenomena.

This book is composed of 20 Chapters written by well-recognized world experts in this field. Chapters 1–4, 8, and 10 review various nanomaterials for nanopackaging technologies and their most recent research including nanomaterials for electrical and thermal interconnections and nanosurface manipulation for nanopackaging. Chapter 5 addresses a novel combustion method for nanoparticle synthesis and nanosurface coating. Novel nanomaterials for renewable energy and energy conversion devices are reviewed in Chapters 6 and 7. Chapter 9 addresses passive devices by using nanomaterials, and Chapter 11 reviews structural analysis of nanoelectronics and optical devices. Latest reviews on nano-thermal science are presented in Chapters 12 and 13. Some of the hot issues in the health care arena are the biomedical device and NEMS packaging, and the biosensor device and materials. Their recent development and research are reviewed in Chapters 14–17. Computational molecular and atomistic scale simulations could provide us with a profound insight on the understanding of nano- and bio-related systems which are difficult to realize without high-cost equipment. These modeling efforts are reviewed in Chapters 18 to 20.

We are indebted to all the contributions and efforts of the authors who share their expertise in these vital areas in delivering this book to our readers.

Atlanta, Georgia, USA
Atlanta, Georgia, USA

C.P. Wong
Kyoung-Sik (Jack) Moon

Contents

Nanomaterials for Microelectronic and Bio-packaging	1
C.P. Wong and Kyoung-sik (Jack) Moon	
Nano-conductive Adhesives for Nano-electronics Interconnection	19
Yi Li, Kyoung-sik (Jack) Moon, and C.P. Wong	
Biomimetic Lotus Effect Surfaces for Nanopackaging	47
Yonghao Xiu and C.P. Wong	
Applications of Carbon Nanomaterials as Electrical Interconnects and Thermal Interface Materials	87
Wei Lin and C.P. Wong	
Nanomaterials via NanoSpray Combustion Chemical Vapor Condensation, and Their Electronic Applications	139
Andrew Hunt, Yongdong Jiang, Zhiyong Zhao, and Ganesh Venugopal	
1D Nanowire Electrode Materials for Power Sources of Microelectronics	167
Jaephil Cho	
Mechanical Energy Harvesting Using Wurtzite Nanowires	185
Xudong Wang and Zhong Lin Wang	
Nanolead-Free Solder Pastes for Low Processing Temperature Interconnect Applications in Microelectronic Packaging	217
Hongjin Jiang, Kyoung-sik (Jack) Moon, and C.P. Wong	
Introduction to Nanoparticle-Based Integrated Passives	247
Ranjith John and Ajay P. Malshe	
Thermally Conductive Nanocomposites	277
Jan Felba	

Physical Properties and Mechanical Behavior of Carbon Nano-tubes (CNTs) and Carbon Nano-fibers (CNFs) as Thermal Interface Materials (TIMs) for High-Power Integrated Circuit (IC) Packages: Review and Extension	315
Yi Zhang, Ephraim Suhir, and Claire Gu	
On-Chip Thermal Management and Hot-Spot Remediation	349
Avram Bar-Cohen and Peng Wang	
Some Aspects of Microchannel Heat Transfer	431
Y. Joshi, X. Wei, B. Dang, and K. Kota	
Nanoprobes for Live-Cell Gene Detection	479
Gang Bao, Won Jong Rhee, and Andrew Tsourkas	
Packaging for Bio-micro-electro-mechanical Systems (BioMEMS) and Microfluidic Chips	505
Edward S. Park, Jan Krajniak, and Hang Lu	
Packaging of Biomolecular and Chemical Microsensors	565
Peter J. Hesketh and Xiaohui Lin	
Nanobiosensing Electronics and Nanochemistry for Biosensor Packaging	613
Dasharatham G. Janagama and Rao R. Tummala	
Molecular Dynamics Applications in Packaging	665
Yao Li, Jeffrey A. Hinkley, and Karl I. Jacob	
Nanoscale Deformation and Strain Analysis by AFM/DIC Technique	695
Y.F. Sun and John H.L. Pang	
Nano-Scale and Atomistic-Scale Modeling of Advanced Materials	719
Ruo Li Dai, Wei-Hsin Liao, Chun-Te Lin, Kuo-Ning Chiang, and Shi-Wei Ricky Lee	
Index	759

Contributors

Gang Bao Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, USA, gang.bao@bme.gatech.edu

Avram Bar-Cohen Department of Mechanical Engineering, University of Maryland, College Park, MD, USA, abc@umd.edu

Kuo-Ning Chiang Department of Power Mechanical Engineering, National Tsing Hua University, Hsinchu, Taiwan, knchiang@pme.nthu.edu.tw

Jaephil Cho School of Energy Engineering, Ulsan National Institute of Science & Technology, Ulsan, Korea, jpcho@unist.ac.kr

Ruo Li Dai Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, ruoli.dai@gmail.com

B. Dang IBM Corporation, Armonk, NY, USA, dangbing@us.ibm.com

Jan Felba Institute of Microsystem Technology, Wrocław University of Technology, Wrocław, Poland, jan.felba@pwr.wroc.pl

Claire Gu Department of Electrical Engineering, University of California, Santa Cruz, CA, USA, claire@soe.ucsc.edu

Peter J. Hesketh School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA, peter.hesketh@me.gatech.edu

Jeffrey A. Hinkley Langley Research Center, NASA, Hampton, VA, USA, jeffrey.a.hinkley@nasa.gov

Andrew Hunt nGimat Co., Atlanta, GA, USA, ahunt@ngimat.com

Karl I. Jacob Department of Polymer, Textile and Fiber Engineering, Georgia Institute of Technology, Atlanta, GA, USA, jacob@gatech.edu

Dasharatham G. Janagama Package Research Center (PRC), Georgia Institute of Technology, Atlanta, GA, USA, jdgoud@hotmail.com

Hongjin Jiang Intel Corporation, Chandler, AZ, USA, hongjinjiang@gmail.com

Yongdong Jiang nGimat Co., Atlanta, GA, USA, yjiang@ngimat.com

Ranjit John College of Engineering, University of Arkansas, Fayetteville, AR, USA, rjohn@uark.edu

Y. Joshi School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA, yogendra.Joshi@me.gatech.edu

K. Kota School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA, krishna.kota@me.gatech.edu

Jan Krajniak School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA, USA, jan.krajniak@chbe.gatech.edu

Shi-Wei Ricky Lee Department of Mechanical Engineering, Hong Kong Institute of Science and Technology, Kowloon, Hong Kong, rickylee@ust.hk

Yao Li Department of Polymer, Textile and Fiber Engineering, Georgia Institute of Technology, Atlanta, GA, USA, yli6@gatech.com

Yi Li Intel Corporation, Chandler, AZ, USA, yi.li@intel.com

Wei-Hsin Liao Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, NT, Hong Kong, whliao@cuhk.edu.hk

Chun-Te Lin Department of Power Mechanical Engineering, National Tsing Hua University, Hsinchu, Taiwan, d883783@oz.nthu.edu.tw

Wei Lin School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA, wlin31@gatech.edu

Xiaohui Lin School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA, xlin3@mail.gatech.edu

Hang Lu School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA, USA, hang.lu@chbe.gatech.edu

Ajay P. Malshe College of Engineering, University of Arkansas, Fayetteville, AR, USA, apm2@uark.edu

Kyoung-sik (Jack) Moon School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA, jack.moon@gatech.edu

John H.L. Pang School of Mechanical & Aerospace Engineering, Nanyang Technological University, Singapore, Singapore, mhlpang@ntu.edu.sg

Edward S. Park School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA, USA, edward.park@chbe.gatech.edu

Won Jong Rhee Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, USA, wjrhee93@bme.gatech.com

Ephraim Suhir Department of Electrical Engineering, University of California, Santa Cruz, CA, USA, Bell Laboratories, Physical Sciences and Engineering

Research Division, Murray Hill, NJ, USA (ret); Department of Mechanical Engineering, University of Maryland, College Park, MD, USA; and ERS Co., Los Altos, CA, USA, suhire@aol.com

Y.F. Sun School of Mechanical & Aerospace Engineering, Nanyang Technological University, Singapore, Singapore, sunyaofeng@pmail.ntu.edu.sg

Andrew Tsourkas Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA, atsourk@seas.upenn.edu

Rao R. Tummala Package Research Center (PRC), Georgia Institute of Technology, Atlanta, GA, USA, rao.tummala@ee.gatech.edu

Ganesh Venugopal nGimat Co., Atlanta, GA, USA, gvenugopal@ngimat.com

Peng Wang Department of Mechanical Engineering, University of Maryland, College Park, MD, USA, wangp@umd.edu

Xudong Wang School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA, xudong@engr.wisc.edu

Zhong Lin Wang School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA, zhong.wang@mse.gatech.edu

X. Wei IBM Corporation, Armonk, NY, USA, xwei@us.ibm.com

C. P. Wong School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA, cp.wong@mse.gatech.edu

Yonghao Xiu School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA, yxiu@gatech.edu

Yi Zhang Department of Electrical Engineering, University of California, Santa Cruz, CA, USA, lyrixzhang@hotmail.com

Zhiyong Zhao nGimat Co., Atlanta, GA, USA, zzhao@ngimat.com

Nanomaterials for Microelectronic and Bio-packaging

C.P. Wong and Kyoung-sik (Jack) Moon

Abstract This chapter addresses the state-of-art nanoscience and technologies related to the next generation high-density microelectronics and bio-packaging applications, including carbon nanotubes (CNTs) for electrical/thermal devices, lead-free nanoalloys for lead-free interconnection, nano-conductive adhesives, molecular wires for electrical interconnects, low-stress and high thermal conductive flip-chip underfills, high- k dielectric for embedded passives, bio-mimic Lotus effect with both nano-micro surfaces for self-cleaning and molecular dynamic (MD) simulations for nanomaterial study and prediction, etc.

Keywords Nanomaterials · Carbon nanotubes interfaces · Electrical and thermal interconnection · Solar cell packaging · Nanoparticle synthesis · Melting point depression · Nano-embedded capacitors · Nanocomposites · Molecular dynamic simulation

1 Introduction

Microelectronics, photonics, and bio-packaging technology trends have been always relying on increasing the packaging density, resulting from meeting consumer demands for light-weight, compact, highly reliable and multifunctional electronic, biomedical, or communication devices. To address such demanding ultra-high-density packaging technology, nanotechnology, nanoscience, and bioengineering play an important role in this arena. For over a decade, nano/biotechnologies such as carbon nanotubes (CNTs), nanoparticles, molecular self-assembly, nanoimprints, and sensors have been explored and studied for mechanical, physical, chemical, photonic, electronic, and biological properties. The packaging scientists prospect

C.P. Wong (✉)

School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: cp.wong@mse.gatech.edu

beyond nanomaterials or devices, aiming at system-level components (chip to chip or chip to package interconnection, thin-film components, dielectrics, passives, encapsulants, coatings, power sources, etc.) leading to nano/biomodules. This emphasis is expected to lead to a number of commercial applications during the next decades. However, the implementation of nanotechnology in the system-level applications requires a thorough understanding of nano/bioscience and engineering.

In this chapter, several fundamental researches related to nanotechnologies for microelectronics, photonics, and bio-packaging are discussed, including carbon nanotubes (CNTs) for electrical/thermal interconnects, nanosolder materials with melting point depression, molecular wires for high-performance electrical interconnects, high- k nanocomposites, high thermally conductive nanocomposites, and bio-mimic self-cleaning nanosurfaces.

1.1 Interfaces of Carbon Nanotubes (CNTs)/Substrates for Electrical and Thermal Interconnects

Carbon nanotubes (CNTs) have attracted much interest due to their extraordinary electrical, thermal, and mechanical properties, and their wide range of potential applications [1–4]. The quality control of the CNTs grown, e.g., purity (catalyst residual and amorphous carbon), diameter, length, and defects, directly affect the final properties of CNTs. Unfortunately, the results reported so far vary much and reproducible growth of high-quality CNTs is still an issue. Aligned CNT (ACNT) structures make the best use of the 1-D transport properties of individual CNTs and have been proposed as the promising candidate for thermal interface materials (TIM) and electrical interconnect materials, gas sensors and electrodes, etc. Chapter 4 will address the carbon materials in more detail.

However, there are challenges in the chemical vapor deposition (CVD) for ACNT fabrication, including its high growth temperature and few suitable substrate materials for the CVD. To address these key issues, three low-temperature ACNT transfer techniques to position ACNTs onto various surfaces are reported by Georgia Tech team, as shown in Fig. 1, which use in-situ open-ended and surface functionalized, aligned multiwall-CNTs via (1) solder, (2) polymer adhesives, and (3) self-assembled monolayer as transfer media [4, 5]. Introducing a trace amount of oxidants during the CVD process can eliminate amorphous carbons on CNT surfaces, open the CNT tips, and controllably functionalize the tips and walls. Consequently, the CNTs are of high purity and superior electrical conductivity. A number of research papers on CNTs (recently on graphene – a monolayer of graphite, another emerging carbon material) have been published in the past decade resulting in tremendous amount of findings on their material properties, structures, fabrications, manipulation, and applications. Nevertheless, due to a lack of knowledge on the defect control and their interface manipulation, there are still barriers in using CNTs and graphenes in

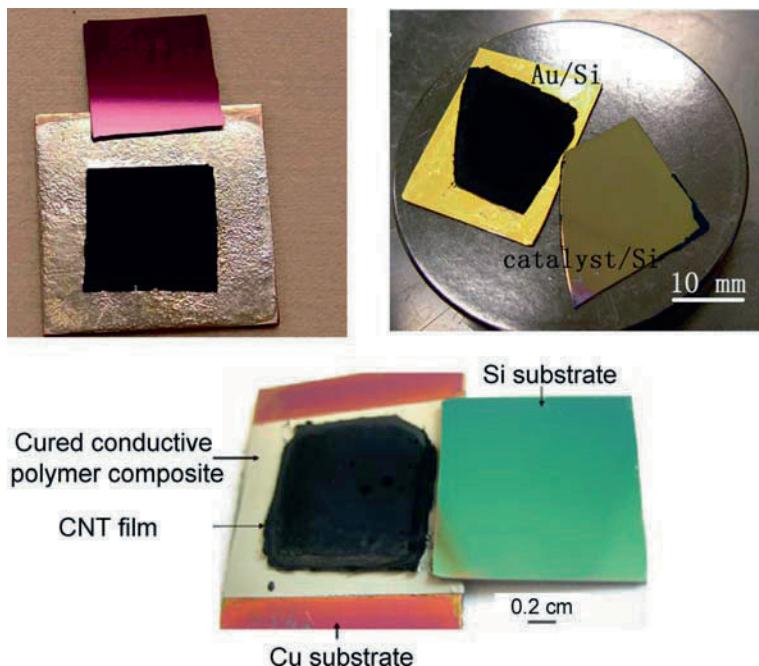


Fig. 1 Low-temperature ACNT transfer technology: solder transfer, adhesive transfer, and chemical transfer (from left to right in sequence)

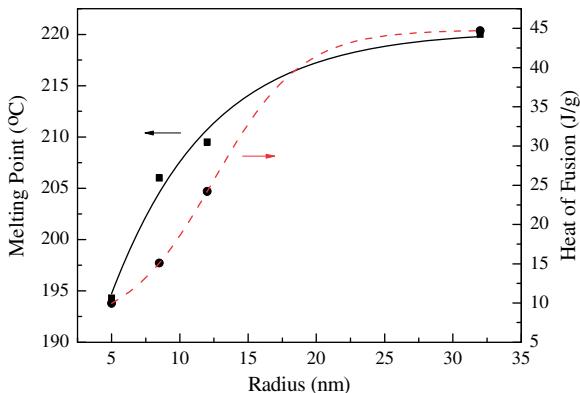
device-level applications. Therefore, more understanding on defects and interfaces should be made in the near future in order to fully utilize their exotic material properties.

1.2 Tin/Silver Alloy Nanoparticle Pastes for Low-Temperature Lead-Free Interconnect Applications

Tin/silver (96.5Sn3.5Ag) is one of the promising lead-free alternatives for a Sn/Pb solder that has been a de facto standard bonding material in microelectronics packaging [6]. However, the melting point (T_m) of 96.5Sn3.5Ag alloy (221 C) is around 40°C higher than that of eutectic Sn/Pb solders. The higher T_m requires the higher reflow temperature in the electronics manufacturing process, which has adverse effects not only on energy consumption, but also on the package reliability. Therefore, studies on lowering the melting point of lead-free metals have been drawing much attention of electronic packaging engineers.

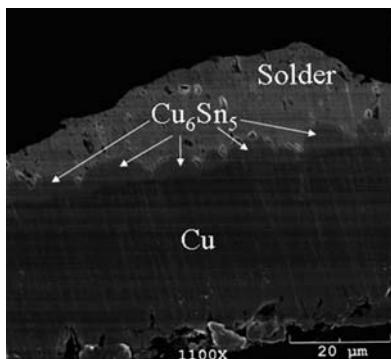
The Georgia Tech team has reported that different sized SnAg alloy nanoparticles were successfully synthesized by the chemical reduction method. An organic

Fig. 2 Relationship between the radius of the as-synthesized SnAg alloy nanoparticles and their corresponding melting points and heat of fusion, respectively



surfactant was used to not only prevent aggregation, but also protect the SnAg alloy nanoparticles from oxidation. Both the particle size-dependent melting point depression and latent heat of fusion have been observed (Fig. 2). It has already been found that surface melting of small particles occurs in a continuous manner over a broad temperature range, whereas the homogeneous melting of the solid core occurs abruptly at the critical temperature T_m [7]. For smaller sized metal nanoparticles, the surface melting is strongly enhanced by curvature effects. Therefore with the decreasing of particle size, both the melting point and latent heat of fusion will decrease. Among all the synthesized particles, the 5 nm (radius) SnAg alloy nanoparticles have a melting point at $\sim 194.3^\circ\text{C}$, which will be a good candidate for the low melting temperature lead-free solders and can solve the issues from the high-temperature reflow for the micron-sized lead-free particles. The wetting properties of the (32 nm in radius) SnAg alloy nanoparticle pastes on the cleaned copper surface were studied in a 230°C oven with an air atmosphere. The cross-section of the sample after reflow is shown in Fig. 3. It was observed that the SnAg

Fig. 3 SEM image of the cross-section of the wetted SnAg alloy nanoparticles (32 nm in radius) on the cleaned copper foil



alloy nanoparticles completely melted and wetted on the cleaned copper foil surface. The energy-dispersive spectroscopy (EDS) results revealed the formation of the intermetallic compound (Cu_6Sn_5), which showed scallop-like morphologies in Fig. 2.

This study has demonstrated the feasibility of the SnAg alloy nanoparticle pastes as a candidate for the low processing temperature lead-free interconnect applications.

The melting point can be dramatically decreased when the size of substances is reduced to nanometer size [8]. To date, the size-dependent melting behavior has been proved both theoretically and experimentally [9–11]. The high ratio of the surface area to volume of nanoparticles has been known as one of the driving forces for the size-dependent melting point depression, which is employed currently in various areas, in particular, where low-temperature bonding, high thermal or electrical conductivity is required. Chapter 8 addresses this as well.

1.3 Enhanced Electrical Properties of Anisotropically Conductive Adhesive with π -Conjugated Molecular Wire Junctions for Enhanced Electrical Properties

As microelectronics requirements are driven toward smaller, higher density, and more cost-effective solutions, anisotropically conductive adhesives (ACAs) are widely used in electronics industry [12, 13]. The ACA joint is established by trapping conductive particles dispersed in the polymer matrix in between integrated circuit and conductive pads when heat and pressure are applied. ACAs can provide electrical conductivity in Z-direction as well a mechanical interconnect between integrated circuit and conductive pads. The most significant advantage of ACAs is fine pitch capability. However, the ACA joints have lower electrical conductivity and poor current-carrying capability due to the restricted contact area and poor interface between conductive particles and IC/conductive pads, compared to the metallurgical joint of the metal solders. In order to enhance the electrical performance of ACA joints, the Georgia Tech team has introduced π -conjugated self-assembled molecular wire junctions to improve the electrical properties of anisotropically conductive adhesives (ACAs) [14] and nonconductive films (NCFs) [15] for the electronic interconnects in previous work. Crucial for successful applications of self-assembly monolayers (SAMs) in ACAs and NCFs are interfacial properties (chemisorption or physisorption), the characteristics of the molecules (conjugated or unconjugated, thermal stability, and so on), surface binding geometry, and packing density. Therefore, current work investigates the effects of those factors on the performance of ACAs in detail by using a series of molecules with different functional groups. The electrical properties of an anisotropic conductive adhesive (ACA) joint were investigated using a submicron-sized (~ 500 nm in diameter) silver (Ag) particle as conductive filler with the effect of π -conjugated self-assembled molecular wires (a)). The ACAs with submicron-sized Ag particles have higher current-carrying

capability (~ 3400 mA) than those with micro-sized Au-coated polymer particles (~ 2000 mA) and Ag nanoparticles (~ 2500 mA). More importantly, by construction of π -conjugated self-assembled molecular wire junctions between conductive particles and integrated circuit (IC)/substrate, the electrical conductivity has increased by one order of magnitude and the current-carrying capability of ACAs has improved by 600 mA (Fig. 4).

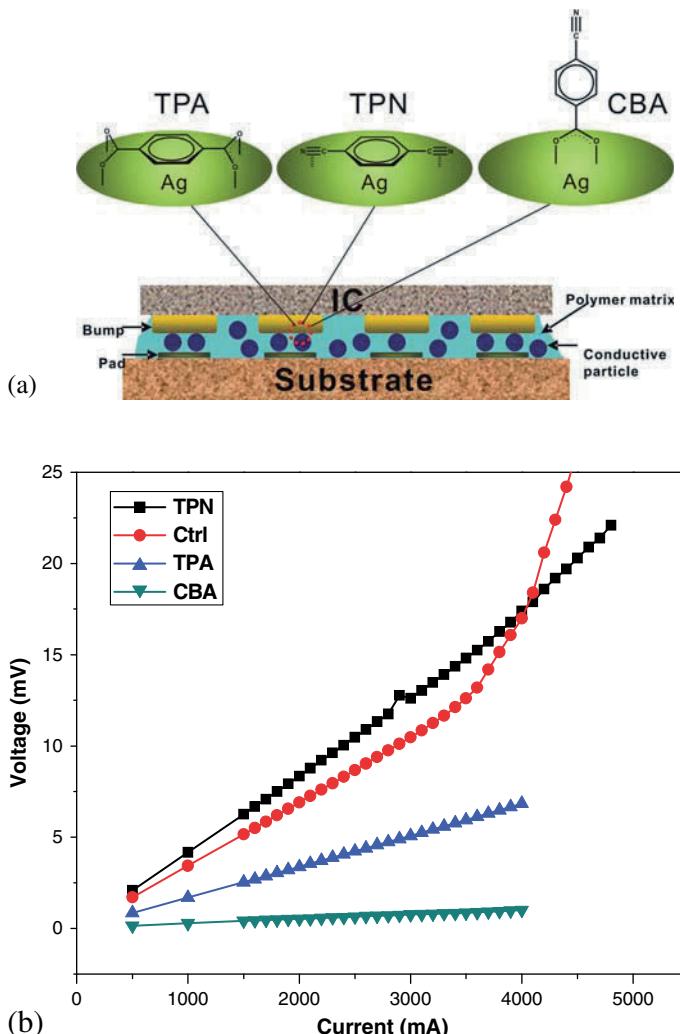


Fig. 4 (a) Cross-section of ACA joints with π -conjugated molecular wire junctions; (b) I - V curve of ACA filled with Ag particles (Ctrl), TPN-treated Ag particles, TPA-treated Ag particles, and CBA-treated particles

1.4 Low-Stress and High Thermal Conductive Underfill for Cu/Low-k Application

The underfill materials have been used to improve the flip-chip device reliability by effectively distributing thermal stress of the package evenly to all solder joints/balls versus concentrating on the farthest to the distance neutral point (DNP), resulting in dramatically enhancing solder joint fatigue life [16–18]. In addition to using low modulus and CTE underfills for reducing the stress, researches also show the thermally conductive underfills can enhance the reliability as well, because of its thermal dissipative capability [18]. Therefore, the high thermal conductive underfill material is expected to play an important role to improve the Cu/low-*k* device reliability, as the thermal management is becoming the bottleneck of electronic packaging technology due to the reduced device profile, increased package density (e.g., 3D packaging), and application of porous materials in Cu/low-*k* device. However, there is no underfill material available, which can absorb such high thermal stress and dissipate heat to protect the next generation Cu/low or ultra-low-*k* and high-power devices. Consequently, a novel high-performance underfill/polymer composite material is required for the highly reliable next generation fine pitch Cu/low-*k* devices.

Optimizing the thermal conductivity and mechanical properties of these high performance underfill materials by surface modification of high thermal conductive inorganic filler has been focused. SiC particles are high thermally conductive but have relatively low amounts of the hydroxyl group on the particle surfaces, hence the difficulty to be involved in the surface reaction and poor compatibility with the epoxy resin. The SiC particle surfaces were thermally coated with a nanolayer of SiO₂ by oxidizing Si atoms on the surface at high temperature, as shown in Fig. 5a. Further surface modification of SiC by γ -glycidoxypropyl-trimethoxysilane (GOP) (Fig. 5b) was applied to graft epoxide groups onto SiC surface. The surface modification improves the chemical bonding density between filler and polymer matrix and reduces the surface phonon scattering, resulting in the thermal conductivity improvement in epoxy/SiC composites. Moreover, results showed that the added silane improved the adhesion strength between the underfill and the silicon wafer. Also, mechanical property measurements indicated the relatively low moduli and low coefficient of thermal expansion, which will help to reduce the global thermal stress of the underfilled flip-chip structure.

1.5 High-Dielectric Constant (*k*) Polymer Nanocomposites for Embedded Capacitor Applications

Embedded passives technology is one of the key emerging techniques for realizing the system integration such as system-in-package (SiP), system-on-package (SOP), which offers various advantages over traditional discrete components, including higher component density, increased functionality, improved electrical performance,

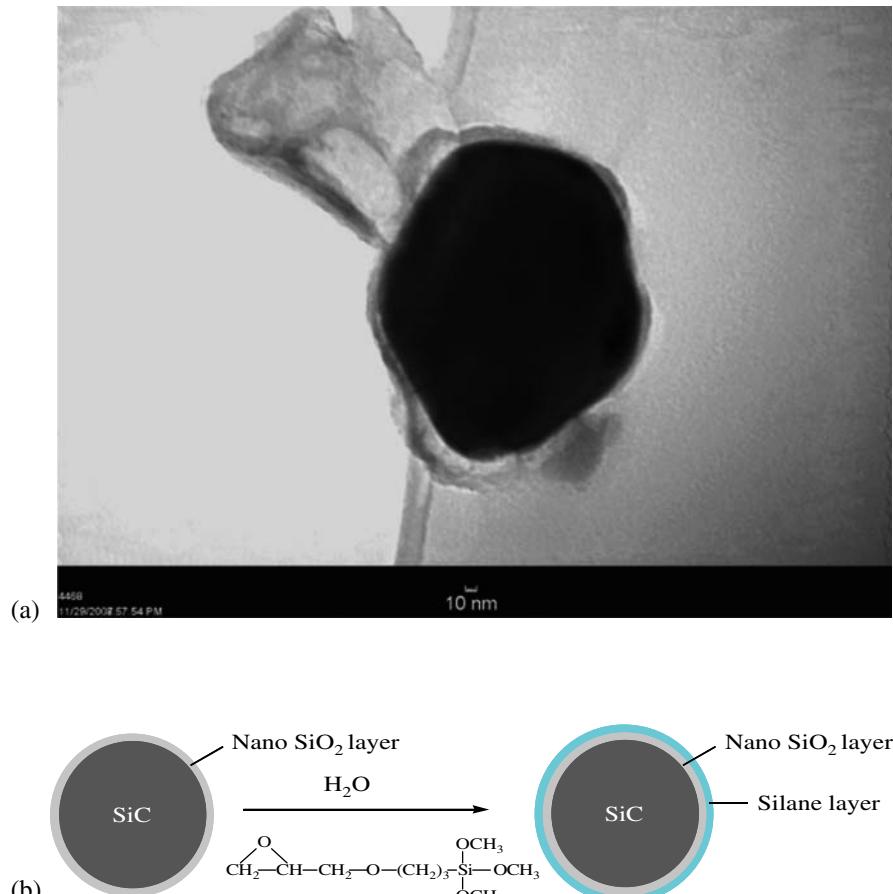


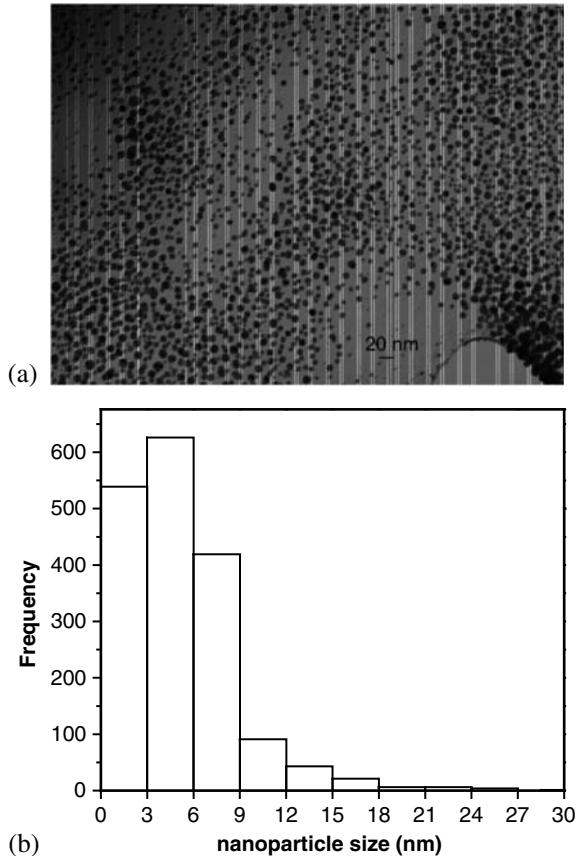
Fig. 5 (a) TEM image of a SiC particle thermally coated by a nanolayer SiO_2 ; (b) surface modification of thermally coated SiC particle by silane

increased design flexibility, improved reliability, and reduced unit cost. Novel materials for embedded capacitor applications are in great demand, for which high dielectric constant (k), low dielectric loss, and process compatibility with printed circuit boards (PCBs) are the most important material prerequisites. Chapter 9 covers the embedded component packaging in more detail.

Dielectric nanocomposites based on nanoparticles with controlled parameters were designed and developed to fulfill the balance between sufficiently high- k and low dielectric loss and satisfy the requirements to be a feasible option for embedded capacitor applications [19]. To this end, people have explored the use of material and process innovations to reduce the dielectric loss sufficiently while maintain the high- k of the nanocomposites [20, 21]. They take advantage of the high- k polymer matrix and the inter-particle barrier layer formed intrinsically by an insulating shell

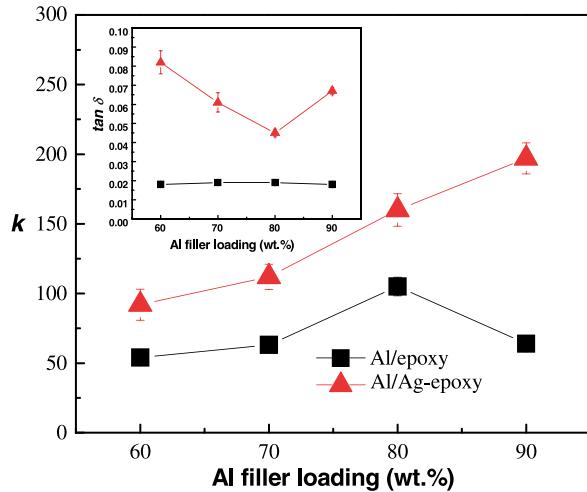
around the conducting metal core to achieve a simultaneously high- k and low dielectric loss conductive filler/polymer composite. First, a silver (Ag)-epoxy mixture was prepared via an in situ photochemical method, in which Ag nanoparticles were generated by chemical reduction of a metallic precursor upon UV irradiation. Figure 6 displays TEM micrograph and histogram of Ag nanoparticles synthesized via an in situ photochemical reduction in epoxy resin. The particle size analysis results show that well-dispersed Ag nanoparticles with the size mostly smaller than 15 nm in polymer matrix were obtained. The average size was calculated as 5.2 nm [22].

Fig. 6 (a) TEM micrograph and (b) histogram of Ag nanoparticles synthesized by an in situ photochemical reduction in epoxy resin



The as-prepared Ag-epoxy mixture was then utilized as a high- k polymer matrix to host self-passivated aluminum (Al) fillers for preparation of Al/Ag-epoxy composites as embedded capacitor candidate materials. Figure 7 shows the dielectric properties of Al/epoxy composites and Al/Ag-epoxy composites as a function of Al filler loading at a frequency of 10 kHz. The Al/Ag-epoxy composites exhibit more than 50% increase in k values compared with Al/epoxy composites with the same filler loading of Al, which demonstrated that the incorporation of well-dispersed Ag

Fig. 7 Dielectric constant (k) values of Al/epoxy and Al/Ag-epoxy composites as a function of Al filler loading (at 10 kHz). Dielectric loss tangent ($\tan \delta$) values of Al/epoxy and Al/Ag-epoxy composites are displayed in the inset



nano particles in the polymer matrix delivers a huge impact on the dielectric constant enhancement. It is also notable that the dielectric loss tangent values (see the inset of Fig. 7) for Al/Ag-epoxy composites with different Al filler loadings are all below 0.1, which is tolerable for some applications such as decoupling capacitors. The low dielectric loss can be attributed to the good dispersion of Ag nanoparticles in the polymer matrix together with the thin self-passivated aluminum oxide layer forming an insulating boundary outside of the Al particles. The results suggest that the metal–polymer nanocomposites via an *in situ* photochemical approach can be employed as a high- k polymer matrix to enhance the k value of the composites while maintaining the relatively low dielectric loss tangent.

The trends of the dielectric properties with the variation of filler loadings behaved differently for Al/epoxy composites and Al/Ag-epoxy composites. This phenomenon can be explained in terms of the origins and mechanisms of the dielectric loss of the heterogeneous conductive filler/polymer composite material. The concentration of silver nanoparticles in the polymer matrix plays a significant role in determining the dielectric properties of the high- k nanocomposites. At low Ag concentration, the dielectric behaviors of Al/Ag-epoxy composites are mainly determined by interfacial polarization, while conduction and electron transport of Ag dominate the Al/Ag-epoxy composites at higher Ag concentration [22].

1.6 Bio-mimetic Lotus Surface

The effect of surface structures on superhydrophobicity (SH) is of much interest due to the dependence of a structure on the attainment of a high water contact angle ($>150^\circ$) and reduced contact angle hysteresis ($<10^\circ$). Superhydrophobicity was first observed on lotus leaves where high water contact angle and low hysteresis cause water droplet that falls on the surface to bead up and roll off the surface,

thereby leading to water repellency and self-cleaning characteristics [23–25]. Such surface properties are also critical for antistiction in microelectro-mechanical system (MEMS) [26], friction reduction [27], and anticorrosion [28] applications. In addition, superhydrophobic surfaces offer much promise for the formation of high-performance micro/nanostructured surfaces with multifunctionality that can be used in optoelectronic [29, 30], photoelectric [31], microelectronic, catalytic, and biomedical applications. In particular, its antireflection application in photovoltaic (PV) solar cell devices for enhancing the efficiency has been considered very important since a PV cell is a critical component for renewable energy developments. Figure 8 shows the typical PV module package structures where incident sunlight transmits through a cover glass and a polymer layer and arrives at the top surface of a silicon cell. Light reflection at the silicon top surface can significantly reduce the cell efficiency.

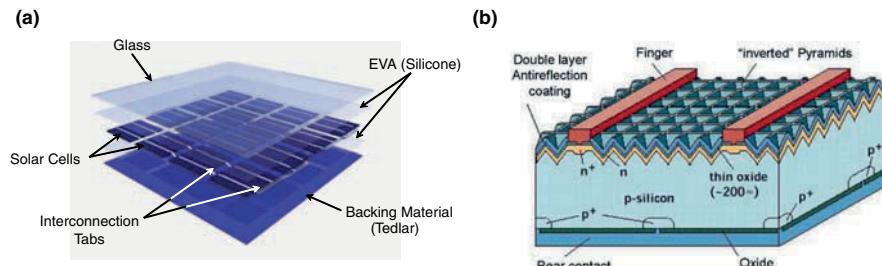


Fig. 8 (a) General structure of a PV module; (b) structure of passivated emitter, real locally diffused (PERL) solar cell with near theoretical efficiency

The Georgia Tech team used a metal-assisted etching technique to form nanoscale roughness and thereby form hierarchical structures by metal-assisted etching of micron-size pyramid-textured surfaces (Fig. 9). These structures yield superhydrophobic surfaces, which may have important application in light harvesting, antireflection, and light-trapping properties of solar cells.

A SiN coating on the pyramid-textured silicon surface has been employed to reduce the reflection. Recently, a nanosurface with two-tier roughness was found to provide not only superhydrophobicity with self-cleaning capability but also very low light reflectivity on a silicon cell even without SiN coating [32]. Because of the presence of surface nanostructures, the surface absorbs most of the incident light, thus reducing reflection, especially in the 300–1000 nm wavelength regime. The nanotextured surface may also increase the path length of light as it travels through the cell, which allows thinner solar cells with reduced cost; furthermore, the surface may trap the weakly absorbed light reflected from the back surface by total internal reflection at the front surface/air interface. On the micro/nanostructured two-scale surfaces (Fig. 10), the weighted light reflectance is further reduced to 3.8%, as shown in Table 1. More studies on nanostructure surfaces are introduced in Chapter 3.

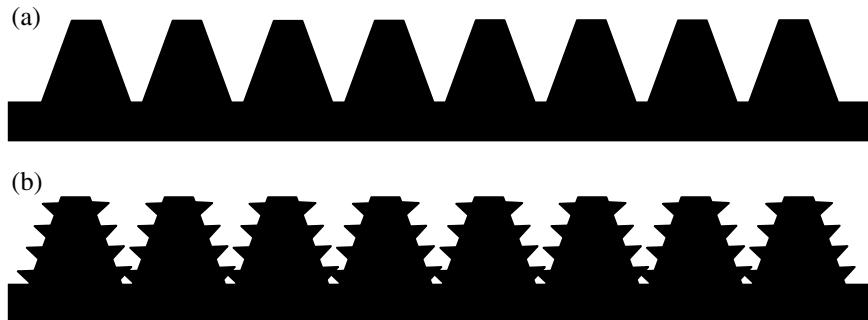


Fig. 9 Schematics of two-tier surface roughness fabrication: (a) only micron-sized roughness and (b) additional nano-sized surface roughness

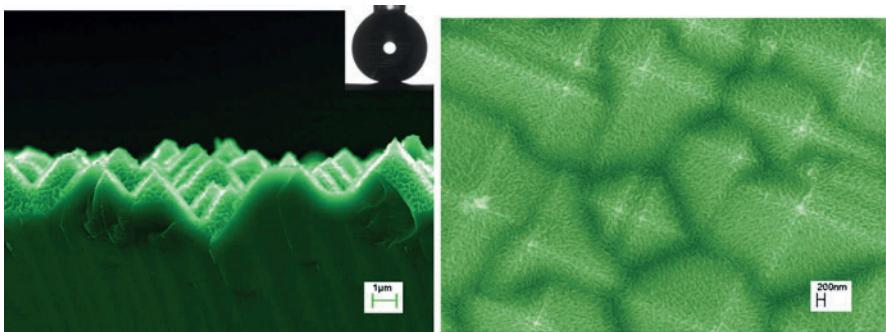


Fig. 10 (a) Cross section and (b) top view of Si pyramid surfaces etched made by Georgia Tech team

Table 1 Weighted reflectance on different textured surfaces

Sample	Weighted reflectance %
Flat Si surface	37.3
Pyramid-textured surface	12.3
Nanotextured surface	6.4
Two-scale textured surface	3.8

1.7 Molecular Dynamic (MD) Simulations in Nanomaterial Study

Although a variety of nanomaterials have been studied and considered in many applications, a number of fundamental questions on their individual behavior have been asked, such as thermal behavior during cooling or heating, how they react with different organic molecules in specific solvents, and thermal transport at complicated interfaces, because the deep understanding of their individual behavior

enables scientists and engineers to manipulate the nanomaterials appropriately and utilize their properties fully. To answer those questions many researchers have been reporting their experimental results, which have been obtained and analyzed by sophisticated equipments. However, those experimental approaches still have technical limitations such as characterization resolutions, noise, and reproducibility.

In particular, for thermal behavior, various equipments such as transmission electron microscopy and electron diffraction [33, 34], x-ray diffraction (XRD) [35], and differential scanning calorimetry (DSC) [36, 37] have been applied to monitor the melting behavior of the nano-sized materials. At the melting point, there are changes in the diffraction pattern associated with the disordering of the structure, thus XRD and electron diffraction have been used to investigate the melting of the nanomaterials. Since melting is associated with the heat of a system, DSC is also used to directly measure the thermal properties of the nanomaterials. However, experimental techniques are normally incapable of resolving the melting temperature of a single nanoparticle, and the distribution in size of nanoparticles often obscures the physics behind melting. Also, experimental measurements have another limitation, i.e., the melting of particles is influenced by the type of the substrate on which the nanoparticles are placed. Moreover, experiments cannot give detailed information on the atomic-level behavior at elevated temperatures since direct observations are extremely difficult. As such, experimental methods have some technical limitations when studying the melting behavior of nanoparticles.

Theoretical models based on the kinetics, thermodynamics, and mechanical properties of solids that are typically bulk crystals have been employed. Those approaches also have limitations since they do not have the information about the individual atoms comprising the nanoparticle [38]. For the nano-scale substances, a more atomistic approach is necessary. In particular, molecular dynamic (MD) simulation is an effective way of investigating the coalescence or melting behavior at an atomic level. It is generally more difficult and more expensive to study the properties of a microstructure experimentally than to perform the corresponding simulation. As such, the MD simulation technique has been applied by several researchers to explore the melting behavior of nanoparticles [39–41], including the melting point depression phenomenon, and the role of extrinsic defects (grain boundaries and free surfaces) in the melting process. It was found that disorder and the melting transition were initiated at the interface in the simulations. More details on MD simulation will be shown in Chapter 18.

The thermal behavior and interaction of single elements or alloy nanoparticles on certain substrates [42–48] are very interesting topics because various applications have already been employing nanomaterials in their systems to improve electrical and thermal properties. For instance, how the individual nanoparticles behave or interact with adjacent particles during heating and cooling as a function of time can be beautifully visualized as shown in Fig. 11.

The modified embedded atom method (MEAM) is popularly used in conjunction with MD simulations to investigate the effect of different particle sizes and temperature ramp-up rates on melting behavior of nanoparticles. The MEAM is an empirical

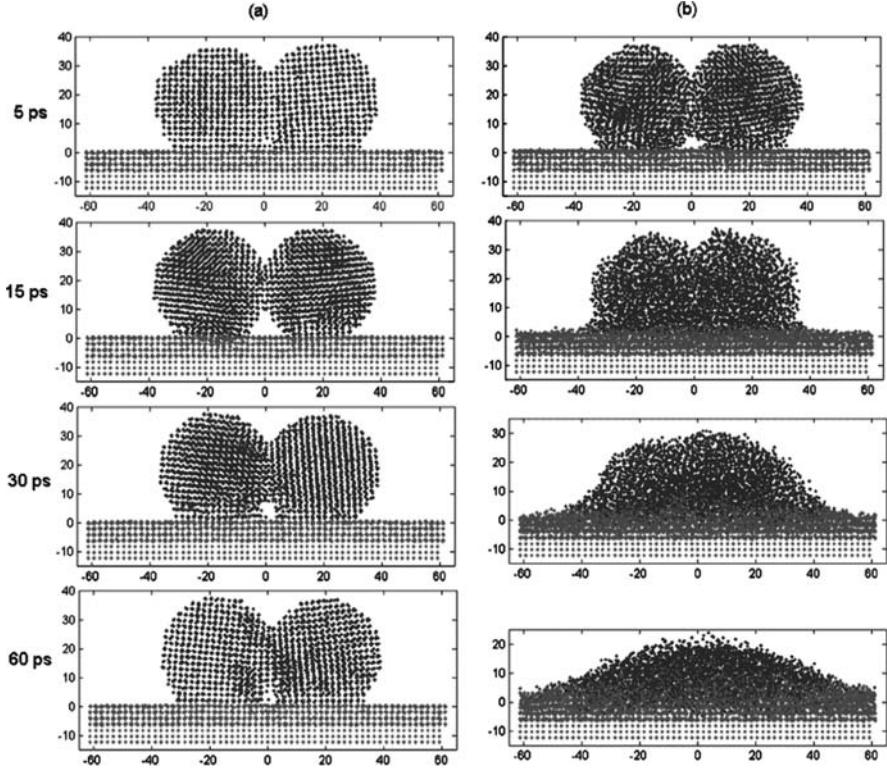


Fig. 11 The x - z plane projection of the system with simulation temperatures of (a) 400 K and (b) 1000 K, and simulation time of 5, 15, 30, and 60 ps, respectively [44]

extension of the embedded atom method (EAM), which can be applied to both metallic and covalent bonding by including angular forces [49, 50]. In those modeling studies, it is not simple to define what temperature will be the melting point of particles and a heating rate is another of this complication. Thus, the evolution of potential energy with respect to temperature is used to obtain the melting temperature of nanoparticles and the two-phase equilibrium method is used to identify the melting points of nanoparticles with different sizes [46]. Also the pair correlation function (PCF) is used to monitor the crystal regularity of particles and study the structure differences of shells and cores of a particle [51]. Comparisons of these researches with thermodynamic models [52] showed that the thermodynamic models were not able to correctly predict the melting points of nanoparticles with very small sizes.

Figure 12 demonstrated that atoms located in the outer shell of the particle lost their structural order before the atoms in core region did. This result confirmed the surface melting behavior of nanoparticles of this size. The MD simulation has been proved as a very powerful tool for exploring nanomaterial systems at atomistic scale,

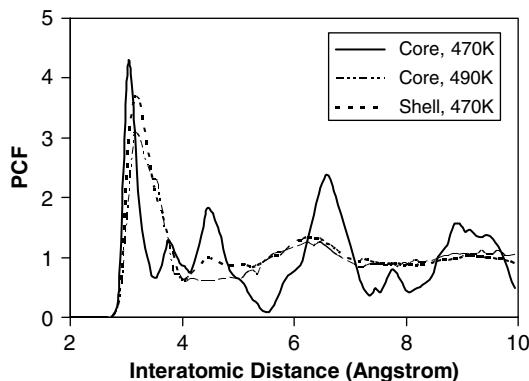


Fig. 12 PCF of a Sn particle; heating rate 0.1 K/ps, diameter 8 nm, shell width is 1 nm, core radius is 3 nm [46]

by itself as well as combined with other simulation methods and has been demonstrating its strong potential in this arena. We can expect that a numerous current and future questions on such a small-scale world could be answered through the powerful computational studies together with sophisticated experimental methods.

2 Conclusive Remarks

While a variety of researches on nanomaterials for nano-bio devices and systems have been carried out, this field is still at its infancy. However, there are many potential applications in electronic, photonic, and bio-medical applications, which will benefit our society. Therefore, much research and development work need to be studied. This is an area where scientists and engineers need to work together to make this goal a reality.

References

1. Zhu L.B. and Wong C.P., Well-aligned open-ended carbon nanotube architectures: an approach for device assembly. *Nano Lett.* 2006; 6: 243.
2. Zhu L.B. and Wong C.P., Aligned carbon nanotube stacks by water-assisted selective etching. *Nano Lett.* 2005; 5: 2641.
3. Jiang H.J., Moon K., and Wong C.P., The preparation of stable metal nanoparticles on carbon nanotubes whose surfaces were modified during production. *Carbon* 2007; 45: 655.
4. Jiang H.J., Moon K., and Wong C.P., Low temperature carbon nanotube film transfer via conductive polymer composites. *Nanotechnology* 2007; 18: 125203.
5. Lin W., Xiu Y., Jiang H., Zhang R., Hildreth O., Moon K., and Wong C.P., Self-assembled monolayer-assisted chemical transfer of in situ functionalized carbon nanotubes. *J. Am. Chem. Soc.* 2008; 30: 9636–9637.
6. Artaki I., Jackson A.M., and Vianco P.T., Evaluation of lead-free solder joints in electronic assemblies. *J. Electron. Mater.* 1994; 23(8): 757.

7. Hu W.Y., Xiao S.G., Yang J.Y., and Zhang Z., Melting evolution and diffusion behavior of vanadium nanoparticles. *Eur. Phys. J. B* 2005; 45(4): 547.
8. Pawlow P.Z., The dependency of melting point on the surface energy of a solid body. (Supplement). *Phys. Chem.* 1909; 65(5): 545.
9. Lai S.L., Guo J.Y., Petrova V., Ramanath G., and Allen L.H., Size-dependent melting properties of small tin particles: nanocalorimetric measurements. *Phys. Rev. Lett.* 1996; 77(1): 99.
10. Bachels T., Guntherodt H.J., and Schafer R., Melting of isolated tin nanoparticles. *Phys. Rev. Lett.* 2000; 85(6): 1250.
11. Zhao S.J., Wang S.Q., Cheng D.Y., and Ye H.Q., Three distinctive melting mechanisms in isolated nanoparticles. *J. Phys. Chem. B* 2001; 105(51): 12857.
12. Li Y., Moon K., and Wong C.P., Electronics without lead. *Science* 2005; 308(141): 9.
13. Li Y. and Wong C.P., Recent advances of conductive adhesives as a lead-free alternative in electronic packaging: materials, processing, reliability and applications. *Mater. Sci. Eng. R Rep.* 2006; R51: 1.
14. Li Y., Moon K., and Wong C.P., Adherence of self-assembled monolayers on gold and their effects for high-performance anisotropic conductive adhesives. *J. Electron. Mater.* 2005; 34: 266.
15. Li Y., Yim M., and Wong C.P., High performance nonconductive film with π -conjugated self-assembled molecular wires for fine pitch interconnect applications. *J. Electron. Mater.* 2007; 36: 549.
16. Tsukada Y., Surface laminar circuit packaging. Proceedings of the 42nd Electronic Components and Technology Conference, 1992:22.
17. Han B. and Guo Y., *J. Electron. Packag.* 1995; 117: 185.
18. Lee W., Han I.Y., Yu J., Kim S.J., and Byun K.Y., "Thermal characterization of thermally conductive underfill for a flip-chip package using novel temperature sensing techniques. *Thermochim. Acta* 2007; 455: 148–155.
19. Ulrich R.K. and Schaper L.W., *Integrated Passive Component Technology*. IEEE Press, Wiley-Interscience, Hoboken, NJ, 2003.
20. Lu J., Moon K., Xu J., and Wong C.P., Synthesis and dielectric properties of novel high-K polymer composites containing in-situ formed silver nanoparticles for embedded capacitor applications. *J. Mater. Chem.* 2006; 16(16): 1543.
21. Lu J. and Wong C.P., Tailored dielectric properties of high-k polymer composites via nanoparticle surface modification for embedded passives applications. IEEE Proceedings of the 57th Electronic Components and Technology Conference, Reno, NV, 2007; pp. 1033–1039.
22. Lu J., Moon K., and Wong C.P., Silver/polymer nanocomposite as a high-k polymer matrix for dielectric composites with improved dielectric performance. *J. Mater. Chem.* 2008; DOI: 10.1039/B807566B.
23. Neinhuis C. and Barthlott W., Characterization and distribution of water-repellent, self-cleaning plant surfaces. *Ann. Bot.* 1997; 79: 667.
24. Barthlott W. and Neinhuis C., The purity of sacred lotus or escape from contamination in biological surfaces. *Planta* 1997; 202: 1.
25. Neinhuis C., Koch K., and Barthlott W., Movement and regeneration of waxes through plant cuticles. *Planta* 2001; 213: 427.
26. Bhushan B. and Jung Y.C., Micro- and nanoscale characterization of hydrophobic and hydrophilic leaf surfaces. *Nanotechnology* 2006; 17: 2758.
27. Cottin-Bizonne C., Barrat J.-L., Bocquet L., and Charlaix E., Low-friction flows of liquid at nanopatterned interfaces. *Nat. Mater.* 2003; 2: 237.
28. Wang S.T., Feng L., and Jiang L., One-step solution-immersion process for the fabrication of stable bionic superhydrophobic surfaces. *Adv. Mater.* 2006; 18: 767.
29. Chattopadhyay S., Li X., and Bohn P.W., In-plane control of morphology and tunable photoluminescence in porous silicon produced by metal-assisted electroless chemical etching. *J. Appl. Phys.* 2002; 91: 6134.

30. Campbell P. and Green M.A., Light trapping properties of pyramidal textured surfaces. *J. Appl. Phys.* 1987; 62: 243.
31. Koynov S., Brandt M.S., and Stutzmann M., Black nonreflecting silicon surfaces for solar cells. *Appl. Phys. Lett.* 2006; 88: 203107.1–203107.3.
32. Xiu Y., Zhang S., Yelundur V., Rohatgi A., Hess D., and Wong C.P., Superhydrophobic and low light reflectivity silicon surfaces fabricated by hierarchical etching. *Langmuir* 2008; 24: 10421–10426.
33. Yasuda H., Mitsuishi K., and Mori H., Particle-size dependence of phase stability and amorphouslike phase formation in nanometer-sized Au-Sn alloy particles. *Phys. Rev. B* 2001; 64: 094101.
34. Lee J. and Mori H., Solid-liquid two-phase structures in isolated nanometer-sized alloy particles. *Phys. Rev. B* 2004; 70: 144105.
35. Shi F.G., Size dependent thermal vibrations and melting in nanocrystals. *J. Mater. Res.* 1994; 9: 1307.
36. Koch C.C., Jang J.S.C., and Gross S.S., The melting point depression of tin in mechanically milled tin and germanium powder mixtures. *J. Mater. Res.* 1989; 4: 557.
37. Zhang M., Yu M., Schietekat F., Olson E., Kwan A., Lai S., Wisleder T., Greene J., and Allen L., Size-dependent melting point depression of nanostructures: Nanocalorimetric measurements. *Phys. Rev. B* 2001; 62: 10548.
38. Zhao M., Zhou X.H., and Jiang Q., Comparison of different models for melting point change of metallic nanocrystals. *J. Mater. Res.* 2001; 16: 3304.
39. Phillipot S.R., Lutsko J.F., Wolf D., and Yip S., Molecular-dynamics study of lattice-defect-nucleated melting in silicon. *Phys. Rev. B* 1989; 40: 2831.
40. Nguyen T., Ho P.S., Kwok T., Nitta C., and Yip S., Grain-Boundary Melting Transition in an Atomistic Simulation Model. *Phys. Rev. Lett.* 1986; 57: 1919.
41. Zhao S.J., Wang S.Q., Zhang T.G., and Ye H.Q., A cluster-induced structural disorder and melting transition in grain-boundary of B2 NiAl: A molecular-dynamics simulation on parallel computer. *J. Phys. Condens. Matter* 2000; 12: L549.
42. Dong H., Zhang Z., and Wong C.P., Molecular dynamics study of a nano-particle joint for potential lead-free anisotropic conductive adhesives applications. *J. Adhes. Sci. Technol.* 2006; 19(2): 87–94.
43. Dong H., Moon K., and Wong C.P., Molecular dynamics study on coalescence of Cu nanoparticles and their deposition on Cu substrate. *J. Electron. Mater.* Nov. 2004; 33: 1326–1330.
44. Dong H., Moon K.S., and Wong C.P., Molecular dynamics study of nanosilver particles for low-temperature lead-free interconnect applications. *J. Electron. Mater.* 2005; 34(1): 40.
45. Dong H., Moon K., Jiang H., Wong C.P., and Hua F., Simulation study of nanoparticle melting behavior for lead free nano solder application. 11th International Symposium on Advanced Packaging Materials: Processes, Properties and Interface, 15–17 March 2006:107.
46. Dong H., Moon K.-S., Jiang H., Baskes M.I., Hua F., and Wong C.P., Computer simulation of nano tin melting behavior: Effect of particle size temperature ramping up rate. American Chemical Society National Meeting & Exposition, March 29, 2006 (Invited).
47. Alavi S., Molecular dynamics simulations of the melting of aluminum nanoparticles. *J. Phys. Chem. A* 2006; 110(4): 1518–1523.
48. Yang Z., Yang X., and Xu Z., Molecular dynamics simulation of the melting behavior of Pt–Au nanoparticles with core–shell structure. *J. Phys. Chem. C* 2008; 112(13): 4937–4947.
49. Baskes M.I., Modified embedded-atom potentials for cubic materials and impurities. *Phys. Rev. B* 1992; 46: 2727.
50. Ravelo R., Aguilar J., and Baskes M.I., Molecular-Dynamics Studies of Thin Films of Sn on Cu. *Mater. Res. Soc. Symp. Proc.* 1998; 492: 43.
51. Baskes M.I., Determination of modified embedded atom method parameters for nickel. *Mater. Chem. Phys.* 1997; 50: 152.
52. Hanszen K., Theoretische Untersuchungen über den Schmelzpunkt kleiner Kugelchen. *Z. Phys.* 1960; 157: 523–553.

Nano-conductive Adhesives for Nano-electronics Interconnection

Yi Li, Kyoung-sik (Jack) Moon, and C.P. Wong

Abstract With the phasing out of lead-bearing solders, electrically conductive adhesives (ECAs) have been identified as one of the environmentally friendly alternatives to tin/lead (Sn/Pb) solders in electronics packaging applications. In particular, with the requirements for fine-pitch and high-performance interconnects in advanced packaging, nanoconductive adhesives are becoming more and more important due to the special electrical, mechanical, optical, magnetic, and chemical properties that nano-sized materials can possess. There has been extensive research for the last few years on materials and process improvement of ECAs, as well as the advances of nanoconductive adhesives that contain nano-filler such as nanoparticles, nanowires, or carbon nanotubes and nanomonomolayer graphenes. In this chapter, recent research trends on electrically conductive adhesives (ECAs) and their related nanotechnologies are discussed, with the particular emphasis on the emerging nanotechnology, including materials development and characterizations, processing optimization, reliability improvement, and future challenges/opportunities identification. The state of the art on nanoisotropic/anisotropic conductive adhesives incorporated with nanosilver, carbon nanotubes, and nanonickel, and their recent studies on those for flexible nano/bioelectronics, transparent electrodes, and jettable processes are addressed in this chapter. Future studies on nanointerconnect materials are discussed as well.

Keywords Electrically conductive adhesives (ECAs) · Isotropically conductive adhesives (ICA) · Percolation threshold · Nanotechnology · Carbon nanotubes (CNT) · Self-assembled monolayer (SAM) · Thermo-compression bonding (TCB) · Sintering · Current-carrying capability · Anisotropic conductive adhesives (ACAs) · Fine-pitch interconnection · Nanoconductive fillers · Transparent electrodes

Y. Li (✉)
Intel Corporation, Chandler, AZ, USA
e-mail: yi.li@intel.com

1 Introduction

Miniaturization of electronic and photonic devices is a demanding trend in current and future technologies including consumer products, military, biomedical, and space applications. In order to achieve this miniaturization, new packaging technologies have emerged such as embedded active and passive components and 3D packaging, where new packaging designs and materials are required.

Fine-pitch interconnection is one of the most important technologies for the device miniaturizations. Metal solders including eutectic tin/lead solder and lead-free alloys are the most common interconnect materials in the electronic/photonic packaging areas.

Conventional eutectic tin/lead solder or lead-free solders cannot meet the requirements for fine-pitch assembly due to their stencil printing resolution limit and bridging issues of solders in fine-pitch bonding, which is an intrinsic characteristic of metal solders.

Recently, polymer-based nanomaterials have been intensively studied to replace the metal-based interconnect materials, and many research efforts to enhance material properties of the electrically conductive adhesives (ECAs) are on going. Furthermore, incorporating nanotechnology into the ECA systems not only enables the ultra-fine-pitch capability but also enhances the electrical properties such as reducing the percolation threshold and improving the electrical conductivity.

In this chapter, recent research and study trends on ECAs and their related nanotechnologies are discussed, and the future of the nanomaterial-based polymer composites for fine-pitch interconnection is reviewed.

Electrically conductive adhesives (ECAs) are composites of polymeric matrices and electrically conductive fillers. The polymeric resin, such as an epoxy, a silicone, or a polyimide, provides physical and mechanical properties such as adhesion, mechanical strength, impact strength, and the metal filler (such as, silver, gold, nickel, or copper) conducts electricity. Metal-filled thermoset polymers were first patented as electrically conductive adhesives in the 1950s [1–3]. Recently, ECA materials have been identified as one of the major alternatives for lead-containing solders for microelectronics packaging applications. ECAs offer numerous advantages over conventional solder technology, such as environmental friendliness, mild processing conditions (enabling the use of heat-sensitive and low-cost components and substrates), fewer processing steps (reducing processing cost), low stress on the substrates, and fine-pitch interconnect capability (enabling the miniaturization of electronic devices) [4–9]. Therefore, conductive adhesives have been used in flat panel displays such as LCD (liquid crystal display), and smart card applications as an interconnect material and in flip-chip assembly, CSP (chip-scale package) and BGA (ball grid array) applications in replacement of the solder. However, no currently commercialized ECAs can replace tin-lead metal solders in all applications due to some challenging issues such as lower electrical conductivity, conductivity fatigue (decreased conductivity at elevated temperature and humidity aging or normal use condition) in reliability testing, limited current-carrying capability, and poor impact strength. Table 1 gives a general comparison between tin-lead solder and a generic commercial ECA [6].

Table 1 Conductive adhesives compared with solder

Characteristic	Sn/Pb solder	ECA (ICA)
Volume resistivity	0.000015 Ω cm	0.00035 Ω cm
Typical junction R	10–15 mW	<25 mW
Thermal conductivity	30 W/m-deg.K	3.5 W/m-deg.K
Shear strength	2200 psi	2000 psi
Finest pitch	300 μ m	<150–200 μ m
Minimum processing temperature	215°C	<150–170°C
Environmental impact	Negative	Very minor
Thermal fatigue	Yes	Minimal

Depending on the conductive filler loading level, ECAs are divided into isotropically conductive adhesives (ICAs), anisotropically conductive adhesives (ACAs) and nonconductive adhesives (NCAs). For ICAs, the electrical conductivity in all x -, y -, and z -directions is provided due to high filler content exceeding the percolation threshold. For ACAs or NCAs, the electrical conductivity is provided only in z -direction between the electrodes of the assembly. Figure 1 shows the schematics of the interconnect structures and typical cross-sectional images of flip-chip joints by ICA, ACA, and NCA materials illustrating the bonding mechanism for all three adhesives.

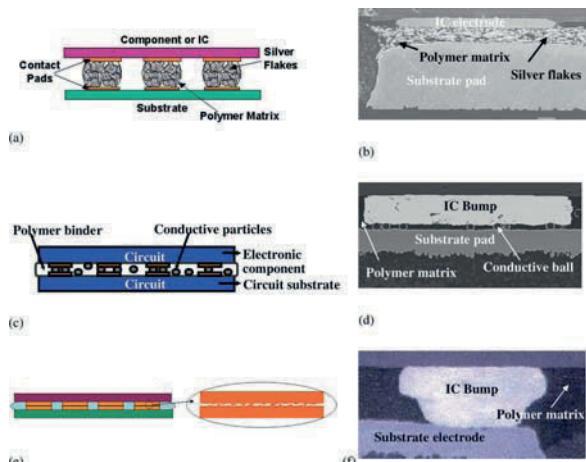


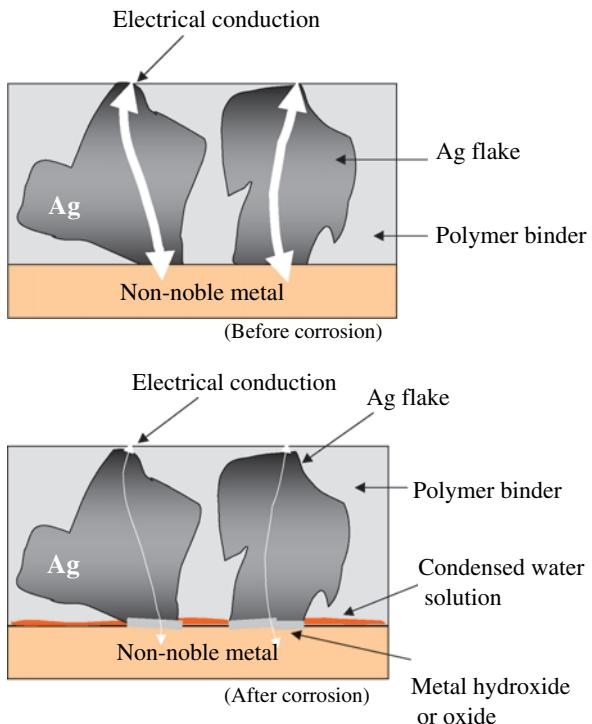
Fig. 1 Schematic illustrations and cross-sectional views of (a, b) ICA, (c, d) ACA, and (e, f) NCA flip-chip bondings

Isotropic conductive adhesives, also called as “polymer solder,” are composites of polymer resin and conductive fillers. The adhesive matrix is used to form a mechanical bond for the interconnects. Both thermosetting and thermoplastic materials are used as the polymer matrix. Epoxy, cyanate ester, silicone, polyurethane, etc., are widely used thermosets, and phenolic epoxy, maleimide acrylic preimidized

polyimide, etc., are the commonly used thermoplastics. An attractive advantage of thermoplastic ICAs is that they are reworkable, i.e., can easily be repaired. A major drawback of thermoplastic ICAs, however, is the degradation of adhesion at high temperature. A drawback of polyimide-based ICAs is that they generally contain solvents. During heating, voids are formed when the solvent evaporates. Most of commercial ICAs are based on thermosetting resins. Thermoset epoxies are by far the most common binders due to the superior balanced properties, such as excellent adhesive strength, good chemical and corrosion resistances, and low cost, while thermoplastics are usually added to allow toughening and rework under moderate heat. The conductive fillers provide the composite with electrical conductivity through physical contact between the conductive particles. The possible conductive fillers include silver (Ag), gold (Au), nickel (Ni), copper (Cu), and carbon in various forms (graphites, carbon nanotubes, etc.), sizes, and shapes. Among different metal particles, silver flakes are the most commonly used conductive fillers for current commercial ICAs because of the high conductivity and the maximum contact between flakes. In addition, silver is unique among all the cost-effective metals by nature of its conductive oxide. Oxides of most common metals are good electrical insulators. Copper powder, for example, becomes a poor conductor after oxidation/aging. Nickel and copper-based conductive adhesives generally do not have good resistance stability. As such, electrical reliability issues of ECA joints, in particular on non-noble metal finishes at elevated temperature and high relative humidity are of critical concern for implementing ECAs into the high-performance device interconnects. Intensive studies on this ECA joint failure led to understanding of the failure mechanisms, that corrosion at the ECA joints was the underlying mechanism rather than a simple oxidation process of the non-noble metal used [10, 11] (Fig. 2). Under the elevated temperature/humidity (for example, 85°C/85%RH), the non-noble metal acts as an anode, and is oxidized by losing electrons, and then turns into metal ion ($M-ne^- = M^{n+}$). The noble metal acts as a cathode, and its reaction generally is $2H_2O + O_2 + 4e^- = 4OH^-$. Then M^{n+} combines with OH^- to form a metal hydroxide, which further oxidizes to metal oxide. After corrosion, a layer of metal hydroxide or metal oxide is formed at the interface. Because this layer is electrically insulating, the contact resistance increases dramatically.

Since this underlying mechanism was reported, many research activities for reliable ECAs have been intensively conducted via various approaches, such as using low moisture absorption resins, incorporating corrosion inhibitors or oxygen scavengers, using sacrificial anode materials and oxide-penetrating sharp particles [10, 12–15]. In addition, many approaches to improve the electrical conductivity of ECAs have been reported, such as increasing the cure shrinkage of the matrix resin, replacing a lubricant of Ag flake with a shorter chain length carboxylic acid, adding an in-situ reducing agent for preventing the Ag flake-surface oxidation and adding low melting point alloys (LMA) for enhancing conductivity [16–21]. With the help of those intensive researches on enhancing the ECA performance, recently, ICAs

Fig. 2 Metal hydroxide or oxide formation after galvanic corrosion



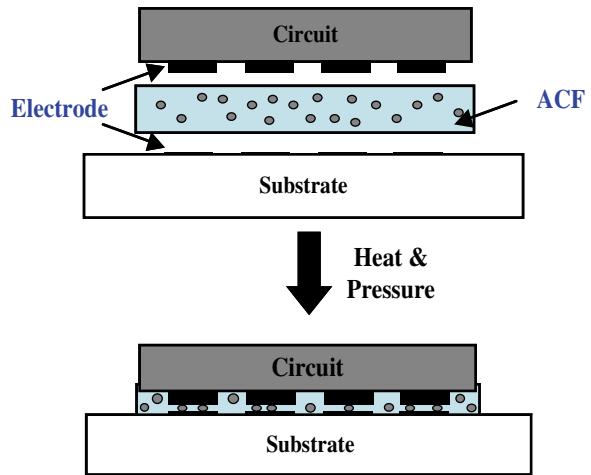
have also been considered as an alternative to tin/lead solders in surface mount technology (SMT) [22, 23], flip chip, [24] and other applications.

Anisotropic conductive adhesives (ACAs) or anisotropic conductive films (ACFs) provide uni-directional electrical conductivity in the vertical or Z-axis. This directional conductivity is achieved by using a relatively low-volume loading of conductive fillers (5–20 vol.-%). The low-volume loading is insufficient for inter-particle contact and prevents conductivity in the X-Y plane of the adhesives. The Z-axis adhesive, in film or paste form, is interposed between the surfaces to be connected. Heat and pressure are simultaneously applied to this stack-up until the particles bridge the two conductor surfaces.

The ACF-bonding process is a thermo-compression bonding as shown in Fig. 3. In case of tape-carrier packages (TCP) bonding, the ACF material is attached on a glass substrate and pre-attached. Final bonding is established by the thermal cure of the ACF resin, typically at 150–220°C, 3–20 s, and 100–200 MPa, and the conductive particle deformation between the electrodes of TCP and the glass substrate by applied bonding pressure.

Interconnection technologies using ACFs are major packaging methods for flat panel display modules to be of high resolution, light weight, thin profile,

Fig. 3 Thermo-compression bonding using ACF



and low power consumption [25], and are already successfully implemented in the forms of outer lead bonding (OLB), flex to PCB bonding (FPCB), reliable direct chip attach such as chip-on-glass (COG), chip-on-film (COF) for flat panel display modules [26–29], including liquid crystal display (LCD), plasma display panel (PDP), and organic light-emitting diode display (OLED). As for the small and fine-pitched bump of driver ICs to be packaged, fine-pitch capability of ACF interconnection is much more desired for COG, COF, and even OLB assemblies. There have been advances in development works for improved material system and design rule for ACF materials to meet fine-pitch capability and better adhesion characteristics of ACF interconnection for flat panel displays.

In addition to the LCD industry, ACA/ACF is now finding various applications in flex circuits and surface mount technology (SMT) for chip-scale package (CSP), application-specific integrated circuit (ASIC), and flip-chip attachment for cell phones, radios, personal digital assistants (PDAs), sensor chip in digital cameras, and memory chip in lap-top computers. In spite of the wide applications of ACA/ACF, there are some key issues that hinder their implementations as high-power devices. The ACA/ACF joints generally have lower electrical conductivity, poor current-carrying capability, and electrical failure during thermal cycling.

To meet the requirements for future fine-pitch and high-performance interconnects in advanced packaging, ECAs with nanomaterial or nanotechnology attract more and more interests due to the specific electrical, mechanical, optical, magnetic, and chemical properties. There has been extensive research on nanoconductive adhesives, which contain nano-filler such as nanoconductive particles, nanowires, or carbon nano tube (CNT). This chapter will provide a comprehensive review of most recent research results on nanoconductive adhesives.

2 Recent Advances on Nanoisotropic Conductive Adhesive (Nano-ICA)

2.1 ICAs with Silver Nanowires

Wire-type nano1D materials (such as silver nanowires and CNT) have been addressed for their interconnect applications because of the flexibility and lowering the percolation threshold in composites. Those advantages bring stress absorbing as well as light-weight interconnections. Silver nanowires are one of the important 1D materials for interconnect due to their high electrical conductivity. An ICA filled with nanosilver wires has been reported and compared to two other ICAs filled with micrometer-sized (roughly 1 μm and 100 nm, respectively) silver particles for the electrical and mechanical properties [30]. It was found that, at a low filler loading (e.g., 56 wt%), the bulk resistivity of ICA filled with the Ag nanowires was significantly lower than the ICAs filled with 1 μm or 100 nm silver particles. The better electrical conductivity of the ICA-filled nanowires was contributed to the fewer number of contact (lower contact resistance) between nanowires, more stable conductive network, and more significant contribution from the tunneling effects among the nanowires.

It was also found that, at the same filler loading (e.g., 56 wt%), the ICAs filled with Ag nanowires showed similar shear strength to those of the ICAs filled with 1 μm and 100 nm, respectively. However, to achieve the same level of electrical conductivity, the filler loading must be increased to at least 75 wt% for the ICA filled with micro-sized Ag particles, and the shear strength of these ICAs is then decreased (lower than that of the ICA filled with 56 wt% nanowires) due to the higher filler loading.

Chen et al. reported synthetic routes of Ag nanowires for conductive adhesive applications and demonstrated that the electrical conductance of a multicomponent system (mixture of different geometry fillers) was better than that of a single-component system (e.g., sphere or flake themselves). A possible mechanism was described in the report as these nanowires act as bridges to establish perfect linkage among particles, and the probability of contact and contact area becomes more than the cases without wires [31].

2.2 Effect of Nano-sized Silver Particles to the Conductivity of ICAs

Nano-sized fillers are being introduced into the ECA to replace the micro-sized silver flakes in recent research. Lee et al. studied the effects of nano-sized filler to the conductivity of conductive adhesives by substituting micro-sized Ag particles with nano-sized Ag colloids either in part or as a whole to a polymeric system (polyvinyl acetate - PVAc) [32]. It was found that, when nano-sized silver particles were added into the system at 2.5 wt% each increment, the resistivity

increased in almost all the cases, except when the quantity of micro-sized silver was slightly lower than the threshold value. At that point, the addition of about 2.5 wt% brought a significant decrease in resistivity. Near the percolation threshold, when the micro-sized silver particles were still not connected, the addition of a small amount of nano-sized silver particles helped to build the conductive network and lowers the resistivity of the composite. However, when the filler loading was above the percolation threshold and all the micro-sized particles were connected, the addition of nano-particles seemed only to signify the relative contribution of contact resistance between the particles. Due to its small size, for a fixed amount of addition, the nano-sized silver colloid contained a larger number of particles when compared with micro-sized particles. This large number of particles should be beneficial to the connection between particles. However, it also inevitably increased the contact resistance. As a result, the overall effect was an increase in resistivity upon the addition of nano-sized silver colloids. They also studied the effects of temperature on the conductivity of ICAs. Heating the composite to a higher temperature could reduce the resistivity significantly. This is likely due to the high surface activity of nano-sized particles. For micro-sized paste, this temperature effect was considered negligible. The interdiffusion of silver atoms among nano-sized particles helped to reduce the contact resistance quite significantly and the resistivity reached $5 \times 10^{-5} \Omega \text{ cm}$ after treatment at 190°C for 30 min.

Ye et al. reported that the addition of nanoparticles showed a negative effect on electrical conductivity [33]. They proposed two types of contact resistance, i.e., restriction resistance due to small contact area and tunneling resistance when nanoparticles are included in the system. It was believed that the conductivity of micro-sized Ag particle-filled adhesives could be dominated by constriction resistance, while the nanoparticle containing conductive adhesives is controlled by tunneling and even thermionic emission. Fan et al. also observed a similar phenomenon (adding nano-size particles reduced both electrical and thermal conductivities) [34].

The overall resistance (R_{total}) of the isotropic conductive adhesive (ICA) formulation is the sum of the resistance of fillers (R_{fillers}), the resistance between fillers ($R_{\text{btw fillers}}$), and the resistance between filler and pads ($R_{\text{filler to bond pad}}$) (Equation 1). In order to decrease the overall contact resistance, the reduction of the number of contact points between the particles may be obviously effective. Incorporation of nano-fillers increases the contact resistance and reduces the electrical performance of the ICAs. The number of contacts between the small particles is larger than that between the large particles as shown in Fig. 4a, b. A recent study by Jiang et al. showed that nano-silver particles could exhibit sintering behavior at curing temperature of ICAs [35, 36]. If nanoparticles are sintered together, then the contact between fillers will be fewer. This will lead to smaller contact resistance (Fig. 4c). By using effective surfactants for the dispersion and effective capping those nano-sized silver fillers in ECAs, obvious sintering behavior of the nano-filters can be achieved. The sintering of nano-silver fillers improved the interfacial properties of conductive fillers and polymer matrices and reduced the contact resistance between

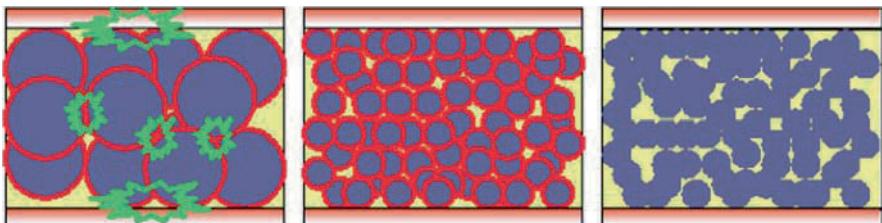


Fig. 4 Schematic illustration of particles between the metal pads [35]

fillers. The dispersion and interdiffusion of silver atoms among nano-sized particles could be facilitated and the resistivity of ICA could be reduced to $5 \times 10^{-6} \Omega \text{ cm}$.

$$R_{\text{total}} = R_{\text{btw fillers}} + R_{\text{filler to bond pad}} + R_{\text{fillers}} \quad (1)$$

2.3 ICA Filled with Aggregates of Nano-sized Ag Particles

To improve the mechanical properties under thermal cycling conditions while still maintaining an acceptably high level of electric conductivity, Kotthaus et al. studied an ICA material system filled with aggregates of nano-sized Ag particles [37]. The idea was to develop a new filler material which did not sacrifice the mechanical property of the polymer matrix to such a great extent. A highly porous Ag powder was attempted to fulfill these requirements. The Ag powder was produced by the inert gas condensation method. The powders consist of sintered networks of ultra-fine particles in the size range 50–150 nm. The mean diameter of these aggregates could be adjusted down to some microns. The as-sieved powders were characterized by low level of impurity content, an internal porosity of about 60%, a good ability for resin infiltration.

Using Ag IGC instead of Ag flakes is more likely to retain the properties of the resin matrix because of the infiltration of the resin into the pores. Measurements of the shear stress-strain behavior indicated that the thermo-mechanical properties of bonded joints may be improved by a factor up to two, independent of the chosen resin matrix.

Resistance measurements on filled adhesives were performed within a temperature range from 10 to 325 K. The specific resistance of a Ag IGC-filled adhesive is about $10^{-2} \Omega \text{ cm}$ and did not achieve the typical value of commercially available adhesives of about $10^{-4} \Omega \text{ cm}$. The reason may be that Ag IGC particles are more or less spherical whereas Ag flakes are flat. So, the decrease of the percolation threshold because of the porosity of Ag IGC is overcompensated by the disadvantageous shape and the intrinsically lower specific conductivity. For certain applications where mechanical stress plays an important role, this conductivity may be sufficient and therefore the porous Ag could be suitable as a new filler material for conductive adhesives.

2.4 Nano-Ni Particle-Filled ICA

Sumitomo Electric Industries, Ltd. (SEI) has developed a new liquid-phase deposition process using plating technology. This new nanoparticle fabrication process achieves purity greater than 99.9% and allows easy control of particle diameter and shape. The particles' crystallite size calculated from the result of X-ray diffraction measurement is 1.7 nm, which leads to an assumption that the particle size of primary particles is extremely small. When the particle size of nickel and other magnetic metals becomes smaller than 100 nm, they are changed from the multidomain particles to the single-domain particles, and their magnetic properties change. That is, if the diameter of nickel particles is around 50 nm, each particle acts like a magnet that has only a pair of magnetism, and magnetically connects with each other to form the chain-like clusters. When the chain-like clusters are applied to conductive paste, electrical conduction of the paste is expected to be better than the existing paste. The developed chain-like nickel particles were mixed with pre-defined amount of polyvinylidene difluoride (PVdF) that acts as an adhesive. Then, *N*-methyl-2-pyrrolidone was added to this mixture to make conductive paste. This paste was applied on a polyimide film and then dried to make a conductive sheet. Specific volume resistivity of the fabricated conductive sheet was measured by the quadruple method. The same measurement was also conducted on the conductive sheet that used the paste made of conventional spherical nickel particles. Measurement of the sheet resistance immediately after paste application defined that the developed chain-like nickel powders had low resistance of about one-eighth of that of the conventionally available spherical nickel particles. This result showed that, when the newly developed chain-like nickel particles were applied as conductive paste, high conductivity can be achieved without pressing the sheet.

2.5 Nano-conductive Adhesives for Via-Filling Applications in Organic Substrates

The demand of higher density organic substrates is increasing recently. Traditionally, greater wiring densities are achieved by reducing the dimensions of vias, lines, and spaces, increasing the number of wiring layers, and utilizing blind and buried vias. However, each of these approaches possess inherent limitations, for example, those related to drilling and plating of high aspect ratio vias, reduced conductance of narrow circuit lines, and increased cost of fabrication related to additional wiring layers. One method of extending wiring density beyond the limits imposed by these approaches is to form more metal-to-metal Z-axis interconnection of sub-composites during lamination to form a composite structure. Conductive joints can be formed during lamination using an electrically conductive adhesive. As a result, one is able to fabricate structures with vertically terminated vias of arbitrary depth. Replacement of conventional plated through holes with vertically terminated

vias opens up additional wiring channels on layers above and below the terminated vias and eliminates via stubs, which cause reflective signal loss.

Conductive adhesives usually have high filler loading that weaken the overall mechanical strength. Therefore, reliability of the conductive joint formed between the conductive adhesive and the metal surface to which it is mated is of prime importance. Conductive adhesives can have broad particle size distributions. Larger particles can be a problem when filling smaller holes (e.g., diameter of 60 μm or less), resulting in voids.

Das et al. developed conductive adhesives using controlled-sized particles, ranging from nanometer scale to micrometer scale, and used them to fill small diameter holes to fabricate Z-axis interconnections in laminates for interconnect applications [38]. A variety of metals, including Cu, Ag, and low melting point (LMP) alloys with particle sizes ranging from 80 nm to 15 μm were used to make the epoxy-based conductive adhesives. Among all, silver-based adhesives showed the lowest volume resistivity and highest mechanical strength. It was found that with increasing curing temperature, the volume resistivity of the silver-filled paste decreased due to sintering of metal particles. Sinterability of silver adhesive was further evaluated using high-temperature/pressure lamination and shows a continuous metallic network when laminated at 365°C.

As a case study, they used a silver-filled conductive adhesive as a Z-axis interconnecting construction for a flip-chip plastic ball grid array package with a 150 μm die pad pitch. High aspect ratio, small diameter ($\sim 55 \mu\text{m}$) holes were successfully filled. Silver-filled adhesives were electrically and mechanically better than Cu and LMP-filled adhesives. All adhesives maintained high tensile strength even after 1000 cycles thermal cycling (-55 to 120°C). Conductive joints were stable after three times reflowing, 1000 cycles thermal cycling, pressure cooker test (PCT), and solder shock. The adhesive-filled joining cores were laminated with circuitized subcomposites to produce a composite structure. High-temperature/pressure lamination was used to cure the adhesive in the composite and provide stable, reliable Z-interconnections among the circuitized subcomposites.

2.6 Nano-ICAs Filled with CNT

2.6.1 Electrical and Mechanical Characterization of CNT-Filled ICAs

Carbon nanotubes are a new form of carbon, which was first identified in 1991 by Sumio Iijima of NEC, Japan [39]. Nanotubes are sheets of graphite rolled into seamless cylinders. Besides growing single wall nanotubes (SWNTs), nanotubes can also have multiple walls (MWNTs) – cylinders inside the other cylinders, which act as multichannel transport of electrical and thermal for thermal interface materials (TIMs) applications. The carbon nanotube can be 1–50 nm in diameter and 10–100 μm or up to a few centimeters in length, with each end “capped” with half of a fullerene dome consisting of five- and six-member rings. Along the sidewalls and cap, additional molecules can be attached to functionalize the

nanotube to adjust its properties. CNTs are chiral structures with a degree of twist in the way that the graphite rings join into cylinders. The SWCNT chirality determines whether a nanotube will conduct in a metallic or semiconducting manner. Carbon nanotubes possess many unique and remarkable properties. The measured electrical conductivity of metallic carbon nanotubes is in the order of 104 S/cm. The thermal conductivity of carbon nanotubes at room temperature can be as high as 6600 W/m K [40]. The Young's modulus of carbon nanotube is about 1 TPa. The maximum tensile strength of carbon nanotube is close to 30 GPa, some reported at TPa [41]. Since carbon nanotubes have very low density and long aspect ratios, they have the potential of reaching the percolation threshold at very low weight percent loading in the polymer matrix.

Epoxy-based conductive adhesives filled with MWNTs have been recently developed [42]. It was found that ultrasonic mixing process helped disperse CNTs in the epoxy more uniformly and made them contact better, and thus lower electrical resistance was achieved. The contact resistance and volume resistivity of the conductive adhesive decreased with increasing CNT loading. The percolation threshold for the MWNTs used was less than 3 wt%. With 3 wt% loading, the average contact resistance was comparable with solder joints. It was also found that the performance of CNTs-filled conductive adhesive was comparable with solder joints in high frequency. By replacing metal particle fillers with CNTs in the conductive adhesive, a higher percentage of mechanical strength was retained. For example, with 0.8 wt% of CNT content, the 80% of shear strength of the polymer matrix was retained, while conventional metal-filled conductive adhesives only retain less than 28% of shear strength of the polymer matrix.

Experiments conducted by Qian et al. [43] show 36–42% and 25% increases in elastic modulus and tensile strength, respectively, in polystyrene (PS)/CNT composites. The TEM observations in their experiments showed that cracks propagated along weak CNT–polymer interfaces or relatively low CNT density regions and caused failure. If the outer layer of MWNTs can be functionalized to form strong chemical bonds with the polymer matrix, the CNT/polymer composites can be further reinforced in mechanical strength and have controllable thermal and electrical properties.

2.6.2 Effect of Adding CNT to the Electrical Properties of ICAs

Effect of adding CNT to the electrical conductivity of silver-filled conductive adhesive with various CNT loadings were reported [44]. It was found that the CNT could enhance the electrical conductivity of the conductive adhesives greatly when the silver filler loading was still below percolation threshold. For example, 66.5 wt% silver-filled conductive adhesive without CNT had a resistivity of $10^4 \Omega \text{ cm}$, but showed a resistivity of $10^{-3} \Omega \text{ cm}$ after adding 0.27 wt% CNT. Therefore, it is possible to achieve the same level of electrical conductivity by adding a small amount of CNT to replace the silver fillers.

2.6.3 Composites Filled with Surface-Treated CNTs

Although CNTs have exceptional physical properties, incorporating CNTs into other materials has been inhibited by the surface chemistry of carbon. Problems such as phase separation, aggregation, poor dispersion within a matrix, and poor adhesion to the host must be overcome. Zyyvex claimed that they have overcome these restrictions by developing a new surface treatment technology that optimizes the interaction between CNTs and the host matrix [45]. A multifunctional bridge was created between the CNT sidewalls and the host material or solvent. The power of this bridge was demonstrated by comparing the fracture behavior of the composites filled with untreated and surface-treated nanotubes. It was observed that the untreated nanotubes interacted poorly with the polymer matrix, and thus left behind voids in the matrix after fracture. However, for composite filled with treated nanotubes, the nanotube remained in the matrix even after the fracture, indicating strong CNT interaction with the matrix. Due to their superior dispersion in the polymer matrix, the treated nanotubes achieved the same level of electrical conductivity at much lower loadings than the untreated nanotubes [45].

2.7 Inkjet Printable Nano-ICAs and Inks

Areas for printing very fine-pitch matrix (e.g., very fine-pitch paths, antennas) are very attractive. But there are special requirements for inkjet printing materials, namely the most important ones are low viscosity and very homogenous structure (such as molecular fluid) to avoid sedimentation and separation during the process. Additionally, for electrical conductivity of printed structure, the liquid has to contain conductive particles, with nano-sized dimension to avoid the printing nozzle blocking and prevent sedimentation phenomenon. The nano-sized silver seems to be one of the best candidates for this purpose, especially when its particle size dimensions will be less than 10 nm.

Inkjetting is an accepted technology for dispensing small volumes of material (50–500 pl). Currently, traditional metal-filled conductive adhesives cannot be processed by inkjetting (due to their relatively high viscosity and the size of filler material particles). Smallest droplet size achievable by traditional dispensing techniques is in the range of 150 μm , yielding proportionally larger adhesive dots on the substrate. Electrically conductive inks are available on the market with metal particles (gold or silver) <20 nm suspended in a solvent at 30–50 wt%. After deposition, the solvent is evaporated and electrical conductivity is enabled by a high metal ratio in the residue. Some applications include a sintering step [46, 47].

There are many requirements for an inkjettable, Ag particle-filled conductive adhesive. The silver particles must not exceed a maximum size determined by the diameter of the injection needle used. At room temperature, the adhesive should resist sedimentation for at least 8, preferably 24 hours. A further requirement by the end user on the adhesive's properties was a two-stage curing mechanism. In the

first curing step, the adhesive surface is dried. In this state, the product may be stored for several weeks. The second curing step involves glueing the components with the previously applied adhesive. By heating and applying pressure the adhesive is cured. Thus, the processing operation is similar to that required for soldering. A conductivity in the range of $10^{-4} \Omega \text{ cm}$ in the bulk material is required. An adhesive less prone to sedimentation was formulated by using suitable additives. Furthermore, the formation of filler agglomerations during deflocculation and storage was reduced. This effect was achieved by making the additives adhere to the filler particle surfaces. This requires a very sensitive balance. If the insulation between individual silver particles becomes very strong, overall electrical conductivity is significantly reduced.

Jana Kolbe et al. demonstrated feasibility of an inkjettable, isotropically conductive adhesive in the form of a silver-loaded resin with a two-step curing mechanism [48, 49]. In the first step, the adhesive was dispensed (jetted) and precured leaving a “dry” surface. The second step consisted of assembly (wetting of the second part) and final curing. The attainable droplet sizes were in the range of $130 \mu\text{m}$, but could be further reduced by using smaller (such as $50 \mu\text{m}$) and more advanced nozzle shapes.

Nano-Ag particles for nano-ink applications are coated with capping agents for dispersion, which should be removed before the solidification process such as sintering. Otherwise, the particles cannot fuse or sinter each other. The capping agent removal usually requires heating the printed pattern to evaporate solvent and induce sintering. Wakuda et al. recently demonstrated room temperature sintering Ag nanoparticles that are protected by a dispersant in an air atmosphere, where the dodecylamine dispersant was removed in methanol and the Ag nanoparticles were densely sintered. The sintered Ag wires showed low resistivity of $7.3 \times 10^{-5} \Omega \text{ cm}$, which can be used for inkjet printing [50].

3 Recent Advances of Nano-ACA/ACF

3.1 Low-Temperature Sintering of Nano-Ag-Filled ACA/ACF

One of the concerns for ACA/ACF is the higher joint resistance since interconnection using ACA/ACF relies on mechanical contact, unlike the eutectic solder, which forms the metallurgical bonding during reflow. An approach to minimize the joint resistance of ACA/ACF is to make the conductive fillers fuse each other and form metallic joints such as metal solder joints. However, to fuse metal fillers in polymers does not appear feasible, since a typical organic printed circuit board ($T_g \sim 125^\circ\text{C}$), on which the metal filled polymer is applied, cannot withstand such a high temperature; the melting temperature (T_m) of Ag, for example, is around 960°C . Studies have showed that T_m and sintering temperatures of materials could be dramatically reduced by decreasing the size of the materials [51–53]. It has been reported that the surface premelting could be a primary mechanism of the T_m depression of

the fine nanoparticles (<50 nm). For nano-sized particles, sintering behavior could occur at much lower temperatures, as such, the use of the fine metal particles in ACAs would be promising for high electrical performance of ACA joints by eliminating the interface between metal fillers. The application of nano-sized particles can also increase the number of conductive fillers on each bond pad and result in more contact area between fillers and bond pads as illustrated in Fig. 5. For the sintering reaction in a certain material system, temperature and duration are the most important parameters, in particular, the sintering temperature.

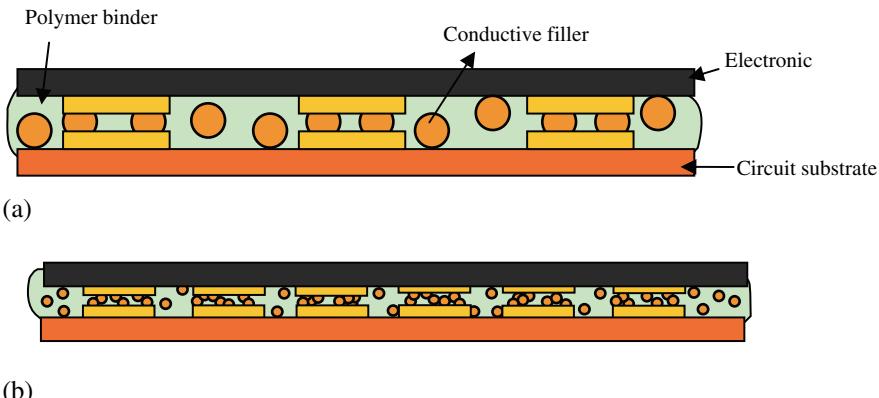


Fig. 5 Schematic illustrations of ACA/ACF with (a) micro-sized conductive fillers and (b) nano-sized conductive fillers

Current–resistance (I – R) relationship of the nano-Ag-filled ACA joint is shown in Fig. 6. As can be seen from the figure, with increasing curing temperatures, the resistance of the ACA joints decreased significantly, from 10^{-3} to $5 \times 10^{-5} \Omega$. Also, ACAs cured at higher temperature exhibited higher current-carrying capability than ACAs cured at low temperature. This phenomenon suggested that more sintering of nano-Ag particles and subsequently superior interfacial properties between fillers and metal bond pads were achieved at higher temperatures [53], yet the x – y direction of the ACF maintains an excellent insulation property for electrical insulation.

3.2 Self-Assembled Molecular Wires for Nano-ACA/ACF

To enhance the electrical performance of ACA/ACF materials, self-assembled molecular wires (SAMW) have been introduced into the interface between metal fillers and metal-finished bond pad of ACAs [54, 55]. These organic molecules adhere to the metal surface and form physico-chemical bonds, which allow electrons to flow, as such, it reduces electrical resistance and enables a high current flow. The unique electrical properties are due to their tuning of metal work functions by those organic monolayers. The metal surfaces can be chemically modified

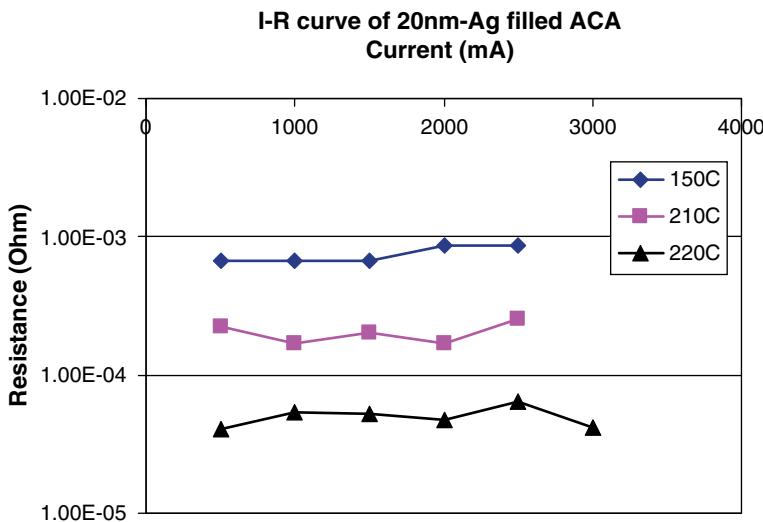
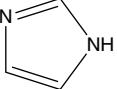


Fig. 6 Current–resistance (*I*–*R*) relationship of nano-Ag-filled ACAs with different curing temperatures [53]

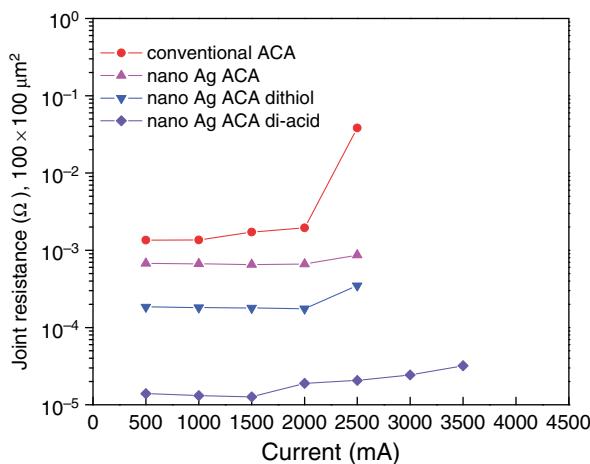
by the organic monolayer and the reduced work functions can be achieved by using suitable organic monolayer coatings. An important consideration when examining the advantages of organic monolayers pertains to the affinity of organic compounds to specific metal surfaces. Table 2 gives the examples of molecules preferred for maximum interactions with specific metal finishes; although only molecules with symmetrical functionalities for both head and tail groups are shown, molecules and derivatives with different head and tail functional groups are possible for interfaces concerning different metal surfaces.

Different organic molecular wires, dicarboxylic acid, and dithiol have been introduced into ACA/ACF joints. For SAM-incorporated ACA with micron-sized gold/polymer or gold/nickel fillers, lower joint resistance and higher maximum allowable current (highest current applied without inducing joint failure) was achieved for low-temperature curable ACA (<100°C). For high curing temperature ACA (150°C), however, the improvement was not as significant as low curing temperature ACAs, due to the partial desorption/degradation of organic monolayer coating at the relatively high temperature [56]. However, when dicarboxylic acid or dithiol was introduced into the interface of nano-silver-filled ACAs, significantly improved electrical properties could be achieved at high-temperature curable ACA/ACF, suggesting the coated molecular wires did not suffer degradation on silver nanoparticles at the curing temperature (Fig. 7). The enhanced bonding could attribute to the larger surface area and higher surface energy of nanoparticles, which enabled the monolayers to be more readily coated and relatively thermally stable on the metal surfaces [57].

Table 2 Potential organic monolayer interfacial modifiers for different metal finishes

Formula	Compound	Metal finishes
R-S-H	Thiols	Au, Ag, Cu, Pt, Zn
R-COOH	Carboxylates	Fe/Fe _x O _y , Ti/TiO ₂ , Ni, Al, Ag
R-C≡N	Cyanides	Au, Ag, Pt
R-N=C=O	Isocyanates	Pt, Pd, Rh, Ru
	Imidazole and triazole derivatives	Cu
R-SiOH	Organosilane derivatives	SiO ₂ , Al ₂ O ₃ , quartz, glass, mica, GeO ₂ , etc.

R denotes alky and aromatic groups

**Fig. 7** Electrical properties of nano-Ag-filled ACA with dithiol or dicarboxylic acid [57]

3.3 Silver Migration Control in Nano-silver-Filled ACA

Silver is the most widely used conductive fillers in ICAs and exhibits exciting potentials in nano-ACA/ACF due to many unique advantages of silver. Silver has the highest room temperature electrical and thermal conductivity among all the conductive metals. Silver is also unique among all the cost-effective metals by nature of its conductive oxide (Ag₂O). In addition, silver nanoparticles are relatively easy to be formed into different sizes (a few nanometers to 100 nm) and shapes (such as spheres, rods, wires, disks, flakes) and well dispersed in a variety of polymeric matrix materials. Also the low-temperature sintering makes silver one of the promising candidate as conductive fillers in nano-ACA/ACF. However, silver migration has long been a reliability concern in electronic industry. Metal migration is an

electrochemical process, whereby, metal (e.g., silver), in contact with an insulating material, in a humid environment and under an applied electric field, leaves its initial location in ionic form and deposits at another location [58]. It is considered that a threshold voltage exists above which the migration starts. Such migration may lead to a reduction in electrical spacing or cause a short circuit between interconnections. The migration process begins when a thin continuous film of water forms on an insulating material between oppositely charged electrodes. When a potential is applied across the electrodes, a chemical reaction takes place at the positively biased electrode where positive metal ions are formed. These ions, through ionic conduction, migrate toward the negatively charged cathode and over time, they accumulate to form metallic dendrites. As the dendrite growth increases, a reduction of electrical spacing occurs. Eventually, the dendrite silver growth reaches the anode and creates a metal bridge between the electrodes, resulting in an electrical short circuit [59].

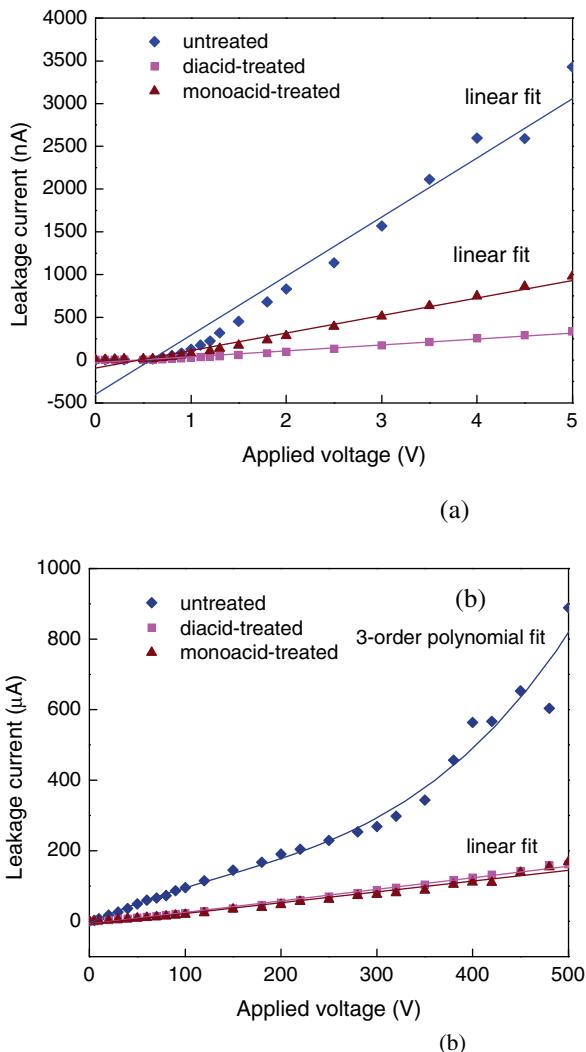
Although other metals may also migrate under specific environment, silver is more susceptible to migration, mainly due to the high solubility of silver ion, low activation energy for silver migration, high tendency to form dendrite shape, and low possibility to form a stable passivation oxide layer [60–62]. The rate of silver migration is increased by (1) an increase in the applied potential; (2) an increase in the time of the applied potentials; (3) an increase in the level of relative humidity; (4) an increase in the presence of ionic and hydroscopic contaminants on the surface of the substrate; and (5) a decrease in the distance between electrodes of the opposite polarity.

To reduce silver migration and improve the reliability, several methods have been reported. The methods include (1) alloying the silver with an anodically stable metal such as palladium [61] or platinum [63] or even tin [64]; (2) using hydrophobic coating over the PWB to shield its surface from humidity and ionic contamination [65], since water and contaminates can act as a transport medium and increase the rate of migration; (3) plating of silver with metals such as tin, nickel, or gold, to protect the silver filler and reduce migration; (4) coating the substrate with polymer [66]; (5) applying benzotriazole (BTA) and its derivatives [67]; (6) employing siloxane epoxy polymers as diffusion barriers due to the excellent adhesion of siloxane epoxy polymers to conductive metals [68]; (7) chelating silver fillers in ECAs with molecular monolayers [69]. As an example shown in Fig. 8 [70], with carboxylic acids and forming chelating compounds with silver ions, the silver migration behavior (leakage current) could be significantly reduced and controlled.

3.4 ACF with Straight-Chain-Like Nickel Nanoparticles

Sumitomo Electric recently developed a new concept ACF using nickel nanoparticles with a straight-chain-like structure as conductive fillers [72]. They applied the formulated straight-chain-like nickel nanoparticles and solvent in a mixture of epoxy resin on a substrate film. Then the particles were made to orient toward the vertical direction of the film surface and fixed in a resin by evaporating the solvent, and this approach is similar to the elastomeric conductive polymer interconnects (ECPI), dispersed Ni-coated glass spheres in a silicone matrix, which has

Fig. 8 Leakage current–voltage relationship of nano-Ag conductive adhesives at (a) low voltages and (b) high voltages [71]



been developed by AT&T Bell Laboratory [73]. In the estimation using 30 μm -pitch IC chips and glass substrates (the area of Au bumps was 2000 μm^2 , the distance of space between neighboring bumps was 10 μm), the new ACF showed excellent reliability of electrical connection after high-temperature, high-humidity (60°C/90%RH) test and thermal cycle test (between -40 and 85°C). The samples were also exposed to high-temperature, high-humidity (60°C/90%RH) for insulation ability estimation. Although the distance between two electrodes was only 10 μm , ion migration did not occur and insulation resistance has been maintained at over 1 G Ω for 500 h. This result showed that the new ACF has superior insulation reliability and has potential to be applied in very fine interconnections.

3.5 Nanowire ACF for Ultra-fine-pitch Flip-Chip Interconnection

To satisfy the reduced I/O pitch and avoid electric shorting, a possible solution is to use high-aspect-ratio metal post. Nanowires exhibit high possibilities due to the small size and extremely high aspect ratio. In literature, nanowires could be applied in FET (field effective transistor) sensor for gas detection, magnetic hard-disk, nanoelectrodes for electrochemical sensor, thermal-electric device for thermal dissipation and temperature control, etc. [71,74,75]. To prepare nanowires, it is important to define nanostructures on the photoresist. Many expensive methods such as e-beam, X-ray, or scanning probe lithography have been used but the length of nanowires cannot be achieved to micro-meter size. A less expensive alternative is electrodeposition of metal into nano-porous template such as anodic aluminum oxide (AAO) [76] or block-copolymer self-assembly template. The disadvantages of block-copolymer template include thin thickness (that means short nanowires), nonuniform distribution and poor parallelism of nano-pores. However, AAO has benefits of higher thickness ($>10\ \mu\text{m}$), uniform pore size and density, larger size, and very parallel pores. Lin et al. [77] developed a new ACF with nanowires. They used AAO templates to obtain silver and cobalt nanowires array by electrodeposition. And then low viscous polyimide (PI) was spread over and filled into the gaps of nanowires array after surface treatment. The bi-metallic Ag/Co nanowires could remain parallel during fabrication by magnetic interaction between cobalt and applied magnetic field. The silver and cobalt nanowires/polyimide composite films could be obtained with a diameter of about 200 nm for nanowire and maximum film thickness up to 50 μm . The X-Y insulation resistance is about 4–6 G Ω and Z-direction resistance including the trace resistance (3 mm length) is less than 0.2 Ω . They also demonstrated the evaluation of this nanowire composite film by stress simulation. They found that the most important factor for designing nanowire ACF was the volume ratio of nanowires. However, actually the ratio of nanowires cannot be too small to influence the electric conductance. They concluded that it is important to get a balance between electric conductance and thermal-mechanical performance by increasing film thickness or decreasing modulus of polymer matrix.

3.6 An In Situ Formation of Nano-conductive Fillers in ACA/ACF

One of the challenging issues in the formation of nano-filler ACA/ACF is the dispersion of nano-conductive fillers in ACA/ACF. A lot of research has been going on in recent years to address the dispersion issue of nano-composite because nano-fillers due to their high surface reactivity and high electrostatic force tend to agglomerate. For the fine-pitch electronic interconnects using nano-ACA/ACF, the dispersion issues need to be solved. The efforts usually include the physical approaches such as sonication and chemical approaches such as surfactants. Recently, a novel ACA/ACF incorporated with in-situ formed nano-conductive particles is proposed for the next generation high-performance fine-pitch electronic packaging applications [78, 79]. This novel interconnect adhesive combines

the electrical conduction along the z -direction (ACA-like) and the ultra fine-pitch (<100 nm) capability. Instead of adding the nano-conductive fillers in the resin, the nanoparticles can be *in situ* formed during the curing/assembly process. By using *in situ* formation via chemical reduction of nanoparticles, during the polymer curing process, the filler concentration and dispersion could be controlled and the drawback of surface oxidation of the nano-fillers could be easily overcome [80].

3.7 CNT-Based Conductive Nanocomposites for Transparent, Conductive, and Flexible Electronics

The electrically conductive and optically transparent coating is an upcoming potential for various applications such as transistors [81–84], diodes [85], sensors [86, 87], optical modulators [88] or as conductive backbone for electrochemical polymer coating [89], smart windows [90], photovoltaics [91], solar cells [92,93] and organic light-emitting diode (OLED) [94, 95].

With the popularity of flex circuits or substrates, in particular for large-scale flexible flat panel displays, there are demands for flexible interconnect materials, especially under highly mechanical bending conditions of printed electronics. ITO (indium tin oxide), fluorine tin oxide (FTO), and conductive polymers such as polyaniline (PANI) have been de facto options as transparent and conductive coating materials. However, those materials have no flexibility in nature, which is critical for flexible devices such as bendable displays and sensors. Thus, materials with mechanical flexibility as well as electrical conductivity and transparency can find various applications. The CNT incorporated polymer composite is one of the candidates because its tremendously high aspect ratio can provide a significantly low percolation threshold in electrical conduction. Its tiny size (<10 nm in diameter (SWCNT), smaller than $1/4\lambda$ of visible light) lets light pass through and the polymer matrix renders flexibility. Therefore, intensive efforts to materialize transparent, conductive, and flexible CNT-based polymer composite have been dedicated.

Kaempgen et al. found that single wall CNTs (SWCNTs) are more suitable for transparent conductive coatings than multi wall CNTs (MWCNTs) [96]. This is likely due to a significantly higher diameter of each MWCNT, which increases the light absorption (Fig. 9). The longer CNTs increase the conductivity for the same transparency with the decreased number of contacts because the electrical transport is dominated by contact resistances between the CNTs [97–101]. The uniform dispersion is a key factor to manufacture reproducible coatings or composite films.

Some CNT-based polymer composites already met the requirements for various applications. Figure 10 compares the resistivity (1 k Ω /sq) of the CNT coating with 90% light transmittance to those of various target applications [96], where it can be seen that currently 90% transparent CNT coating can meet the requirements for many applications such as touch screen and electrostatic discharge, although

Fig. 9 Transmittance vs. resistivity plot for sprayed CNT layers and ITO [96]

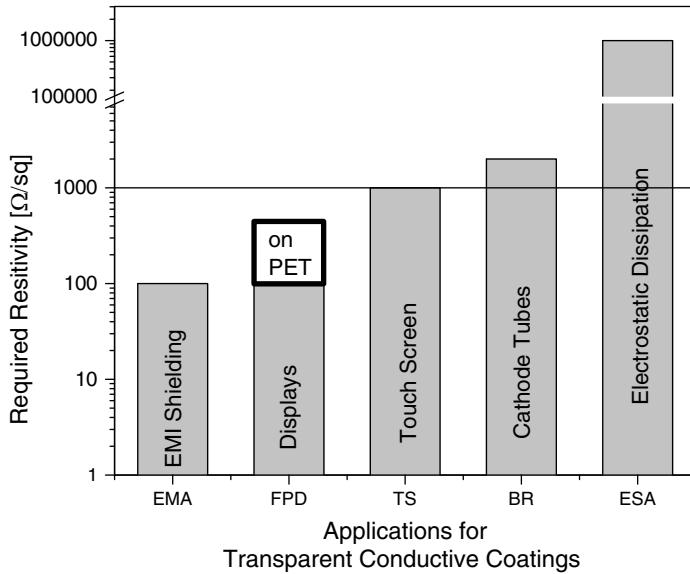
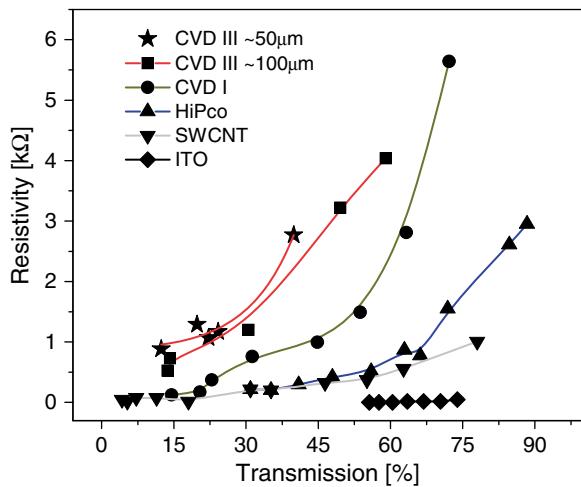


Fig. 10 Required surface resistivity in typical applications for transparent conductive coatings compared to the performance of thin conductive CNT composites [96]

only shielding of electromagnetic interference (EMI) is out of reach at least for transparent coatings. There are trade-off values between transparency and surface resistance of the CNT composites with regard to their applications. Recently, a CNT film on polymer or glass with 80% light transmittance exhibited a resistance of 20 Ω/sq [102]. With extensive research and development of CNT nanocomposite, more reliable conductive and transparent adhesives will find versatile applications in microelectronics.

4 Concluding Remarks

In this chapter, recent advances in nano-electrically conductive adhesives (ECAs) were reviewed and in particular, innovative approaches by using cutting-edge nanotechnologies were presented. The nanotechnology-based ECAs will be playing a very important role in future nanoelectronics, nanophotonics, nano-biomedical devices etc.

Although Ag filler is known as one of best conductive materials for ECA, its high cost and intrinsic propensity for silver migration are of concern for low-cost consumer products or high-performance device-packaging interconnects. As such, copper will be a silver alternative as ECA filler, as copper is the second highest electrically conductive element after silver and has less migration tendency. However, copper is very vulnerable to oxygen and corrosion, and unlike Ag, its oxide is a poor electrical conductor. Therefore, in order to achieve all the benefits from copper, much research effort on preventing copper from oxidation/corrosion is needed and the use of SAM/surfactant as mentioned in Table 2 may provide some solutions.

Tremendous evolution in display technologies has been demonstrated over the past decades, and future display technologies need very flexible and high-performance interconnect materials for nano- or bio-medical device packaging. In general, bending or physical deformation of ECAs can degrade the electrical conductivity. Therefore, novel flexible electrically conductive materials to respond to repeated bending or high elongations will be required for future flexible display technologies. In addition, transparent conductive adhesive is needed for some applications such as organic LED (OLED) application. Ouyang et al. [103] have demonstrated a transparent conducting polymer glue and the feasibility of fabricating organic electronic devices through a lamination process by using the PEDOT:PSS electric glue. By combining this technology with the large-area continuous coating process of organic films, it is possible to fabricate large-area organic electronic devices through a continuous roll-to-roll coating process. This electric glue made from a conducting polymer provides new applications for conducting polymers; it will revolutionize the fabrication process of flexible electronic devices.

Although some researches on enhancing current-carrying capability of ECAs have been carried out, there is no commercial ECA material to replace metallic solders yet. Thus, increasing current-delivering capability of the ECA will be a continuing research effort. Employing molecular wires in the ECA joints or conductive elements and manipulating conductive joint work functions with high thermal stability, are some of the promising approaches for the high-current ECA challenges.

With particle size of ECAs becoming smaller in particular for nano-ACAs, there is serious implication from the point of view of the manufacturing process. When ACAs are used to attach the metal pads on two mating surfaces, the warpage and coplanarity of these surfaces need to be carefully controlled to ensure all the metal pads on the mating surfaces are making contact to the conductive spheres in the ACA simultaneously. Large and compliant (deformable) conductive particles will ease the manufacturing process because these particles can accommodate the warpage and

noncoplanarity. With conductive particles becoming even smaller in nano-ACAs, the manufacturing will become more challenging.

The ECAs will be continuing to evolve to meet future technological needs, as such, much attention is needed on the breakthrough technology developments by engineers and scientists including materials scientists, chemists and physicists, etc. We all should bear in mind that high-performance interconnection materials for advanced nanoelectronics, photonics, bio-medical devices, and systems packaging are the key for future microelectronics.

References

1. Wolfson H. and Elliot G., Electrically conducting cements containing epoxy resins and silver. US Patent 2,774,747, 1956.
2. Matz K.R., Electrically conductive cement and brush shunt containing the same. US Patent 2,849,631, 1958.
3. Beck D.P., Printed electrical resistors. US Patent 2,866,057, 1958.
4. Li Y., Moon K., and Wong C.P., Science 2005; 308: 1419–1420.
5. Li Y. and Wong C.P., Mater. Sci. Eng. R 2006; 51: 1–35.
6. Liu J., (Ed.), Conductive adhesives for Electronics Packaging. Electrochemical Publications Ltd., Port Erin, Isle of Man, 1999; Chapter 1.
7. Hwang J.S. (Ed.), Environment-Friendly Electronics: Lead-free Technology. Electrochemical Publications Ltd., Port Erin, Isle of Man, 2001; Chapter 1, pp. 4–10.
8. Murray C.T., Rudman R.L., Sabade M.B., and Pocius A.V., Mater. Res. Bull. 2003; 28: 449–454.
9. Lau J., Wong C.P., Lee N.C., and Lee S.W.R., In: Electronics Manufacturing: with Lead-Free, Halogen-Free, and Conductive-Adhesive Materials. McGraw-Hill, New York, 2002.
10. Lu D., Tong Q.K., and Wong C.P., IEEE Trans. Compon. Packag. Manuf. Technol., Part C 1999; 22(3): 228–232.
11. Persson K., Nylund A., Liu J., and Olefjord I., Proceedings of the 7th European Conference on Applications of Surface and Interface Analysis, Gothenburg, June 1997.
12. Lu D. and Wong C.P., J. Appl. Polym. Sci. 1999; 74: 399–406.
13. Li Y., Moon K., and Wong C.P., J. Adhes. Sci. Technol. 2005; 19(16): 1427–1444.
14. Matienzo L.J., Egitto F.D., and Logan P.E., J. Mater. Sci. 2003; 38: 4831–4842.
15. Li H., Moon K., and Wong C.P., J. Electron. Mater. 2004; 33: 106–113.
16. Lu D., Tong Q.K., and Wong C.P., IEEE Trans. Compon. Packag. Manuf. Technol. Part C 1999; 22: 223–227.
17. Li Y., Moon K., and Wong C.P., IEEE Trans. Compon. Packag. Manuf. Technol. 2006; 29(1): 173–178.
18. Li Y., Moon K., Whitman A., and Wong C.P., IEEE Trans. Compon. Packag. Manuf. Technol. 2006; 29(4): 758–763.
19. Gallagher C., Matijasevic G., and Maguire J.F., Proceedings of 47th IEEE Electronic Components and Technology Conference 1997; 554–560.
20. Roman J.W. and Eagar T.W., Proceedings of the International Society for Hybrids and Microelectronics Society, San Francisco, CA, 1992; 52.
21. Li Y., Moon K., Li H., and Wong C.P., Proceedings of 54th IEEE Electronic Components and Technology Conference, Las Vegas, Nevada, June 1–4, 2004; 1959–1964.
22. Cavasin D., Brice-Heams K., and Arab A. Proceedings of 53rd Electronic Components and Technology Conference 2003; 1404–1407.
23. Kisiel R., J. Electron. Packag. 2002; 124: 367.

24. de Vries H., van Delft J., and Slob K., *IEEE Trans. Compon. Packag. Manuf. Technol.* 2005; 28: 499.
25. Watanabe I., Gotoh Y., and Kobayashi K. *Proc. Asia Display/IDW*, Nagoya, Japan 2001; 553–556.
26. Nishida H., Sakamoto K., and Ogawa H., *IBM J. Res. Dev.* 1998; 42: 517.
27. Williams D.J., Whalley D.C., Boyle O.A., and Ogunjimi A.O., *Solder. Surf. Mount Tech.* 1993; 5: 4.
28. Liu J., Tolvgard A., Malmodin J., and Lai Z., *IEEE Trans. Compon. Packag. Manuf. Technol.* 1999; 22: 186.
29. Clot P., Zeberli J.F., Chenuz J.M., Ferrando F., and Styblo D., *Proceedings of the International Electronics Manufacturing Technology Symposium 24th IEEE/CPMT*, Austin, TX, 1999; 36.
30. Wu H., Wu X., Liu J., Zhang G., Wang Y., Zeng Y., and Jing J., *Development of a novel isotropic conductive adhesive filled with silver nanowires. J. Compos. Mater.* 2006; 40(21): 1961–1968.
31. Chen C., Wang L., Li R., Jiang G., Yu H., and Chen T., *J. Mater. Sci.* 2007; 42(9): 3172.
32. Lee H.S., Chou K.S., and Shih Z.W., *Effect of nano-sized silver particles on the resistivity of polymeric conductive adhesives. Int. J. Adhes. Adhes.* 2005; 25: 437–441.
33. Ye L., Lai Z., Liu J., and Tholen A., *Effect of Ag particle size on electrical conductivity of isotropically conductive adhesives. IEEE Trans. Electron. Packag. Manuf.* 1999; 22(4): 299–302.
34. Fan L., Su B., Qu J., and Wong C.P., *Electrical and thermal conductivities of polymer composites containing nano-sized particles. Proceedings of Electronic Components and Technology Conference*. IEEE, NJ, 2004; 148–154.
35. Jiang H.J., Moon K., Lu J., and Wong C.P., *J. Electron. Mater.* 2005; 34: 1432–1439.
36. Jiang H.J., Moon K., Li Y., and Wong C.P., *Surface functionalized silver nanoparticles for ultrahigh conductive polymer composites. Chem. Mater.* 2006; 18(13): 2969–2973.
37. Kotthaus S., Günther B.H., Haug R., and Schafer H., *Study of isotropically conductive bondings filled with aggregates of nano-sized Ag-particles. IEEE Trans. Compon. Packag. Manuf. Technol. Part A* 1997; 20(1): 15–20.
38. Das R.N., Lauffer J.M., and Egitto F.D., *Proceedings of Electronic Components and Technology Conference*. IEEE, 2006; 112–118.
39. Iijima S., *Nature* 1991; 354: 56.
40. Berber S., Kwon Y.K., and Tomànek D., *Unusually high thermal conductivity of carbon nanotubes. Phys. Rev. Lett.* 2000; 84(20): 4613–4616.
41. Yu M.F., Files B.S., Arepalli S., and Ruoff R.S., *Tensile loading of ropes of single wall carbon nanotubes and their mechanical properties. Phys. Rev. Lett.* 2000; 84(24): 5552–5555.
42. Li J. and Lump J.K., *Electrical and mechanical characterization of carbon nanotube filled conductive adhesive. In: Proceedings of Aerospace Conference*. IEEE, NJ, 2006; 1–6.
43. Qian D., Dickey E.C., Andrews R., and Rantell T., *Load transfer and deformation mechanisms in carbon nanotube polystyrene composites. Appl. Phys. Lett.* 2000; 76: 2868.
44. Lin X.C. and Lin F., *Improvement on the properties of silver-containing conductive adhesives by the addition of carbon nanotube. In: Proceedings of High Density Microsystem Design and Packaging*. IEEE, NJ, 2004; 382–384.
45. Rutkofsky M., Banash M., Rajagopal R., and Chen J., *Using a carbon nanotube additive to make electrically conductive commercial polymer composites. Zyxex Corporation Application Note 9709: <http://www.zyxex.com/Documents/9709.PDF>* 28; 2006.
46. Kamyshny A., Ben-Moshe M., Aviezer S., and Magdassi S., *Ink-jet printing of metallic nanoparticles and microemulsions. Macromol. Rapid Commun.* 2005; 26: 281–288.
47. Cibis D. and Currie U., *Inkjet printing of conductive silver paths. In: 2nd International Workshop on Inkjet Printing of Functional Polymers and Materials*. Eindhoven, The Netherlands, 2005.

48. Kolbe J., Arp A., Calderone F., Meyer E.M., Meyer W., Schaefer H., and Stuve M., Inkjettable conductive adhesive for use in microelectronics and microsystems technology. In: Proceedings of IEEE Polytronic Conference. IEEE, NJ, 2005; 1–4.
49. Moscicki A., Felba J., Sobierajski T., Kudzia J., Arp A., and Meyer W., Electrically conductive formulations filled nano size silver filler for ink-jet technology. In: Proceedings of IEEE Polytronic Conference. IEEE, NJ, 2005; 40–44.
50. Wakuda D., Hatamura M., and Saganuma K., *Chem. Phys. Lett.* 2007; 441: 305–308.
51. Moon K., Dong H., Maric R., Pothukuchi S., Hunt A., Li Y., and Wong C.P., *J. Electron. Mater.* 2005; 34: 132–139.
52. Matsuba Y., Erektutoronikusu Jisso Gakkaishi 2003; 6(2): 130–135.
53. Li Y., Moon K. and Wong C.P., *J. Appl. Polym. Sci.* 2006; 99(4): 1665–1673.
54. Li Y., Moon K., and Wong C.P. In: Proceedings of 54th IEEE Electronic Components and Technology Conference. IEEE, NJ, 2004; 1968–1974.
55. Li Y. and Wong C.P. In: Proceedings of 55th IEEE Electronic Components and Technology Conference. IEEE, NJ, 2005; 1147–1154.
56. Li Y., Moon K., and Wong C.P., *J. Electron. Mater.* 2005; 34(3): 266–271.
57. Li Y., Moon K., and Wong C.P., *J. Electron. Mater.* 2006; 34(12): 1573–1578.
58. Davies G. and Sandstrom J. *Circuits Manuf.* 1976; 56–62.
59. Harsanyi G. and Ripka G., *Electrocomp. Sci. Technol.* 1985; 11: 281–290.
60. Giacomo G.A., In: J. McHardy and F. Ludwig (Eds.) *Electrochemistry of Semiconductors and Electronics: Processes and Devices*. Noyes Publications, Park Ridge, NJ, 1992; pp. 255–295.
61. Manepalli R., Stepniak F., Bidstrup-Allen S.A., and Kohl P.A., *IEEE Trans. Adv. Packag.* 1999; 22: 4–8.
62. Giacomo D., In: *Reliability of Electronic Packages and Semiconductor Devices*. McGraw-Hill, New York, 1997; Chapter 9.
63. Wassink R., *Hybrid. Circ.* 1987; 13: 9–13.
64. Shirai Y., Komagata M., and Suzuki K. In: 1st International IEEE Conference on Polymers and Adhesives in Microelectronics and Photonics. IEEE, NJ, 2001; 79–83.
65. Der Marderosian A. Ratheon Co. Equipment Division, Equipment Development Laboratories 1978; 134–141.
66. Schonhorn H. and Sharpe L.H., Prevention of surface mass migration by a polymeric surface coating. US Patent 4377619, 1983.
67. Brusic V., Frankel G.S., Roldan J., and Saraf R., *J. Electrochem. Soc.* 1995; 142: 2591–2594.
68. Wang P.I., Lu T.M., Murarka S.P., and Ghoshal R., US Pending Patent 20050236711, 2005.
69. Li Y. and Wong C.P. US Patent Pending, 2006.
70. Li Y. and Wong C.P., Monolayer-protection for electrochemical migration control in silver nanocomposite. *Appl. Phys. Lett.* 2006; 81: 112112.
71. Prinz G.A., *Science* 1998; 282: 1660.
72. Toshioka H., Kobayashi M., Koyama K., Nakatsugi K., Kuwabara T., Yamamoto M., and Kashihara H., *SEI Tech. Rev.* 2006; 62: 58–61.
73. Chang D.C. and Smith E.L., US patent #5206585, 1993.
74. Lieber C.M., Nanowire nanosensors for high sensitive and selective detection of biological and chemical species. *Science* 2001; 293: 1289–1292.
75. Martin C.R. and Menon V.P., Fabrication and evaluation of nanoelectrode ensembles. *Anal. Chem* 1995; 67: 1920–1928.
76. Xu J.M., *Appl. Phys. Lett.* 2001; 79: 1039–1041.
77. Russell T.P., Ultra-high density nanowire array grown in self-assembled di-block copolymer template. *Science* 2000; 290: 2126–2129.
78. Li Y., Moon K., and Wong C.P., In: Proceedings of 56th IEEE Electronic Components and Technology Conference. IEEE, NJ, 2006; 1239–1245.
79. Li Y., Zhang Z., Moon K., and Wong C.P., Ultra-fine pitch wafer level ACF (anisotropic conductive film) interconnect by in-situ formation of nano fillers with high current carrying capability. US Patent Pending, 2006.

80. Moon K., Jiang H., Zhang R., and Wong C., In-situ formed nanoparticles in polymer matrix for low pressure bonding, GTRC Invention Disclosure No. 4443, 2008.
81. Snow E.S., Novak J.P., Lay M.D., Houser E.H., Perkins F.K., and Campbell P.M., *J. Vac. Sci. Technol. B* 2004; 22(4): 1990.
82. Lay M.D., Novak J.P., and Snow E.S., *Nano Lett.* 2004; 4(4): 603.
83. Shiraishi M., Takenobu T., Iwai T., Iwasa Y., Kataura H., and Ata M., *Chem. Phys. Lett.* 2004; 394: 110.
84. Meitl M.A., Zhou Y., Gaur A., Jeon S., Usrey M.L., Strano M.S., and Rogers J.A., *Nano Lett.* 2004; 4(9): 1643.
85. Zhou Y., Gaur A., Hur S., Kocabas C., Meitl M.A., Shim M., and Rogers J.A., *Nano Lett.* 2004; 4(10): 2031.
86. Li Z., Dharap P., Nagarajaiah S., Barrera E.V., and Kim J.D., *J. Adv. Mat.* 2004; 16: 640.
87. Abraham J.K., Philip B., Witchurch A., Varadan V.K., and Reddy C.C., *Smart Mater. Struct.* 2004; 1045: 13.
88. Wu Z., Chen Z., Du X., Logan J.M., Sippel J., Nikolou M., Kamaras K., Reynolds J.R., Tanner D.B., Hebard A.F., and Rinzler A.G., *Science* 2004; 305: 1273.
89. Ferrer-Anglada N., Kaempgen M., Skákalová V., Dettlaf-Weglowska U., and Roth S., *Diam. Relat. Mater.* 2004; 13(2): 256.
90. Hu L., Zhoa Y-L., Ryu K., Zhou C., Stoddart J.F., and Grüner G. *Advanced Materials* 2008; 20: 5939–5946.
91. Gruner G., *J. Mater. Chem.* 2006; 16: 3533–3539.
92. Ago H., Petritsch K., Shaffer M.S.P., Windle A.H., and Friend R.H., *Adv. Mater.* 1999; 11: 1281.
93. Du Pasquier A., Unalan H.E., Kanwal A., Miller S., and Chhowalla M., *Appl. Phys. Lett.* 2005; 87: 203511.
94. Shan B. and Cho K., *Phys. Rev. Lett.* 2005; 94: 236602.
95. Mitschke U. and Bauerle P., *J. Mater. Chem.* 2000; 10: 1471.
96. Kaempgen M., Duesberg G.S., and Roth S., *Appl. Surf. Sci.* 2005; 252: 425–429.
97. Kaiser A.B., Dusberg G., and Roth S., *Phys. Rev. B* 1998; 57: 1418.
98. Liu K., Roth S., Duesberg G.S., Kim G.-T., and Schmid M., *Progress in Molecular Nanostructures*, H. Kuzmany, J. Fink, M. Mehring, S. Roth (Eds.). American Institute of Physics, New York, 1998; AIP Conference Proceedings 442, 61–64.
99. Shirashi M., *Synth. Met.* 2002; 9198: 1.
100. Duesberg G.S., Graham A.P., Kreupl F., Liebau M., Seidel R., Under E., and Hölein W., *Diam. Rel. Mater.* 2004; 13: 354.
101. Lu K.L., Lago R.M., Chen Y.K., Green M.L.H., Harris P.J.E., and Tsang S.C., *Carbon* 1996; 34: 814.
102. Trancik J.E., Barton S.C., and Hone J., *Nano Lett.* 2008; 8(4): 982–987.
103. Ouyang H.Y. and Yang Y., *Adv. Mater.* 2006; 18: 2141–2144.

Biomimetic Lotus Effect Surfaces for Nanopackaging

Yonghao Xiu and C.P. Wong

Abstract This chapter addresses fundamentals on biomimetic nano Lotus surfaces, their chemistry, and nanostructure texture physics. In addition, preparation methodologies for various nano-textured superhydrophobic surfaces, their applications, and future researches are presented.

Keywords Superhydrophobic surface · Lotus surface · Nano texture · Nanostructure · Bio packaging · Biomimetic nanosurface · Bio-assay · Solar cell · Microfluidics · MEMS

1 Introduction

Lotus effect or superhydrophobicity was first described by Barthlott in 1996 [1, 2]. When he studied lotus leaf surfaces, he found that the surface is rough instead of smooth. Combined with hydrophobic surface layers, a superhydrophobic state was observed. Water droplets falling on the surface bead up and roll off the surface instantaneously. Superhydrophobicity requires that both a surface hydrophobic layer and rough surface structures are present. This phenomenon is quite common in the biology arena, e.g., on some insects and plants leaves, shown in Fig. 1. Artificial superhydrophobic surfaces were prepared before the phenomenon was widely acknowledged. In 1986, Sacher et al. reported that very water-repellent films were obtained by depositing hexamethyldisiloxane on Si chips in a high-energy plasma at low temperature [3]. Drops of water had contact angles as high as 180° and rolled off the slightly inclined surfaces, which suggested a near-zero roll off angle. Morra et al. had prepared a superhydrophobic surface by O₂ plasma etching on PTFE surfaces in 1989 [4]. Actually after 5 min of treatment, water drops rolled easily across the surface. Due to the absence of hysteresis, no obstacle to their movement occurred. The first super water- and oil-repellent surface was

Y. Xiu (✉)

School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: yxiu@gatech.edu

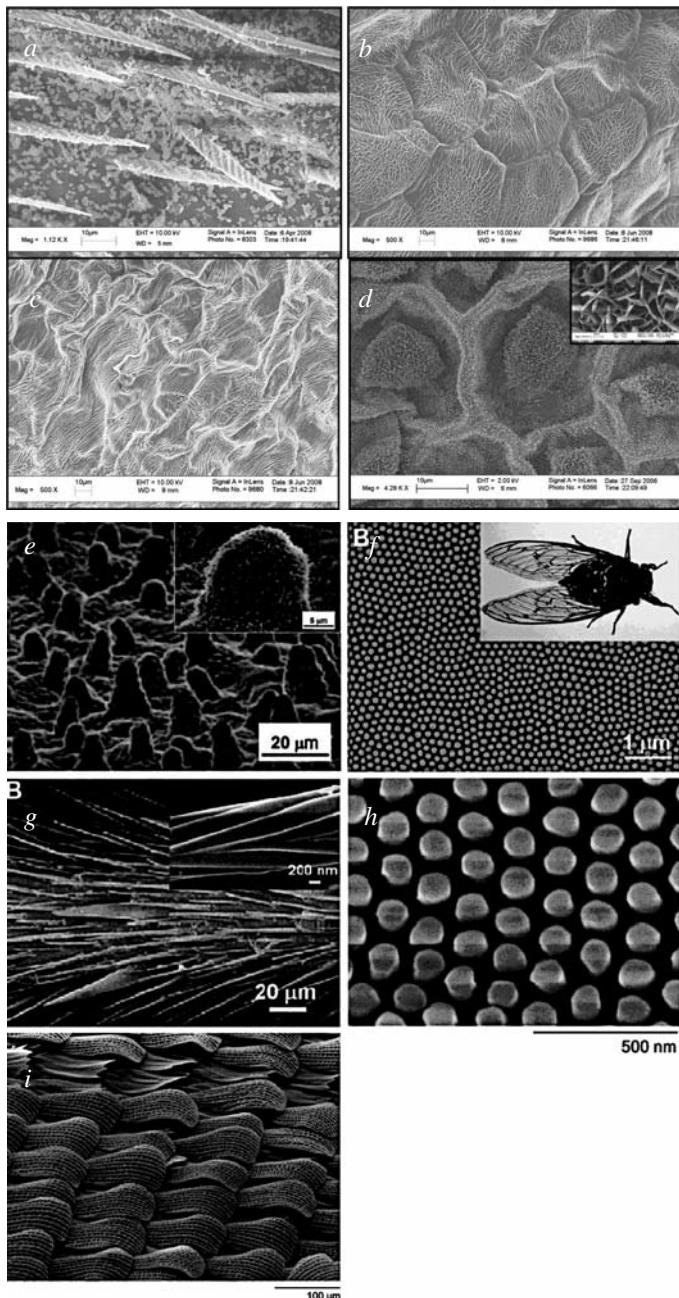


Fig. 1 SEM images of different bio-surface structures: (a) Wasp wing, (b) Azaleas petal, (c) rose petal, (d) Colocasia esculenta leaf surface, (e) micro- and nanostructures on the lotus leaf (*Nelumbo nucifera*) [6], (f) cicada wing [7], the leg with numerous oriented spindly microsetae, (inset) nanoscale groove structure on a seta, (h) surface of mosquito (*culex pipiens*) eye [9], (i) the wings of *Papilio ulysses*; the way the tiles are displayed together with the detail of the texture confer anisotropy to the texture [10]

developed in Tsujii's group [5] on an anodically oxidized aluminum surface after a hydrophobic treatment.

The effect of roughness on surface wetting is scale-dependent. Besides the surface hydrophobicity, the surface roughness can be more important in achieving dewetting surfaces even without the most hydrophobic materials (e.g., fluorine containing polymers). On flat surfaces oil repellency is difficult to achieve. However, with appropriate surface roughness with special geometry designs, a superoleophobic surface may result. This signifies the importance of surface wetting/dewetting control via flexible, fine-tuned surface structure design at micro- and nanoscales.

The microelectronic industry is developing very fast. Microsystems include automotive, computer and business equipment, communications, consumer, industrial and medical, military, and aerospace. Packaging provides electrical connections for signal transmission, power input, and voltage control. It also provides for thermal dissipation and the physical protection required for reliability. Today assembly and packaging are a limiting factor in both cost and performance for electronic systems. Wireless and mixed signal devices, bio-chips, optoelectronics, and MEMS have placed new requirements on packaging and assembly. With the device size reduction, it poses even harsher requirement for the packaging which gives it protection not only for reliability but also for signal propagation (high speed), thermal dissipation, and power insulation.

There are several areas that require a water/moisture resistant surface layer. For polymer packaging, reduction in moisture absorption and diffusion is the most important in achieving system reliability. For MEMS, hermetic sealing is very important to prevent capillary force-induced failure. In addition, moisture diffusion can cause corrosion of the fine metal lines. Since package must function in humid environments, methods to reduce moisture absorption or diffusion would be essential. In recent years, the superhydrophobic surface is attracting intense research interest due to the promising applications in self-cleaning and moisture/water repellency, such as in anticorrosion, antistiction of MEMS device packaging, microfluidics, solar cell packaging, bio-assay. These applications will be discussed in detail in Section 6.

2 Thermodynamics Aspect of Superhydrophobic Surfaces

2.1 Contact Angle on a Rough Surface

On a flat surface, the change in surface free energy, ΔG , accompanying a small displacement of the liquid such that the change in area of solid covered, ΔA , is

$$\Delta G = \Delta A_{SL}(\gamma_{SL} - \gamma_{SV}) + \Delta A_{LV}\gamma_{LV} \quad (1)$$

$$\Delta G = \Delta A(\gamma_{SL} - \gamma_{SV}) + \Delta A\gamma_{LV}\cos(\theta - \Delta\theta) \quad (2)$$

At equilibrium,

$$\Delta = 0$$

And according to Young's equation

$$\gamma_{SL} - \gamma_{SV} + \gamma_{LV}\cos\theta = 0$$

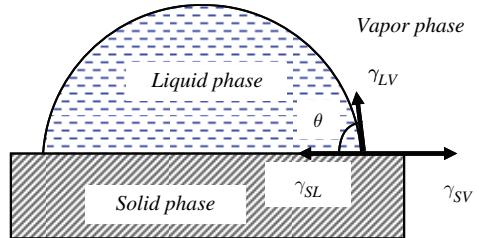
or

$$\cos\theta_Y = \frac{\gamma_{SV} - \gamma_{SL}}{\gamma_{LV}} \quad (3)$$

Here, γ_{SV} , is the surface tension at solid/vapor interface, γ_{SL} , is the interface tension at solid/liquid interface; and γ_{LV} , is the surface tension at liquid/vapor interface.

On a smooth surface, this Young's equation (3) determines contact angles of liquid droplet as shown in Fig. 2.

Fig. 2 The illustration of relationship between the surface and interface tensions according to Young's equation



On a rough surface, however, Young's equation no longer holds for the apparent contact angles due to the introduction of a new factor: surface roughness. Water droplet/solid interface contact can be either in Wenzel state or in Cassie state (shown in Fig. 3), depending on the surface hydrophobicity and geometries of the surface structures.

Usually two equations were used for the calculation of the apparent contact angles corresponding to the two regimes. One is Wenzel equation as shown in Equation (4):

$$\cos\theta_A = r \cos\theta_Y \quad (4)$$

This equation described the relation between the surface roughness r and the apparent contact angle, where r is determined by the ratio of total surface area to the projected surface area. This equation is valid for a complete solid/liquid contact interface.

For heterogeneous surfaces, Cassie equation can be used to describe the heterogeneity effect on the apparent contact angle [11],

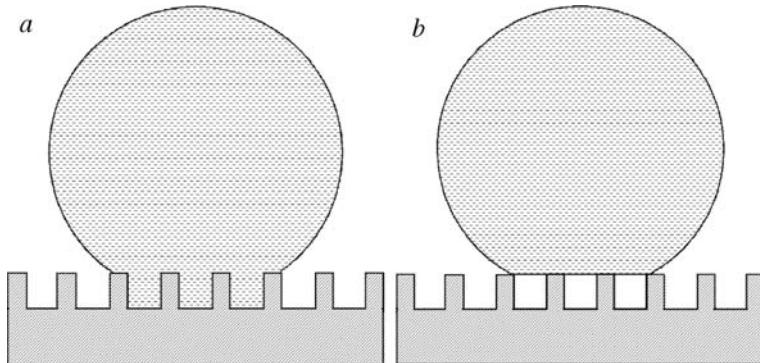


Fig. 3 The two contact regimes of water droplet with rough surfaces: **(a)** Wenzel contact regime, **(b)** Cassie contact regime

$$\cos\theta_A = f \cos \theta + f - 1 \quad (5)$$

where θ_A is the apparent contact angle on the rough surface, f is the projected surface fraction of solid, and θ_Y is the contact angle on a flat surface as per Young's equation.

The surface roughness r can be described by Wenzel roughness factor:

$$r = \frac{A_{\text{Actual}}}{A_{\text{projected}}} \quad (6)$$

The development of Wenzel and Cassie equations is shown below. For water droplets on rough surfaces, according to Equation (2),

$$\Delta G = \Delta A r (\gamma_{SL} - \gamma_{SV}) + \Delta A \gamma_{LV} \cos(\theta - \Delta\theta)$$

And the equation reduced to

$$r(\gamma_{SL} - \gamma_{SV}) + \gamma_{LV} \cos \theta_A = 0 \quad (7)$$

Therefore,

$$\cos \theta_A = r \frac{\gamma_{SL} - \gamma_{SV}}{\gamma_{LV}} = r \cos \theta_Y$$

This is the Wenzel equation shown in Equation (4) [12].

Similarly for Cassie equation, considering the fraction of solid/air and solid/liquid interface,

$$f_1(\gamma_{S1V} - \gamma_{S1L}) + f_2(\gamma_{S2V} - \gamma_{S2L}) - \gamma_{LV} \cos \theta_A = 0$$

or

$$\gamma_{LV} \cos \theta_A = f_1(\gamma_{S1V} - \gamma_{S1L}) + f_2(\gamma_{S2V} - \gamma_{S2L}) \quad (8)$$

When S_2 is also air, $\gamma_{S2V} = 0$ and $f_1 + f_2 = 1$, Equation (8) reduced to Cassie equation (Equation 5)

If the micropillar surface is composed of a second scale roughness, then the Cassie equation can be written as [13, 14],

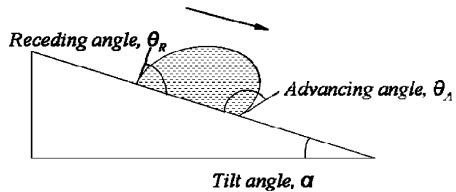
$$\cos \theta_A = rf \cos \theta_Y + f - 1 \quad (9)$$

where r is the Wenzel roughness factor of the structures at the solid–liquid contact interface.

2.2 Contact Angle Hysteresis

The roll-off feature of superhydrophobic surfaces can be characterized by the roll-off angle which is defined by the tilt angle at which a water droplet sitting on top starts to move; it can also be characterized by the contact angle hysteresis which is the difference between the advancing angle and the receding angle when the water droplet starts to move. They are illustrated in Fig. 4.

Fig. 4 The illustration of the tilt angle and the advancing and receding angles



The relation between the contact angle hysteresis and the tilt angle can be illustrated in the following equation:

$$\sin \alpha = \frac{\gamma d(\cos \theta_R - \cos \theta_A)}{mg} \quad (10)$$

where d is the diameter of the moving droplet, m is the mass of the droplet.

To achieve roll-off superhydrophobicity, the contact angle hysteresis must be controlled. Only when the hysteresis is small can the surface be self-cleaning (usually $<10^\circ$). If the surface contact angle hysteresis is high, the droplet will be difficult to roll off, instead, it will just stick to the rough surface. Therefore self-cleaning effect cannot be effectively achieved.

3 Low Surface Energy Materials

A variety of materials have been used to prepare superhydrophobic surfaces including both organic and inorganic materials. For polymeric materials, which are generally inherently hydrophobic, fabrication of surface roughness is the primary focus.

For inorganic materials, which are generally hydrophilic, a surface hydrophobic treatment is performed after the surface structures are fabricated.

Most polymeric materials can be used to fabricate superhydrophobic surfaces. For instance, plasma etching has been employed to generate roughness on polymer surfaces to achieve superhydrophobicity; polyethylene was etched by CF₄/O₂ plasmas [15] or direct catalytic polymerization process [16], etc. In addition, transparent and superhydrophobic surfaces can be achieved on poly(ethylene terephthalate) surfaces [17]. Examples of inorganic materials that have been used to create superhydrophobic surfaces are silica, alumina, hematite, boehmite, titania, and copper can be superhydrophobic if the roughened surfaces are treated appropriately with a hydrophobic material [18, 19]. In this sense, even for inorganic, the top surface can be covered with an organic layer or monolayer for the reduced surface energy.

Low energy materials are particularly important in achieving stable superhydrophobic surfaces; among these materials, fluoropolymer/fluorocarbon is one of the lowest. The F–C bond is one of the strongest covalent bonds. In addition, the bulky F atoms cover well the C atoms beneath the surface (compared to H atoms), thus preventing chemical attack of the weaker C–C bonds. These facts lead to the significant potential of fluorocarbon layers for self-cleaning coatings. Indeed, Teflon has been roughened to produce superhydrophobic surfaces. Short-chain fluorocarbon silanes have been also investigated to form monolayers on a rough surface to achieve superhydrophobicity (10 Å thickness is enough to give a hydrophobic layer without any effect of the underlying substrates) [20]. Studies have shown that the surface energy with CF₃ groups is lower than that of CF₂ covered surfaces [21]. Another very important low surface energy material is silicone. The siloxane backbone shows very high bond energy and the side CH₃ gives the low surface energy.

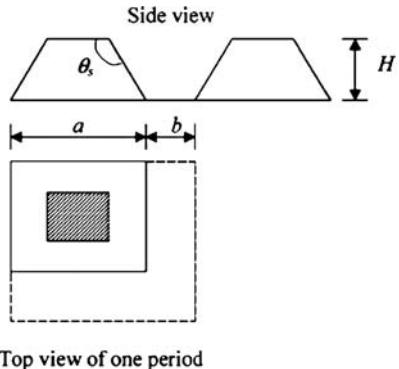
4 Surface Structure Effect

For superhydrophobic surfaces, the surface structure/roughness is essential; the geometrical shape is also critical in achieving the desired surface superhydrophobicity. Patankar investigated the energy barrier between the Cassie and the Wenzel states for surface geometries of inclined structures. He demonstrated that an energy barrier to prevent the transition from a Cassie high energy state to a Wenzel low energy state as shown in Fig. 5 may not necessarily be present; that is, the transition may be spontaneous.

A Cassie drop will be formed only if a liquid–vapor interface can connect from pillar to pillar without touching the valleys. Such an interface can establish equilibrium at the top corners of the pillars only if $\theta_s \leq \theta_e \leq 180^\circ$. For other values of θ_e a Cassie drop is not possible and the valleys will always be filled with liquid.

A pyramidal Si surface represents a good model surface for the investigation of this structure inclination effect. It was shown mathematically that the inclination angle is very important in maintaining the Cassie state. As the inclination angle θ_s increases, the difficulty in achieving a Cassie superhydrophobic state increases. As

Fig. 5 Side and top views of geometrical roughness obtained by periodically placed pillars with inclined side walls [26]



predicted, the liquid–vapor interface on the composite surface cannot be maintained and the droplet is always in the Wenzel state. However, by implementing a second scale of surface roughness on the pyramid surface, the liquid–vapor interface can be maintained on the pyramid top and a Cassie state can be effectively achieved.

A superoleophobic surface has been also reported by using the structure shown in Fig. 6a. In order to achieve oil repellency, a re-entrant (or overhang) structure design was employed. On this surface, a high contact angle can be achieved even though the surface is hydrophilic ($\theta < 90^\circ$). The reentrant structure was also fabricated in electrospun polymer/polyhedral oligomeric silsesquioxane (POSS) composite fibers as shown in Figs. 6b and 7.

Similar results have been also reported by fabrication of a nanonail structure on Si surfaces using plasma etching techniques [23]. Instead of attempting to guarantee that the Cassie–Baxter state always remains the minimum-energy state, this

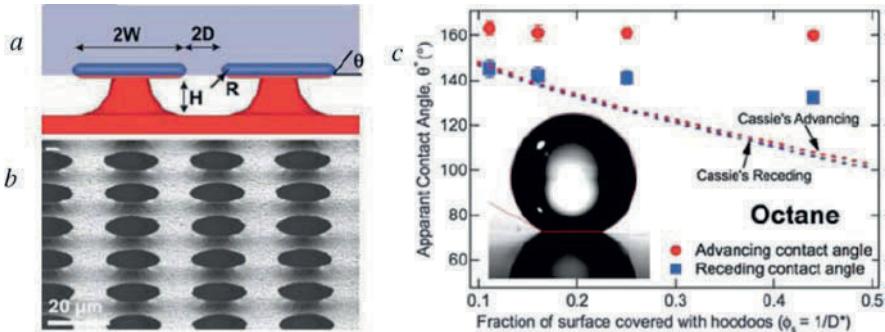
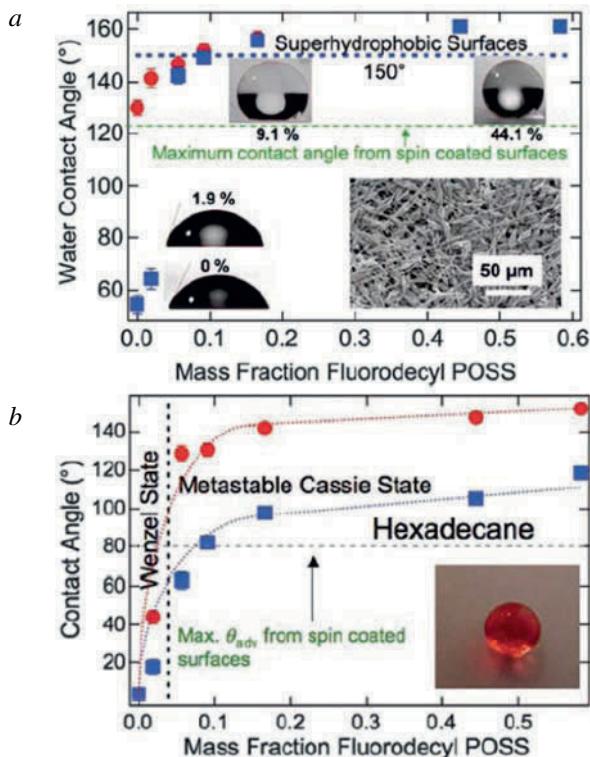


Fig. 6 (a) Cartoon highlighting the formation of a composite interface on surfaces with re-entrant topography. The geometric parameters R , D , H , and W characterizing these surfaces are also shown. The blue surface is wetted, while the red surface remains nonwetted when in contact with a liquid whose equilibrium contact angle is $\theta_Y (< 90^\circ)$; (b) SEM micrographs for so-called micro-hoodoo surfaces having circular flat caps. The sample is viewed from an oblique angle of 30° ; (c) contact angles for octane on silanized micro-hoodoos as a function of $f(\phi_s)$ [22]

Fig. 7 (a) θ_{adv} (dots) and θ_{rec} (squares) for water on electrospun polymer surfaces. The inset shows a SEM micrograph for an electrospun surface containing 9.1 weight % POSS; (b) θ_{adv} (dots) and θ_{rec} (squares) for hexadecane on the electrospun surfaces, the inset of (a) shows a drop of hexadecane (dyed with oil red O) on a 44 weight % fluorodecyl POSS electrospun surface [22]



approach created a system in which the height of the energy barrier that separates the metastable Cassie–Baxter state from the stable Wenzel state is designed to be sufficiently high to achieve effective pinning of the liquid in the desired nonwetting state. The idea is to create a special type of three-dimensional surface topography that normally inhibits transitions from the Cassie–Baxter state to the Wenzel state but that would allow such transitions under the influence of external stimuli, such as electrical fields. A particular example of such topography is the “nanonail” structure in Fig. 8.

Porous Si surfaces from Au-assisted electroless etching process have shown the ability to achieve an oil-repellent surface [24, 25]. These structures show similar overhang character to that for nanonail structures as shown in Fig. 9; droplet contact angles of water and various organic solvents are shown in Fig. 10.

Achievement of high contact angles for both polar (water) and non-polar droplets (e.g., hexadecane) requires the design of surface structures similar to those shown in Fig. 10. In order to achieve mechanical stability and optimized hierarchical surface structures for enhanced self-cleaning on superhydrophobic surfaces, it is critical to understand the fundamental surface physics, especially the surface–structure effect on superhydrophobicity to improve control over the surface structures (geometrical

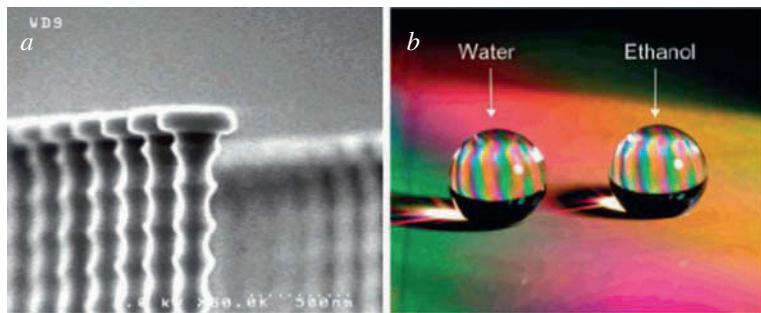


Fig. 8 (a) Scanning electron microscopy (SEM) image of 2- μm -pitch nanonails [23]. The nail head diameter D is ~ 405 nm, the nail head thickness h is ~ 125 nm, and the nanonail stem diameter d is ~ 280 nm; (b) a nanonail-covered substrate in action. Droplets of two liquids with the very different surface tensions, 72 mN/m (water) and 21.8 mN/m (ethanol), sit next to each other on the 2- μm -pitch nanonail substrate

Fig. 9 Porous Si surface fabricated by Au-assisted electroless etching. (a) Top view SEM image and (b) schematic cross-sectional profile of water in contact with the porous Si surface [24]

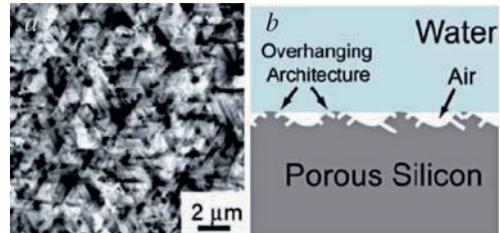
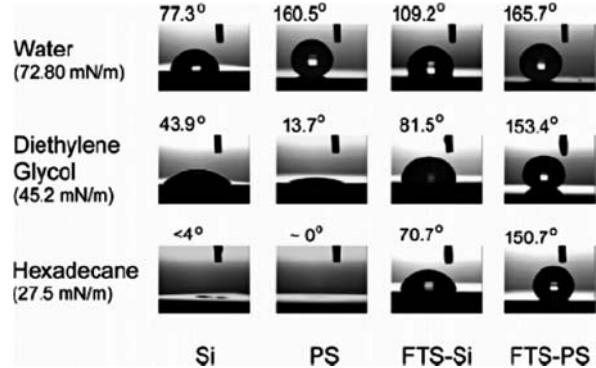


Fig. 10 Static contact angles of water, diethylene glycol, and hexadecane on flat silicon (Si), porous silicon (PS) with tilted pores, flat Si coated with FTS (FTS-Si), and porous silicon with tilted pores coated with FTS (FTS-PS) [24]



shapes, pitch, height, aspect ratio, hierarchical design, etc). These design criteria can then be used to achieve the requirements necessary for various applications.

As described previously, for surface superhydrophobicity, there are two requirements. The first is high contact angle ($>150^\circ$), which relates the contact angle to the surface structure (expressed by the Wenzel surface roughness factor r) and can

be described by the improved Cassie–Baxter equation [13, 26]. The second parameter is the contact angle hysteresis, which is defined by the difference between the surface advancing contact angle and the surface receding contact angle.

5 Approaches to Preparing Superhydrophobic Surface Coatings

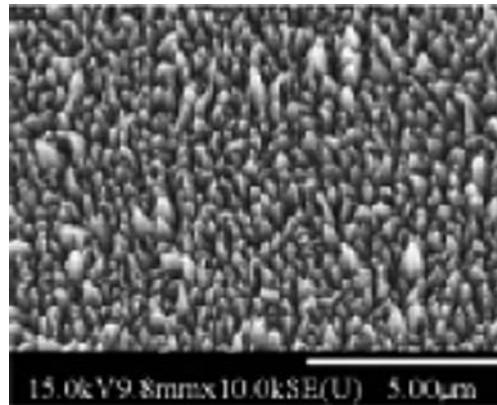
5.1 Plasma Etching Techniques

Plasma techniques are widely employed in microelectronics processing (Si surface pattern etching, cleaning, removal of photoresists, and dielectric film deposition such as inorganic or polymer materials) [27]. Plasma etching techniques are also capable of surface roughness generation and therefore offer a method to fabricate superhydrophobic surfaces.

5.1.1 Plasma Etching of Fluoropolymers

Using plasma etching techniques, polymer materials ranging from PTFE to silicones to polyethylene can be etched to produce a rough surface. To roughen PTFE, the technique of radio-frequency sputtering has been reported [4, 28]. For long sputtering times, a superhydrophobic Teflon surface was prepared by use of an O₂ plasma; the resulting surface showed a water contact angle of 168° [29]. Plasma ion beam treatment using O₂ and Ar have been also reported in preparation of superhydrophobic PTFE coatings [30] with a surface morphology as shown in Fig. 11.

Fig. 11 The SEM surface morphology of plasma-treated PTFE surface by Ar/O₂ (2 sccm/2 sccm) 1.5 KeV [30]



By using a simple plasma-based technique that combines etching and plasma polymerization on Si substrates, superhydrophobic surfaces have been fabricated by tailoring their surface chemistry and surface topology [31]. This technique showed the capability of large area processibility and good reproducibility. Using SF₆ as an etchant, surface roughness can be produced by over-etching the photoresist layer.

Overall, the method to generate superhydrophobic surfaces consisted of three steps. First, a 2.3- μm -thick photoresist polymer layer (microposit S1813 supplied by Shipley Company) is deposited by spin coating on a Si wafer. Then, photoresist layers are etched using an inductively coupled plasma with a magnetic pole-enhanced ICP (MaPEICP) source. Third, the CFx layer is deposited on etched photoresist/Si wafers from C₄F₈ using the same plasma reactor but with the 13.56 MHz r.f. capacitive plasma mode. The key step of this process is etching of the surface topography into the substrate to create high roughness before deposition of the fluorocarbon coating. Superhydrophobicity is achieved after the C₄F₈ plasma polymer deposition process.

5.1.2 Plasma Etching of Si with Micromask

SF₆ and CHF₃ plasmas were used to investigate the surface roughness generation on Si surfaces for surface wettability modification [32]. Due to the nonuniform removal of photoresist (Microposit S1813, Shipley, spincoating at 2,000 rpm), the residue that remains on the surface acts as localized masks to create roughness on the Si surface. After the surface roughness was formed as shown in Fig. 12, a CFx layer was deposited on top of the rough surface using plasma-enhanced chemical vapor deposition (PECVD). The surfaces treated by the CHF₃ plasma are all superhydrophobic up to 10 min of etching, while SF₆ etching will not produce superhydrophobic after 2.5 min etching; these results are due to differences in Si etch rate with the two fluorocarbon gases. After 5 min etching for SF₆, the surface contact angle becomes 116°, which is approximately the contact angle on a flat surface for CFx coatings.

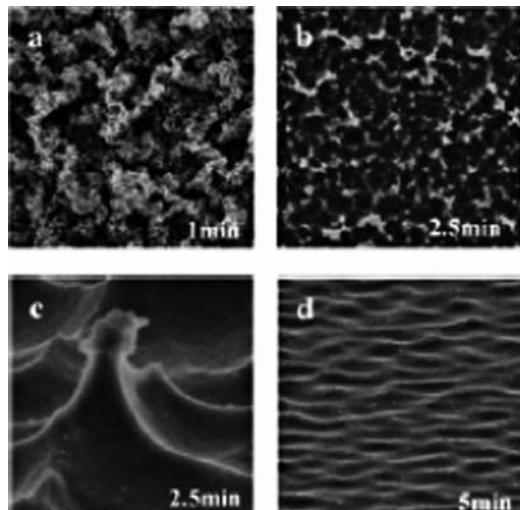


Fig. 12 SEM image of Si surface etched by SF₆ for (a) 1, (b, c) 2.5, and (d) 5 min. Image (c) is a magnified view (with an observation angle of 45°) of one residual photoresist particle on a peaked feature. The areas in images (a), (b), and (d) are 20 × 20 μm ; the area of image (c) is 2 × 2 μm [32]

A detailed analysis of roughness generation was conducted in photoresist-masked isotropic Si plasma etch process performed in an inductively coupled plasma (ICP) etcher with SF₆, by systematically changing the pressure (2.5–70 mTorr), SF₆ flow rate (50–300 sccm), platen power (0–16 W), and ICP power (1,000–3,000 W) [33]. A high ICP power (3,000 W) and a high flow rate (300 sccm) result in the lowest normalized bottom roughness. High pressures and low flow rates increase the roughness. Increasing the ion energy by applying a radio frequency platen power of 16 W has a great impact on the roughness. Actually the bottom surface becomes completely smooth. The roughness morphology at these process conditions shows a crystallographic anisotropy, where (111) planes are revealed. When an oxide mask was used, the surface roughness was reduced compared to that for a photoresist mask. When there is no mask, the surface roughness is very low which suggests an effect due to the photoresist mask layer (which redeposits during etching) on the surface roughness.

5.1.3 Plasma Etching of Poly(Dimethylsiloxane) (PDMS)

Polydimethylsiloxane is a very low surface energy material with a surface energy of 21–22 (mJ/m²), slightly higher than that of *poly(tetrafluoroethylene)* (PTFE) (~18 mJ/m²). However, after plasma treatment, the surface may show a much higher surface energy. For instance, superhydrophobic PDMS has been fabricated by an SF₆ plasma treatment followed by fluorocarbon film deposition that generated hydrophobic high aspect ratio columnar-like nanostructures as shown in Fig. 13 [34].

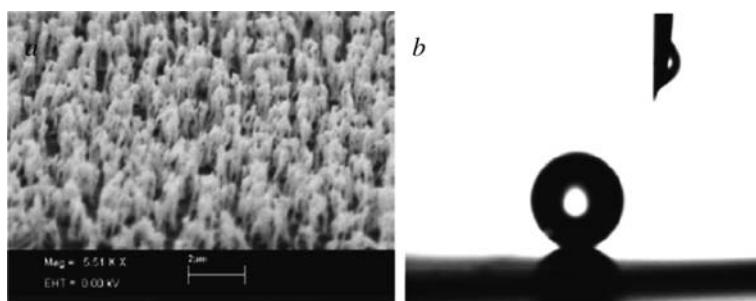


Fig. 13 SEM image of a PDMS elastomer (Sylgard 184) surface after a 6 min SF₆ plasma treatment. 1.45 μm-high nano-columns are shown. (b) Image of a water droplet rolling off the surface in (a) after being conformally coated with a 20 nm-thick FC film [34]

In addition, under appropriate plasma processing conditions (SF₆ plasma etching for a sufficient time to generate the appropriate surface roughness), superhydrophobic and simultaneously transparent surfaces can be produced [34], which is of great practical importance. In particular, nanostructured PDMS surfaces are visibly transparent for treatment times up to 2 min (height of column: 130 nm), although they become opaque for etching times longer than 4 min.

5.1.4 Plasma Etching of Polyethylene Terephthalate (PET) (Domain-Selective Plasma Etching) and Low-Density Polyethylene (LDPE, Different Etching Rate for Crystalline and Amorphous Phase)

To establish superhydrophobicity of PET, a method consisting of a two-step process comprised of nano-texturing by an oxygen plasma treatment and subsequent deposition of a hydrophobic coating by means of low-temperature chemical vapor deposition or plasma-enhanced chemical vapor deposition was employed to form ultra water-repellent polymer sheets [17, 35, 36]. After oxygen etching, many protrusions are observed on the polymer surface. In general, noncrystalline domains are more readily etched than the crystalline domains. PET contains crystalline and amorphous domains and the protrusions of plasma-etched PET were probably formed as a result of domain-selective plasma etching. This surface nano-texture remained after deposition of the hydrophobic coatings using organosilane precursors; surface topologies are shown in Fig. 14. The surface-modified substrate was transparent (the polymer is optically transparent) and ultra water-repellent, showing a water contact angle greater than 150°.

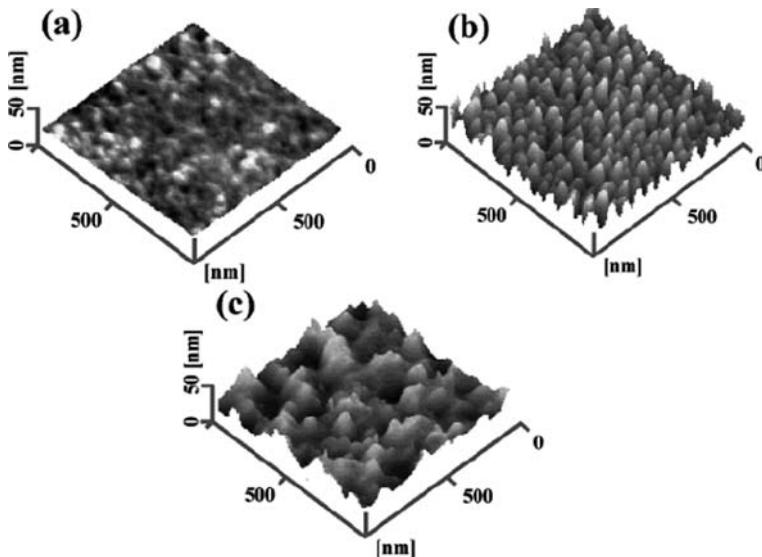
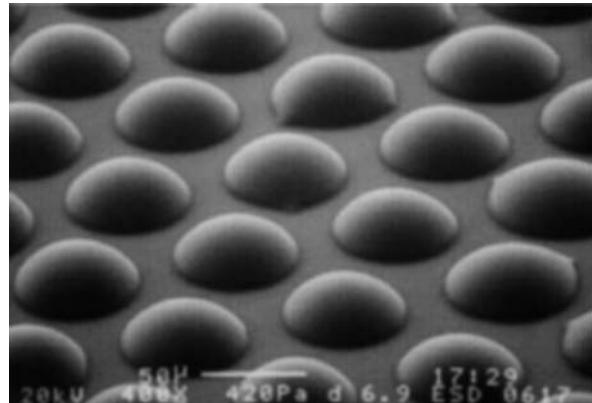


Fig. 14 AFM of PET (a) untreated, (b) O₂ plasma treatment, and (c) FAS coating after O₂ plasma [36]

By selectively patterning a surface (deposited fluorocarbon polymer) a micro-patterned polymer substrate with superhydrophobic/superhydrophilic domains as shown in Fig. 15 can be prepared. Such surfaces are extremely useful in chemical and biological sciences to keep surfaces from retaining or being fouled by samples and reagents.

To achieve superhydrophobicity on LDPE surfaces, a similar process has been employed due to the fact that LDPE also shows noncrystalline and crystalline

Fig. 15 Environmental scanning electron microscope (ESEM) image of micro-patterned water droplets on a superhydrophobic/superhydrophilic surface prepared by a multi-step drying process [37]



domains [38–40]. Two routes to achieve superhydrophobicity are possible. The first corresponds to a one-step synthesis (CF₄ plasma modification of LDPE). The second route involves two steps (O₂ plasma treatment followed by CF₄ plasma deposition onto LDPE). However, these surfaces tend to lose their superhydrophobic behavior after water vapor condensation.

Superhydrophobicity can also be achieved by polystyrene (PS) etching in CF₄–O₂ [41]. By controlling the etching parameters, the surface wetting characteristics can be tailored from sticky to slippery superhydrophobicity. Using this process, the optical transparency can be maintained on the PS surfaces (Fig. 16).

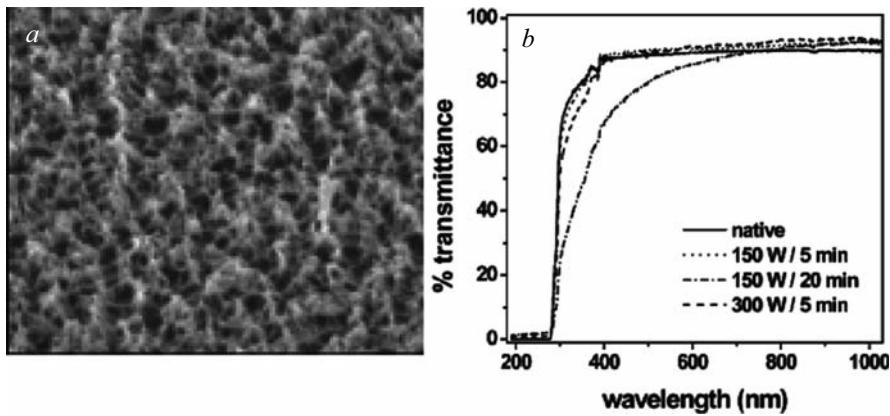
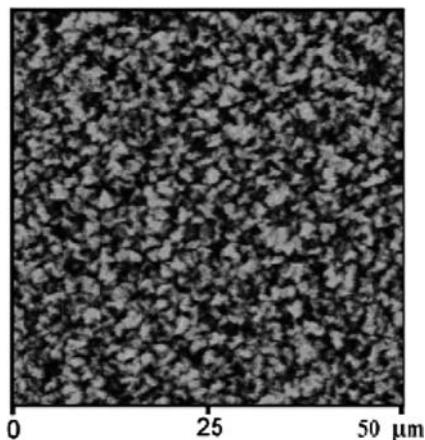


Fig. 16 (a) SEM image of PS samples treated with a CF₄–O₂ discharge with 17% O₂ for 5 min [41]; 300 W 5-min plasma etching; (b) transmittance of the prepared surfaces

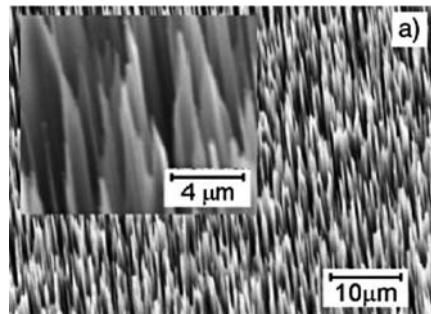
Woodward et al. used a two-step CF₄ plasma etching and fluorination process followed by curing to prepare a fluorinated rough polybutadiene surface (Fig. 17) [42]. This surface exhibited superhydrophobicity with a water contact angle of $\sim 175^\circ$ and contact angle hysteresis of 0.4° [41].

Fig. 17 AFM height image of polybutadiene in CF₄ plasma (power 50 W) for 900 s [42]



Dorrer et al. reported a method to manipulate superhydrophobicity/superhydrophilicity by precise control of the surface chemistry using various polymer coatings on Si nanograss surfaces generated by anisotropic plasma etching (Fig. 18) [43]. When a hydrophobic polymer poly(heptadecafluorodecylacrylate) (PFA) was used to coated Si nanograss surfaces, superhydrophobicity (179° , no observable hysteresis) was achieved and the surface showed “condensation resistance” to water vapor.

Fig. 18 Electron microscopy images of a nanograss surface coated with a PFA thin film [43]



5.2 Chemical Vapor Deposition (CVD) Process

Silicone nanofilaments have been formed from trichloromethylsilane (TCMS) precursor via a gas phase reaction route at room temperature and pressure without a carrier gas [44]. Silanes with hydrolysable groups (chloride) react with water to form silanols, which then couple to hydroxyl groups at the surfaces, and subsequently polymerize to a silicone layer, depending on the reaction conditions. When dry reaction conditions with only traces of water were maintained, monolayers or

nearly monolayer thick coatings resulted. When equimolar amounts of liquid TCMS and water vapor were used in the reaction vessel, a superhydrophobic surface was achieved; the coating then consisted of polymethylsilsesquioxane nanofilaments. This technique can be used to coat a variety of surfaces such as cotton fiber, wood, polyethylene, ceramics, titanium and aluminum. In addition, on glass surfaces, a transparent coating can be achieved as shown in Fig. 19. When the as-deposited surface was treated in an O₂ plasma followed by a treatment with fluoroalkylsilane, oil-repellent surfaces resulted [45] as shown in Fig. 19d.

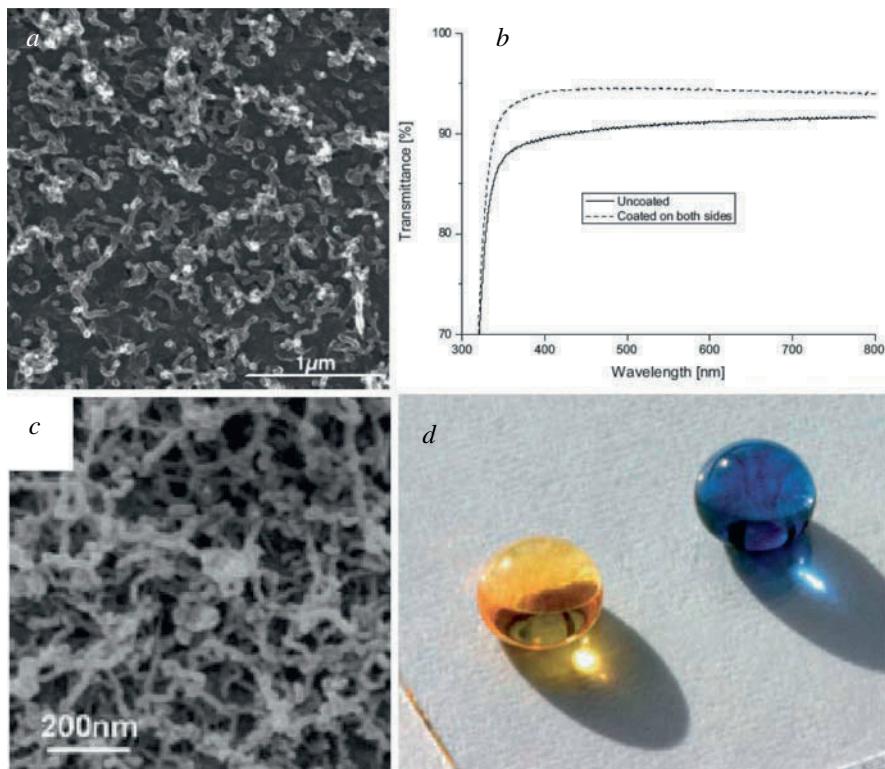
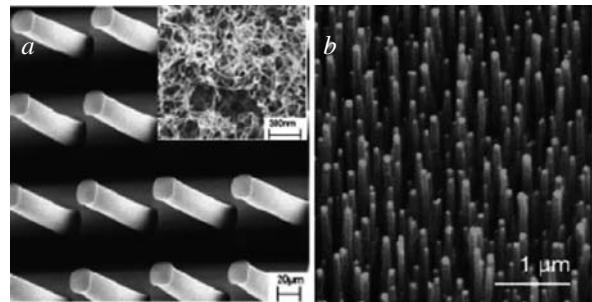


Fig. 19 Scanning electron microscopy images of silicone nanofilaments on Si. **(b)** UV-vis spectra of coated and uncoated glass slides [44]; **(d)** 10-ml drop of colored hexadecane (yellow) and colored water (blue) on a PFOTS-modified silicone nanofilament coating on glass [45]

Zhu et al. reported a method of superhydrophobic materials preparation by using two-tier CNT surfaces formed by growing patterned CNT arrays on CNT covered Si surfaces via CVD (Fig. 20a) [46]. Because of the presence of nanoroughness on the CNT film/array surface, a two-tier roughness can be achieved. Therefore, after coating the surfaces with a fluorocarbon film, superhydrophobicity was achieved with a contact angle of 167° and a hysteresis of 1°. A CNT nanoforest has been also prepared by PECVD (Fig. 20b) [47]. After the surface was coated with PTFE,

Fig. 20 (a) Typical SEM images of CNT arrays on Si substrates; (b) SEM images of carbon nanotube forests prepared by PECVD with nanotube diameter of 50 nm and a height of 2 μm [46, 47]



superhydrophobicity was achieved down to the microscopic level. Even micrometer-sized water droplets can be suspended on top of the nanotube forest.

5.3 Sol–Gel Process

Sol-gel processes can also produce rough surfaces on a variety of oxides such as silica, alumina, and titania [48, 49]. This approach is a very versatile process for the preparation of superhydrophobic thin films or bulk materials. The sol-gel process can form a flat surface coating, xerogel coating, or aerogel coating depending on the process conditions; both xerogel and aerogel show rough or fractal surfaces.

Silica aerogels can contain up to 15 wt% of absorbed air and water vapor. In order to decrease the pore affinity to absorption and thus make aerogels less hydrophilic, a number of strategies have been used, such as avoiding the presence of terminal hydroxyl groups by co-gelling certain Si precursors containing at least one non-polar chemical group, i.e., $\text{CH}_3\text{-Si}$ [50]. Superhydrophobic silica aerogels can also be obtained by surface modification of standard silica gels with a heavily fluorinated silyl chloride followed by supercritical fluid solvent removal [48]. After either ambient drying or supercritical drying, both surfaces show superhydrophobic properties (Fig. 21).

Attempts have been made to synthesize highly flexible and superhydrophobic silica aerogels using methyltrimethoxysilane (MTMS) as a precursor by a two-step acid–base sol–gel process [51] (Figs. 22 and 23). The aerogels consist of cross-linked networks of silica polymer chains extended in three dimensions. Due to the presence of non-polar alkyl groups (i.e., methyl) attached to the silica polymer chains, the inter-chain cohesion is minimized resulting in an elastic and flexible three-dimensional network. Increased dilution of the MTMS precursor with methanol solvent yielded a silica network with a low degree of polymerization which exhibited higher flexibility. In contrast, for lower dilution of the MTMS precursor, an extensive polymerization resulted in dense and rigid structures. Because of the new property, i.e., flexibility, imparted to the aerogel, it can be bent into a variety of shapes and acts as a good shock absorber as well. A water droplet contact angle of 164° can be achieved on such aerogel surfaces.

Fig. 21 Drop of water on top of fluorinated aerogel, inset shows a close-up [48]

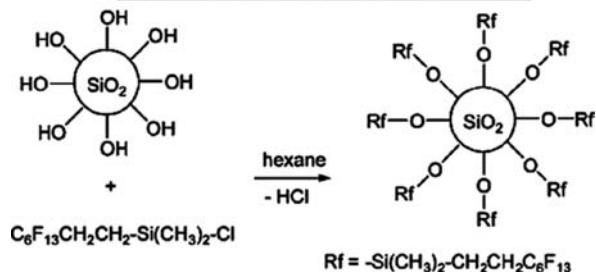
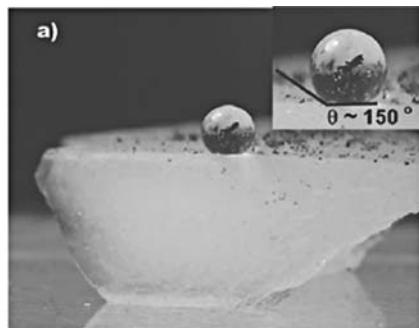
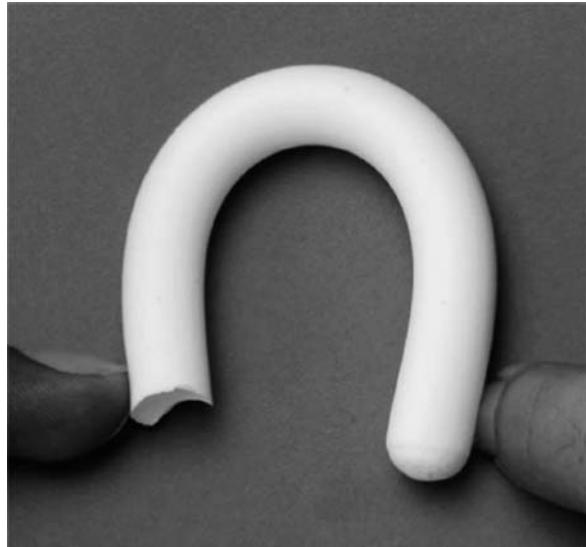
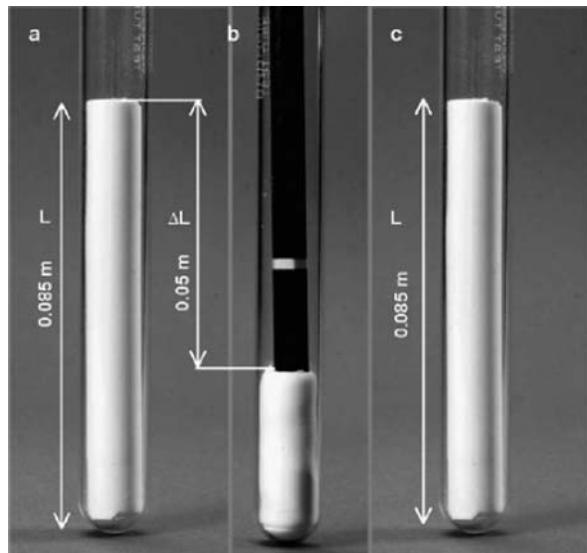


Fig. 22 Flexible silica aerogels prepared at two different MeOH/MTMS molar ratio of 35 [51]



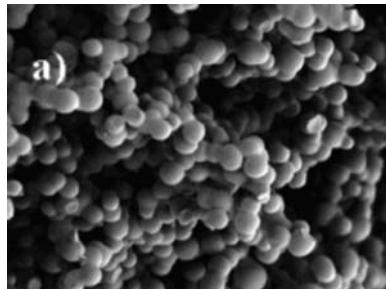
Another method to produce superhydrophobic silica materials is the phase separation method in which a hardening process “freezes” a phase separation that occurs concurrently with hardening of one of the phases [52]. This method produced co-continuous materials consisting of a solid phase and a liquid phase. When the liquid

Fig. 23 Photograph of the three states of a flexible aerogel sample: (a) without stress, (b) with stress, and (c) after releasing the applied stress [51]



was removed, a porous structure remained. The materials system was comprised of condensed organotriethoxysilane in a mixture of organic solvent and water. Such sol-gel materials show a great promise as stronger superhydrophobic surfaces and as bulk materials and are relatively inexpensive to produce. The reaction occurs through hydrolysis of the ethoxy groups and polymerization of the resulting silanol groups. Polymerization causes a decrease in dipole moment, leading to hydrophobic phase separation. The dried material has the organic groups on its surface, causing the foams to be superhydrophobic (Fig. 24).

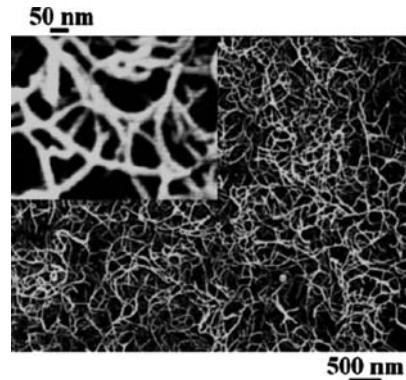
Fig. 24 Electron micrographs of gold-coated foams: MTEOS with 1.1 M ammonia, heated to 300°C [52]



Despite abrasion of these sol-gel surfaces, the high water droplet contact angle can be maintained. This result is significant for applications wherein handling and mechanical abrasion are inevitable. Even at temperatures as high as 300°C, no loss of superhydrophobicity is detected.

Gao et al. reported a method to fabricate a perfect superhydrophobic surface ($\theta_A/\theta_R = 180^\circ/180^\circ$) by immersion of a Si wafer into $\text{CH}_3\text{SiCl}/\text{toluene}$ solution;

Fig. 25 SEM images of a superhydrophobic surface from CH_3SiCl_3 [53]



rinsing with toluene, ethanol, and water; and drying the surface coating [53]. The surface morphology generated is shown in Fig. 25.

Alumina is a hard material that is abrasion resistant. Using a sol-gel process, nanoscale surface structures can be generated by various methods (Fig. 26). Tadanaga et al. showed that by immersing the porous alumina gel films prepared by the sol-gel method in boiling water, alumina thin films with a roughness of 20–50 nm were formed [49]. As a result, high optical transparency can be achieved. After a surface hydrophobic treatment, a superhydrophobic surface with a contact angle of 165° resulted.

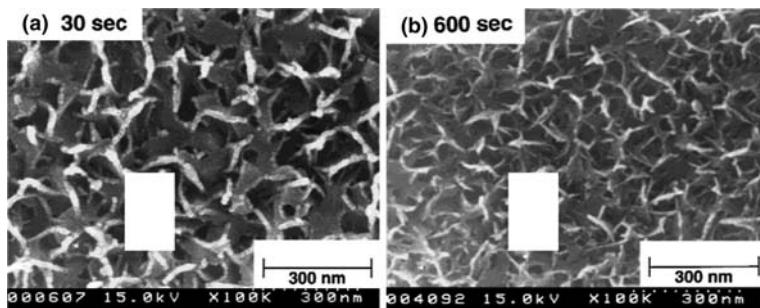


Fig. 26 FE-SEM photographs of the surface of the Al_2O_3 films immersed in boiling water for (a) 30 s and (b) 600 s [49]

5.4 Wet Chemical Etching

In addition to plasma etching, wet chemical etching has been also employed to generate surface roughness, frequently with using Si as a substrate. Wet chemical etching can be used for a variety of applications in both polishing and roughening. If the etch conditions are controlled properly, a rough surface can be prepared. For example, KOH etching of Si(100) can be controlled to achieve a pyramid-textured surface.

5.4.1 Metal-Assisted Etching

The metal-assisted etching process on Si surfaces involves (1) deposition of metal nanoparticles (Au, Pt, Ag, Pd, etc.) on the Si surface, and (2) etching of the metal coated Si in HF/oxidant solutions [54]. During this process, Si is continuously etched away at the site of the metal/Si contacts and surface roughness formed by control of the metal particle density. Deposition methods including direct metal deposition by sputtering or electron beam evaporation and electrochemical deposition of metal particles from reactive metal precursors. Figure 27 shows the result of using different metals for assisted etching on Si surfaces. Metal has been deposited onto Si surfaces by reduction of metal precursors such as HAuCl₄ [25], AgNO₃ [55], Ni(NO₃)₂ [56]. E-beam evaporation or DC sputtering can also be used to deposit metal onto Si surfaces for this etching process [54].

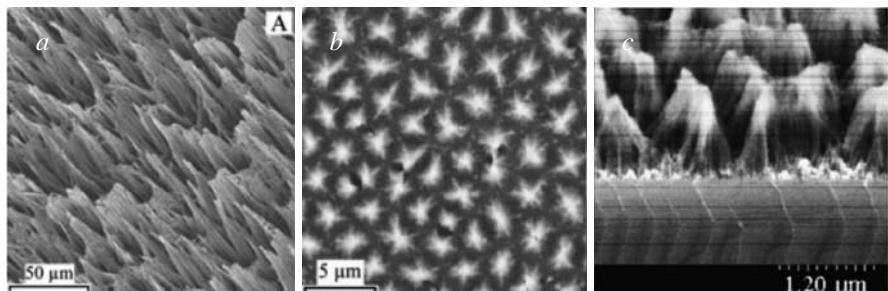


Fig. 27 SEM images of etched Si surfaces (a) from AgNO₃/HF etching, 0.02/5.0; (b) from Ni(NO₃)₂/HF etching 0.08/5.0 [56]; and (c) evaporated Pt on Si, etching in HF/H₂O₂ [54]

Superhydrophobic surfaces with hierarchical structures on Si using metal-assisted etching of Si pyramid surfaces (Fig. 28) [57, 58] by Xiu et al. The surface can be made as not only superhydrophobic but also non-reflecting for light incident on to the surface (Fig. 29). Therefore, multifunctional surfaces are possible. This

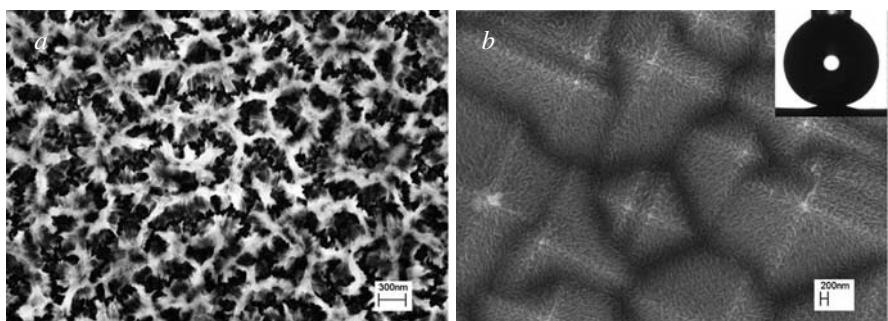


Fig. 28 (a) Si (100) surface etched by Au-assisted etching and (b) Si micro-pyramid surface etched by Au-assisted etching; inset: water droplet contact angle of $\sim 165^\circ$

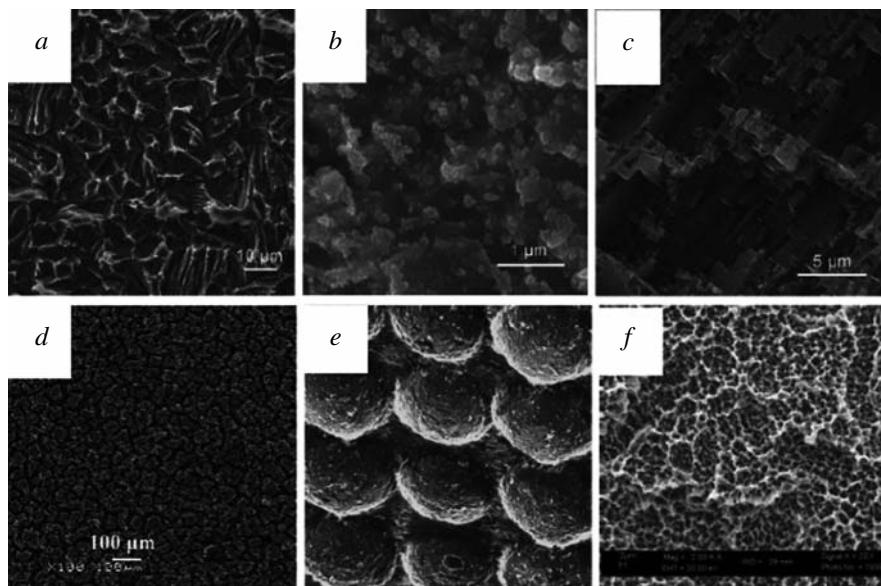


Fig. 29 SEM images of different etched surfaces: (a) zinc surface etched with 4.0 mol/L HCl solution for 90 s at room temperature [59]; (b) copper surface etched with a modified Livingston's dislocation etchant for 24 h at ambient temperature [59]; (c) Al etched by Beck's dislocation etchant for 10 s; (d) Al alloy (2024Al, 92.8 wt% Al, 5.5 wt% Cu) etching by NaOH [60]; (e) Cu surface [61], Cu etching was carried out in a potassium persulfate solution rough etched top and pits; (f) Ti surfaces from sandblasting and acid etching [62]

surface is promising to be employed in solar cell surface texturing and passivation when used with the appropriate chemical passivation on the surface.

5.4.2 Dislocation Selective Chemical Etching and Other Etching Methods

Dislocation selective chemical etching can be performed on the surface of polycrystalline metals such as aluminum, copper and zinc to prepare rough surfaces as shown in Fig. 29 [59–62]. The key to this etching technique is the use of a dislocation selective etchant that preferentially dissolves the dislocation sites in the grains. Water droplet contact angles higher than 150° and roll-off angles of less than 10° were obtained when surfaces were treated with fluoroalkylsilane.

Superhydrophobic surfaces can also be achieved by Al alloy etching using NaOH. Etching mainly occurred in the Al domain because of preferential dislocation etching; roughness is due to the remaining Cu and other elements on the surface. The as prepared surface shows a contact angle of 154° and sliding angle of $\sim 3^\circ$ after hydrophobic treatment [60]. On Cu surfaces, the microstructures can be prepared by photolithography and subsequent etching using a potassium persulfate solution [61]. A Ti surface with low surface energy was prepared by sandblasting and acid etching (HCl/H₂SO₄) after a hydrophobic treatment [62] (Fig. 29f).

The initial advancing contact angle on the surface is 139.9° without any hydrophobic treatment, which is the highest contact angle considering that there is no more hydrophobic treatment on the surface. However, the receding contact angle is $<5^\circ$ and repeated measurement gave a contact angle of 0° .

Electrochemical anodization can also be used to generate surface roughness. For example, Ti was treated by electrochemical anodization to form a sponge-like nanostructured TiO_2 film on the surface [63] (Fig. 30a). After a hydrophobic surface treatment, the advancing angle on the surface was 160.1° and the hysteresis was only 0.8° .

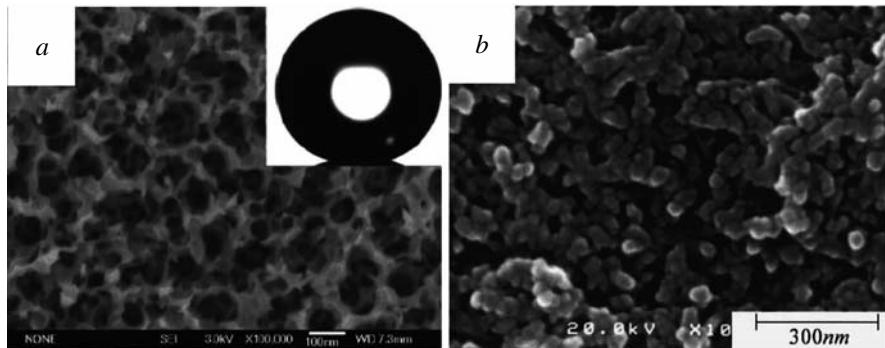


Fig. 30 Anodized surfaces: (a) Ti anodization, TiO_2 formed on surface [63]; (b) Al-anodized surface [5]

Al surfaces can also be anodized to prepare a fractal surface [5] (Fig. 30b). After a fluorinated monoalkyl phosphate (F-MAP) treatment, the surface showed a water droplet contact angle of 163° and rapeseed oil contact angle of 150° . The fluorination chain length was very important in this case and the degree of oil repellency for the silane coupling agent was much less than that of F-MAP.

5.5 Other Methods

5.5.1 Monodisperse Nanoparticles

Colloidal crystals can be formed on a substrate from monodispersed nanospheres, e.g., SiO_2 and polystyrene spheres, as a result of assembly into arrays of hexagonal close packed or FCC layers. In this scenario, the as-formed surface is rough due to the regular arrangement of the monodispersed spheres. Although this surface is not superhydrophobic even after a surface treatment with hydrophobic surface modification agents [64], superhydrophobicity can be achieved by imparting nanoscale structures on/in the surface to achieve multiscale roughness as shown in Fig. 31. Furthermore, this surface can serve as an ideal model surface for the investigation of the effect of surface roughness on superhydrophobicity. Monodispersed silica particles can be used to form a thin layer on a glass slide [14]. After deposition of a Au layer on top of the spheres followed by heat treatment to form Au

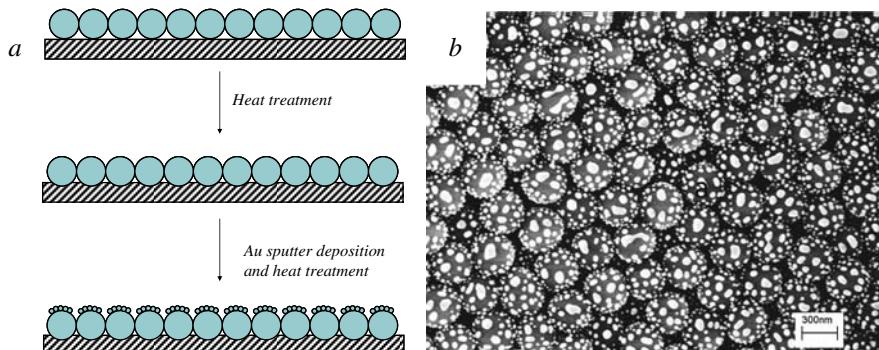


Fig. 31 (a) Illustration of the hierarchical roughness formation by monodisperse particle and Au deposition; (b) SEM image of the surface prepared from the process [14]

nanoparticles, surface nanoroughness was generated. By controlling the Au deposition conditions, the necessary nanoroughness can be effectively designed into the surface. In addition, non-close-packed hexagonally ordered structures were also reported using a templating technique [65]. By using low reflective index and low surface energy materials such as fluorinated polymers, a self-cleaning and antireflection coating can be prepared. Such tunability of the structures suggests specific ways by which multifunctional surface properties such as superhydrophobicity, antireflection, mechanical stability, and durability can be achieved.

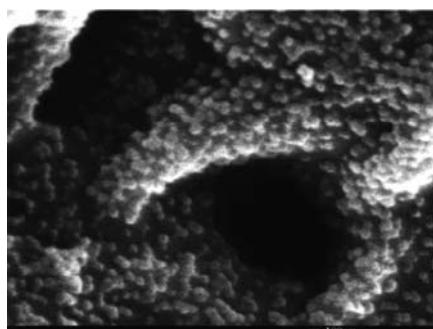
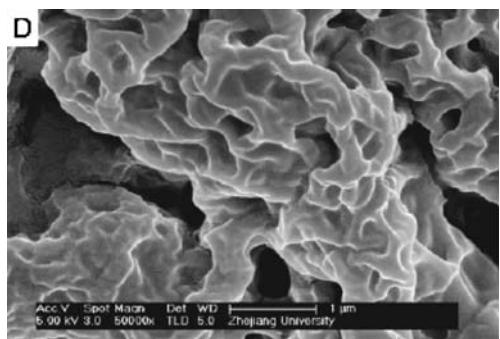
5.5.2 Layer-by-Layer Self-assembly

Layer-by-layer (LbL) self-assembly, which is based on the alternating physisorption of oppositely charged building blocks, represents a method to immobilize polyelectrolytes, colloid nanoparticles, and biomacromolecules, such as enzymes, extracellular matrices (ECM), and deoxyribonucleic acid (DNA) [66]. The sensitivity of the LbL multilayer toward its environment (e.g., pH, ionic strength) further provides new approaches to adjust the layered nanostructure with a tailored composition and architecture. The use of the LbL technique for constructing superhydrophobic coatings from poly(acrylic acid) (PAA) and polyethyleneimine (PEI), especially using the silver ion (Ag^+) to enhance the exponential growth of the PEI/PAA multi-layers has been explored [67]. It was found that the addition of Ag^+ can significantly improve the surface roughness as shown in Fig. 32. On such a surface, a high contact angle of 172° can be achieved after surface fluorination by (tridecafluoroctyl)triethoxysilane.

Zhai et al. prepared a polyelectrolyte multilayer surface by LbL assembly and then overcoated the surface with silica nanoparticles to mimic the hierarchical scale that is present on lotus leaf surfaces [68] (Fig. 33). Superhydrophobicity was achieved after surface fluorination with a fluoroalkyl silane.

Honeycomb-like microporous polymer films with 500 nm to 50 μm pores (diameter) have been prepared by casting polymer solutions under humid conditions (Fig. 34). If a fluoropolymer was used, superhydrophobicity was achieved [69]. In

Fig. 32 SEM image of (PAA/PEI-Ag+) 7.5 film [67]



Treatments for Forming Superhydrophobic Films from Polyelectrolyte Multilayers

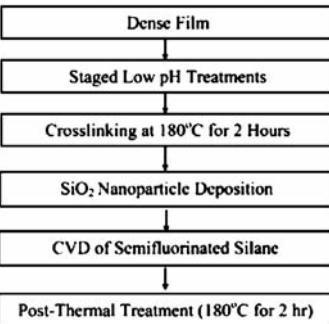
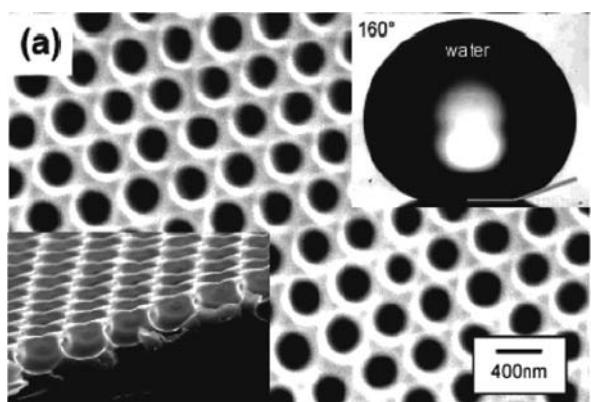


Fig. 33 SEM image of the fully treated polyelectrolyte structure (a 2 h immersion in a pH 2.7 solution followed by a 4 h immersion in a pH 2.3 solution, with no water rinse) with silica nanoparticles [68]

Fig. 34 Scanning electron micrograph of the 300-nm-sized honeycomb patterned film (top image and cross section) and the water contact angle on this film [69]



this process, water droplets condensed on the solution film surface act as templates. Depending on the drying rate of solvents, the pore size can be controlled by controlling the water droplet size. This surface can also be optically transparent under the proper formation conditions.

6 Applications

Superhydrophobicity is relatively a new research field in recent years describing the extreme water repellency of rough hydrophobic surfaces. However, the use of this technique is for a long time. Early products include breathable fabrics and some membrane filters, which use the same effect as superhydrophobic surfaces, but these are rarely referred to as being superhydrophobic. For example, Gore-Tex® is superhydrophobic which was patented in 1972.

Another important application for superhydrophobic technique may be in the field of oil/water separation [70]. Feng et al. demonstrated that a mesh coated with PTFE showed two extremities in wettability for water and oil. For water, it is extremely nonwetting surface (superhydrophobic), while for oil, it changed to complete wetting (superoleophilic). The extreme difference in wettability for water and oil showed the possibility in separation of oil and water. When a mixture of oil and water is put on the mesh film, water will remain on the upper part of the film, while the oil will penetrate through the mesh and be collected at the underneath.

Application of superhydrophobic coating in biomedical-related field is of great significance. One possible application may be in the self-cleaning coating of biomedical devices and equipments to prevent bacteria/virus growth and adhesion. The superhydrophobic surfaces have the great advantage of preventing water droplet staying on the surface and at the same time, preventing water vapor from condensing onto the surface. This will effectively prevent bacteria growth on the surface, due to the fact that any organism growth requires sufficient water/moisture. In addition, the much reduced adhesion on the superhydrophobic surface will make bacteria hard to deposit/reside on the surface. Therefore, the propagation of the micro-organisms is prohibited. Combination of these two effects, superhydrophobic surface coatings are very promising for the application in the biomedical device/equipment medical hygiene concerns. For the application in the bio-compatible surfaces, the superhydrophobic surface is based on the bio-inert principle, similar to PTFE and silicone because of the very low surface energy. However, they can be much cheaper because the bulk materials selection is flexible and cheap while the surface will resemble PTFE or silicone for the reduced surface energy, therefore, adhesion. Superhydrophobic surface coatings have remarkable chemical inertness due to the much reduced surface energy, therefore, providing a minimized chronic foreign body reaction. This will help reduce the risk of pain-related complications for long-term patient comfort. In addition, this coating helps reduce the risk of infection by reducing potential bacteria colonization sites which is a problem for currently used polymer coatings.

The friction reduction in underwater coatings is a new field for the superhydrophobic coatings. Speedo™ is one proved example of using hydrophobic suite to reduce the friction when swimming in water. Regarding superhydrophobic surfaces, the drag reduction comes from (1) the hydrophobic surface (low surface energy) with reduced friction; (2) the rough surfaces can form air cushion when the surface is soaked in water, therefore the solid/water contact can be minimized, which will result in the much reduced friction originated from the water/solid contact. The application of superhydrophobic coatings onto the outside surface of submarine or ship will be promising in the reduction of the friction in motion. Therefore, energy consumption will be greatly reduced. This is beneficial not only in energy conservation but also in the control of CO₂ reduction which is a critical issue in global warming.

6.1 Anticorrosion

Corrosion-induced failure plays an important role in microelectronic devices and package failures. Failures of microelectronic devices and packages not only cause the malfunction of the devices but also sometimes lead to catastrophic consequences for the whole systems. Corrosion in microelectronic packaging depends on the package type, electronic materials, fabrication and assembly processes, and environmental conditions such as moisture condensation, ionic or organic contaminants, temperature, residual and thermal stress and electrical bias, etc. With the ever-reducing feature sizes of microelectronic components and devices, they are more susceptible to corrosion-induced failures. The better performance and reliability requirements drive to improve corrosion resistance of packaging systems. There are basically three types of packaging for microelectronic components and devices. They are ceramic, metal, and plastic packaging. Ceramic and metal packages are hermetic packages mainly used in the military, aerospace, and automobile microelectronic devices for high reliability. Plastic packaging is non-hermetic package but become more and more widely used because of the low cost and better manufacturability. Compared to other package types, plastic-packaging systems have more corrosion-related problems because the polymeric materials used in plastic package systems are more vulnerable to the permeation of moisture into the inside dies, wires, bond pads, lead frames, and solder joints. Therefore, corrosion is becoming a more serious issue in modern microelectronic packages reliability.

Corrosion involves essentially electrochemical processes except for those oxidized at the elevated temperature under the dry environment. The basic requirements for electrochemical corrosion include electrically conductive anode, cathode, interconnecting electrolyte (humidity environment), and driving force (electrochemical potential). Impurity ions such as halides compounds which can be dissociated in water (electrolytes) are very dangerous resources in corrosion of electronics packaging impurity ions due to the formation of paths of electrochemical

corrosion. Preventing corrosion is actually a matter of preventing moisture condensation on interest areas such as die interfaces, wires, bond pads, solder joints. Superhydrophobic coatings, which are extremely water-repellent, show great potential for the anticorrosion applications in microelectronic devices.

Recent discoveries have linked the mechanism for the self-cleaning of a lotus plant to a microscopic morphology leading to superhydrophobic surfaces. This water-repellent surface suggests huge opportunities in the area of corrosion inhibition for metal components. A simple superhydrophobic-coating technique was developed by Barkhudarov et al. focusing on aluminum corrosion protection with a rough, highly porous organosilica aerogel-like film [71]. When silica-coated Al surfaces were submerged in 5 wt% NaCl in D₂O, neutron reflection measurement was conducted to measure the increase of the oxidation layer and the decrease of Al layer. It was found that the superhydrophobic layer can effectively reduce the corrosion process as shown in Fig. 35.

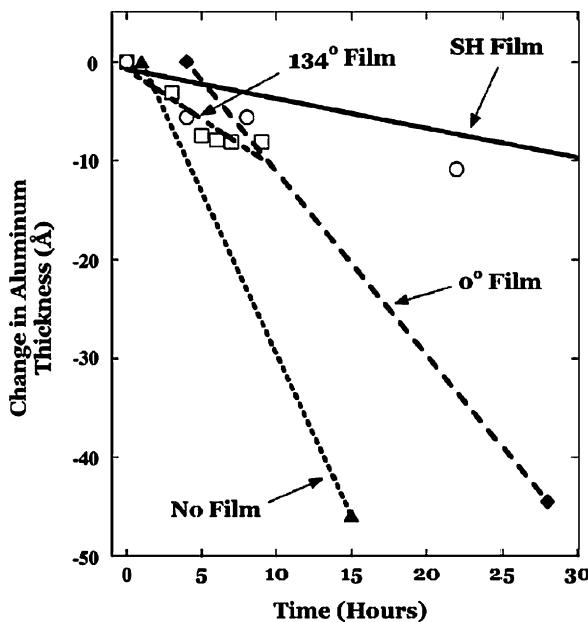


Fig. 35 The change in thickness of the aluminum layer vs. time for samples protected by films of varying contact angle and a sample with only native Al₂O₃ layer (without protective film). At each data point, the initial aluminum layer thickness for that sample was subtracted in order to leave only information about changes in thickness. The *solid* and *dashed* lines represent linear fits of the aluminum layer thickness decrease. Note: the figure does not show all data points used to obtain the linear fit for the SH film sample, which was measured up to 186 h (This is why the SH line does not appear to be an accurate fit for the points shown in this figure) [71]

6.2 Applications in MEMS Antistiction

Microelectromechanical systems (MEMS) have been used extensively to perform basic signal transduction operations in miniaturized sensors and actuators. However, autoadhesion, or spontaneous sticking (stiction) between MEMS structures, remains a major limitation in bringing this new class of engineering devices to the broader market. Free-standing mechanical structures fabricated from polycrystalline Si may strongly adhere to each other when brought into contact, due to the hydrogen bonding between surface hydroxyl groups [72]. In a high humidity ambient, this problem is exacerbated by adsorption of water and capillary condensation. Water capillaries or condensation can cause catastrophic failure of MEMS devices [73]. Low surface energy coatings for antistiction of MEMS are required for many practical MEMS devices [72–74].

Two primary methods have been invoked to prevent these failures. One is a surface treatment with low surface energy coatings, e.g., octadecyl trichlorosilane (OTS), to inhibit surface condensation of water [75]. The other method involves the fabrication of structured surfaces in order to reduce the contact area and thereby reduce the adhesion force between the free standing structures and the substrate [76].

Superhydrophobic silica surfaces were proposed and demonstrated for the first time in MEMS antistiction by AFM adhesion tests using tipless probes as shown in Fig. 36. This approach greatly reduces the adhesion forces between the tipless probe and the superhydrophobicly coated substrate, which showed that it is promising to reduce stiction failures of MEMS devices. Water does not condense on this superhydrophobic surface thus eliminating capillary forces. Furthermore, the

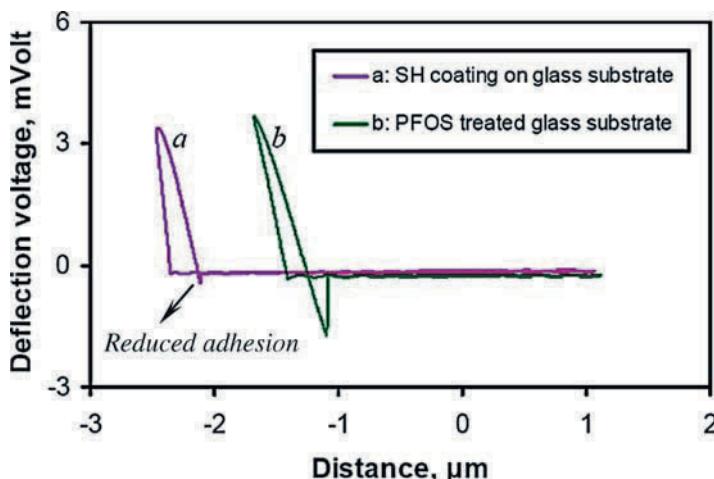


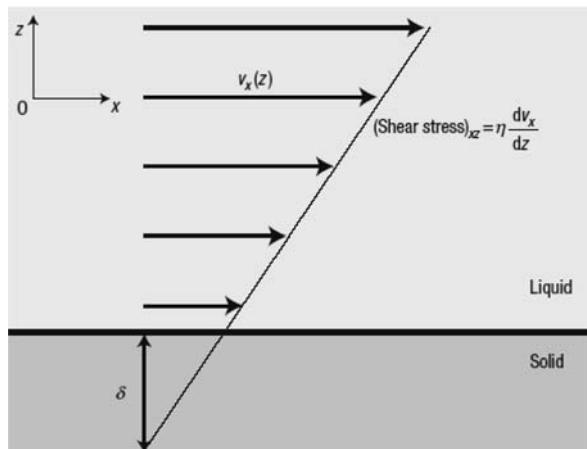
Fig. 36 Surface–probe interactions as determined from AFM force measurements between a tipless Si_3N_4 probe and either a superhydrophobic film spin-casted from ethanol:sol ratio of 5:1 or a PFOS-treated glass slide

adhesion force between the free standing structures and the substrate can also be effectively reduced by decreasing the contact area due to the presence of surface nanostructures.

6.3 Slip Flow in Microfluidics

In fluid mechanics, usually a non-slip boundary condition is assumed that the fluid velocity at a fluid–solid interface equals that of the solid surface as shown in Fig. 37. Recent development in microfluidics showed that the violation of the assumption may exist. For surfaces with large surface-to-volume ratio, surface properties may dramatically affect the flow resistance. The possibility of engineering slip properties in a controlled way is therefore crucial for microfluidic applications. Mechanisms behind the boundary slip include surface roughness and structural characteristics of roughness, roughness-induced dewetting on hydrophobic surfaces, dissolved gas, and bubbles on the surface. For hydrophilic rough surfaces, the no-slip boundaries can be achieved. However, for a hydrophobic surface, roughness may increase the slip due to the transition to a superhydrophobic state [77]. Liquid cannot enter the valleys of the rough surface, instead, water only stay at the top of them, with a minimum solid/liquid contact area. The as-formed air bubbles help lubricating the flow, reducing the friction of the flow process.

Fig. 37 Definition of the slip length. The linear velocity (v_x) profile in the flowing, Newtonian fluid (characterized by a Newtonian viscosity η and a constant shear stress) does not vanish at the solid boundary. Extrapolation into the solid at a depth δ is necessary to obtain a vanishing velocity, as assumed by the macroscopic “no-slip” boundary condition [77]

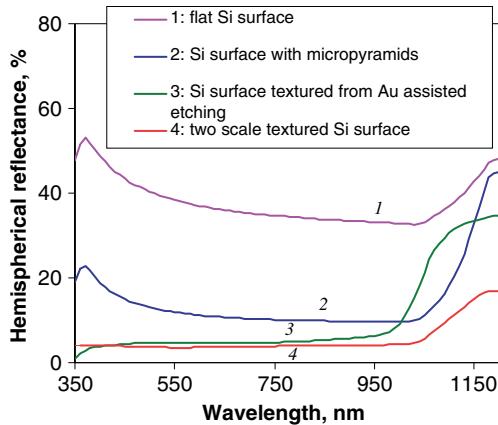


6.4 Solar Cells

Superhydrophobic surfaces have been prepared by using a variety of materials. In addition to the establishment of superhydrophobicity, these materials have been further functionalized to realize electro-active surfaces with controlled wetting properties [23], transparent surfaces [44], adhesive surfaces [78], photocatalytic surfaces

[79], and photonic bandgap layers [80]. However, due to its widespread use in electronic and photovoltaic devices, incorporation of superhydrophobic properties into Si surfaces may extend and enhance the more traditional applications of these devices. Metal-assisted etching of Si surfaces has been employed to define specific structures on Si surfaces [54, 81]. After deposition of a thin layer of a metal such as Au, Pt, or Pt/Pd onto p- or n-type Si surfaces, nanostructures can be formed by immersing this substrate into various mixtures of hydrofluoric acid/oxidant/solvent. This approach has been used to prepare photoluminescent surfaces [82] and “black” anti-reflecting Si surfaces [54]. In fact, the reflectivity can be reduced to nearly zero if the appropriate structures are formed. The use of etched Si surfaces for photovoltaic applications has been also investigated, and a solar cell efficiency of 9.31% was reported [83]. On the hierarchically etched Si surfaces, the weighted light reflection loss can be reduced to below 4% without any other antireflection coatings as shown in Fig. 38. The etched surface showed a surface hierarchical structure shown in Fig. 39

Fig. 38 Light reflection on Si nano-textured surfaces generated by Au-assisted etching. Au (5-nm thickness)-assisted etching in HF:H₂O₂:H₂O of 1:5:10 (v/v/v) for 2 min [58]



6.5 Bio-Assay

In biotechnology, one is interested in controlling droplets containing biologically relevant molecules (DNA and proteins), minimizing contamination. Already the construction of complementary DNA (cDNA)-microarrays prepared by spotting techniques requires the specific wetting properties of the substrates. Glass slides, the substrate most commonly used, are usually only mildly hydrophobized, such that the drops of drying cDNA solution produce unwanted ring-like structures, a nuisance known in the field as the “doughnut-effect” (see the following Fig. 40a). This effect is related to the well-studied “coffee-stain effect”; an evaporation-driven convection mechanism drives dissolved or dispersed particles inside the drop to the edge of the drop, where evaporation is fastest [84] therefore, after drying, the deposited material is not a spot, but a ring-like shape. Superhydrophobicity is advantageous in

Fig. 39 Cross sections of Si pyramid surfaces etched by Au-assisted etching in HF/H₂O₂/H₂O (v/v/v 1:5:10) [58]

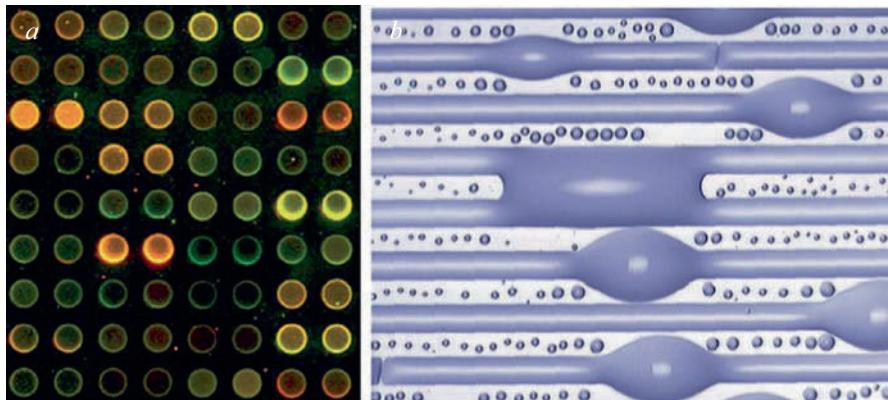
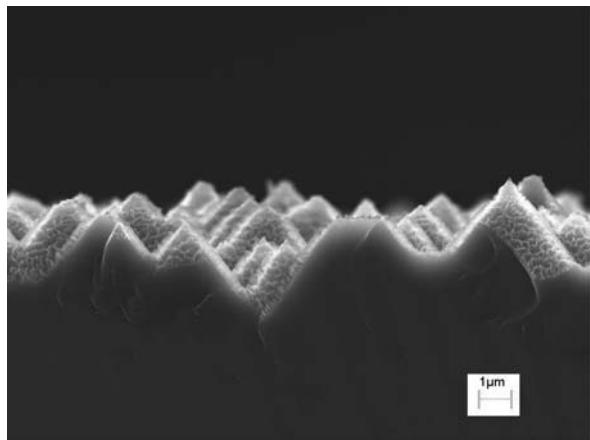


Fig. 40 (a) Complementary DNA microarray on a silanized glass plate. The doughnut-like shapes of the deposits are clearly visible; (b) liquid channels on a hydrophilically–hydrophobically patterned substrate: a possible pathway to surface-tension-controlled microfluidics [84]

this application: the almost fully spherical drops on a superhydrophobic surface can shrink exactly the same as a drop in free air. In addition, the position and shape of spotted drops can be steered by combining hydrophilic and hydrophobic surface patterns. The usefulness of this idea is shown in a real biosystem: some desert beetles capture their drinking water by a hydrophobically–hydrophilically structured back. In the same way, by prepatternning of a substrate, hydrophilic prespotting (anchoring) on an otherwise hydrophobic or superhydrophobic surface, results in new possibilities for improvements in spotting and analyzing DNA and proteins by avoiding wall contact.

Beyond the possibility of improving deposit shapes on substrates, the possibility of a guided motion of droplets on superhydrophobic surfaces offers the chance to develop a droplet-based microfluidics system, in contrast to the classical concept

based on microfluidic channels. Driving the liquid along the channels and making them merged at predefined locations offers a novel way to mix reactants or steer biochemical reactions, defining the concept of a “liquid microchip” or “surface-tension confined microfluidics.” One advantage of the open structures over capillaries, in addition to their ease of cleaning, is that blocking of the capillary by unforeseen chemical reactions cannot occur.

Microfluidics can also be based on droplets on superhydrophobic surfaces alone; because the drops have very low contact areas with the substrate, they are easy to move by external fields such as electrostatic forces or surface acoustic waves. Systems that make use of a droplet-based actuation mechanism are being developed by many researchers. The aim here is to control droplet positioning and motion on the substrates with as little surface contact as possible, and making the droplet-based system a programmable reactor, by which the liquid positions are prescribed and tuned externally.

6.6 Superhydrophobic Surfaces for Outdoor Applications

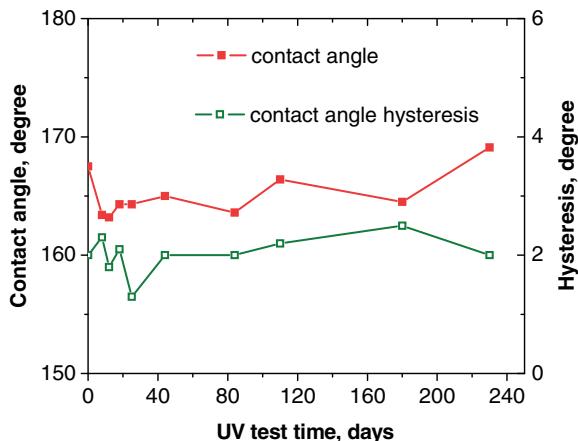
For the self-cleaning applications, most involve the outdoor environment. This pose another engineering requirement: weathering resistance which mostly involves the UV exposure. Therefore a UV-resistant superhydrophobic surface is required.

For polymeric materials, UV degradation (unzipping of polymer chain and photo-oxidation) is inevitable due to the intrinsic defects in the polymer chain and impurities (acting as catalyst) inside the polymer. Formation of carbonyl or hydroxyl groups on the surface ensues. This happened first on the surface which means that before any mechanical property loss, surface properties changed first from hydrophobic to hydrophilic. Currently no efficient method has been reported for the prevention of polymer surface degradation under UV irradiation.

Xiu et al. reported a method using inorganic materials such as silica with the combination of fluoroalkyl silane treatment for the UV-resistant superhydrophobic surfaces [85]. The surface was tested for $\sim 5,500$ h without obvious degradation in superhydrophobicity as shown in Fig. 41. This method has used small stable molecules such as fluoroalkyl silanes for the surface layer instead of polymers, which prevent the possible chain unzipping reactions and the subsequent loss of superhydrophobicity.

In summary, superhydrophobic surfaces have been successfully prepared by using a variety of materials and the mechanism for self-cleaning and water repellency has been better understood. However, one of the biggest problems facing all self-cleaning or contaminant-free surface applications is degradation and failures under harsh environment. For biotechnology applications, this is not so relevant; the surfaces will often be used for analytic purposes and hence designed as disposables. If one thinks of contaminant-free surfaces for use in complicated conditions, the issues of robustness under mechanical abrasion (e.g., handing of the coated surfaces) has not yet been resolved. And yet a different matter is the use of self-cleaning

Fig. 41 Stability of a PFOS-treated rough silica thin film surface [85]



surfaces in outdoor applications. In addition, stability under high humidity condition also needs improvement by designing small features with cost-effective approaches. The dew point may be reduced on the superhydrophobic surface with small features. Oil or surfactant repellent or superoleophobic surface may also be desirable which further add more difficulties to the design of superhydrophobicity. As coatings improve and prices fall down the use of superhydrophobic surfaces is likely to increase, particularly as many of the drawbacks are now well-understood and some can be minimized by choosing specific coatings.

References

1. Wagner T., Neinhuis C., Barthlott W., Wettability and contaminability of insect wings as a function of their surface sculptures. *Acta Zoologica* 1996; 77: 213–215.
2. Neinhuis C., Barthlott W., Characterization and distribution of water-repellent, self-cleaning plant surfaces. *Annals of Botany* 1997; 79: 667–677.
3. Sacher E., Sapieha J.K., Schrieber H.P., Wertheimer N.R., McIntyre N.S., Silanes, Surfaces, and Interfaces, Vol. I. Taylor & Francis: London, 1986.
4. Morra M., Occhiello E., Garbassi F., Contact-angle hysteresis in oxygen plasma treated poly(tetrafluoroethylene). *Langmuir* 1989; 5: 872–876.
5. Shibuichi S., Yamamoto T., Onda T., Tsujii K., Super water- and oil-repellent surfaces resulting from fractal structure. *Journal of Colloid and Interface Science* 1998; 208: 287–294.
6. Nosonovsky M., Bhushan B., Biologically inspired surfaces: Broadening the scope of roughness. *Advanced Functional Materials* 2008; 18: 843–855.
7. Sun T.L., Feng L., Gao X.F., Bioinspired surfaces with special wettability. *Accounts of Chemical Research* 2005; 38: 644–652.
8. Gao X.F., Jiang L., Water-repellent legs of water striders. *Nature* 2004; 432: 36.
9. Gao X.F., Yan X., Yao X., Xu L., Zhang K., Zhang J.H., Yang B., Jiang L., The dry-style antifogging properties of mosquito compound eyes and artificial analogues prepared by soft lithography. *Advanced Materials* 2007; 19: 2213–2217.
10. Berthier S., Iridescent Colors of Insects. Springer: New York, 2007.

11. Cassie A.B.D., Baxter S., Wettability of porous surfaces. *Transactions of the Faraday Society* 1944; 40: 0546–0550.
12. Wenzel R.N., Resistance of solid surfaces to wetting by water. *Industrial and Engineering Chemistry* 1936; 28: 988–994.
13. Marmur A., Wetting on hydrophobic rough surfaces: to be heterogeneous or not to be? *Langmuir* 2003; 19: 8343–8348.
14. Xiu Y., Zhu L., Hess D.W., Wong C.P., Biomimetic creation of hierarchical surface structures by combining colloidal self-assembly and Au sputter deposition. *Langmuir* 2006; 22: 9676–9681.
15. Lu X.Y., Zhang J.L., Zhang C.C., Han Y.C., Low-density polyethylene (LDPE) surface with a wettability gradient by tuning its microstructures. *Macromolecular Rapid Communications* 2005; 26: 637–642.
16. Han W., Wu D., Ming W.H., Niemantsverdriet H., Thune P.C., Direct catalytic route to superhydrophobic polyethylene films. *Langmuir* 2006; 22: 7956–7959.
17. Teshima K., Sugimura H., Inoue Y., Takai O., Takano A., Ultra-water-repellent poly(ethylene terephthalate) substrates. *Langmuir* 2003; 19: 10624–10627.
18. Tadanaga K., Morinaga J., Matsuda A., Minami T., Superhydrophobic-superhydrophilic micropatterning on flowerlike alumina coating film by the sol-gel method. *Chemistry of Materials* 2000; 12: 590–592.
19. Nakajima A., Fujishima A., Hashimoto K., Watanabe T., Preparation of transparent superhydrophobic boehmite and silica films by sublimation of aluminum acetylacetone. *Advanced Materials* 1999; 11: 1365–1368.
20. Israelachvili J., *Intermolecular and Surface Forces*, 2nd ed. Academic Press: New York, 1991.
21. Nishino T., Meguro M., Nakamae K., Matsushita M., Ueda Y., The lowest surface free energy based on -CF₃ alignment. *Langmuir* 1999; 15: 4321–4323.
22. Tuteja A., Choi W., Ma M., Mabry J.M., Mazzella S.A., Rutledge G.C., McKinley G.H., Cohen R.E., Designing superoleophobic surfaces. *Science* 2007; 318: 1618–1622.
23. Ahuja A., Taylor J.A., Lifton V., Sidorenko A.A., Salomon T.R., Lobaton E.J., Kolodner P., Krupenkin T.N., Nanonails: A simple geometrical approach to electrically tunable super-lyophobic surfaces. *Langmuir* 2008; 24: 9–14.
24. Cao L.L., Price T.P., Weiss M., Gao D., Super water- and oil-repellent surfaces on intrinsically hydrophilic and oleophilic porous silicon films. *Langmuir* 2008; 24: 1640–1643.
25. Cao L.L., Hu H.H., Gao D., Design and fabrication of micro-textures for inducing a superhydrophobic behavior on hydrophilic materials. *Langmuir* 2007; 23: 4310–4314.
26. Patankar N.A., Transition between superhydrophobic states on rough surfaces. *Langmuir* 2004; 20: 7097–7102.
27. Lieberman M.A., Lichtenberg A.J., *Principles of Plasma Discharges and Materials Processing*, 2nd Translator. Wiley-Interscience: London, 2005.
28. Stelmashuk V., Biederman H., Slavinska D., Zemek J., Trchova M., Plasma polymer films rf sputtered from PTFE under various argon pressures. *Vacuum* 2005; 77: 131–137.
29. Shiu J.Y., Kuo C.W., Chen P., Fabrication of tunable superhydrophobic surfaces. *Proceedings of SPIE-The International Society for Optical Engineering* 2005; 5648: 325–332.
30. Byun D., Lee Y., Tran S.B.Q., Nugyen V.D., Kim S., Park B., Lee S., Inamdar N., Bau H.H., Electrospray on superhydrophobic nozzles treated with argon and oxygen plasma. *Applied Physics Letters* 2008; 92: 093507.
31. Larsen K.P., Petersen D.H., Hansen O., Study of the roughness in a photoresist masked, isotropic, SF₆-based ICP silicon etch. *Journal of the Electrochemical Society* 2006; 153: G1051–G1058.
32. Lejeune M., Lacroix L.M., Bretagnol F., Valsesia A., Colpo P., Rossi F., Plasma-based processes for surface wettability modification. *Langmuir* 2006; 22: 3057–3061.
33. Teshima K., Sugimura H., Inoue Y., Takai O., Takano A., Transparent ultra water-repellent poly(ethylene terephthalate) substrates fabricated by oxygen plasma treatment and subsequent hydrophobic coating. *Applied Surface Science* 2005; 244: 619–622.

34. Tserepi A.D., Vlachopoulou M.E., Gogolides E., Nanotexturing of poly(dimethylsiloxane) in plasmas for creating robust super-hydrophobic surfaces. *Nanotechnology* 2006; 17: 3977–3983.
35. Teshima K., Sugimura H., Inoue Y., Takai O., Takano A., Wettability of poly(ethylene terephthalate) substrates modified by a two-step plasma process: Ultra water-repellent surface fabrication. *Chemical Vapor Deposition* 2004; 10: 295.
36. Teshima K., Sugimura H., Inoue Y., Takai O., Takano A., Transparent ultra water-repellent poly(ethylene terephthalate) substrates fabricated by oxygen plasma treatment and subsequent hydrophobic coating. *Applied Surface Science* 2005; 244: 619–624.
37. Teshima K., Sugimura H., Takano A., Inoue Y., Takai O., Ultrahydrophobic/ultrahydrophilic micropatterning on a polymeric substrate. *Chemical Vapor Deposition* 2005; 11: 347–352.
38. Fresnais J., Chapel J.P., Poncin-Epaillard F., Synthesis of transparent superhydrophobic polyethylene surfaces. *Surface & Coatings Technology* 2006; 200: 5296–5305.
39. Fresnais J., Benyahia L., Poncin-Epaillard F., Dynamic (de)wetting properties of superhydrophobic plasma-treated polyethylene surfaces. *Surface Interface Analysis* 2006; 38: 144–149.
40. Fresnais J., Benyahia L., Chapel J.P., Poncin-Epaillard F., Polyethylene ultrahydrophobic surface: synthesis and original properties. *European Physical Journal: Applied Physics* 2004; 26: 209–214.
41. Di Mundo R., Palumbo F., d'Agostino R., Nanotexturing of polystyrene surface in fluorocarbon plasmas: from sticky to slippery superhydrophobicity. *Langmuir* 2008; 24: 5044–5051.
42. Woodward I., Schofield W.C.E., Roucoules V., Badyal J.P.S., Super-hydrophobic surfaces produced by plasma fluorination of polybutadiene films. *Langmuir* 2003; 19: 3432–3438.
43. Dorrer C., Ruhe J., Wetting of silicon nanograss: from superhydrophilic to superhydrophobic surfaces. *Advanced Materials* 2008; 20: 159–164.
44. Artus G.R.J., Jung S., Zimmermann J., Gautschi H.P., Marquardt K., Seeger S., Silicone nanofilaments and their application as superhydrophobic coating. *Advanced Materials* 2006; 18: 2758–2762.
45. Zimmermann J., Rabe M., Artus G.R.J., Seeger S., Patterned superfunctional surfaces based on a silicone nanofilament coating. *Soft Matter* 2008; 4: 450–452.
46. Zhu L.B., Xiu Y.H., Xu J.W., Tamirisa P.A., Hess D.W., Wong C.P., Superhydrophobicity on two-tier rough surfaces fabricated by controlled growth of aligned carbon nanotube arrays coated with fluorocarbon. *Langmuir* 2005; 21: 11208–11212.
47. Lau K.K.S., Bico J., Teo K.B.K., Chhowalla M., Amarantunga G.A.J., Milne W.I., McKinley G.H., Gleason K.K., Superhydrophobic carbon nanotube forests. *Nano Letters* 2003; 3: 1701–1705.
48. Roig A., Molins E., Rodriguez E., Martinez S., Moreno-Manas M., Vallribera A., Superhydrophobic silica aerogels by fluorination at the gel stage. *Chemical Communications* 2004; 20: 2316–2317.
49. Tadanaga K., Katata N., Minami T., Formation process of super-water-repellent Al₂O₃ coating films with high transparency by the sol-gel method. *Journal of the American Ceramic Society* 1997; 80: 3213–3216.
50. Doshi D.A., Shah P.B., Singh S., Branson E.D., Malanoski A.P., Watkins E.B., Majewski J., van Swol F., Brinker C.J., Investigating the interface of superhydrophobic surfaces in contact with water. *Langmuir* 2005; 21: 7805–7811.
51. Rao A.V., Bhagat S.D., Hirashima H., Pajonk G.M., Synthesis of flexible silica aerogels using methyltrimethoxysilane (MTMS) precursor. *Journal of Colloid and Interface Science* 2006; 300: 279–285.
52. Shirtcliffe N.J., McHale G., Newton M.I., Perry C.C., Intrinsically superhydrophobic organosilica sol-gel foams. *Langmuir* 2003; 19: 5626–5631.
53. Gao L.C., McCarthy T.J., A perfectly hydrophobic surface ($\theta(A)/\theta(R)=180$ degrees/180 degrees). *Journal of the American Chemical Society* 2006; 128: 9052.

54. Koynov S., Brandt M.S., Stutzmann M., Black nonreflecting silicon surfaces for solar cells. *Applied Physics Letters* 2006; 88: 203107/1.
55. Peng K.Q., Wu Y., Fang H., Zhong X.Y., Xu Y., Zhu J., Uniform, axial-orientation alignment of one-dimensional single-crystal silicon nanostructure arrays. *Angewandte Chemie International Edition* 2005; 44: 2737–2742.
56. Peng K.Q., Yan Y.J., Gao S.P., Zhu J., Dendrite-assisted growth of silicon nanowires in electroless metal deposition. *Advanced Functional Materials* 2003; 13: 127–132.
57. Xiu Y., Zhu L., Hess D.W., Wong C.P., Hierarchical silicon etched structures for controlled hydrophobicity/superhydrophobicity. *Nano Letters* 2007; 7: 3388–3393.
58. Xiu Y., Zhang S., Yelundur V., Rohatgi A., Hess D.W., Wong C.P., Superhydrophobic and low light reflectivity silicon surfaces fabricated by hierarchical etching. *Langmuir* 2008; 24: 10421–10426.
59. Qian B.T., Shen Z.Q., Fabrication of superhydrophobic surfaces by dislocation-selective chemical etching on aluminum, copper, and zinc substrates. *Langmuir* 2005; 21: 9007–9009.
60. Guo Z.G., Zhou F., Hao J.C., Liu W.M., Effects of system parameters on making aluminum alloy lotus. *Journal of Colloid and Interface Science* 2006; 303: 298–305.
61. Shirtcliffe N.J., McHale G., Newton M.I., Perry C.C., Wetting and wetting transitions on copper-based super-hydrophobic surfaces. *Langmuir* 2005; 21: 937–943.
62. Rupp F., Scheideler L., Olshanska N., de Wild M., Wieland M., Geis-Gerstorfer J., Enhancing surface free energy and hydrophilicity through chemical modification of microstructured titanium implant surfaces. *Journal of Biomedical Materials Research Part A* 2006; 76A: 323–334.
63. Lai Y.K., Lin C.J., Huang J.Y., Zhuang H.F., Sun L., Nguyen T., Markedly controllable adhesion of superhydrophobic spongelike nanostructure TiO₂ films. *Langmuir* 2008; 24: 3867.
64. Shiu J.Y., Kuo C.W., Chen P.L., Mou C.Y., Fabrication of tunable superhydrophobic surfaces by nanosphere lithography. *Chemistry of Materials* 2004; 16: 561–564.
65. Sun C.H., Jiang P., Jiang B., Broadband moth-eye antireflection coatings on silicon. *Applied Physics Letters* 2008; 92: 051107.
66. Zheng H., Okada H., Nojima S., Suye S., Hori T., Layer-by-layer assembly of enzymes and polymerized mediator on electrode surface by electrostatic adsorption. *Science and Technology of Advanced Materials* 2004; 5: 371–376.
67. Ji J., Fu J.H., Shen J.C., Fabrication of a superhydrophobic surface from the amplified exponential growth of a multilayer. *Advanced Materials* 2006; 18: 1441–1444.
68. Zhai L., Cebeci F.C., Cohen R.E., Rubner M.F., Stable superhydrophobic coatings from polyelectrolyte multilayers. *Nano Letters* 2004; 4: 1349–1353.
69. Yabu H., Shimomura M., Single-step fabrication of transparent superhydrophobic porous polymer films. *Chemistry of Materials* 2005; 17: 5231–5234.
70. Feng L., Zhang Z.Y., Mai Z.H., Ma Y.M., Liu B.Q., Jiang L., Zhu D.B., A super-hydrophobic and super-oleophilic coating mesh film for the separation of oil and water. *Angewandte Chemie International Edition* 2004; 43: 2012–2014.
71. Barkhudarov P.M., Shah P.B., Watkins E.B., Corrosion inhibition using superhydrophobic films. *Corrosion Science* 2008; 50: 897–902.
72. Ashurst W.R., Yau C., Carraro C., Lee C., Kluth G.J., Howe R.T., Maboudian R., Alkene based monolayer films as anti-stiction coatings for polysilicon MEMS. *Sensors and Actuators: A, Physical* 2001; 91: 239–248.
73. Mastrangelo C.H., Hsu C.H., Mechanical stability and adhesion of microstructures under capillary forces: Part I: basic theory. *IEEE Journal of Microelectromechanical Systems* 1993; 2: 33–43.
74. Mastrangelo C.H., Hsu C.H., Mechanical stability and adhesion of microstructures under capillary forces – Part II: experiments. *Journal of Microelectromechanical Systems* 1993; 2: 44–55.

75. Kulkarni S.A., Mirji S.A., Mandale A.B., Thermal stability of self-assembled octadecyl-trichlorosilane monolayers on planar and curved silica surfaces. *Thin Solid Films* 2006; 496: 420–425.
76. Lee C.C., Hsu W., Method on surface roughness modification to alleviate stiction of microstructures. *Journal of Vacuum Science & Technology B* 2003; 21: 1505–1510.
77. Cottin-Bizonne C., Barrat J.L., Bocquet L., Low-friction flows of liquid at nanopatterned interfaces. *Nature Materials* 2003; 2: 237–240.
78. Mahdavi A., Ferreira L., Sundback C., Nichol J.W. et al., A biodegradable and biocompatible gecko-inspired tissue adhesive. *Proceedings of the National Academy of Sciences* 2008; 105: 2307–2312.
79. Zhang X.T., Jin M., Liu Z.Y., Nishimoto S., Saito H., Murakami T., Fujishima A., Preparation and photocatalytic wettability conversion of TiO₂-based superhydrophobic surfaces. *Langmuir* 2006; 22: 9477–9479.
80. Wang J.X., Wen Y.Q., Feng X.J., Control over the wettability of colloidal crystal films by assembly temperature. *Macromolecular Rapid Communications* 2006; 27: 188–192.
81. Li X., Bohn P.W., Metal-assisted chemical etching in HF/H₂O₂ produces porous silicon. *Applied Physics Letters* 2000; 77: 2572–2574.
82. Chattopadhyay S., Li X.L., Bohn P.W., In-plane control of morphology and tunable photoluminescence in porous silicon produced by metal-assisted electroless chemical etching. *Journal of Applied Physics* 2002; 91: 6134–6140.
83. Peng K.Q., Xu Y., Wu Y., Yan Y.J., Lee S.T., Zhu J., Aligned single-crystalline Si nanowire arrays for photovoltaic applications. *Small* 2005; 1: 1062–1067.
84. Deegan R.D., Bakajin O., Dupont T.F., Huber G., Nagel S.R., Witten T.A., Capillary flow as the cause of ring stains from dried liquid drops. *Nature* 1997; 389: 827–829.
85. Xiu Y.H., Hess D.W., Wong C.P., UV and thermally stable superhydrophobic coatings from sol-gel processing. *Journal of Colloid and Interface Science* 2008; 326: 465–470.

Applications of Carbon Nanomaterials as Electrical Interconnects and Thermal Interface Materials

Wei Lin and C.P. Wong

Abstract A brief review on CNT structure, electrical property, heat transport, and synthesis methods, CNT applications as electrical interconnects and TIMs, CNT integration into circuits and packaging, with focus on CNT transfer technology, are presented. In comparison with CNT/polymer composites, nanographite/polymer composites are more promising for TIM applications, which is discussed together with current thermal measurement techniques, commercialized and non-commercialized.

Keywords Carbon nanotubes · Graphene · Thermal interface materials · Nano interconnect · Thermal property measurements

1 Introduction

As we have been experiencing, revolutionary increase in speed and reliability of microprocessors has been successfully achieved in the past 60 years. The faster and higher performance of microprocessors is based on increased transistor density. As originally proposed, Moore's law stated that the number of transistors in semiconductor devices or integrated circuits (ICs) would double approximately every 2 years [1]. This prediction has been realized, largely due to device scaling, characteristic of fine-pitch (pitch=distance between two adjacent electrodes or electrical traces) interconnects. Copper interconnects, introduced in 1998, are now routinely used with the minimum feature size down to 65 nm; 45-nm node can be found in some commercial devices; new technologies for smaller nodes are being intensively studied. However, the electrical resistivity of copper interconnects increases with a decrease in dimensions due to grain-boundary and electron surface scattering [2]. As current density further increases, the electromigration issue for metal interconnects

W. Lin (✉)

School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: wlin31@gatech.edu

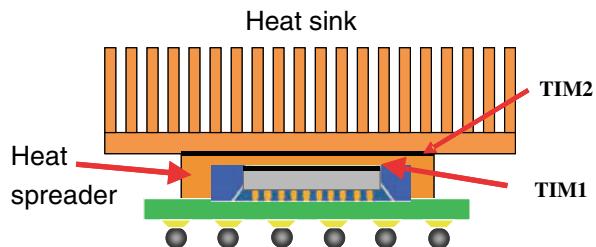
becomes more severe. In these regards, carbon nanotubes (CNTs) have been proposed as a future interconnecting material due to their ultra-high current carrying capacity (10^9 A/cm 2), thermal stability, and high resistance to electromigration [3–11]. Today, the main challenges of CNT interconnects (circuits) are (1) purification of metallic or semiconducting CNTs; (2) selective positioning of CNTs; (3) effective and reliable contacts at CNT junctions [12,13]. Although the purification issue has almost been addressed by recent development of various CNT separation methods, CNT positioning and contact reliability issues are still unaddressed [14–18]. Appropriately, graphene, an alternative to CNTs in the family of carbon nanomaterials, has been proposed to be capable of offering a solution to the problem of selective positioning for interconnects.

In spite of electrical performances, increasing microprocessor performance is associated with an increased cooling demand; in other words, more heat dissipation is required. It has been reported that a reduction in the device operation temperature corresponds to an exponential increase in reliability and life expectancy of a device [19]. To control device temperature within operation limits is critical. Thermal design power (TDP, the maximum sustained power dissipated by the microprocessor) has in the past increased steadily with increasing microprocessor performance [20]. Although multicore microprocessors should alleviate the growth in TDP with increased performance, thermal non-uniformity, usually referred to as the “hot spot” issue, where the local power density could be >300 W/cm 2 , must be paid more attention to in circuit design and operation [20]. In fact, heat flux from a chip is not uniform; more than 90% of the total power is dissipated from the core [21]. Although microprocessors, ICs, and other sophisticated electronic components are designed to operate at temperatures above ambient temperature (generally in the range of 60–100°C), the thermal problem for chip cooling is not about maintaining the average temperature of the chip below the design point, but is about maintaining the temperature of the hottest spot below the design point. The hot spot issue makes the heat dissipation near the chip more difficult.

Effective heat dissipation has become a key issue for further development of high-performance semiconductor devices. Efforts on improving thermal management have led to new cooling technology developments, including turbulator-enhanced air cooling, hybrid air–water cooling, water cooling, direct liquid cooling, etc. [22]. As chip power keeps increasing, novel cooling techniques will be needed to maintain chip temperatures within functional temperature limits. The development of liquid cooling technologies such as liquid microchannel/minichannel cooling represents a major evolution in this field [19,23]. No matter what kinds of cooling systems are used, usually heat sinks are necessarily implemented into the packaging, either directly on the backside of a chip or on the heat spreader which sits between the chip and the heat sink, disperses heat, and averages heat fluxes [21]. However, surface asperities greatly limit the actual contact between solid surfaces (e.g., die–heat spreader and heat spreader–heat sink), undermining the thermal conduction. Therefore, a thin layer of thermal interface materials (TIMs), e.g., TIM1 between the chip and the heat spreader and TIM2 between the heat spreader and the heat sink, has been introduced to fill the gap between the asperities to reduce

the contact thermal resistance. Schematic illustration of a typical TIM assembly architecture is shown in Fig. 1.

Fig. 1 Schematic illustration of TIM architecture in flip chip bonding



Unavoidably, a portion of the allowable temperature drop occurs when heat flows from the chip to the heat spreader through TIM1 and from the heat spreader to the heat sink through TIM2. Such temperature drops might not be a big problem in advanced cooling systems but are crucial for air cooling systems. We note that thermal management is more than a technical challenge; economic challenge is often the reason why new cooling technologies find barriers against introduction. The cost, simplicity, and fit considerations of convection methods underscore the desire to cool systems and components with air convection. In an air cooling system, heat dissipation from the chip is by conduction through the many solid layers and interfaces of the package into the heat sink, conduction through the base and into the fins of the heat sink, and finally convection from the fins of the heat sink to the cooling air stream. The temperature drop caused by thermal conduction between the chip and heat spreader is predicted to become comparable to the maximum allowed temperature drop (temperature difference at the chip and heat sink). In other words, the chip–heat spreader thermal resistance will be comparable to the total allowed resistance by 2010, eliminating the possibility for a temperature drop at the heat sink [24]. Heat spreaders are usually high thermal conductivity materials, such as Cu, diamond, AlN, and SiC. Thus, the spreaders themselves cause only a very small fraction of the total thermal resistance. Effective reduction of resistance between the chip and the spreader (the largest possible fraction of system resistance in the near future) is therefore limited to the development of high thermal conductivity TIMs and the suppression of the interfacial thermal resistance between TIMs and mating surfaces. Development of novel TIMs is thus crucial to meet thermal performance requirements for future generations of high-performance IC chips. Heat dissipation challenges create opportunities for fundamental research in materials and thermal management strategies. Specifically, it has been suggested that future cooling approaches may be based on micro- and nanotechnologies [25]. For thermal management applications, the distinctive thermal properties of carbon nanotubes (CNTs) and graphene attract much attention and give rise to new opportunities in thermal management of microelectronic devices and packaging systems.

In this chapter, a brief review is first given on understandings of CNT structure, electrical property, heat transport, and synthesis methods (Sections 2 and 3). Next, CNT applications as interconnects (Section 4) and TIMs (Section 5) are presented, followed by CNT integration into circuits and packaging, with focus on CNT transfer technology (Section 6). Based on the basic understanding of CNT nanoelectronics, a bright future of graphene nanoelectronics is foreseen in Section 7. In comparison with CNT/polymer composites, nanographite/polymer composites are more promising for TIM applications, which will be discussed in Section 8. Section 9 gives a brief summary of current thermal measurement techniques, commercialized and non-commercialized. Finally, in Section 10, future research needs for CNT and graphene applications as interconnects and TIMs in nanoelectronics are discussed.

2 Structure and Properties of Carbon Nanotubes

2.1 Carbon Nanotube Structure

Carbon nanotubes (CNTs) are carbon allotropes, with cylindrical hollow structures and extremely high length/diameter aspect ratio of up to 10^7 [26,27]. Generically, CNTs can be divided into two distinct types – single-walled CNT (SWNT) and multi-walled CNT (MWNT), depending on the number of their walls. An SWNT has only one shell with a small diameter (usually < 2 nm). An MWNT, basically similar to an SWNT in the individual wall structure, consists of two or more concentric cylindrical shells of rolled graphene sheets coaxially arranged around a central hollow with a constant separation between the layers [28–30]. Typically, MWNT diameters are in the range of 2–30 nm.

An SWNT is conveniently illustrated by rolling a graphene sheet of honeycomb structure along the chiral vector C_h [31,32]. The circumference of any SWNT can be described in terms of the chiral vector:

$$C_h = m\alpha_1 + n\alpha_2 \quad (1)$$

where α_1 and α_2 are unit vectors; m and n are numbers of steps along the corresponding unit vectors. The chiral angle, θ , determined relative to the direction defined by α_1 , uniquely defines an SWNT structure. Chiral nanotubes have general (n, m) values and a chiral angle between 0 and 30° . Zigzag nanotubes correspond to $(m, 0)$ or $(0, n)$ and have a chiral angle of 0° ; armchair nanotubes have (n, n) and a chiral angle of 30° . Both zigzag and armchair SWNTs have a mirror plane and thus are considered as achiral. The diameter of an SWNT is given by

$$d = \frac{|C_h|}{\pi} = \frac{\sqrt{3}a_{c-c}\sqrt{m^2 + mn + n^2}}{\pi} \quad (2)$$

where a_{c-c} is the C–C bond length (1.42 Å). Figure 2 shows the vector C_h constructed for $(m, n) = (3, 1)$.

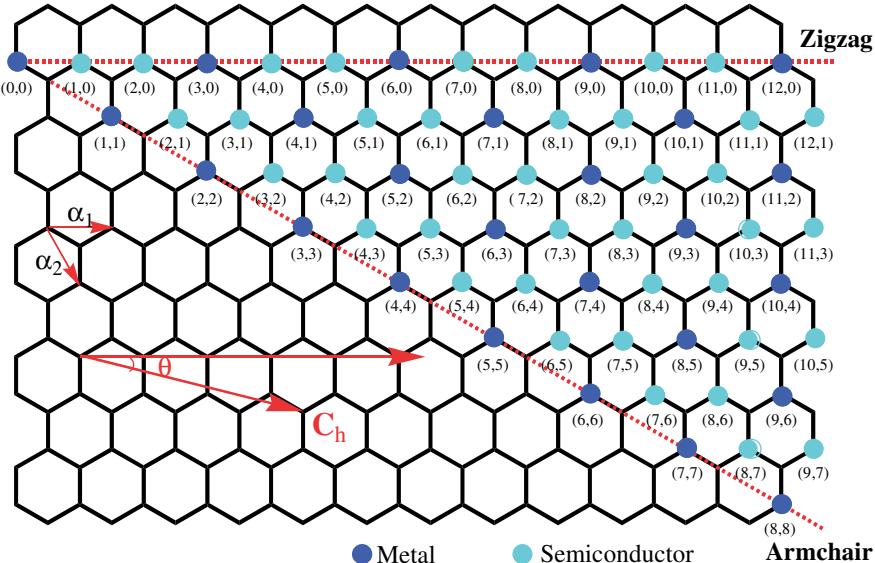


Fig. 2 Formation of a carbon nanotube from a 2D graphene sheet along the vector that specifies a chiral nanotube

The formation of CNTs is determined by free energy. Carbon has four electrons in total in its valence orbitals, which construct an sp^2 -hybridized conjugated structure in the honeycomb-like graphene sheets. However, graphene sheets of a finite size have many edge atoms with dangling bonds possessing high-energy states. When graphene sheets wrap up to become CNTs (although this may not be the real mechanism of CNT synthesis), curvature-induced strain energy is compensated by the reduction in energy by eliminating edge dangling bonds [33]. Although perfect open-ended SWNTs are seamless and fully consist of six-membered rings, real CNTs are far from being exempt of local defects. Typically, synthesized CNTs are capped with high strain energy at the tips. Hexagonal graphitic unit can be locally replaced with heptagonal and pentagonal units, corresponding to positive and negative Gaussian curvatures in the two-dimensional (2D) graphene lattice [34]. In MWNTs, the most distinct defect is the lattice misfit between adjacent walls, also called interfacial dislocation [35,36]. The consequence of this interplanar stacking disorder weakens the electronic coupling between the layers compared with graphite stacks.

2.2 Electronic Structure and Electrical Properties

Recall that a simple approximation based on a tight-binding calculation for the electronic structure of a 2D graphene sheet is expressed as [33]

$$E_{2D}(k_x, k_y) = \pm \gamma_0 \left\{ 1 + 4 \cos\left(\frac{\sqrt{3}k_x a}{2}\right) \cos\left(\frac{k_y a}{2}\right) + 4 \cos^2\left(\frac{k_y a}{2}\right) \right\}^{1/2} \quad (3)$$

where k_x , k_y , and γ_0 are wave vectors and the nearest-neighbor overlap integral, respectively; $a = 0.246$ nm is the in-plane lattice constant. Based on this calculation, the valence and conductance bands touch and are degenerated at the six K (k_F) points at the 2D Brillouin zone corner where the Fermi level in reciprocal space is located. Therefore, graphene is a zero-gap semiconductor. This has proved to be in good agreement with the first-principles results [37].

When graphene is rolled over to form a CNT, the periodic boundary condition of $C_h \cdot k = 2\pi q$ is imposed to eliminate the wave vectors $k = (k_x, k_y)$, where q is an integer. In this way, 1D energy bands can be obtained by slicing the 2D energy dispersion relations [31]. From tight-binding approximations, all CNTs are conductive, but have different gap widths. SWNTs with

$$(n - m) = 3q \quad (4)$$

(where q is an integer) are metallic and have band gap as small as a few meV, while the remainders with

$$(n - m) = 3q \pm 1 \quad (5)$$

are semiconducting and have a higher band gap energy. This suggests that one third of the SWNTs grown are metallic, while two thirds are semiconducting. The band gap for a semiconducting nanotube is given by [38]

$$E_g = 2a_{c-c}\gamma_0/D \quad (6)$$

where D is the nanotube diameter. Therefore, the band gap of a 1-nm semiconducting nanotube is between 0.7 and 0.9 eV, which is in good agreement with scanning tunneling microscopy (STM) measurements [38,39].

MWNTs can also be metallic or semiconducting, analogous to SWNTs, determined by the band gap width. Electron energy loss spectroscopy (EELS) has been experimentally used to probe the local electronic structure of MWNTs. The low electron energy loss spectroscopy (LEELS) contains information about the joint density of states (DOS) related to the individual and collective excitation (plasmons) processes of the valence electrons [29]. For MWNTs, smaller tubes show lower carrier density and greater average carrier localization due to the increasing

curvature. Therefore, MWNTs, which usually have large diameters, are metallically conductive.

Fortunately, owing to the strong C–C bond and rolled graphitic structure, CNTs with even extremely small diameters as 1 nm do not show the Peierls instability – a common phenomenon for most of the 1D materials, which exhibit a metal-to-insulator transition when they come down to nanoscale [40]. When the length of a 1D conductor is smaller than its electron mean free path, the electron transport is ballistic. This has become the most significant property of CNTs. For a perfect metallic SWNT with 100% transparent contacts, low-bias transport (e.g., $V_b < 0.1$ V) is ballistic over length scales that can reach micrometers, due to extremely weak acoustic phonon scattering in the 1D conductor. At high bias, strong excitation of optical phonons by electrons leads to energy dissipation in the tube and current saturation.

For ballistic transport, neglecting electron spin and sublattice degeneracy, the quantum resistance for a single channel is a constant:

$$R_Q = h/2e^2 \quad (7)$$

where h is the Planck's constant and e is the electron charge. Taking into account electron spin and sublattice degeneracy, the corresponding conductance is $G_0 = 2e^2/h$, which is frequently used for metallic SWNTs in literatures.

A single metallic SWNT can carry up to 25 μA of current, which corresponds to a current density of 10^9 A/cm² [41,42]. MWNTs have similar conductance and current density when only the outer wall/shell contributes to the current transport [43–45]. However, multichannel quasiballistic conduction with extremely high conductance of $460 G_0$ and large current-carrying capacity was found in an open-ended MWNT [46]. This implies that CNT conductance and current-carrying capacity will be dramatically improved by integrating open-ended MWNTs which have multichannels in metallic contact with the electrodes.

It was proposed that the conductance of an MWNT or an SWNT is determined by two factors: the conducting channels per shell and the number of shells. The number of shells for an individual MWNT is diameter-dependent [47]:

$$N_{\text{shell}} = 1 + \frac{D_{\text{outer}} - D_{\text{inner}}}{2\delta} \quad (8)$$

where $\delta = 0.34$ nm is the intershell distance, corresponding to the interlayer distance in graphite, and D_{outer} and D_{inner} are the external and the internal tube diameter, respectively. Assuming that the metallic to semiconducting tube ratio is 1/2, the number of conducting channels per shell can be written as [48]

$$\tilde{N}_{\text{chan/shell}} = \begin{cases} (ad + b)/3, & d > 6 \text{ nm} \\ 2r, & d < 6 \text{ nm} \end{cases} \quad (9)$$

where $a = 0.1836 \text{ nm}^{-1}$, $b = 1.275$, r is the metallic nanotube ratio, and d is the diameter of the nanotube shell. One conducting channel provides either quantized conductance G_0 or ohmic conductance ($G_i = G_0 l/\lambda$), depending upon the tube length l . For the low-bias situation that is suitable for interconnect applications, the diameter-dependent channel conductance for one shell is [47]

$$G_{\text{shell}}(d,l) = \begin{cases} G_0 \tilde{N}_{\text{chan/shell}}, & l \leq \lambda \\ G_i \tilde{N}_{\text{chan/shell}}, & l > \lambda \end{cases} \quad (10)$$

where λ is the electron mean free path. For example, the electrical conductance of a metallic SWNT ($\tilde{N}_{\text{chan/shell}} = 2$) is

$$G_{\text{shell}}(d,l) = \begin{cases} 2G_0, & l \leq \lambda \\ 2G_i, & l > \lambda \end{cases} \quad (11)$$

An MWNT consists of several shells, each with its own d , λ , and $\tilde{N}_{\text{chan/shell}}$. Therefore the total conductance comes from the contribution by each shell:

$$G_{\text{MWNT}} = \sum_{N_{\text{shell}}} G_{\text{shell}}(d,l) \quad (12)$$

An SWNT rope or MWNT can be viewed as a parallel assembly of individual SWNTs. Naeemi et al. derived physical models for the conductivity of MWNT interconnects [48]. The results indicate that for long interconnects (hundreds of micrometers) MWNTs may have conductivities several times larger than that of copper or even SWNT bundles. Furthermore, the more robust structure of MWNTs renders them better mechanical stability, which is an advantage for interconnect applications over SWNTs.

As mentioned before, a CNT structure is not perfect. When a few local defects are present, electron scattering is distinctly increased. In actual operation, conductance may be even lower due to electron-electron coupling, intertube coupling, and contact resistance at CNT/substrate interface. Therefore, it is considered promising to use nanotube bundles/arrays aligned in parallel for interconnects.

2.3 Separation of Metallic and Semiconducting SWNTs

In terms of CNT nano devices, distinct properties of metallic and semiconducting SWNTs lead to distinct applications. For example, metallic SWNTs can function as leads in a nanoscale circuit while semiconducting SWNTs could perform as Schottky-type field-effect transistors [43,49–53]. Since SWNT samples produced via various techniques are generally mixtures of metallic and semiconducting

SWNTs, large-scale separation of metallic and semiconducting SWNTs in a nanotube sample is important and has attracted much recent attention [43, 54–70]. Current separation techniques can be divided roughly into the following three categories:

- (1) *Destructive separations.* In 2001, Collins et al. found that Joule heat could burn off metallic SWNTs in an SWNT fabrication, generating entire arrays of nanoscale field-effect transistors based solely on the fraction of semiconducting SWNTs [43]. Miyata et al. reported selective oxidation of SWNTs by hydrogen peroxide. An enrichment in metallic SWNTs of higher than 80% was achieved in the final product [68]. Recently, Huang et al. found that upon laser irradiation in air, metallic SWNTs in a carbon nanotube thin film can be destroyed in preference to their semiconducting counterparts when the wavelength and power intensity of the irradiation are appropriate and the carbon nanotubes are not heavily bundled [67].
- (2) *Non-destructive physical separations.* Krupke et al. took advantage of the difference in the relative dielectric constants of metallic and semiconducting SWNTs with respect to the solvent and used an alternating current dielectrophoresis process to separate metallic from semiconducting SWNTs [55].
- (3) *Non-destructive chemical separations.* Selective chemical functionalization of SWNT sidewalls, covalently or non-covalently, is of particular interest since it facilitates the solubilization/suspension of the functionalized SWNTs in certain organic solvents during the final separation step [54]. Sun's group reported an effective separation process based on the selective interaction of porphyrins toward semiconducting SWNTs [56]. Dyke et al. realized the separation by a controlled reaction process. 4-*tert*-Butylbenzenediazonium salt and 4-(2'-hydroxyethyl)-benzenediazonium salt were used to react with metallic and semiconducting SWNTs in a dispersion, respectively, and sequentially [64]. The as-functionalized SWNTs were finally separated based on a chromatographic separation process. Maeda et al. achieved a metallic SWNT enriched (87%) tetrahydrofuran solution based on selective attachment of amine to metallic SWNT surfaces [57]. This process is characteristic of a large-scale separation; however, the enrichment of the metallic SWNTs needs further improvement. Toyoda et al. synthesized a long-alkyl-chain benzenediazonium compound that reacted to a metallic SWNT surface selectively and then precipitated the unreacted semiconducting SWNTs with an organic solvent [58]. This technique was able to separate semiconducting from metallic SWNTs, whereas it required a careful kinetic control over the chemical reaction. Mioskowski's group reported separation of semiconducting from metallic SWNTs by selective functionalization of semiconducting SWNTs with azomethine ylides [66]. Bao's group reported a spin-assisted process to prepare self-sorted, aligned SWNT networks for thin-film transistors [65]. The selective deposition was based on selective interaction between the semiconducting SWNTs and the functional groups, i.e., amine groups, on the silane monolayer on a silicon wafer. Recently, Tanaka

et al. reported an interesting and effective approach of separating semiconducting from metallic SWNTs. They froze, thawed, and squeezed a piece of gel containing SWNTs and sodium dodecyl sulfate. The solution squeezed out contained 70% pure metallic SWNTs and the gel left contained 95% pure semiconducting SWNTs [70]. Field-effect transistors constructed from the separated semiconducting SWNTs were shown to function without any electrical breakdown.

2.4 Heat Transport

Great progress has been achieved on understanding the extraordinarily high thermal conductivity of individual SWNTs, both theoretically and experimentally [71–76]. Molecular dynamics simulations of an SWNT by Berber et al. indicated that the thermal conductivity of an SWNT can be as high as 6600 W/m-K at room temperature [71]. Pop et al. presented a method for extracting the thermal conductivity of an individual SWNT from high-bias electrical measurements in the temperature range from 300 to 800 K by reverse fitting the data to an existing electrothermal transport model [76]. Their results showed that the thermal conductivity for an SWNT of 2.6 μm long and 1.7 nm diameter was nearly 3500 W/m-K at room temperature.

The high thermal conductivity of SWNTs is based on the large phonon mean free path, i.e., on the order of micron at low temperatures [72,73,77]. Although both phonon and electron contribute to thermal conductance, for SWNTs at low temperatures, or even at room temperature, electron contribution is negligible compared to that from phonons.

For graphite and related systems, the heat capacity can be written as

$$C = C_{\text{ph}} + C_{\text{e}} \quad (13)$$

where C_{ph} and C_{e} are the contributions from phonons and electrons, respectively. If $T \ll \hbar v / \kappa_B R$, C_{ph} can be written as [78]

$$C_{\text{ph}} = \frac{3L\kappa_B^2 T}{\pi \hbar v} \times 3.292 \quad (14)$$

where κ_B is the Boltzmann's constant, v is the velocity of sound ($\sim 10^4$ m/s), and L is the tube length. The electronic specific heat for a metallic SWNT is also linear with temperature for $T \ll \hbar v_F / \kappa_B R$:

$$C_{\text{e}} = \frac{4\pi L\kappa_B^2 T}{3\hbar v_F} \quad (15)$$

where v_F is the Fermi velocity ($\sim 10^6$ m/s).

Therefore, the ratio between phonon and electron contributions to the specific heat is

$$\frac{C_{\text{ph}}}{C_e} \approx \frac{v_F}{v} = 100 \quad (16)$$

Experiments showed a linear increment of thermal conductance of SWNTs with increased temperature at low temperatures [74,77]. However, at high temperatures, e.g., from 300 to 800 K, the thermal conductivity decays with temperature due to strong scattering. The length effect of the thermal conductivity of SWNTs has been measured by a 3ω method, showing that thermal conductivity of SWNTs increases with length, which is consistent with theoretical predictions [79].

For MWNTs, heat transport is complex due to the intershell interaction. The strong phonon coupling between the MWNT shells should cause a behavior similar to that of graphite. A thermal relaxation technique has been used to measure the specific heat of MWNTs from 0.6 to 210 K; results indicate that MWNTs have a similar specific heat to that of SWNT ropes and graphite [80]. Thermal conductivity of MWNTs has been experimentally measured [81–83]. Kim et al. developed a microfabricated suspended device hybridized with MWNTs ($\sim 1 \mu\text{m}$) to allow the study of thermal transport where no substrate contact was involved [81]. The thermal conductivity and thermoelectric power of a single CNT were measured, and the observed thermal conductivity was $>3000 \text{ W/m-K}$ at room temperature. Choi et al. obtained the thermal conductivity of individual MWNTs (outer diameter of $\sim 45 \text{ nm}$) by employing a 3ω method [82,83]. The thermal conductivity was reported to be $650\text{--}830 \text{ W/m-K}$ at room temperature. Given that MWNTs are more easily synthesized and assembled and possess higher rigidity, MWNTs have been considered more promising, compared with SWNTs, to be applied as TIMs in next-generation electronic packaging.

In an SWNT rope, phonons propagate along individual tube axes as well as between parallel tubes, leading to phonon dispersion in both longitudinal and transverse directions. The net effect of dispersion is a significant reduction of the specific heat at low temperatures compared to an isolated nanotube [84]. Yang et al. investigated the thermal conductivity of MWNT films prepared by microwave CVD using a pulsed photothermal reflectance technique [85]. The average thermal conductivity of carbon nanotube films, with film thickness from 10 to $50 \mu\text{m}$, was $\sim 15 \text{ W/m-K}$ at room temperature and independent of tube length. However, by taking into account a small volumetric fraction of CNTs, the effective nanotube thermal conductivity can reach 200 W/m-K .

Finally, we should note that the influence of structural defects on the heat transport in the 1D structure of CNTs could be anomalously high, much larger than in bulk materials. Yi et al. measured the thermal conductivity of millimeter-long aligned MWNTs [86]. At room temperature, the thermal conductivity of these samples was only $\sim 25 \text{ W/m-K}$, due to a large number of CNT defects. However,

thermal conductivity recovered to a high value after the aligned MWNTs were annealed at 3000°C. Further discussion about MWNT films for TIMs will be presented in Section 5.

3 CNT Growth

Reliable synthesis techniques capable of yielding high-purity CNTs in desirable quantities and at selective positions on growth substrates are critical to realize CNT applications in electronic packaging. Till now, various synthesis techniques have been developed to produce CNTs for specific research uses or applications, mainly including arc-discharge, laser ablation, and various chemical vapor deposition (CVD) processes, each of them having its own advantages and disadvantages. Since there are a large number of scientific literatures dealing with various CNT syntheses, the readers are referred to other monographs on CNTs and the related references in this section for detailed information. Only a brief summary and comparison is presented in this section.

3.1 Arc-Discharge and Laser Ablation Methods

In 1992, Ebbesen and Ajayan adapted the standard arc-discharge technique used for fullerene synthesis to a gram-level synthesis of MWNTs under a helium atmosphere [28]. In this process, carbon atoms are evaporated with inert gas plasma characterized by high electric currents passing between opposing carbon electrodes (cathode and anode). Usually the carbon anode contains a small percentage of metal catalyst such as cobalt, nickel, or iron. The results show that the purity and yield depend sensitively on the gas pressure in the reaction vessel. The length of the synthesized MWNTs is several micrometers with diameters ranging from 2 to 20 nm. However, such synthesized CNTs are inevitably accompanied by the formation of carbon particles that are attached to the nanotube walls. A subsequent purification process is necessary to achieve high-purity nanotubes. In 1993, Bethune and coworkers reported a large-scale synthesis of SWNTs by an arc-discharge method using a carbon electrode that contained ~2 at.% cobalt [87]. At high temperatures, carbon and metal catalysts are co-vaporized into the arc, leading to the formation of carbon nanotubes with very uniform diameter (~1.2 nm). However, fullerenes (by-products of the arc-discharge process) also form readily in this process. In order to obtain high-purity SWNTs, a purification process is therefore necessary. In 1996, the Smalley group reported the synthesis of high-quality SWNTs with yields > 70% using a laser ablation method [88]. This method utilized double pulse lasers to evaporate graphite rods doped with 1.2 at.% of a 50:50 mixture of Co and Ni powder, which was placed in a tube furnace heated to 1200°C in flowing argon at 500 Torr; this process was followed by heat treatment in vacuum at 1000°C to sublime C₆₀ and other

small fullerenes. The resulting SWNTs were quite uniform, had a diameter of ~ 1.38 nm, and formed ropes consisting of tens of individual SWNTs closely packed into hexagonal crystal structures that were stabilized through van der Waals forces.

The success in producing large quantities of CNTs by arc discharge offers wide availability of CNTs for fundamental studies and exploration of potential applications. However, there are several concerns associated with these two growth methods: first, energy consumption is large; second, only powdered samples with tangled CNT bundles are produced, thereby making purification, manipulation, and assembly difficult; third, by-products are usually unavoidable, including fullerenes, graphitic polyhedrons, and amorphous carbon in the form of particles or overcoats on nanotubes; fourth, patterned growth of CNTs is a great challenge [89].

3.2 Chemical Vapor Deposition (Common Thermal CVD, Plasma-Enhanced CVD, and Liquid Injection CVD)

Both MWNT and SWNT syntheses have been well developed using chemical vapor deposition (CVD) processes. Generally, a CVD process for CNT synthesis involves decomposition of carbon sources, e.g., gaseous precursors or volatile compounds of carbon, catalyzed by metal nanoparticles at high temperatures. The mechanism for CNT growth has been generally assumed to be a dissociation–diffusion–precipitation process during which elemental carbon is formed on the surface of a metal particle followed by diffusion and precipitation in the form of cylindrical graphite [90,91]. However, the detailed nanotube growth mechanisms are still not well understood.

CVD processes have great advantages in controllable tuning of CNT length, diameter, morphology, alignment, doping, and CNT adhesion on substrates. This is why CVD is the most attractive for CNT applications in nanoscale electronics. Various CVD processes have been developed, including common thermal CVD [92], plasma-enhanced CVD (PECVD) [93,94], floating catalyst CVD (e.g., HiPCO) [95], liquid injection CVD (LJ-CVD) [96], hot filament CVD (HFCVD) [97], etc. A brief comparison between thermal CVD, PECVD, and LJ-CVD is listed in Table 1.

Table 1 A brief comparison between CNTs synthesized by thermal CVD, PECVD, and LJ-CVD

	Growth rate	Quality	Packaging density	Cost
Thermal CVD	Low	High	Low	High
PECVD	Medium	Medium	Medium	Medium
LJ-CVD	High	Low	High	Medium or low

The critical parameters in CVD growth of CNTs are the carbon precursor, catalyst composition, reactor chamber pressure, and growth temperature. The effects of these parameters on CNT growth have been investigated extensively. Growth temperature, typically ranging from 400°C [98,99] to 3600°C [100], has a significant effect on CNT growth. Generally, PECVD can lower the growth temperature, compared to thermal CVD, because plasma dissociates hydrocarbon molecules more efficiently and promotes surface diffusion of carbon on the catalyst droplet, thereby allowing CNT growth to occur at lower temperatures than are possible with thermal CVD [101,102]. Interestingly, a thermal CVD process for CNT growth (even SWNT growth) at as low as 450°C was recently reported [103]. PECVD also has particular advantages in good alignment of CNT arrays, which is significant for CNT application as vertical interconnects [104]. One drawback of the current synthesis of aligned CNT (ACNT) arrays by thermal CVD or PECVD methods is the relatively low ACNT packing density on substrates. This density limitation mainly comes from the low density of catalyst nanoparticles deposited on a substrate. Usually, a thin layer of catalyst is deposited on a substrate surface, e.g., by an e-beam evaporation process, for CVD or PECVD synthesis. The thickness of the catalyst layer after deposition is typically 0.2–6 nm. A proper thickness of the catalyst layer is one prerequisite for formation and separation of catalyst particles at high temperatures

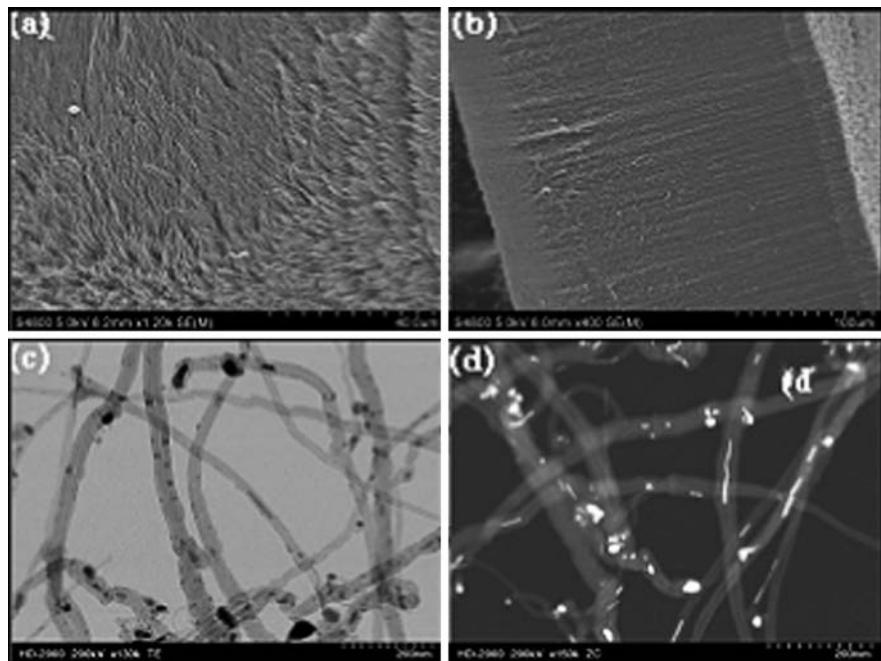


Fig. 3 SEM images of aligned CNT arrays synthesized by LJ-CVD: (a) side view; (b) top view; (c) TEM image of randomly dispersed individual CNTs; (d) Z-scan TEM image of randomly dispersed individual CNTs with bright spots indicating metal catalysts

with small distribution in diameter. For thermal CVD process, ACNT packing density on silicon substrate typically lies between 300 and 1000 per μm^2 . This means that, for ACNTs with average diameter of 15 nm, ACNT coverage on silicon substrate is < 10%. Although such a surface coverage is sufficient to, in theory, provide excellent electrical or thermal conductance, defects are unavoidable in real CNTs, which greatly degrade the intrinsic properties of ACNTs. Therefore, for electrical or thermal applications, higher packing density of ACNTs on substrates is desired.

The advantage of liquid injection CVD (LJ-CVD) over thermal CVD and PECVD is to grow ACNT arrays with a higher packing density [96]. LJ-CVD does not need a catalyst layer on the substrate; instead, the catalyst is generated *in situ* in the reaction chamber by decomposition of a metal catalyst precursor, i.e., ferrocene dissolved in xylene. Shown in Fig. 3a and b are the SEM (scanning electron microscope) images of ACNTs synthesized by LJ-CVD. A TEM image in Fig. 3c shows the diameters of CNTs ranging from 20 to 30 nm. Since the metal catalyst was formed continuously during the CNTs growth in LJ-CVD, it is randomly embedded in the cavity of CNTs, as evidenced by the Z-scan TEM image (Fig. 3d). From TGA results, the average metal content in CNTs arrays is about 10–15 wt%. Currently, helical CNT arrays have been synthesized by LJ-CVD, indicating more flexibility of LJ-CVD in controlling CNT morphology [105].

4 Carbon Nanotubes for Interconnect Applications

The International Technology Roadmap for Semiconductors (ITRS) predicts that in 2010 the current density will reach the order of magnitude of 10^6 A/cm^2 , a value which can only be supported by CNTs. Wei et al. [42] measured the reliability of two individual MWNTs as electrical interconnects, by forcing a constant current through the CNTs at 250°C. The CNTs showed no degradation after 350 h at current densities exceeding 10^{10} A/cm^2 . Naeemi and Meindl [106] suggested that mono- or bi-layer metallic SWNTs may be promising candidates for short local interconnects. We should note that, at small bias voltages, i.e., for interconnect applications, mean free path in CNTs is large; however, at large bias voltages, i.e., transistor applications, electron–phonon scatterings in CNTs degrade the transconductance [7]. Therefore, transistor applications require assembly of ultra-short CNTs, which is still a big challenge for GSI. In this section, a brief review is given on current progresses in vertical ACNT interconnects.

Vertical interconnects pass through vias or contact holes in the dielectrics to connect horizontal interconnects to the source, drain, or gate metallization of transistors. Researchers from Fujitsu and Infineon have investigated vertical CNT interconnects extensively [3,8,107–110]. In ref. 3, holes were formed using conventional photolithography and subsequent wet etching with a buffered HF solution. Catalyst was selectively introduced into the vias by deposition followed by a lift-off process. However, CNTs grown inside the trenches are of low packing density and entangled. Therefore, Li et al. [4] introduced an alternative bottom-up approach to

separate ACNTs, as shown in Fig. 4. They grew vertical ACNTs on patterned catalyst spots and then used SiO_2 to fill the gaps. After chemical mechanical polishing, the ACNT tips shot out and were ready for next step of metallization. In order to improve the CNT/substrate interfacial strength after CVD growth and simultaneously achieve low-resistance ohmic contact, Nihei et al. [5] modified the catalyst composition by using a layer of titanium below the nickel catalyst layer. It was postulated that titanium carbide formed during the growth at the CNT/Ti interface. Actually, conductive contacts between the graphitic network of CNTs and metallic electrodes are of great importance in electronic devices. To date, most results indicate that the properties of CNT-interconnected devices are dominated by the electronic behavior at the CNT–metal contacts [111]. Yang et al. [85] reported using intense electron beam irradiation to form covalent bonding between CNTs and metal particles; metallic conductivity of the junctions with low resistance was formed. Nihei et al. [5] has also predicted that uniformly small diameters of CNTs and a higher CNT packaging density would be beneficial. In this regard, a conventional CVD process needs improvement. Normally, catalyst particles for CVD growth of CNT are formed from a deposited thin catalyst film by a high-temperature treatment that causes segregation of the film into small particles or islands. There is no guarantee that the CNTs would be uniformly small in diameter with high packing density on growth substrates. Awano et al. [8,109,112] have reported a size-classified catalyst nanoparticle CVD process for controlled growth of CNTs with uniform diameter and high density at low temperatures. Laser ablation is used to evaporate catalyst particles from a catalyst metal target. The resulting nanoparticles are size-separated by a low-pressure differential mobility analyzer in terms of their movement within an electric field. The size-classified nanoparticles are then deposited onto substrates in a deposition chamber. Finally, the substrates with deposited nanoparticles are placed in a hot-filament CVD chamber for nanotube growth. The nanoparticles present will not coalesce on the substrate during CNT growth due to the low CNT growth temperature.

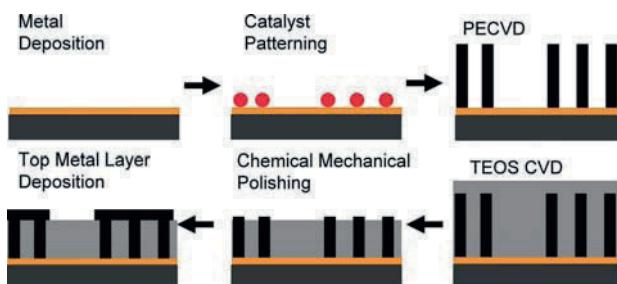


Fig. 4 Schematic process of the bottom-up approach for ACNT interconnects [4]

ACNT growth on copper substrates was a big challenge. Awano et al. [109] put a 5-nm Ta barrier layer between the catalyst layer and the copper wiring film and used the process described in ref. 8 to grow ACNTs on the Cu film for interconnection. A schematic of a CNT via fabrication process is shown in Fig. 5. The top contact

layers of 50-nm Ti and 300-nm Cu were deposited to form connections to the CNTs. These metal layers were then patterned, and the Co particles on the SiO₂ layer were simultaneously removed during the etching step. Figure 6 shows a SEM image of CNT vias with a diameter of 2 μm grown by a size-classified catalyst nanoparticle CVD process [8]. Via resistance for a 2-μm-diameter CNT structure was measured with a four-point probe using Kelvin patterns and a TiN contact layer. The measured resistance of CNT vias is ~0.59 Ω at room temperature. This value is of the same order as the theoretical value of W plugs but is still one order of magnitude higher than the theoretical value of Cu vias. Further improvement is needed on CNT quality and packing density control.

Fig. 5 Schematic process of ACNT via fabrication [109]

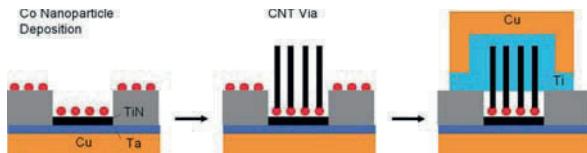
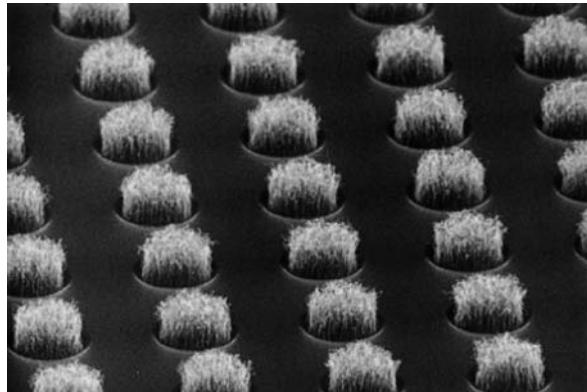


Fig. 6 A SEM image of the ACNT via array produced by the size-classified catalyst nanoparticles



5 Carbon Nanotubes as Thermal Interface Materials (TIMs)

5.1 A Brief Review of TIMs and Requirements for Next-Generation TIMs

TIMs play a key role in developing thermal solutions and have been extensively researched in the past decades [21,113–131]. Developing new TIMs is the route toward higher packaging system reliability. Commonly used TIMs include polymers filled with fillers of high thermal conductivity, solders, and filled phase change

materials (PCMs). Although polymer-based TIMs, e.g., boron nitride (BN)/polyethylene glycol (PEG), BN/silicone, have their advantages in good adhesion, high compliance, easy processing, and low cost, their thermal conductivities at low filler loadings are relatively low. Increasing filler loading improves the thermal conductivity of TIM layer itself (~ 7 W/m-K) but brings down its compliance and increases bond line thickness (BLT), which consequently increases contact thermal resistance. Additionally, reliability is also an issue for polymer-based TIMs because compliance of polymers degrades with operation time. In comparison, solder TIMs such as In, AuSn, InPb, and InSnBi have high thermal conductivities (50–100 W/m-K), low melting point (<170°C), and high compliance and are now actively being pursued as TIMs. Currently, thermal resistance of a solder TIM layer (R_{TIM} , including the resistance at the two contact interfaces) is typically 5–20 mm²-K/W. However, they have the drawbacks such as high CTE (coefficient of thermal expansion), low compliance, high cost, relatively low fatigue strength, voids issue, and complexity in application. PCMs have high fluidity during processing, making use of the latent heat of phase change during heating up, which dissipates much more heat generated by the die. However, PCMs, compared with solder TIMs, are not so popular due to a reliability issue because sometimes PCMs leak (flow out) from the interfaces during operation. In general, none of the above TIMs meet the requirements for next-generation TIMs; implementation of nanotechnology seems unavoidable.

5.2 Carbon Nanotube/Polymer Composites

Amorphous polymers such as epoxy resin have extremely low thermal conductivities (usually ~ 0.2 W/m-K). One challenging issue in polymer TIM composites is that thermal conductivity of the composite should be enhanced with limited filler loading levels. CNT, as the filler, theoretically is able to effectively enhance the thermal conductivity of polymer composites. Great efforts have been dedicated, however, with only moderate success [132–140]; representative data from the references are summarized in Table 2, showing that no significant enhancement has been achieved yet. Researchers have been trying to understand such a phenomenon [138,141–148]. One important factor is structural defects in CNTs that act as scattering centers for phonons. The other widely accepted finding is that a large interfacial thermal resistance across the nanotube–polymer interface causes a significant degradation in the thermal conductivity enhancement.

Thermal resistance at an interface – Kapitza resistance, R_K – represents a barrier to heat flow associated with the differences in phonon spectra of two phases in contact. For high structural mismatch and weak contact at an interface, R_K is large.

In thermal transport, R_K is defined by

$$R_K = \frac{a_K}{K_m} = \frac{1}{\Lambda} \quad (17)$$

Table 2 Summary of published data on effective thermal conductivity of CNT/polymer composites

Matrix	CNT	CNT loading	CNT length (nm)	Outer wall diameter (nm)	K_e/K_m
Epoxy	SWNT	1 wt%	100–2000	Broad distribution	2.25
Epoxy	SWNT	1 vol%	77–257	1.1	< 1.08
Epoxy	SWNT	0.5 vol%	77–257	1.1	1.30
Silicon	SWNT	3.8 wt%	—	—	1.65
Silicon	MWNT	0.4 vol%	300	12	2.16
Epoxy	MWNT	0.2 vol%	~5000	< 2	1.06
Epoxy	SWNT	0.2 vol%	~1000	15	1
Epoxy	MWNT	0.2 vol%	~5000	< 2	1.01

K_e and K_m refer to the thermal conductivities of the composite and the polymer matrix, respectively.

$$J_Q = \Lambda \Delta T \quad (18)$$

where J_Q , Λ , R_K , K_m , and a_K are heat flux, interfacial thermal conductance, interfacial thermal resistance, thermal conductivity of the matrix, and Kapitza radius, respectively.

In a CNT/polymer nanocomposite system, the R_K plays a much more important role than in common composites with micro-sized fillers, not only because the interface phase accounts for as high as 15 vol% but also because the a_K is around 5–20 nm, comparable to or even larger than the tube diameter [149,150]. Experimental measurements showed the R_K in CNT suspension is around $8 \cdot 10^{-8}$ m²·K/W [150]; this value is of the same order of magnitude as those in other composite materials [149]. Such a large R_K arises from the weak coupling between the rigid tube structure and the soft polymer matrix [143]. For a nanotube, there are four acoustic modes: a longitudinal mode, corresponding to motion of atoms along the tube axis; a twist mode, corresponding to torsion around the tube axis; two degenerate transverse modes – bending and squeezing modes, corresponding to atomic displacements perpendicular to the nanotube axis. The longitudinal and twisting modes involve only the movement of carbon atoms tangential to the tube surface and thus are expected to couple very weakly to the soft material around the tube, while the bending and squeezing modes involve out-of-plane motion and thus are more strongly coupled with the soft material around. Due to the weak coupling, sharp contrast in the phonon spectra at the interface between CNTs and an amorphous polymer matrix leads to strong backscattering of thermal energy waves at the interface.

Based on the theory above, an effective approach to reducing R_K is to provide a stronger coupling between CNTs and a polymer matrix by covalent bonding. Therefore, proper surface functionalization of CNTs should help improve the interaction between the CNT fillers and the polymer matrix and consequently result in a significant increase in the interfacial thermal conductance, [144] as shown in Fig. 7a.

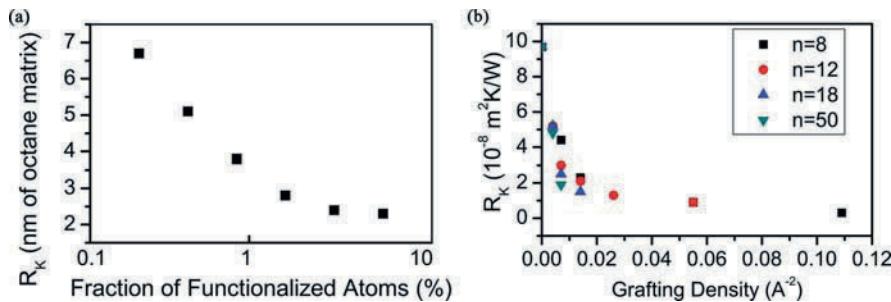


Fig. 7 Effects of surface functionalization (a) and grafting density of organic chains (b) on the interfacial thermal resistance of CNT/polymer matrix [145]

Atomistic-based modeling was conducted to calculate the R_K and found that grafting organic chains or polymer chains to the surface of CNTs with covalent chemical bonds increases the conductance across the interface [145]. R_K was found to be a function of several parameters, e.g., grafting density and length of the linear chain covalently bonded to the CNT surface. Increasing chain length and grafting density lowers the interfacial thermal resistance, as shown in Fig. 7b.

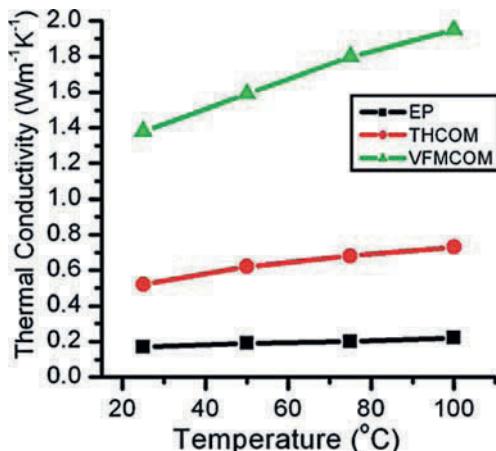
Some modeling and experiment have been performed to examine the conclusions [138,146]. However, conventionally, length and diameter of functionalized CNTs are not controllable. Evaluation of the effects of surface functionalization on the R_K , and consequently on the overall effective thermal conductivity of the composite, should exclude the influence brought about by the uncontrollable length and diameter of CNTs. This problem becomes a barrier for understanding the thermal transport in CNT/polymer systems, especially the R_K .

Progress has recently been made in understanding the influence of interface bonding between CNTs and a polymer matrix on the overall thermal conductivity of the composite. Lin et al. embedded ACNTs into an epoxy matrix and used variable frequency microwave (VFM) radiation to assist the curing process; the interface between the ACNTs and the epoxy matrix was enhanced [151]. In fact, microwave irradiation has attracted much interest in synthetic organic chemistry due to its special role in dramatically increasing reaction rates and the capability of inducing chemical reactions which cannot proceed by thermal heating alone [152–156]. It is postulated that during the curing process in the VFM chamber, the interfacial bonding between the CNTs and the epoxy matrix was dramatically improved. The enhanced interface effectively improved the overall thermal conductivity of the composite, as shown in Fig. 8.

5.3 Aligned Carbon Nanotube (ACNT) Based Thermal Interface Materials (TIMs)

Vertically aligned CNTs utilize the superior longitudinal thermal conductivity of individual nanotubes and exhibit the overall thermal conductivities of $\sim 80 \text{ W/m-K}$

Fig. 8 Thermal conductivities of pure epoxy, ACNT/epoxy composite cured by thermal heating (THCOM), and ACNT/epoxy composite cured in a VFM chamber (VFMCOM)



or higher [116,157–162]. Fabrications and characterizations of ACNT TIMs have been a recent research focus. Hu et al. used a 3ω method to test the thermal contact resistance between a 13- μm -thick ACNT array and the surface of a free mating substrate. The results showed that the contact resistances were 17 and 15 $\text{mm}^2\text{-K/W}$, respectively, under the pressures of 0.040 and 0.100 MPa [163]. Wang et al. used a non-contact technique (photothermal) to measure the thermal conductivity of MWNT arrays but revealed only $\sim 27 \text{ W/m-K}$ for MWNTs [164]. Ngo et al. used electrodeposited copper as a gap filler to enhance the stability and thermal transport of carbon nanofibers [165]. They reported the interfacial thermal resistance of $25 \text{ mm}^2\text{-K/W}$ under a pressure of 0.414 MPa with a 1D reference bar method. Using a similar characterization method, Xu et al. reported a minimum thermal interface resistance of $19.8 \text{ mm}^2\text{-K/W}$ for an ACNT array on a silicon substrate [116]. By using phase change material (PCM) with ACNTs, they produced a minimum resistance of $5.2 \text{ mm}^2\text{-K/W}$, as shown in Fig. 9. Xu et al. used a photothermal metrology to evaluate the thermal conductivity of aligned CNT arrays grown on silicon substrates by PECVD [166]. The effective thermal resistance was $12\text{--}16 \text{ mm}^2\text{-K/W}$, which is comparable to the resistance of commercially available thermal grease. Tong et al. used a phase-sensitive transient thermoreflectance technique to measure the thermal resistance of the two interfaces on each side of a vertically aligned CNT array as well as the ACNT itself [167]. They concluded that the interface between the free-end CNT tips and the opposing substrate in contact dominated the interfacial resistance. Cola et al. fabricated an interface material comprised of a metal foil with CNTs synthesized on both sides of the foil. This fabrication lowered the interfacial resistance to less than $10 \text{ mm}^2\text{-K/W}$ [160]. Furthermore, they grew ACNTs on Si and copper substrates and fabricated them to make a two-sided ACNT TIM layer. By this assembly, they reported a minimum resistance of $4 \text{ mm}^2\text{-K/W}$.

A brief comparison between the thermal interfacial resistance for currently used TIMs and ACNT TIMs is listed in Table 3. Although solder TIMs have the lowest

Fig. 9 Thermal resistance of copper–silicon interfaces with a CNT array (Cu–PCM–CNT–Si), compared with PCM (Cu–PCM–Si), as a function of pressure [116]

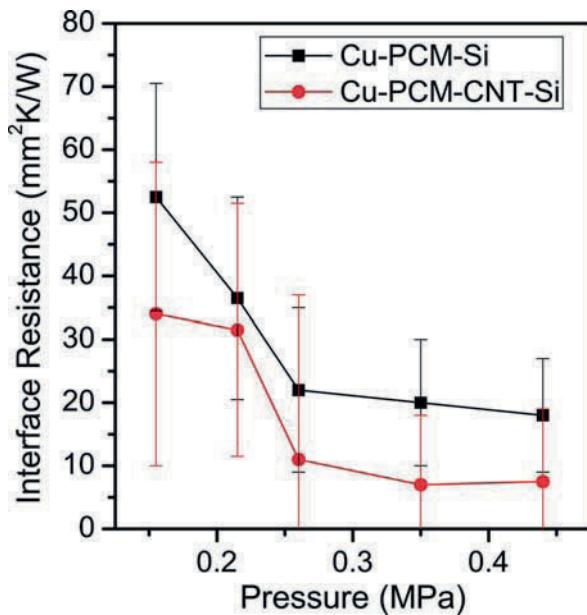


Table 3 A comparison of the best data reported on TIM resistance, reduced to the equivalent BLT of 1 mil

Polymer composites	Filled PCMs	Solders	Graphite paper	ACNT
7	8	2	10	4–20

thermal resistivity among all the commonly used TIMs so far, bond line thickness (BLT) of solder TIMs cannot be as small as polymer-based TIMs due to the high CTE and processing complexity; therefore, the final total thermal interface resistance of solder TIMs is not so low as expected, which hinders application of solder TIMs as next-generation TIMs. The best data reported so far on ACNT TIMs, by Fisher's research group at Purdue University, lie in the range of 4–20 mm²-K/W [116,157,160,163,168]. However, it should be further understood why different synthesis methods or catalyst compositions give the same CNT quality and why the same synthesis method gives totally different CNT diameter, diameter distribution, and substrate coverage. In fact, a high reproducibility in ACNT synthesis is a prerequisite for ACNT TIM applications, which requires further understanding on CNT growth mechanism and surface science involved in the growth substrate during a CVD process. This will be discussed in detail in Section 5.5.

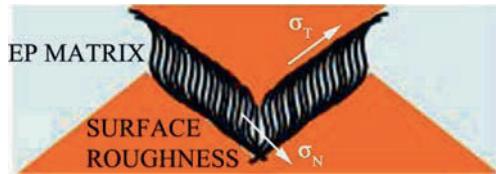
What indeed determines the interfacial thermal resistance for an ACNT TIM?

First, the intrinsic thermal resistance of the ACNTs is an important factor. Theoretically, a longer CNT possesses a better thermal conductivity. However, in reality, this may not be true. Unavoidably, more defects present in a long

CNT will lead to thermal conductance degradation. Therefore, the shorter the ACNT array, the better the thermal conductance. There is no doubt that improving CNT structural perfection by high-temperature annealing would be beneficial. However, common growth substrates such as silicon, copper, and quartz cannot sustain an effective annealing temperature as high as, for instance, 2400°C. Therefore, microwave annealing has been considered an alternative method of CNT post-growth annealing [169].

Second, look at the question again: “what is a TIM?” The answer seems so simple: a layer between two mating substrates (let us say “A” and “B”) for more effective heat dissipation. Yes. Then “what is the picture of the real micro-structure at the A-TIM-B interface?” Or let us ask the question in a different way: “Should A and B be separated by the TIM layer or not?” As a matter of fact, an ideal TIM should go only into the asperities created by the surface roughness of A and B while maintaining as much direct A–B solid contacts as possible. The direct A–B contacts are the main paths where heat flux flows. In this sense, no matter how low the thermal resistance of the TIM is, since an ACNT TIM unavoidably blocks part of the direct A–B contacts, it introduces extra resistance to the direct A–B contacts. This is evident when we refer to a series model for thermal resistance calculations. Of course, ACNTs filled in the asperities reduce the resistance at the non-contact sites. In these regards, a short ACNT TIM layer is preferred. Too small a thickness is, however, not desirable, not only because of the difficulty to grow uniform ACNTs with such a small length, but also because of the surface waviness and the contact pressure. During the pressure-involved TIM assembling, ACNTs become cranked under the pressure at the contact interface with dramatically increased CNT coverage [94]. Figure 10 shows an ideal CNT status at the interface. For a properly rough (waved) surface, when a proper pressure is imposed, more actual contact area between the CNTs and the mating surface can be achieved and consequently better thermal conductance.

Fig. 10 An illustration of the ACNT/EP TIM status at the mating surfaces



Third, there are still chances that a direct A–B contact exhibits a higher resistance than the case with a TIM in between that separates A and B; this means that a classical series model for resistance calculation fails. This is possible when we consider that fundamental of interfacial thermal resistance. The thermal resistance at the A–B contact sites originates from phonon spectrum mismatch. Different A–B combinations show different thermal resistance. In this sense, a short ACNT TIM layer between A and B may play a role of mitigating the A–B interfacial phonon

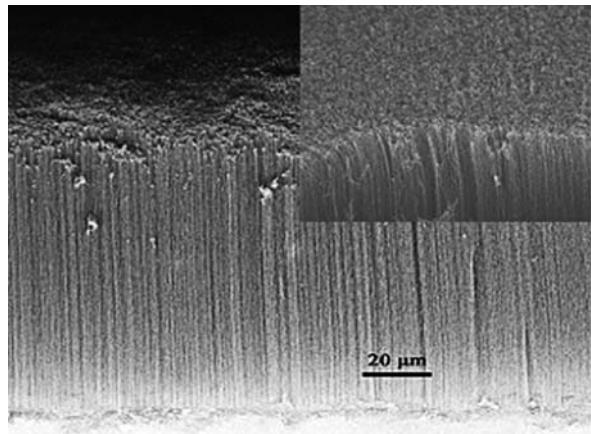
scattering by avoiding a sharp A–B contact. In this case, the interfacial modification at the ACNT/substrate interface is important and will be discussed in detail in Section 6.2.

5.4 ACNT TIM Synthesis on Bulk Copper Substrates

Although ACNT synthesis on silicon as the growth substrate has been widely investigated, direct synthesis of a ACNT layer on the backside of a silicon device is not compatible with a current front-end semiconductor fabrication process due to the high temperature (typically $> 650^{\circ}\text{C}$) required for ACNT syntheses by, for example, chemical vapor deposition (CVD). Great efforts have been made on low-temperature ACNT syntheses; however, overall speaking, the CNT qualities by low-temperature CVD processes are low [102,170–212]. Thus, a silicon substrate is not the growth substrate of choice for real-life ACNT TIM applications; instead, ACNT synthesis on a bulk copper substrate, i.e., a copper lid, is preferred. A few papers have reported ACNT syntheses on various metal substrates; however, the direct ACNT synthesis on a bulk copper substrate is still a great challenge [213–219]. Here, we emphasize “bulk” copper because ACNT synthesis on a thin copper coating on a silicon wafer has been demonstrated for interconnects as discussed before. However, a thin copper coating by, for example, e-beam evaporation, DC sputtering, and electroplating on a smooth substrate is different from a bulk copper substrate in terms of purity, roughness, and crystallinity. Wang et al. developed a process to grow ACNTs on a copper foil surface by water-vapor-assisted CVD [220]. However, the quality and the structure of the “ACNTs” on copper were not clearly described. Recently, Yin et al. reported growing multi-walled CNTs on an oxygen-free copper substrate, however, with a poor CNT alignment [221]. Integrations of ACNTs grown on copper into thermal management were reported by Fisher’s research group [116,157]. Nevertheless, the ACNTs in these two references exhibited different average diameters, diameter distributions, and CNT coverage on the growth substrates, even though the same synthetic conditions were used. Awano et al. synthesized ACNT arrays on TiN-coated copper plates for interconnect demonstration; however, from Fig. 6, we see that the as-synthesized ACNTs show a poor alignment and uniformity. Moreover, Raman spectra of the ACNTs were not given. Recently, Lin and Wong [222] reported a remarkable progress on rapid synthesis of high-quality ACNTs on bulk copper substrates through a common thermal CVD process. By introducing a well-controlled conformal Al_2O_3 support layer on the bulk copper substrate by atomic layer deposition (ALD) prior to the deposition of the iron catalyst layer, they reproducibly synthesized ACNTs of good alignment and high quality on the copper substrate. A SEM image of the synthesized ACNT on copper is shown in Fig. 11.

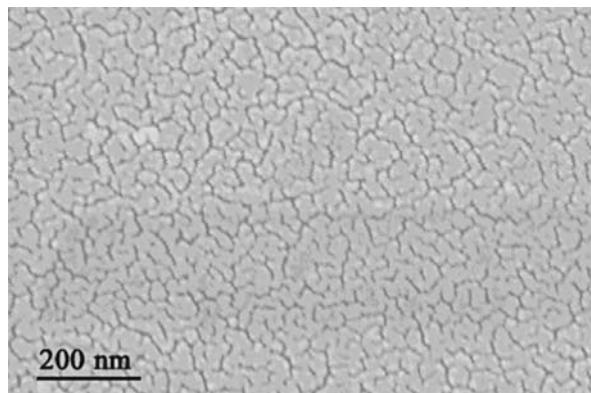
Al_2O_3 support layers of 10- to 15-nm thick have been used to accelerate ACNT growth due to their capability of decomposing hydrocarbons and the interfacial interactions between Al_2O_3 and iron nanoparticles during the CVD syntheses of

Fig. 11 A side-view SEM image of the ACNT TIM synthesized on a polished bulk copper substrate (*inset*: a further magnification)



ACNTs [92,223,224]. Generally, literatures describe the conditions for the formation of an Al_2O_3 deposit layer on a target substrate by e-beam evaporation; however, the continuity of the deposit has rarely been mentioned. During an e-beam evaporation process, a continuous deposit layer forms on a substrate surface through nucleation and growth steps. At an early stage of deposition, heterogeneous nucleation may become rate controlling. This nucleation mechanism leads to isolated aggregates of condensate on the substrate when the deposit layer is thin [225]. This discontinuity is briefly discussed by de los et al. [225,226]. In fact, the discontinuity of the deposit is the right reason why a thin layer of gold deposit on a bare silicon surface by e-beam evaporation was recently employed to selectively etch silicon wafers (Fig. 12) [227]. Therefore, when the Al_2O_3 deposit layer is too thin to become continuous, isolated nanoaggregates are formed [226,228]. In this case, the bottom copper surface is partially exposed to the top iron catalyst and/or the carbon source at the inter-grain voids during the CVD process and the ACNT synthesis fails. There are at least two reasons that account for the synthesis failure. First,

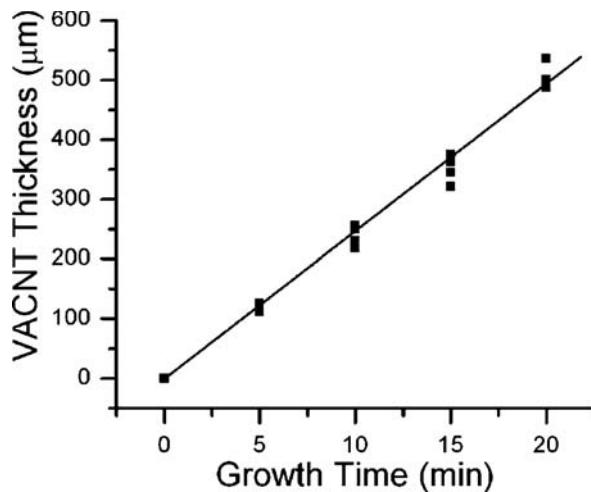
Fig. 12 Surface morphology of a 10-nm gold layer on Si (100) deposited by e-beam evaporation



copper deactivates the CNT growth. Pure copper does not catalyze decomposition of hydrocarbons as the interaction of hydrocarbons with copper surface does not lead to the rupture of the carbon–carbon bonds [229]. In addition, solubility of carbon in solid copper is extremely low [230,231]. Thus, when the copper surface is exposed to the top iron particles during the CVD process, the local copper-rich catalyst composition is capable of decomposing hydrocarbons, however, without giving rise to ordered graphitic structures [229,232]. Second, iron diffuses into copper when the iron particles are in contact with the copper surface at the inter-grain voids. Iron diffusion coefficient in single-crystalline copper is higher than $10^{-12} \text{ cm}^2/\text{s}$ at 1000 K, at least three orders of magnitude higher than that in silicon dioxide ($< 10^{-15} \text{ cm}^2/\text{s}$) [233,234]. Moreover, it is expected that a faster iron diffusion through as well as an enrichment at the grain boundaries in the polycrystalline copper has been used in Lin’s study. In comparison with e-beam evaporation, ALD, also known as atomic layer epitaxy, is a special modification of CVD with the capability of producing films with excellent conformality and precisely controlled thicknesses [235–237]. Furthermore, unlike the thickness monitoring during e-beam evaporation, the thickness of the Al_2O_3 layer by ALD is accurate and insensitive to the surface roughness of copper substrates; as such, it guarantees a reproducible ACNT synthesis.

In comparison with the diffusion-controlled regime ($>740^\circ\text{C}$) of ACNT synthesis on SiO_2/Si surface [238], kinetics-controlled growth on the copper substrate was observed (Fig. 13).

Fig. 13 A plot of ACNT thickness versus growth time, indicating kinetics-controlled ACNT growth on the bulk copper substrate



6 Assembling Technologies of ACNTs

Generally, integration of CNTs into microsystems includes vertical and horizontal positioning. For horizontal interconnects, CNTs must be grown horizontally. However, there has been limited success in achieving controlled horizontal assembly

or growth of CNTs [14,15]. In this section, low-temperature assembling technologies for vertical positioning of ACNTs are discussed.

6.1 Physical Transfer/Anchoring

Transfer of ACNT films has been reported using aqueous treatments to debond the films from the growth substrates and place them onto another substrate surface [239]. Although this transfer separates the growth and assembly of ACNTs, effective contact of ACNTs on the substrate is limited and the interfacial strength is weak. Zhu et al. proposed a solder transfer methodology for ACNTs, similar to flip-chip technology as illustrated schematically in Fig. 14 [223]. The substrates for this process can be copper or FR-4 boards coated with copper. A thin layer of solder paste, e.g., SnPb and SnAgCu, is stencil-printed on the substrate surface. After reflow at a proper temperature, the solidified solder layer is polished to $\sim 30 \mu\text{m}$. The ACNTs on the growth substrate are then flipped to the solder-coated substrate and reflowed to form electrical and mechanical connections. This process is straightforward to implement and overcomes the serious obstacles of integrating ACNTs into electronic packaging by offering low process temperatures and improved adhesion of ACNTs to substrates. This CNT transfer process can be expanded to assemble fine-pitch CNT bundles. Figures 15 and 16 show the transferred CNT film and bundles, respectively, onto copper substrates. Kordas et al. demonstrated a simple and scalable assembly process similar to the above solder transfer technique to fabricate ACNT microfins onto the chip for an air cooling system, as shown in Fig. 17 [161]. For a 1-mm² test chip to reach the same temperature, the applied power can be $\sim 1 \text{ W}$ larger when the ACNT sink is implemented than that of the situation with a bare chip. Testing indicated that the CNT fin structure allowed heat dissipation of ~ 30 and $\sim 100 \text{ W/cm}^2$ more power at 100°C from a hot chip for the cases of

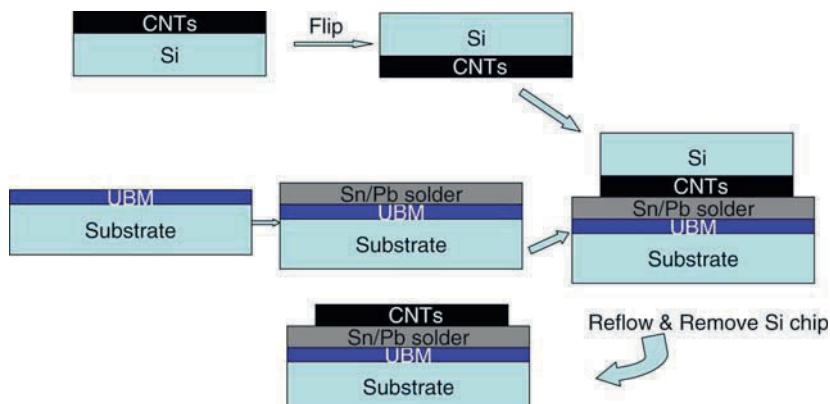


Fig. 14 An illustration of the solder transfer process

Fig. 15 A CNT film anchored on a copper plate by the solder transfer process

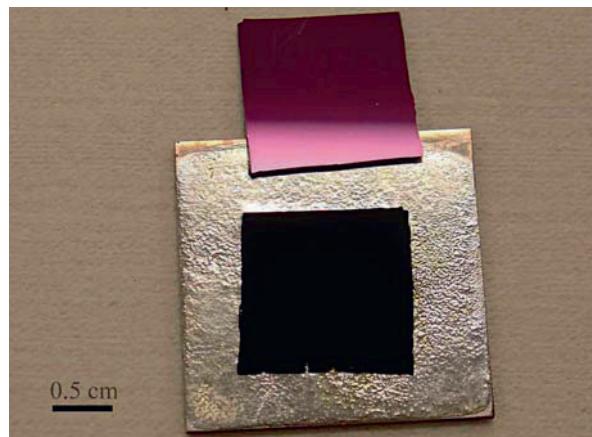


Fig. 16 CNT bundles anchored on a copper plate by the solder transfer process

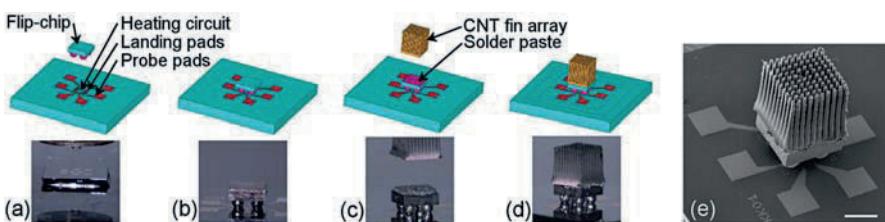
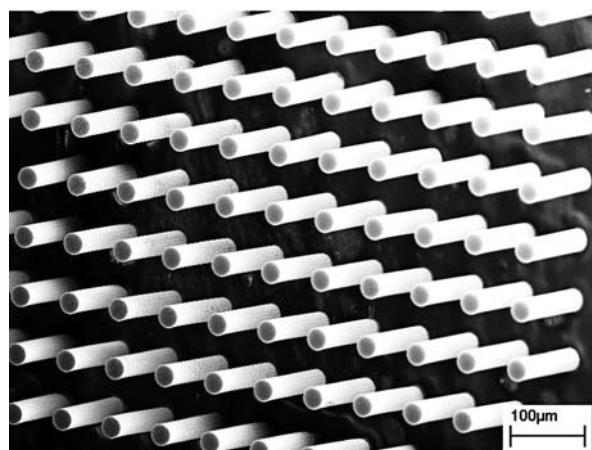


Fig. 17 Schematic illustration of the ACNT microfins fabrication process (a-d) and the image of the fabricated microfins as a heat sink on a flip-chip assembly [161]

natural and forced convections, respectively. The cooling performance of the nanotube fin structures combined with their low weight, mechanical robustness, and ease of fabrication makes them possible candidates for on-chip thermal management applications. Kim et al. used a low-temperature ceramic cement to anchor ACNTs on silicon substrates with a mild curing temperature [240]. By filling the cement with gold particles, conductivity was enhanced, although it was not clearly determined whether the contact between the ACNTs and the gold-filled cement was ohmic. They also demonstrated the transfer of ACNTs onto silicon surface coated with a thin gold layer. The high baking temperature, e.g., 800°C, however, was not favored by electronic packaging. Sundén et al. reported a microwave-assisted transfer of ACNTs onto thermoplastic polymer substrates [156]. Jiang et al. proposed a low-temperature transfer process using conductive polymer composites [241]. An ohmic contact was formed between a CNT film and a highly conductive polymer composite while a semiconductor joint was formed between a CNT film and a high-resistivity polymer composite. Recently, Gan et al. reported a large-scale bonding of ACNTs to metal substrates with the assistance of high-frequency induction heating. This technique provides a potential approach to reproducible large-scale fabrication of ACNTs for various applications [242].

6.2 Chemical Transfer

Recently, a chemical transfer technology was developed by Lin et al. (Fig. 18) [243]. They in situ functionalized ACNT surface and the opened tips with certain functional groups during the CVD process by introducing a trace amount of H₂O₂ as the oxidant into the chamber. The synthesis temperature could be lowered by 50°C (700°C). X-ray photoelectron spectroscopy (XPS) scans indicate the existence of oxygen-containing functional groups on ACNTs (Fig. 19). Fig. 20 (trace a) shows the FTIR spectrum of the in situ functionalized ACNTs. The peak at ~1572 cm⁻¹ is attributed to the C=C asymmetric stretching in graphite-like CNT structure. Consistent with the XPS results, oxygen-containing functional groups were found. Peaks at 3445, 1229, 1179, and 1060 cm⁻¹ are attributed to O—H, C—O, C—O—C,

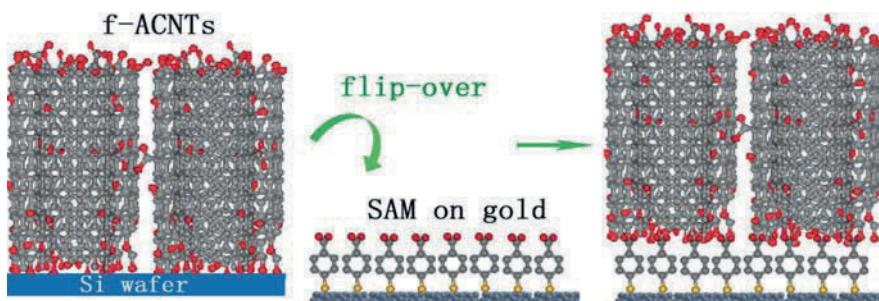
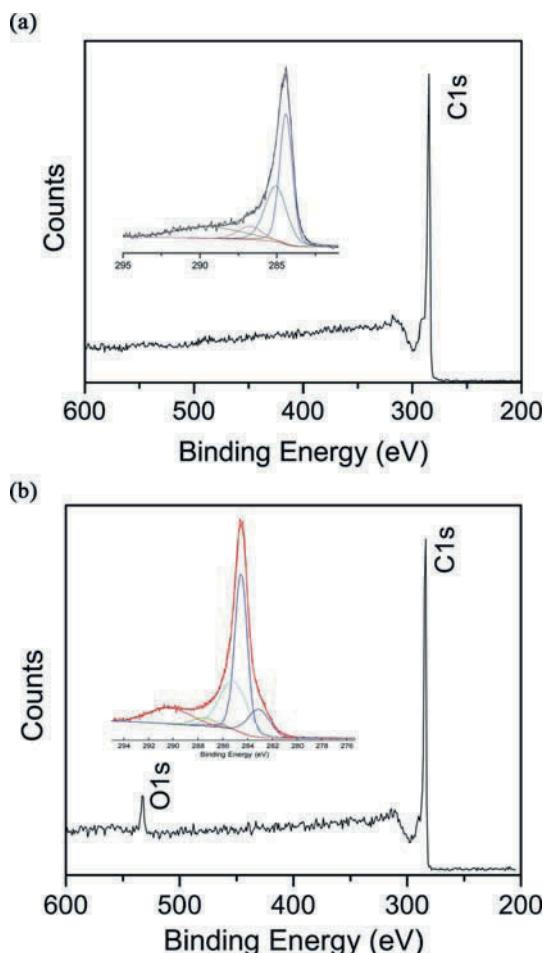


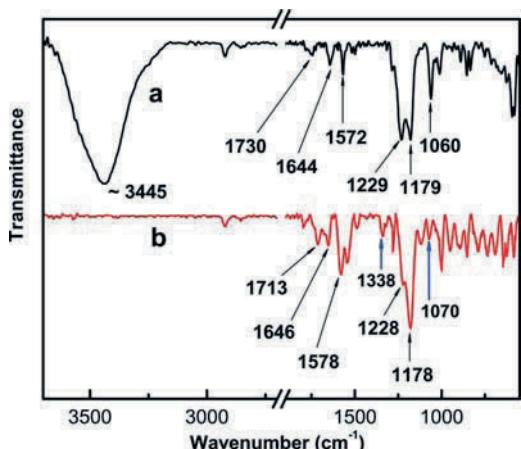
Fig. 18 Schematic illustration of the chemical transfer process

Fig. 19 XPS spectra of pristine ACNT (**a**) and f-ACNT (**b**). Inserted are corresponding high-resolution C 1s spectra



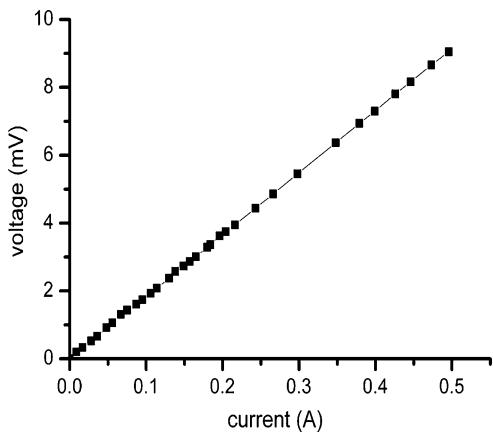
and C—OH stretchings, respectively. These oxygen-containing functional groups, particularly the hydroxyl groups and the carboxylic acid groups, make ACNTs reactive with other functional groups. To prove the reactivity, they used thionyl chloride to react with the *in situ* functionalized ACNTs, the FTIR spectrum of the resultant product s-CNT being shown in Fig. 20 (trace b). The *in situ* functionalized ACNTs were directly transferred to a gold-coated Si surface (or quartz surface) coated with a self-assembled monolayer of conjugated molecules as the bridging agent for mechanical anchoring and electron transport. This transfer technology has numerous benefits in that it covalently bonds the CNTs to the substrate, minimizes the damage to the CNTs during functionalization, can be tailored to attach CNTs to a wide variety of substrates, is carried out at low temperatures, and can be easily

Fig. 20 FTIR spectra of the f-ACNTs (**a**) and the s-CNTs (**b**)



reworked. More importantly, the chemical bonds formed at the CNT/substrate interfaces reduce the interfacial contact resistance. Ohmic contact at the ACNT/substrate interface with low resistance is shown in Fig. 21. Preliminary data show that the transfer process dramatically reduces the thermal interfacial resistance as well.

Fig. 21 The room-temperature I - V curve of the f-ACNT/gold interconnect via the chemical transfer



In our previous discussion, we summarized representative work on ACNT TIMs that have been reported so far. There are at least two main issues in common that inhibited real-life applications of the ACNT “TIMs”: (1) the thermal contact resistance at the CNT/mating substrate interface was large and dictated the overall thermal resistance of the ACNT TIM [244], even when a relatively high pressure was imposed; (2) the interfacial adhesion was extremely weak. One reason for the large contact resistance is mainly that, according to the very recent

report by Panzer, the fraction of CNT–substrate contacts is low [159]. More importantly, in the molecular scale or the atomic scale, the CNTs are not in “contact” with the mating substrate surfaces because there is no chemical bonding or special association, which is also the reason for the weak interfacial adhesion. Thus, a challenge arises as: how can people modify the CNT/mating substrate interface to effectively reduce the thermal contact resistance and simultaneously improve the interfacial adhesion? When heat transfers at the interface between a nanotube and a planar substrate, constriction resistance is developed due to the nanometer-scale contact area and enhanced phonon-boundary scattering at the nanocontacts [245]. Diao et al. [246] used MD simulations to study the ACNT/Si contact resistance. Their results showed that open-ended ACNTs at the interface reduced contact resistance; contact area and the number of bonds formed between the CNT and the Si substrate are the key to the interfacial thermal conductance. Therefore, Lin et al. extended their chemical transfer (or “chemical anchoring”) technology to ACNT TIM assembling [247]. They grew ACNTs on a copper substrate and introduced a thermally stable cross-linked molecular phonon coupler (MPC) to chemically

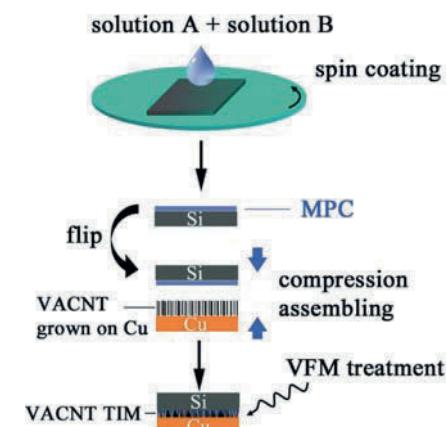


Fig. 22 Schematic illustration of the Si/ACNT TIM/Cu assembling process

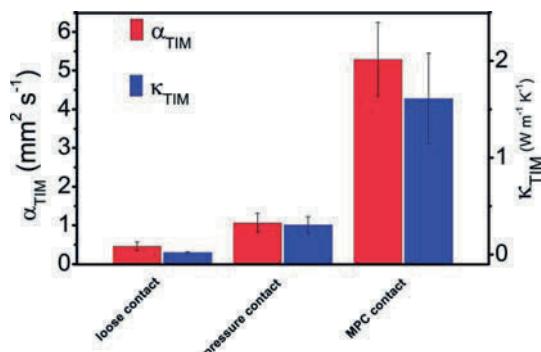


Fig. 23 A comparison of the thermal diffusivities (α_{TIM}) and the equivalent thermal conductivities (κ_{TIM}) among the loose contact, the pressure contact, and the MPC contact ACNT TIM assemblies

bond the ACNTs to the mating Si substrate. The ACNT TIM assembling process is schematically shown in Fig. 22. The MPC at the mating interface facilitates phonon transport across the ACNT/Si interface. The developed ACNT TIM assembling process not only improves the equivalent thermal conductivity of the ACNT TIM by two orders of magnitude (Fig. 23) but also dramatically enhances the interfacial adhesion.

7 Graphene Nanoelectronics

Semiconductor technology has pushed the size of transistors and electrical interconnects toward a few nanometers, at which scale, quantum effects become important. Take the physical limits on CMOS scaling for example. Gate tunneling current, subthreshold channel leakage current, device parameter variability/tolerances, interconnect latency, static and dynamic power dissipation, and so on, are great challenges. Although some innovations have, to a certain extent, addressed some of these challenges such as strain in channel, 3–5 channel FETs, high-permittivity gate dielectrics (e.g., HfSiON), metal gates (e.g., TaSiN), nano lithography techniques, a revolutionary resolution is considered necessary. Graphene, due to its intrinsic peculiar electronic properties, quantum effects, and extremely high in-plane thermal conductivity, has shown potential advantages to bring such a revolution to conventional Si semiconductors. Graphene, a single atomic layer of coplanarized hexagonally bonded carbon atoms, has recently generated enormous interests in academy and industry [248–262]. It has been proposed that graphene nanoelectronics may provide a solution to the critical challenges faced by current and future semiconductor technology due to its unusual properties that, so far as we know, include (1) an intrinsically zero band gap but non-zero (tunable) band gap due to influences of external factors such as the growth substrate and the shape effect [263–268]; (2) a large mean free path of carrier transport comparable to that of SWNTs, allowing potential applications as ballistic transistors [269–278]; (3) an extraordinarily high carrier mobility on the order of $10,000 \text{ cm}^2/\text{V}\cdot\text{s}$ even at room temperature [277,279–281]; (4) constant velocity of holes and electrons that does not depend on their kinetic energy [251]; (5) quantum Hall effect [281–283]; (6) extremely high in-plane thermal conductivity [284,285]; (7) 3D scalability compared with CNTs.

For potential application of graphene interconnects, research work reported is still quite limited. Recently, fabricating electronic circuits consisting of both transistors and interconnects in the same continuous graphene sheet has been proposed. Naeemi and Meindl used compact physical models to study the electrical conductance of graphene nanoribbons as functions of chirality, ribbon width, Fermi level, and the type of electron scatterings at the edges [286]. It was proposed that for widths below 8 nm, single-layer graphene nanoribbons can potentially outperform copper wires with unity aspect ratio. Shao et al. reported the high-temperature electrical resistance of graphene [287]. It was found that as temperature increases,

the electrical resistance of graphene interconnects drops. They explained this phenomenon by the thermal generation of electron–hole pairs at elevated temperatures and carrier scattering by acoustic phonons.

8 Graphite Nanosheet/Polymer Composites – Very Promising TIM Composite Materials

As usual, like any advanced material discovered and prepared, one of its potential applications will be to incorporate it into polymer matrices to make high-performance composites [288–290]. Exfoliated graphite, a stack of tens to hundreds of graphene layers, has been used as an effective filler in polymer matrices for thermal conductivity improvement [291–294]. Exfoliated graphite with a few layers of graphene and a high aspect ratio, also named “graphite nanosheet (GNS),” is of special interest [117,258]. Recently, Sun’s group reported their successful preparation of GNS with a high aspect ratio and incorporation of the GNS into an epoxy (EP) matrix [295]. The as-prepared GNS/EP composite displayed an extremely high in-plane thermal conductivity (~ 80 W/m·K at 33 vol% GNS loading). Compared with carbon nanotube/epoxy nanocomposites, there is no doubt that the GNSs are considered more effective fillers for improving thermal conductivity. This great enhancement, hypothetically, resulted from the reduced GNS/polymer interfacial thermal resistance. However, this hypothesis is contrary to the results by Hung et al. where a strong influence of interfacial resistance was estimated [294]. Lin et al. used a conventional effective medium model to analyze Sun’s results and showed good validity of the model in predicting thermal conductivities (both in-plane and through-thickness) of the GNS/polymer nanocomposites [296]. A high in-plane thermal conductivity, i.e., 1500 W/m·K, of a GNS was chosen in their modeling, on the basis of the following considerations. The in-plane thermal conductivity of a single graphene sheet was estimated to be as high as ~ 5000 W/m·K [284,285]. Given the fact that a multi-walled carbon nanotube (MWNT), which is treated as a concentrically wrapped GNS [29], possesses a similar longitudinal thermal conductivity (>3000 W/m·K) [81] to a single-walled carbon nanotube (~ 3500 W/m·K) [76], it was believed that a GNS possesses a high K_{f11} in the same order of magnitude with K_{f11} of a single graphene sheet. Previous experimental data on carbon nanotubes (CNTs) also showed that the longitudinal thermal conductivities of MWNT arrays were ~ 80 W/m·K [116,157,161]. Since the packing densities of these MWNT arrays are $\sim 5\%$ and tube–tube interactions degrade the effective thermal conductivities [297], they estimated the equivalent longitudinal thermal conductivity of an individual MWNT to be at least 1600 W/m·K. Therefore, the GNS, an unrolled form of an MWNT, should possess a K_{f11} of approximately 1600 W/m·K. Further proof came from the in-plane thermal conductivity of the commercial pyrolytic graphite sheets (PGS) provided by Panasonic, Inc. As the PGS density approaches the theoretical mass density of graphite, the in-plane thermal conductivity of the PGS becomes 1500–1700 W/m·K. To be conservative, 1500 W/m·K was taken in their

modeling. The modeling results indicated the influence of GNS/polymer interfacial thermal resistance on the overall thermal transport, in analogy to CNT/polymer composites. Based on their modeling results, it was proposed that further exfoliation of GNS to achieve a large aspect ratio and GNS/polymer interfacial modification would lead to further enhancement of the thermal transport of GNS/polymer nanocomposites.

9 Thermal Property Measurement Techniques

Various thermal property measurement techniques have been applied to evaluate ACNT TIMs, including laser flash technique, steady-state TIM testers based on ASTM 5470, thermoreflectance technique, photoacoustic technique, photothermal technique, and 3ω technique. For conventional TIMs, that is, TIMs with relatively high thermal resistance, the ASTM-based steady-state measurement technique is the most popular and well-accepted technique. It has almost become a standard in industry. However, generally, the ASTM-based technique has very limited capability of measuring CNT TIMs with small thickness (e.g., a few microns) and low thermal resistance (e.g., $< 10 \text{ mm}^2\text{-K/W}$). Moreover, CNT/substrate contact resistance is needed for scientific understanding of interfacial phonon transport, which cannot be measured directly by the ASTM-based technique. Therefore, the other measurement techniques have been employed to study ACNT TIMs; we will go through them briefly in this section.

9.1 Laser Flash Technique (LFA 447, Netzsch, Inc.)

Laser flash technique (LFA) is a sensitive, non-contact light flash method and has been widely used for measuring thermal diffusivity and specific heat of dense materials [298–302]. In an LFA setup, a high-performance xenon flash lamp, a sample, and a liquid-nitrogen-cooled Sb infrared detector are vertically arranged. During testing, the entire bottom surface of the sample is illuminated homogeneously by a pulsed laser produced by the flash lamp. Both the released energy of flash lamp and the length of the heating pulse can be adjusted via the 32-bit MS windows software named Nanoflash. An integrated furnace and a thermocouple positioned in the sample carrier are used for accurate control of sample temperatures. The temperature on the top surface of the sample is monitored by the infrared detector versus time after the light pulse. The thermal diffusivity of the sample is calculated by fitting the as-obtained data (temperature–time) with various models integrated in the analysis software of the LFA device. The thermal diffusivity measuring range is 0.01–1000 mm^2/s for LFA 447, with a reproducibility of approximately $\pm 3\%$. In addition to the thermal diffusivity, by employing a comparative method, the specific heat of the sample can also be determined, with a reproducibility of approximately $\pm 5\%$.

Typically, a thin layer of gold (100 nm) is sputtered onto the top and bottom surfaces. The gold layer on the bottom surface is to ensure a fast heat homogenization within the bottom surface so as to avoid measurement errors introduced by the heat spot issue. The gold layer on the top surface is to make sure that the infrared signal collected by the detector comes only (or mostly) from the very top surface. For thermal measurement of CNT/polymer composites, which are dense materials, a well-polished sample with a proper size, surface flatness, and roughness is required. For an ACNT TIM sample, since it is porous rather than dense, LFA could not be used to measure the intrinsic thermal properties of the ACNT array. Instead, a substrate/ACNT TIM/substrate trilayer assembly is used. The measurement result is, therefore, the equivalent thermal diffusivity of the ACNT TIM layer that consists of the ACNT array and the two contact interfaces.

9.2 Steady-State Measurement

one-dimensional steady-state measurement is based on ASTM 5470, in which, in simple terms, the TIM is filled between two blocks with coplanar surfaces [116,119,303–306]. One block is heated and the other cooled. Therefore, a constant heat flux is supplied through the copper (or aluminum) blocks across the TIM sample. The overall interface thermal conductance (θ_{TIM}) is determined by extrapolating the temperature jump across the sample. ASTM 5470 based TIM testers are already commercialized, e.g., TIM tester 1300 from AnalysisTech Inc. Two advantages of steady-state measurements over other techniques are as follows: (1) it is convenient to control the pressure (compression) imposed during the testing; (2) it is suitable for TIM reliability evaluation. However, in the steady-state measurement method, it is difficult to determine the thermal resistance between the ACNTs and the substrates. Similar to the laser flash technique, only a combined resistance is measured. The contact resistance may be obtained by extrapolating the resistance–TIM thickness curve toward zero thickness; however, the precision is low.

9.3 Thermoreflectance Technique

Phase-sensitive transient thermoreflectance technique was first developed by Ohsono et al. to determine the thermal conductance of the interface between a silicon substrate and the thermally grown silicon dioxide on the silicon substrate [307]. Thermoreflectance technique has been used in the nanosecond regime to characterize the thermal properties of thin films [308,309], in the picosecond regime to measure both the thermal properties of thin films and their interfaces [310–312], and in the femtosecond regime to measure electron and phonon processes [313,314]. Experimentally, a sample is heated by a high-power pulsed heating laser with intensity sinusoidally modulated at an angular frequency [125,159]. A second low-power

probe laser is focused on one surface of the sample. The heat flux oscillation caused by the heating laser propagates through the sample resulting in a temperature oscillation in the sample and at the interfaces. The reflected laser beam of the probe laser is captured by a photodetector and the intensity signal is sent to an amplifier to extract the signal oscillation frequency. On the basis that the intensity and the phase of the reflected beam are dependent on the temperature of the reflecting surface, the phase shift and/or the intensity provides a dynamic measurement of the temperature response of the sample. A multiparameter search algorithm based on a least square fit to the experimental data is used to obtain the thermal resistances/conductances of the interfaces and the thermal diffusivity of the TIM itself. The key advantage of this transient technique is its capability of distinguishing the interface thermal conductances from various interfaces and TIM itself.

9.4 Photoacoustic Technique

An alternative technique to measure the thermal resistances of thin films and the various interfaces discretely in a multilayered structure is photoacoustic (PA) technique. Experimentally, a sample is mounted at the bottom of a PA cell, which is typically made of sapphire [157,160,168,315,316]. Sapphire has low reflectance and high transmittance for the infrared laser used; therefore, most of the reflected laser energy from the sample surface transmits out of the cell. A modulated heating laser is directed onto the sample surface. An oscillatory change of the temperature in the gas cell is induced due to the heat conduction through the contact with the sample surface and, as well, the mechanical work imposed on the gas medium due to the vibration of the sample surface. The oscillatory temperature change causes an oscillatory pressure change of the gas, which can be sensed as an acoustic wave. A condenser microphone mounted near the inside wall of the cell senses the acoustic signal and transfers it to a lock-in amplifier, where the amplitude and phase of the acoustic signal are measured. Because the transient temperature in the gas is related to the thermal properties of the sample, measuring the pressure allows determination of the thermal quantities such as the interfacial thermal resistance and the thermal diffusivity. Though it is thought that PA technique provides high accuracy and a convenient measurement procedure, PA devices are still not commercially available.

Other techniques such as photothermal technique and 3ω technique are not introduced here; readers are recommended to refer to related literature for details.

10 Future Needs

From the discussions above, we expect promising applications of carbon nanotubes as vertical electrical interconnects and TIMs for microelectronic packaging. However, a number of issues need to be addressed:

- (1) *Structural defect control* of ACNTs during CVD syntheses or post-growth treatments.
- (2) Effective increase in CNT *packing density*. Recent progress has been reported in increased ACNT packing density, i.e., from ~ 1 to $\sim 20\%$, and consequently increased current-carrying capability [317]. However, a densification process with less complexity is in need for real-life applications. The influence of densification on thermal dissipation needs to be investigated.
- (3) For ACNT TIMs, ACNT/substrate *interface modification* is significant for improving phonon transport by reducing phonon scattering. The above-mentioned MPC-assisted chemical modifications need to be optimized and are underway.
- (4) Although graphene nanoelectronics is approaching conceptually, a number of technological issues will emerge and need to be resolved effectively, such as 3D interconnection, controlled doping, band gap tuning, and gate fabrications.
- (5) Undoubtedly, a great deal of attention will be paid to GNS/polymer composites or graphene/polymer composites for applications as TIMs.

References

1. Moore GE, Progress in Digital Integrated Electronics. International Electron Devices Meetings 1975: 11–13.
2. Steinlesberger G, Engelhardt M, Schindler G, Steinhogl W, von Glasow A, Mosig K, Bertagnolli E, Electrical assessment of copper damascene interconnects down to sub-50 nm feature sizes. *Microelectron Eng* 2002;64: 409–416.
3. Kreupl F, Graham AP, Duesberg GS, Steinhogl W, Liebau M, Unger E, Honlein W, Carbon nanotubes in interconnect applications. *Microelectron Eng* 2002;64: 399–408.
4. Li J, Ye Q, Cassell A, Ng HT, Stevens R, Han J, Meyyappan M, Bottom-up approach for carbon nanotube interconnects. *Appl Phys Lett* 2003;82: 2491–2493.
5. Nihei M, Horibe M, Kawabata A, Awano Y, Simultaneous formation of multiwall carbon nanotubes and their end-bonded ohmic contacts to Ti electrodes for future ULSI interconnects. *Jpn J Appl Phys Part 1 Regul Pap Short Notes Rev Pap* 2004;43: 1856–1859.
6. Naeemi A, Sarvari R, Meindl JD, Performance comparison between carbon nanotube and copper interconnects for gigascale integration (GSI). *IEEE Electron Device Lett* 2005;26: 84–86.
7. Naeemi A, Meindl JD, Impact of electron-phonon scattering on the performance of carbon nanotube interconnects for GSI. *IEEE Electron Device Lett* 2005;26: 476–478.
8. Awano Y, Carbon nanotube technologies for LSI via interconnects. *IEICE Trans Electron* 2006;E89C: 1499–1503.
9. Park M, Cola BA, Siegmund T, Xu J, Maschmann MR, Fisher TS, Kim H, Effects of a carbon nanotube layer on electrical contact resistance between copper substrates. *Nanotechnology* 2006;17: 2294–2303.
10. Nieuwoudt A, Massoud Y, Understanding the impact of inductance in carbon nanotube bundles for VLSI interconnect using scalable modeling techniques. *IEEE Trans Nanotechnol* 2006;5: 758–765.
11. Xu T, Wang Z, Miao J, Chen X, Tan CM, Aligned carbon nanotubes for through-wafer interconnects. *Appl Phys Lett* 2007;91: 042108.

12. Kaushik BK, Goel S, Rauthan G, Future VLSI interconnects: optical fiber or carbon nanotube – a review. *Microelectron Int* 2007;24: 53–63.
13. Graham AP, Duesberg GS, Seidel R, Liebau M, Unger E, Kreupl F, Honlein W, Towards the integration of carbon nanotubes in microelectronics. *Diam Relat Mater* 2004;13: 1296–1300.
14. Huang SM, Cai XY, Liu J, Growth of millimeter-long and horizontally aligned single-walled carbon nanotubes on flat substrates. *J Am Chem Soc* 2003;125: 5636–5637.
15. Zhang YG, Chang AL, Cao J, Wang Q, Kim W, Li YM, Morris N, Yenilmez E, Kong J, Dai HJ, Electric-field-directed growth of aligned single-walled carbon nanotubes. *Appl Phys Lett* 2001;79: 3155–3157.
16. Chai Y, Xiao ZY, Chan PCH, Electron-shading effect on the horizontal aligned growth of carbon nanotubes. *Appl Phys Lett* 2009;94: 043116.
17. Lan C, Zakharov DN, Reifenberger RG, Determining the optimal contact length for a metal/multiwalled carbon nanotube interconnect. *Appl Phys Lett* 2008;92: 213112.
18. Nihei M, Horibe M, Kawabata A, Awano Y, Simultaneous formation of multiwall carbon nanotubes and their end-bonded ohmic contacts to Ti electrodes for future ULSI interconnects. *Jpn J Appl Phys* 2004;43: 1856–1859.
19. Schmidt R, Challenges in electronic cooling – opportunities for enhanced thermal management techniques – Microprocessor liquid cooled minichannel heat sink. *Heat Transf Eng* 2004;25: 3–12.
20. Mahajan R, Chiu CP, Chrysler G, Cooling a microprocessor chip. *Proc IEEE* 2006;94: 1476–1486.
21. Prasher R, Thermal interface materials: historical perspective, status, and future directions. *Proc IEEE* 2006;94: 1571–1586.
22. Chu RC, The challenges of electronic cooling: past, current and future. *J Electron Packag* 2004;126: 491–500.
23. Azar K, Power Consumption and Generation in the Electronics Industry – a Perspective, 20th IEEE SEMI-Therm Symposium. San Jose, CA, 2004:201–212.
24. Prasher RS, Nano and micro technology based next generation package-level cooling solutions. *Intel Technol J* 2005;9: 285–292.
25. Schelling PK, Shi L, Goodson KE, Managing heat for electronics. *Mater Today* 2005;8: 30–35.
26. Iijima S, Helical microtubules of graphitic carbon. *Nature* 1991;354: 56–58.
27. Ajayan PM, Iijima S, Smallest carbon nanotube. *Nature* 1992;358: 23.
28. Ebbesen TW, Ajayan PM, Large-scale synthesis of carbon nanotubes. *Nature* 1992;358: 220–222.
29. Ajayan PM, Ebbesen TW, Nanometre-size tubes of carbon. *Rep Prog Phys* 1997;60: 1025–1062.
30. Ajayan PM, Nanotubes from carbon. *Chem Rev* 1999;99: 1787–1799.
31. Hamada N, Sawada S, Oshiyama A, New one-dimensional conductors – graphitic micro-tubules. *Phys Rev Lett* 1992;68: 1579–1581.
32. Dresselhaus MS, Dresselhaus G, Saito R, Carbon-fibers based on C-60 and their symmetry. *Phys Rev B* 1992;45: 6234–6242.
33. Dresselhaus MS, Dresselhaus G, Avouris P, Carbon Nanotubes – Synthesis, Structure, Properties and Applications. Springer-Verlag, New York, 2001.
34. Mackay AL, Terrones H, Diamond from graphite. *Nature* 1991;352: 762.
35. Zhang XF, Zhang XB, Vantendeloo G, Amelinckx S, Debeek MO, Vanlanduyt J, Carbon nano-tubes – their formation process and observation by electron-microscopy. *J Cryst Growth* 1993;130: 368–382.
36. Tsang SC, Chen YK, Harris PJF, Green MLH, A simple chemical method of opening and filling carbon nanotubes. *Nature* 1994;372: 159–162.
37. Mintmire JW, Dunlap BI, White CT, Are fullerene tubules metallic. *Phys Rev Lett* 1992;68: 631–634.

38. Odom TW, Huang JL, Kim P, Lieber CM, Atomic structure and electronic properties of single-walled carbon nanotubes. *Nature* 1998;391: 62–64.
39. Wildoer JWG, Venema LC, Rinzler AG, Smalley RE, Dekker C, Electronic structure of atomically resolved carbon nanotubes. *Nature* 1998;391: 59–62.
40. Peierl RF, Quantum Theory of Solids. Clarendon, Oxford, 1995.
41. Yao Z, Kane CL, Dekker C, High-field electrical transport in single-wall carbon nanotubes. *Phys Rev Lett* 2000;84: 2941–2944.
42. Wei BQ, Vajtai R, Ajayan PM, Reliability and current carrying capacity of carbon nanotubes. *Appl Phys Lett* 2001;79: 1172–1174.
43. Collins PC, Arnold MS, Avouris P, Engineering carbon nanotubes and nanotube circuits using electrical breakdown. *Science* 2001;292: 706–709.
44. Bachtold A, Strunk C, Salvetat JP, Bonard JM, Forro L, Nussbaumer T, Schonenberger C, Aharonov-Bohm oscillations in carbon nanotubes. *Nature* 1999;397: 673–675.
45. McEuen PL, Bockrath M, Cobden DH, Yoon YG, Louie SG, Disorder, pseudospins, and backscattering in carbon nanotubes. *Phys Rev Lett* 1999;83: 5098–5101.
46. Li HJ, Lu WG, Li JJ, Bai XD, Gu CZ, Multichannel ballistic transport in multiwall carbon nanotubes. *Phys Rev Lett* 2005; 95.
47. Haruehanroengra S, Wang W, Analyzing conductance of mixed carbon-nanotube bundles for interconnect applications. *IEEE Electron Device Lett* 2007;28: 756–759.
48. Naeemi A, Meindl JD, Compact physical models for multiwall carbon-nanotube interconnects. *IEEE Electron Device Lett* 2006;27: 338–340.
49. Tans SJ, Verschueren ARM, Dekker C, Room-temperature transistor based on a single carbon nanotube. *Nature* 1998;393: 49–52.
50. Bachtold A, Hadley P, Nakanishi T, Dekker C, Logic circuits with carbon nanotube transistors. *Science* 2001;294: 1317–1320.
51. Heinze S, Tersoff J, Martel R, Derycke V, Appenzeller J, Avouris P, Carbon nanotubes as Schottky barrier transistors. *Phys Rev Lett* 2002;89: 4.
52. Soh HT, Quate CF, Morpurgo AF, Marcus CM, Kong J, Dai HJ, Integrated nanotube circuits: Controlled growth and ohmic contacting of single-walled carbon nanotubes. *Appl Phys Lett* 1999;75: 627–629.
53. Strano MS, Dyke CA, Usrey ML, Barone PW, Allen MJ, Shan HW, Kittrell C, Hauge RH, Tour JM, Smalley RE, Electronic structure control of single-walled carbon nanotube functionalization. *Science* 2003;301: 1519–1522.
54. Campidelli S, Meneghetti M, Prato M, Separation of metallic and semiconducting single-walled carbon nanotubes via covalent functionalization. *Small* 2007;3: 1672–1676.
55. Krupke R, Hennrich F, von Lohneysen H, Kappes MM, Separation of metallic from semiconducting single-walled carbon nanotubes. *Science* 2003;301: 344–347.
56. Li HP, Zhou B, Lin Y, Gu LR, Wang W, Fernando KAS, Kumar S, Allard LF, Sun YP, Selective interactions of porphyrins with semiconducting single-walled carbon nanotubes. *J Am Chem Soc* 2004;126: 1014–1015.
57. Maeda Y, Kimura S, Kanda M, Hirashima Y, Hasegawa T, Wakahara T, Lian YF, Nakahodo T, Tsuchiya T, Akasaka T, Lu J, Zhang XW, Gao ZX, Yu YP, Nagase S, Kazaoui S, Minami N, Shimizu T, Tokumoto H, Saito R, Large-scale separation of metallic and semiconducting single-walled carbon nanotubes. *J Am Chem Soc* 2005;127: 10287–10290.
58. Toyoda S, Yamaguchi Y, Hiwatashi M, Tomonari Y, Murakami H, Nakashima N, Separation of semiconducting single-walled carbon nanotubes by using a long-alkyl-chain benzenediazonium compound. *Chem-Asian J* 2007;2: 145–149.
59. Wang W, Fernando KAS, Lin Y, Meziani MJ, Veca LM, Cao L, Zhang P, Kimani MM, Sun YP, Metallic single-walled carbon nanotubes for conductive nanocomposites. *J Am Chem Soc* 2008;130: 1415–1419.
60. Mattsson M, Gromov A, Dittmer S, Eriksson E, Nerushev OA, Campbell EEB, Dielectrophoresis-induced separation of metallic and semiconducting single-wall carbon

- nanotubes in a continuous flow microfluidic system. *J Nanosci Nanotechnol* 2007;7(10): 3431–3435.
61. Tanaka T, Jin HH, Miyata Y, Kataura H, High-yield separation of metallic and semiconducting single-wall carbon nanotubes by agarose gel electrophoresis. *Appl Phys Express* 2008; 1.
 62. Ghosh S, Rao CNR, Separation of metallic and semiconducting single-walled carbon nanotubes through fluorous chemistry. *Nano Res* 2009;2: 183–191.
 63. Green AA, Duch MC, Hersam MC, Isolation of single-walled carbon nanotube enantiomers by density differentiation. *Nano Res* 2009;2: 69–77.
 64. Dyke CA, Stewart MP, Tour JM, Separation of single-walled carbon nanotubes on silica gel. Materials morphology and Raman excitation wavelength affect data interpretation. *J Am Chem Soc* 2005;127: 4497–4509.
 65. LeMieux MC, Roberts M, Barman S, Jin YW, Kim JM, Bao ZN, Self-sorted, aligned nanotube networks for thin-film transistors. *Science* 2008;321: 101–104.
 66. Menard-Moyon C, Izard N, Doris E, Mioskowski C, Separation of semiconducting from metallic carbon nanotubes by selective functionalization with azomethine ylides. *J Am Chem Soc* 2006;128: 6552–6553.
 67. Huang HJ, Maruyama R, Noda K, Kajiura H, Kadono K, Preferential destruction of metallic single-walled carbon nanotubes by laser irradiation. *J Phys Chem B* 2006;110: 7316–7320.
 68. Miyata Y, Maniwa Y, Kataura H, Selective oxidation of semiconducting single-wall carbon nanotubes by hydrogen peroxide. *J Phys Chem B* 2006;110: 25–29.
 69. Arnold MS, Green AA, Hulvat JF, Stupp SI, Hersam MC, Sorting carbon nanotubes by electronic structure using density differentiation. *Nat Nanotechnol* 2006;1: 60–65.
 70. Tanaka T, Jin HH, Miyata Y, Fujii S, Suga H, Naitoh Y, Minari T, Miyadera T, Tsukagoshi K, Kataura H, Simple and scalable gel-based separation of metallic and semiconducting carbon nanotubes. *Nano Lett* 2009;9(4): 1497–1500.
 71. Berber S, Kwon YK, Tomanek D, Unusually high thermal conductivity of carbon nanotubes. *Phys Rev Lett* 2000;84: 4613–4616.
 72. Yu CH, Shi L, Yao Z, Li DY, Majumdar A, Thermal conductance and thermopower of an individual single-wall carbon nanotube. *Nano Lett* 2005;5: 1842–1846.
 73. Mingo N, Broido DA, Carbon nanotube ballistic thermal conductance and its limits. *Phys Rev Lett* 2005; 95.
 74. Yamamoto T, Watanabe K, Nonequilibrium Green's function approach to phonon transport in defective carbon nanotubes. *Phys Rev Lett* 2006; 96.
 75. Hone J, Llaguno MC, Biercuk MJ, Johnson AT, Batlogg B, Benes Z, Fischer JE, Thermal properties of carbon nanotubes and nanotube-based materials. *Appl Phys A Mater Sci Process* 2002;74: 339–343.
 76. Pop E, Mann D, Wang Q, Goodson K, Dai HJ, Thermal conductance of an individual single-wall carbon nanotube above room temperature. *Nano Lett* 2006;6: 96–100.
 77. Hone J, Whitney M, Piskoti C, Zettl A, Thermal conductivity of single-walled carbon nanotubes. *Phys Rev B* 1999;59: R2514–R2516.
 78. Benedict LX, Louie SG, Cohen ML, Heat capacity of carbon nanotubes. *Solid State Commun* 1996;100: 177–180.
 79. Wang ZL, Tang DW, Li XB, Zheng XH, Zhang WG, Zheng LX, Zhu YTT, Jin AZ, Yang HF, Gu CZ, Length-dependent thermal conductivity of an individual single-wall carbon nanotube. *Appl Phys Lett* 2007; 91.
 80. Mizel A, Benedict LX, Cohen ML, Louie SG, Zettl A, Budraa NK, Beyermann WP, Analysis of the low-temperature specific heat of multiwalled carbon nanotubes and carbon nanotube ropes. *Phys Rev B* 1999;60: 3264–3270.
 81. Kim P, Shi L, Majumdar A, McEuen PL, Thermal transport measurements of individual multiwalled nanotubes. *Phys Rev Lett* 2001; 8721.
 82. Choi TY, Poulikakos D, Tharian J, Sennhauser U, Measurement of thermal conductivity of individual multiwalled carbon nanotubes by the 3-omega method. *Appl Phys Lett* 2005; 87.

83. Choi TY, Poulikakos D, Tharian J, Sennhauser U, Measurement of the thermal conductivity of individual carbon nanotubes by the four-point three-omega method. *Nano Lett* 2006;6: 1589–1593.
84. Hone J, Batlogg B, Benes Z, Johnson AT, Fischer JE, Quantized phonon spectrum of single-wall carbon nanotubes. *Science* 2000;289: 1730–1733.
85. Yang DJ, Zhang Q, Chen G, Yoon SF, Ahn J, Wang SG, Zhou Q, Wang Q, Li JQ, Thermal conductivity of multiwalled carbon nanotubes. *Phys Rev B* 2002; 66.
86. Yi W, Lu L, Zhang DL, Pan ZW, Xie SS, Linear specific heat of carbon nanotubes. *Phys Rev B* 1999;59: R9015–R9018.
87. Bethune DS, Kiang CH, Devries MS, Gorman G, Savoy R, Vazquez J, Beyers R, Cobalt-catalyzed growth of carbon nanotubes with single-atomic-layerwalls. *Nature* 1993;363: 605–607.
88. Thess A, Lee R, Nikolaev P, Dai HJ, Petit P, Robert J, Xu CH, Lee YH, Kim SG, Rinzler AG, Colbert DT, Scuseria GE, Tomanek D, Fischer JE, Smalley RE, Crystalline ropes of metallic carbon nanotubes. *Science* 1996;273: 483–487.
89. Liu J, Rinzler AG, Dai HJ, Hafner JH, Bradley RK, Boul PJ, Lu A, Iverson T, Shelimov K, Huffman CB, Rodriguez-Macias F, Shon YS, Lee TR, Colbert DT, Smalley RE, Fullerene pipes. *Science* 1998;280: 1253–1256.
90. Amelinckx S, Zhang XB, Bernaerts D, Zhang XF, Ivanov V, Nagy JB, A formation mechanism for catalytically grown helix-shaped graphite nanotubes. *Science* 1994;265: 635–639.
91. Amelinckx S, Bernaerts D, Zhang XB, Vantendeloo G, Vanlanduyt J, A structure model and growth-mechanism for multishell carbon nanotubes. *Science* 1995;267: 1334–1338.
92. Hata K, Futaba DN, Mizuno K, Namai T, Yumura M, Iijima S, Water-assisted highly efficient synthesis of impurity-free single-walled carbon nanotubes. *Science* 2004;306: 1362–1364.
93. Kong J, Soh HT, Cassell AM, Quate CF, Dai HJ, Synthesis of individual single-walled carbon nanotubes on patterned silicon wafers. *Nature* 1998;395: 878–881.
94. Qu L, Dai L, Gecko-foot-mimetic aligned single-walled carbon nanotube dry adhesives with unique electrical and thermal properties. *Adv Mater* 2007;19: 3844–3849.
95. Nikolaev P, Bronikowski MJ, Bradley RK, Rohmund F, Colbert DT, Smith KA, Smalley RE, Gas-phase catalytic growth of single-walled carbon nanotubes from carbon monoxide. *Chem Phys Lett* 1999;313: 91–97.
96. Andrews R, Jacques D, Rao AM, Derbyshire F, Qian D, Fan X, Dickey EC, Chen J, Continuous production of aligned carbon nanotubes: a step closer to commercial realization. *Chem Phys Lett* 1999;303: 467–474.
97. Xu YQ, Flor E, Kim MJ, Hamadani B, Schmidt H, Smalley RE, Hauge RH, Vertical array growth of small diameter single-walled carbon nanotubes. *J Am Chem Soc* 2006;128: 6560–6561.
98. Shyu YM, Hong FCN, Low-temperature growth and field emission of aligned carbon nanotubes by chemical vapor deposition. *Mater Chem Phys* 2001;72: 223–227.
99. Shyu YM, Hong FCN, The effects of pre-treatment and catalyst composition on growth of carbon nanofibers at low temperature. *Diam Relat Mater* 2001;10: 1241–1245.
100. Laplaze D, Alvarez L, Guillard T, Badie JM, Flamant G, Carbon nanotubes: dynamics of synthesis processes. *Carbon* 2002;40: 1621–1634.
101. Choi YC, Bae DJ, Lee YH, Lee BS, Park GS, Choi WB, Lee NS, Kim JM, Growth of carbon nanotubes by microwave plasma-enhanced chemical vapor deposition at low temperature. *J Vac Sci Technol A Vac Surf Films* 2000;18: 1864–1868.
102. Hofmann S, Ducati C, Robertson J, Kleinsorge B, Low-temperature growth of carbon nanotubes by plasma-enhanced chemical vapor deposition. *Appl Phys Lett* 2003;83: 135–137.
103. Devaux X, Vergnat M, On the low-temperature synthesis of SWCNTs by thermal CVD. *Physica E* 2008;40: 2268–2271.
104. Bower C, Zhu W, Jin SH, Zhou O, Plasma-induced alignment of carbon nanotubes. *Appl Phys Lett* 2000;77: 830–832.

105. Wang W, Yang KQ, Gaillard J, Bandaru PR, Rao AM, Rational synthesis of helically coiled carbon nanowires and nanotubes through the use of tin and indium catalysts. *Adv Mater* 2008;20: 179–182.
106. Naeemi A, Meindl JD, Monolayer metallic nanotube interconnects: Promising candidates for short local interconnects. *IEEE Electron Device Lett* 2005;26: 544–546.
107. Nihei M, Kawabata A, Kondo D, Horibe M, Sato S, Awano Y, Electrical properties of carbon nanotube bundles for future via interconnects. *Jpn J Appl Phys Part 1 Regul Pap Short Notes Rev Pap* 2005;44: 1626–1628.
108. Hoenlein W, Kreupl F, Duesberg GS, Graham AP, Liebau M, Seidel R, Unger E, Carbon nanotubes for microelectronics: status and future prospects. *Mater Sci Eng C Biomimetic Supramol Syst* 2003;23: 663–669.
109. Awano Y, Sato S, Kondo D, Ohfuti M, Kawabata A, Nihei M, Yokoyama N, Carbon nanotube via interconnect technologies: size-classified catalyst nanoparticles and low-resistance ohmic contact formation. *Phys Status Solidi A Appl Mat* 2006;203: 3611–3616.
110. Graham AP, Duesberg GS, Hoenlein W, Kreupl F, Liebau M, Martin R, Rajasekharan B, Pamler W, Seidel R, Steinhoegl W, Unger E, How do carbon nanotubes fit into the semiconductor roadmap?. *Appl Phys A Mater Sci Process* 2005;80: 1141–1151.
111. Rodríguez-Manzo JA, Banhart F, Terrones M, Terrones H, Grobert N, Ajayan PM, Sumpter BG, Meunier V, Wang M, Bando Y, Golberg D, Heterojunctions between metals and carbon nanotubes as ultimate nanocontacts. *Proc Natl Acad Sci U S A* 2009; 106.
112. Sato S, Kawabata A, Nihei M, Awano Y, Growth of diameter-controlled carbon nanotubes using monodisperse nickel nanoparticles obtained with a differential mobility analyzer. *Chem Phys Lett* 2003;382: 361–366.
113. Gwinn JP, Webb RL, Performance and testing of thermal interface materials. *Microelectron J* 2003;34(3): 215–222.
114. Prasher RS, Surface chemistry and characteristics based model for the thermal contact resistance of fluidic interstitial thermal interface materials. *J Heat Transf Trans ASME* 2001;123: 969–975.
115. Prasher RS, Shipley J, Prstic S, Koning P, Wang JL, Thermal resistance of particle laden polymeric thermal interface materials. *J Heat Transf Trans ASME* 2003;125: 1170–1177.
116. Xu J, Fisher TS, Enhancement of thermal interface materials with carbon nanotube arrays. *Int J Heat Mass Transf* 2006;49: 1658–1666.
117. Yu AP, Ramesh P, Itkis ME, Belyarova E, Haddon RC, Graphite nanoplatelet-epoxy composite thermal interface materials. *J Phys Chem C* 2007;111: 7565–7569.
118. Aoyagi Y, Leong CK, Chung DDL, Polyol-based phase-change thermal interface materials. *J Electron Mater* 2006;35: 416–424.
119. Chung DDL, Thermal interface materials. *J Mater Eng Perform* 2001;10: 56–59.
120. Howe TA, Leong CK, Chung DDL, Comparative evaluation of thermal interface materials for improving the thermal contact between an operating computer microprocessor and its heat sink. *J Electron Mater* 2006;35: 1628–1635.
121. Liu ZR, Chung DDL, Calorimetric evaluation of phase change materials for use as thermal interface materials. *Thermochimica Acta* 2001;366: 135–147.
122. Maguire L, Behnia M, Morrison G, Systematic evaluation of thermal interface materials - a case study in high power amplifier design. *Microelectron Reliab* 2005;45: 711–725.
123. Prasher RS, Rheology based modeling and design of particle laden polymeric thermal interface materials. *IEEE Trans Compon Packag Technol* 2005;28(2): 230–237.
124. Singhal V, Siegmund T, Garimella SV, Optimization of thermal interface materials for electronics cooling applications. *IEEE Trans Compon Packag Technol* 2004;27: 244–252.
125. Tong T, Zhao Y, Delzeit L, Kashani A, Meyyappan M, Majumdar A, Dense, vertically aligned multiwalled carbon nanotube arrays as thermal interface materials. *IEEE Trans Compon Packag Technol* 2007;30: 92–100.
126. Abadi P, Leong CK, Chung DDL, Factors that govern the performance of thermal interface materials. *J Electron Mater* 2009;38(1): 175–192.

127. Aoyagi Y, Chung DDL, Antioxidant-based phase-change thermal interface materials with high thermal stability. *J Electron Mater* 2008;37(4): 448–461.
128. De Mey G, Pilarski J, Wojcik M, Lasota M, Banaszczyk J, Vermeersch B, Napieralski A, Influence of interface materials on the thermal impedance of electronic packages. *Int Commun Heat Mass Transf* 2009;36: 210–212.
129. Kanuparthi S, Subbarayan G, Siegmund T, Sammakia B, An efficient network model for determining the effective thermal conductivity of particulate thermal interface materials. *IEEE Trans Compon Packag Technol* 2008;31: 611–621.
130. Lin C, Chung DDL, Graphite nanoplatelet pastes vs. carbon black pastes as thermal interface materials. *Carbon* 2009;47: 295–305.
131. Liu X, Zhang Y, Cassell AM, Cruden BA, Implications of catalyst control for carbon nanotube based thermal interface materials. *J Appl Phys* 2008; 104.
132. Thostenson ET, Ren ZF, Chou TW, Advances in the science and technology of carbon nanotubes and their composites: a review. *Compos Sci Technol* 2001;61: 1899–1912.
133. Borca-Tasciuc T, Mazumder M, Son Y, Pal SK, Schadler LS, Ajayan PM, Anisotropic thermal diffusivity characterization of aligned carbon nanotube-polymer composites. *J Nanosci Nanotechnol* 2007;7: 1581–1588.
134. Choi ES, Brooks JS, Eaton DL, Al-Haik MS, Hussaini MY, Garmestani H, Li D, Dahmen K, Enhancement of thermal and electrical properties of carbon nanotube polymer composites by magnetic field processing. *J Appl Phys* 2003;94: 6034–6039.
135. Velasco-Santos C, Martinez-Hernandez AL, Fisher FT, Ruoff R, Castano VM, Improvement of thermal and mechanical properties of carbon nanotube composites through chemical functionalization. *Chem Mater* 2003;15: 4470–4475.
136. Liu CH, Huang H, Wu Y, Fan SS, Thermal conductivity improvement of silicone elastomer with carbon nanotube loading. *Appl Phys Lett* 2004;84: 4248–4250.
137. Huang H, Liu CH, Wu Y, Fan SS, Aligned carbon nanotube composite films for thermal management. *Adv Mater* 2005;17: 1652–1656.
138. Bryning MB, Milkie DE, Islam MF, Kikkawa JM, Yodh AG, Thermal conductivity and interfacial resistance in single-wall carbon nanotube epoxy composites. *Appl Phys Lett* 2005; 87.
139. Gojny FH, Wichmann MHG, Fiedler B, Kinloch IA, Bauhofer W, Windle AH, Schulte K, Evaluation and identification of electrical and thermal conduction mechanisms in carbon nanotube/epoxy composites. *Polymer* 2006;47: 2036–2045.
140. Bonnet P, Sireude D, Garnier B, Chauvet O, Thermal properties and percolation in carbon nanotube-polymer composites. *Appl Phys Lett* 2007; 91.
141. Shenogina N, Shenogin S, Xue L, Kebinski P, On the lack of thermal percolation in carbon nanotube composites. *Appl Phys Lett* 2005; 87.
142. Nan CW, Liu G, Lin YH, Li M, Interface effect on thermal conductivity of carbon nanotube composites. *Appl Phys Lett* 2004;85: 3549–3551.
143. Shenogin S, Xue LP, Ozisik R, Kebinski P, Cahill DG, Role of thermal boundary resistance on the heat flow in carbon-nanotube composites. *J Appl Phys* 2004;95: 8136–8144.
144. Shenogin S, Bodapati A, Xue L, Ozisik R, Kebinski P, Effect of chemical functionalization on thermal transport of carbon nanotube composites. *Appl Phys Lett* 2004;85: 2229–2231.
145. Clancy TC, Gates TS, Modeling of interfacial modification effects on thermal conductivity of carbon nanotube composites. *Polymer* 2006;47: 5990–5996.
146. Liu CH, Fan SS, Effects of chemical modifications on the thermal conductivity of carbon nanotube composites. *Appl Phys Lett* 2005; 86.
147. Nan CW, Shi Z, Lin Y, A simple model for thermal conductivity of carbon nanotube-based composites. *Chem Phys Lett* 2003;375: 666–669.
148. Ju S, Li ZY, Theory of thermal conductance in carbon nanotube composites. *Phys Lett A* 2006;353: 194–197.
149. Nan CW, Birringer R, Clarke DR, Gleiter H, Effective thermal conductivity of particulate composites with interfacial thermal resistance. *J Appl Phys* 1997;81: 6692–6699.

150. Huxtable ST, Cahill DG, Shenogin S, Xue LP, Ozisik R, Barone P, Usrey M, Strano MS, Siddons G, Shim M, Kebblinski P, Interfacial heat flow in carbon nanotube suspensions. *Nat Mater* 2003;2: 731–734.
151. Lin W, Moon KS, Wong CP, A combined process of in-situ functionalization and microwave treatment to achieve ultra-small thermal expansion of aligned carbon nanotube/polymer nanocomposites for thermal management. *Adv Mater* 2009;21: 2421–2424.
152. Young DD, Nichols J, Kelly RM, Deiters A, Microwave activation of enzymatic catalysis. *J Am Chem Soc* 2008;130: 10048–10049.
153. Noh HS, Moon KS, Cannon A, Hesketh PJ, Wong CP, Wafer bonding using microwave heating of parylene intermediate layers. *J Micromech Microeng* 2004;14: 625–631.
154. Jiang HJ, Moon KS, Zhang ZQ, Pothukuchi S, Wong CP, Variable frequency microwave synthesis of silver nanoparticles. *J Nanopart Res* 2006;8: 117–124.
155. Moon KS, Li Y, Xu JW, Wong CP, Lead-free interconnect technique by using variable frequency microwave. *J Electron Mater* 2005;34: 1081–1088.
156. Sunden E, Moon JK, Wong CP, King WP, Graham S, Microwave assisted patterning of vertically aligned carbon nanotubes onto polymer substrates. *J Vac Sci Technol B* 2006;24: 1947–1950.
157. Cola BA, Xu J, Cheng CR, Xu XF, Fisher TS, Hu HP, Photoacoustic characterization of carbon nanotube array thermal interfaces. *J Appl Phys* 2007; 101.
158. Huang H, Liu CH, Wu Y, Fan SS, Aligned carbon nanotube composite films for thermal management. *Adv Mater* 2005;17: 1652–1653.
159. Panzer MA, Zhang G, Mann D, Hu X, Pop E, Dai H, Goodson KE, Thermal properties of metal-coated vertically aligned single-wall nanotube arrays. *J Heat Transf Trans ASME* 2008; 130.
160. Cola BA, Xu XF, Fisher TS, Increased real contact in thermal interfaces: A carbon nanotube/foil material. *Appl Phys Lett* 2007; 90.
161. Kordas K, Toth G, Moilanen P, Kumpumaki M, Vahakangas J, Uusimaki A, Vajtai R, Ajayan PM, Chip cooling with integrated carbon nanotube microfin architectures. *Appl Phys Lett* 2007; 90.
162. Zhu LB, Hess DW, Wong CP, Assembling Carbon Nanotube Films as Thermal Interface Materials Electronic Components and Technology Conference: IEEE, 2007:2006–2010
163. Hu XJ, Padilla AA, Xu J, Fisher TS, Goodson KE, 3-omega measurements of vertically oriented carbon nanotubes on silicon. *J Heat Transf Trans ASME* 2006;128: 1109–1113.
164. Wang XW, Zhong ZR, Xu J, Noncontact thermal characterization of multiwall carbon nanotubes. *J Appl Phys* 2005; 97.
165. Ngo Q, Cruden BA, Cassell AM, Sims G, Meyyappan M, Li J, Yang CY, Thermal interface properties of Cu-filled vertically aligned carbon nanofiber arrays. *Nano Lett* 2004;4: 2403–2407.
166. Xu Y, Zhang Y, Suhir E, Wang XW, Thermal properties of carbon nanotube array used for integrated circuit cooling. *J Appl Phys* 2006; 100.
167. Tong T, Zhao Y, Delzeit L, Majumdar A, Kashani A, Third ASME Integrated Nanosystems Conference. New York, 2004.
168. Amama PB, Cola BA, Sands TD, Xu XF, Fisher TS, Dendrimer-assisted controlled growth of carbon nanotubes for enhanced thermal interface conductance. *Nanotechnology* 2007; 18.
169. Moon KS, Lin W, Jiang HJ, Ko H, Zhu LB, Wong CP, Surface treatment of MWCNT array and its polymer composites for TIM application. Electronic Components and Technology Conference, 2008:234–237.
170. Amama PB, Ogebule O, Maschmann MR, Sands TD, Fisher TS, Dendrimer-assisted low-temperature growth of carbon nanotubes by plasma-enhanced chemical vapor deposition. *Chem Commun* 2006; 2899–2901.
171. Bae EJ, Min YS, Kang D, Ko JH, Park W, Low-temperature growth of single-walled carbon nanotubes by plasma enhanced chemical vapor deposition. *Chem Mater* 2005;17: 5141–5145.

172. Chen KC, Chen CF, Chiang JS, Hwang CL, Chang YY, Lee CC, Low temperature growth of carbon nanotubes on printing electrodes by MPCVD. *Thin Solid Films* 2006;498(1–2): 198–201.
173. Chen KC, Chen CF, Lee JH, Wu TL, Hwang CL, Tai NH, Hsiao MC, Low-temperature CVD growth of carbon nanotubes for field emission application. *Diam Relat Mater* 2007;16: 566–569.
174. Chiang WH, Sankaran RM, Synergistic Effects in Bimetallic Nanoparticles for low temperature carbon nanotube growth. *Adv Mater* 2008;20: 4857–4861.
175. Dai LM, Low-temperature, controlled synthesis of carbon nanotubes. *Small* 2005;1: 274–276.
176. Devaux X, Vergnat M, On the low-temperature synthesis of SWCNTs by thermal CVD. *Physica E* 2008;40(7): 2268–2271.
177. Dubosc M, Minea T, Besland MP, Cardinaud C, Granier A, Gohier A, Point S, Torres J, Low temperature plasma carbon nanotubes growth on patterned catalyst. *Microelectron Eng* 2006;83: 2427–2431.
178. Iwasaki T, Zhong GF, Kawarada H, Low-temperature growth of vertically aligned single-walled carbon nanotubes by radical CVD. *New Diam Front Carbon Technol* 2006;16: 177–184.
179. Jayatissa AH, Guo K, Synthesis of carbon nanotubes at low temperature by filament assisted atmospheric CVD and their field emission characteristics. *Vacuum* 2009;83: 853–856.
180. Kim SM, Zhang Y, Wang X, Teo KBK, Gangloff L, Milne WI, Wu J, Eastman M, Jiao J, Low-temperature growth of single-wall carbon nanotubes. *Nanotechnology* 2007; 18.
181. Kondo D, Sato S, Awano Y, Low-temperature synthesis of single-walled carbon nanotubes with a narrow diameter distribution using size-classified catalyst nanoparticles. *Chem Phys Lett* 2006;422: 481–487.
182. Kyung SJ, Lee YH, Kim CW, Lee JH, Yeom GY, Field emission properties of carbon nanotubes synthesized by capillary type atmospheric pressure plasma enhanced chemical vapor deposition at low temperature. *Carbon* 2006;44: 1530–1534.
183. Li CH, Liu HC, Tseng SC, Lin YP, Chen SP, Li JY, Wu KH, Juang JY, Enhancement of the field emission properties of low-temperature-growth multi-wall carbon nanotubes by KrF excimer laser irradiation post-treatment. *Diam Relat Mater* 2006;15(11–12): 2010–2014.
184. Liao KH, Ting JM, Characteristics of aligned carbon nanotubes synthesized using a high-rate low-temperature process. *Diam Relat Mater* 2006;15: 1210–1216.
185. Min YS, Bae EJ, Oh BS, Kang D, Park W, Low-temperature growth of single-walled carbon nanotubes by water plasma chemical vapor deposition. *J Am Chem Soc* 2005;127: 12498–12499.
186. Rummeli MH, Gruneis A, Loffler M, Jost O, Schonfelder R, Kramberger C, Grimm D, Gemming T, Barreiro A, Borowiak-Palen E, Kalbac M, Ayala P, Hubers HW, Buchner B, Pichler T, Novel catalysts for low temperature synthesis of single wall carbon nanotubes. *Physica Status Solidi B* 2006; 243(13):3101–3105.
187. Tam E, Ostrikov K, Plasma-controlled adatom delivery and (re)distribution: Enabling uninterrupted, low-temperature growth of ultralong vertically aligned single walled carbon nanotubes. *Appl Phys Lett* 2008; 93.
188. Uchino T, Bourdakos KN, de Groot CH, Ashburn P, Kiziroglou ME, Dilliway GD, Smith DC, Metal catalyst-free low-temperature carbon nanotube growth on SiGe islands. *Appl Phys Lett* 2005; 86.
189. Yu GJ, Gong JL, Zhu DZ, He SX, Zhu ZY, Synthesis of carbon nanotubes over rare earth zeolites at low temperature. *Carbon* 2005;43: 3015–3017.
190. Chen M, Chen CM, Shi SC, Chen CF, Low-temperature synthesis multiwalled carbon nanotubes by microwave plasma chemical vapor deposition using $\text{CH}_4\text{-CO}_2$ gas mixture. *Jpn J Appl Phys Part 1 Regul Pap Short Notes Rev Pap* 2003;42: 614–619.

191. Choi YC, Bae DJ, Lee YH, Lee BS, Han IT, Choi WB, Lee NS, Kim JM, Low temperature synthesis of carbon nanotubes by microwave plasma-enhanced chemical vapor deposition. *Synth Met* 2000;108: 159–163.
192. Choi YC, Bae DJ, Lee YH, Lee BS, Park GS, Choi WB, Lee NS, Kim JM, Growth of carbon nanotubes by microwave plasma-enhanced chemical vapor deposition at low temperature. *J Vac Sci Technol A Vac Surf Films* 2000;18(4): 1864–1868.
193. Honda S, Katayama M, Lee KY, Ikuno T, Ohkura S, Oura K, Furuta H, Hirao T, Low temperature synthesis of aligned carbon nanotubes by inductively coupled plasma chemical vapor deposition using pure methane. *Jpn J Appl Phys Part 2 Lett* 2003;42: L441–L443.
194. Jeong HJ, Jeong SY, Shin YM, Han JH, Lim SC, Eum SJ, Yang CW, Kim NG, Park CY, Lee YH, Dual-catalyst growth of vertically aligned carbon nanotubes at low temperature in thermal chemical vapor deposition. *Chem Phys Lett* 2002;361: 189–195.
195. Kang HS, Yoon HJ, Kim CO, Hong JP, Han IT, Cha SN, Song BK, Jung JE, Lee NS, Kim JM, Low temperature growth of multi-wall carbon nanotubes assisted by mesh potential using a modified plasma enhanced chemical vapor deposition system. *Chem Phys Lett* 2001;349: 196–200.
196. Lee CJ, Park J, Kim JM, Huh Y, Lee JY, No KS, Low-temperature growth of carbon nanotubes by thermal chemical vapor deposition using Pd, Cr, and Pt as co-catalyst. *Chem Phys Lett* 2000;327: 277–283.
197. Lee CJ, Son KH, Park J, Yoo JE, Huh Y, Lee JY, Low temperature growth of vertically aligned carbon nanotubes by thermal chemical vapor deposition. *Chem Phys Lett* 2001;338: 113–117.
198. Li MW, Hu Z, Wang XZ, Wu Q, Chen Y, Tian YL, Low-temperature synthesis of carbon nanotubes using corona discharge plasma at atmospheric pressure. *Diam Relat Mater* 2004;13: 111–115.
199. Li YL, Yu YD, Liang Y, A novel method for synthesis of carbon nanotubes: low temperature solid pyrolysis. *J Mater Res* 1997;12: 1678–1680.
200. Liao HW, Hafner JH, Low-temperature single-wall carbon nanotube synthesis by thermal chemical vapor deposition. *J Phys Chem B* 2004;108: 6941–6943.
201. Lin YH, Cui XL, Yen C, Wai CM, Platinum/carbon nanotube nanocomposite synthesized in supercritical fluid as electrocatalysts for low-temperature fuel cells. *J Phys Chem B* 2005;109: 14410–14415.
202. Maruyama S, Kojima R, Miyauchi Y, Chiashi S, Kohno M, Low-temperature synthesis of high-purity single-walled carbon nanotubes from alcohol. *Chem Phys Lett* 2002;360: 229–234.
203. Shao MW, Wang DB, Yu GH, Hu B, Yu WC, Qian YT, The synthesis of carbon nanotubes at low temperature via carbon suboxide disproportionation. *Carbon* 2004;42: 183–185.
204. Shiratori Y, Hiraoka H, Yamamoto M, Vertically aligned carbon nanotubes produced by radio-frequency plasma-enhanced chemical vapor deposition at low temperature and their growth mechanism. *Mater Chem Phys* 2004;87: 31–38.
205. Shyu YM, Hong FCN, Low-temperature growth and field emission of aligned carbon nanotubes by chemical vapor deposition. *Mater Chem Phys* 2001;72: 223–227.
206. Simon F, Kuzmany H, Rauf H, Pichler T, Bernardi J, Peterlik H, Korecz L, Fulop F, Janossy A, Low temperature fullerene encapsulation in single wall carbon nanotubes: synthesis of N@C-60@SWCNT. *Chem Phys Lett* 2004;383: 362–367.
207. Ting JM, Liao KH, Low-temperature, nonlinear rapid growth of aligned carbon nanotubes. *Chem Phys Lett* 2004;396: 469–472.
208. Vohs JK, Brege JJ, Raymond JE, Brown AE, Williams GL, Fahlman BD, Low-temperature growth of carbon nanotubes from the catalytic decomposition of carbon tetrachloride. *J Am Chem Soc* 2004;126: 9936–9937.
209. Wang WL, Bai XD, Xu Z, Liu S, Wang EG, Low temperature growth of single-walled carbon nanotubes: small diameters with narrow distribution. *Chem Phys Lett* 2006;419: 81–85.

210. Wang WZ, Kunwar S, Huang JY, Wang DZ, Ren ZF, Low temperature solvothermal synthesis of multiwall carbon nanotubes. *Nanotechnology* 2005;16: 21–23.
211. Wang XH, Hu Z, Wu Q, Chen Y, Low-temperature catalytic growth of carbon nanotubes under microwave plasma assistance. *Catal Today* 2002;72(3–4): 205–211.
212. Zhong DY, Liu S, Zhang GY, Wang EG, Large-scale well aligned carbon nitride nanotube films: low temperature growth and electron field emission. *J Appl Phys* 2001;89: 5939–5943.
213. Wang BA, Liu XY, Liu HM, Wu DX, Wang HP, Jiang JM, Wang XB, Hu PA, Liu YQ, Zhu DB, Controllable preparation of patterns of aligned carbon nanotubes on metals and metal-coated silicon substrates. *J Mater Chem* 2003;13: 1124–1126.
214. Xu FS, Liu XF, Tse SD, Synthesis of carbon nanotubes on metal alloy substrates with voltage bias in methane inverse diffusion flames. *Carbon* 2006;44: 570–577.
215. Hofmeister W, Kang WP, Wong YM, Davidson JL, Carbon nanotube growth from Cu-Co alloys for field emission applications. *J Vac Sci Technol B* 2004;22: 1286–1289.
216. Karwa M, Iqbal Z, Mitra S, Selective self-assembly of single walled carbon nanotubes in long steel tubing for chemical separations. *J Mater Chem* 2006;16: 2890–2895.
217. Talapatra S, Kar S, Pal SK, Vajtai R, Ci L, Victor P, Shaijumon MM, Kaur S, Nalamasu O, Ajayan PM, Direct growth of aligned carbon nanotubes on bulk metals. *Nat Nanotechnol* 2006;1: 112–116.
218. Gao LJ, Peng AP, Wang ZY, Zhang H, Shi ZJ, Gu ZN, Cao GP, Ding BZ, Growth of aligned carbon nanotube arrays on metallic substrate and its application to supercapacitors. *Solid State Commun* 2008;146: 380–383.
219. Singh MK, Singh PP, Titus E, Misra DS, LeNormand F, High density of multiwalled carbon nanotubes observed on nickel electroplated copper substrates by microwave plasma chemical vapor deposition. *Chem Phys Lett* 2002;354: 331–336.
220. Wang H, Feng JY, Hu XJ, Ng KM, Synthesis of aligned carbon nanotubes on double-sided metallic substrate by chemical vapor deposition. *J Phys Chem C* 2007;111: 12617–12624.
221. Yin XW, Wang QL, Lou CG, Zhang XB, Lei W, Growth of multi-walled CNTs emitters on an oxygen-free copper substrate by chemical-vapor deposition. *Appl Surf Sci* 2008;254: 6633–6636.
222. Lin W, Wong CP, Synthesis of Vertically Aligned Multi-Walled Carbon Nanotubes on Copper Substrates for Applications as Thermal Interface Materials, *MRS 2009 Spring Meeting*.
223. Zhu LB, Sun YY, Hess DW, Wong CP, Well-aligned open-ended carbon nanotube architectures: an approach for device assembly. *Nano Lett* 2006;6: 243–247.
224. Zhu LB, Xiu YH, Hess DW, Wong CP, Aligned carbon nanotube stacks by water-assisted selective etching. *Nano Lett* 2005;5: 2641–2645.
225. Powell CF, Oxley JH, Johan M, Blocher J, Vapor deposition. John Wiley & Sons, Inc., New York; 1966.
226. de los Arcos T, Vonau F, Garnier MG, Thommen V, Boyen HG, Oelhafen P, Duggelin M, Mathis D, Guggenheim R, Influence of iron-silicon interaction on the growth of carbon nanotubes produced by chemical vapor deposition. *Appl Phys Lett* 2002;80: 2383–2385.
227. Xiu YH, Zhang S, Yelundur V, Rohatgi A, Hess DW, Wong CP, Superhydrophobic and low light reflectivity silicon surfaces fabricated by hierarchical etching. *Langmuir* 2008;24: 10421–10426.
228. Pisana S, Cantoro M, Parvez A, Hofmann S, Ferrari AC, Robertson J, The role of precursor gases on the surface restructuring of catalyst films during carbon nanotube growth. *Physica E* 2007;37: 1–5.
229. Rodriguez NM, A review of catalytically grown carbon nanofibers. *J Mater Res* 1993;8: 3233–3250.
230. Deck CP, Vecchio K, Prediction of carbon nanotube growth success by the analysis of carbon-catalyst binary phase diagrams. *Carbon* 2006;44: 267–275.
231. Deng WQ, Xu X, Goddard WA, A two-stage mechanism of bimetallic catalyzed growth of single-walled carbon nanotubes. *Nano Lett* 2004;4: 2331–2335.

232. Krishnankutty N, Park C, Rodriguez NM, Baker RTK, The effect of copper on the structural characteristics of carbon filaments produced from iron catalyzed decomposition of ethylene. *Catal Today* 1997;37: 295–307.
233. Almazouzi A, Macht MP, Naundorf V, Neumann G, Diffusion of iron and nickel in single-crystalline copper. *Phys Rev B* 1996;54: 857–863.
234. Kononchuk O, Korablev KG, Yarykin N, Rozgonyi GA, Diffusion of iron in the silicon dioxide layer of silicon-on-insulator structures. *Appl Phys Lett* 1998;73: 1206–1208.
235. Puurunen RL, Surface chemistry of atomic layer deposition: a case study for the trimethylaluminum/water process. *J Appl Phys* 2005; 97.
236. Rahtu A, Alaranta T, Ritala M, In situ quartz crystal microbalance and quadrupole mass spectrometry studies of atomic layer deposition of aluminum oxide from trimethylaluminum and water. *Langmuir* 2001;17: 6506–6509.
237. Groner MD, Elam JW, Fabreguette FH, George SM, Electrical characterization of thin Al₂O₃ films grown by atomic layer deposition on silicon and various metal substrates. *Thin Solid Films* 2002;413: 186–197.
238. Zhu LB, Hess DW, Wong CP, Monitoring carbon nanotube growth by formation of nanotube stacks and investigation of the diffusion-controlled kinetics. *J Phys Chem B* 2006;110: 5445–5449.
239. Murakami Y, Maruyama S, Detachment of vertically aligned single-walled carbon nanotube films from substrates and their re-attachment to arbitrary surfaces. *Chem Phys Lett* 2006;422: 575–580.
240. Kim MJ, Nicholas N, Kittrell C, Haroz E, Shan HW, Wainerdi TJ, Lee S, Schmidt HK, Smalley RE, Hauge RH, Efficient transfer of a VA-SWNT film by a flip-over technique. *J Am Chem Soc* 2006;128: 9312–9313.
241. Jiang HJ, Zhu LB, Moon KS, Wong CP, Low temperature carbon nanotube film transfer via conductive polymer composites. *Nanotechnology* 2007;18: 4.
242. Gan Z, Liu S, Song X, Chen M, Lv Q, Yan H, Cao H, Large-scale Bonding of Aligned Carbon Nanotube Arrays onto Metal Electrodes. 2009.
243. Lin W, Xiu YG, Jiang HJ, Zhang RW, Hildreth O, Moon KS, Wong CP, Self-assembled monolayer-assisted chemical transfer of in situ functionalized carbon nanotubes. *J Am Chem Soc* 2008;130: 9636–9637.
244. Son Y, Pal SK, Borca-Tasciuc T, Ajayan PM, Siegel RW, Thermal resistance of the native interface between vertically aligned multiwalled carbon nanotube arrays and their SiO₂/Si substrate. *J Appl Phys* 2008; 103.
245. Yu CH, Saha S, Zhou JH, Shi L, Cassell AM, Cruden BA, Ngo Q, Li J, Thermal contact resistance and thermal conductivity of a carbon nanofiber. *J Heat Transf Trans ASME* 2006;128: 234–239.
246. Diao J, Srivastava D, Menon M, Molecular dynamics simulations of carbon nanotube/silicon interfacial thermal conductance. *J Chem Phys* 2008; 128.
247. Lin W, Zhang RW, Moon KS, Wong CP, Molecular phonon couplers at carbon nanotube/substrate interface to enhance interfacial thermal transport. *Carbon* 2009; doi: 10.1016/j.carbon.2009.08.033.
248. Berger C, Song ZM, Li TB, Li XB, Ogbazghi AY, Feng R, Dai ZT, Marchenkov AN, Conrad EH, First PN, de Heer WA, Ultrathin epitaxial graphite: 2D electron gas properties and a route toward graphene-based nanoelectronics. *J Phys Chem B* 2004;108: 19912–19916.
249. Freitag M, Graphene - Nanoelectronics goes flat out. *Nat Nanotechnol* 2008;3: 455–457.
250. Lee BK, Park SY, Kim HC, Cho K, Vogel EM, Kim MJ, Wallace RM, Kim JY, Conformal Al₂O₃ dielectric layer deposited by atomic layer deposition for graphene-based nanoelectronics. *Appl Phys Lett* 2008; 92.
251. Westervelt RM, Applied physics - Graphene nanoelectronics. *Science* 2008;320: 324–325.
252. Xuan Y, Wu YQ, Shen T, Qi M, Capamo MA, Cooper JA, Ye PD, Atomic-layer-deposited nanostructures for graphene-based nanoelectronics. *Appl Phys Lett* 2008; 92.
253. Chen ZH, Lin YM, Rooks MJ, Avouris P, Graphene nano-ribbon electronics. *Physica E* 2007;40(2): 228–232.

254. Eisberg N, Electronics – graphene set to replace, silicon. *Chem Ind* 2008; 11–11.
255. Hass J, Feng R, Li T, Li X, Zong Z, de Heer WA, First PN, Conrad EH, Jeffrey CA, Berger C, Highly ordered graphene for two dimensional electronics. *Appl Phys Lett* 2006; 89.
256. Wu XS, Sprinkle M, Li XB, Ming F, Berger C, de Heer WA, Epitaxial-graphene/graphene-oxide junction: an essential step towards epitaxial graphene electronics. *Phys Rev Lett* 2008; 101.
257. Berger C, Song ZM, Li XB, Wu XS, Brown N, Naud C, Mayou D, Li TB, Hass J, Marchenkov AN, Conrad EH, First PN, de Heer WA, Electronic confinement and coherence in patterned epitaxial graphene. *Science* 2006;312: 1191–1196.
258. Geim AK, Novoselov KS, The rise of graphene. *Nat Mater* 2007;6: 183–191.
259. Katsnelson MI, Novoselov KS, Geim AK, Chiral tunnelling and the Klein paradox in graphene. *Nat Phys* 2006;2: 620–625.
260. Novoselov KS, Geim AK, Morozov SV, Jiang D, Katsnelson MI, Grigorieva IV, Dubonos SV, Firsov AA, Two-dimensional gas of massless dirac fermions in graphene. *Nature* 2005;438: 197–200.
261. Ohta T, Bostwick A, Seyller T, Horn K, Rotenberg E, Controlling the electronic structure of bilayer graphene. *Science* 2006;313: 951–954.
262. Son YW, Cohen ML, Louie SG, Half-metallic graphene nanoribbons. *Nature* 2006;444: 347–349.
263. Han MY, Ozilmez B, Zhang YB, Kim P, Energy band-gap engineering of graphene nanoribbons. *Phys Rev Lett* 2007; 98.
264. Son YW, Cohen ML, Louie SG, Energy gaps in graphene nanoribbons. *Phys Rev Lett* 2006; 97.
265. Castro EV, Novoselov KS, Morozov SV, Peres NMR, Dos Santos J, Nilsson J, Guinea F, Geim AK, Neto AHC, Biased bilayer graphene: semiconductor with a gap tunable by the electric field effect. *Phys Rev Lett* 2007; 99.
266. Giovannetti G, Khomyakov PA, Brocks G, Kelly PJ, van den Brink J, Substrate-induced band gap in graphene on hexagonal boron nitride: ab initio density functional calculations. *Phys Rev B* 2007; 76.
267. McCann E, Asymmetry gap in the electronic band structure of bilayer graphene. *Phys Rev B* 2006; 74.
268. Zanella I, Guerini S, Fagan SB, Mendes J, Souza AG, Chemical doping-induced gap opening and spin polarization in graphene. *Phys Rev B* 2008; 77.
269. Gunlycke D, Lawler HM, White CT, Room-temperature ballistic transport in narrow graphene strips. *Phys Rev B* 2007; 75.
270. Huard B, Sulpizio JA, Stander N, Todd K, Yang B, Goldhaber-Gordon D, Transport measurements across a tunable potential barrier in graphene. *Phys Rev Lett* 2007; 98.
271. Katsnelson MI, Geim AK, Electron scattering on microscopic corrugations in grapheme. *Philos Trans R Soc A* 2008;366: 195–204.
272. Wang XR, Ouyang YJ, Li XL, Wang HL, Guo J, Dai HJ, Room-temperature all-semiconducting sub-10-nm graphene nanoribbon field-effect transistors. *Phys Rev Lett* 2008; 100.
273. Fiori G, Iannaccone G, Simulation of graphene nanoribbon field-effect transistors. *IEEE Electron Device Lett* 2007;28: 760–762.
274. Liang GC, Neophytou N, Lundstrom MS, Nikonorov DE, Ballistic graphene nanoribbon metal-oxide-semiconductor field-effect transistors: a full real-space quantum transport simulation. *J Appl Phys* 2007; 102.
275. Liang GC, Neophytou N, Nikonorov DE, Lundstrom MS, Performance projections for ballistic graphene nanoribbon field-effect transistors. *IEEE Trans Electron Devices* 2007;54: 677–682.
276. Liang X, Fu Z, Chou SY, Graphene transistors fabricated via transfer-printing in device active-areas on large wafer. *Nano Lett* 2007;7: 3840–3844.

277. Obradovic B, Kotlyar R, Heinz F, Matagne P, Rakshit T, Giles MD, Stettler MA, Nikonorov DE, Analysis of graphene nanoribbons as a channel material for field-effect transistors. *Appl Phys Lett* 2006; 88.
278. Yan QM, Huang B, Yu J, Zheng FW, Zang J, Wu J, Gu BL, Liu F, Duan WH, Intrinsic current-voltage characteristics of graphene nanoribbon transistors and effect of edge doping. *Nano Lett* 2007;7: 1469–1473.
279. Hwang EH, Adam S, Das Sarma S, Carrier transport in two-dimensional graphene layers. *Phys Rev Lett* 2007; 98.
280. Lemme MC, Echtermeyer TJ, Baus M, Kurz H, A graphene field-effect device. *IEEE Electron Device Lett* 2007;28: 282–284.
281. Zhang YB, Tan YW, Stormer HL, Kim P, Experimental observation of the quantum Hall effect and Berry's phase in graphene. *Nature* 2005;438: 201–204.
282. Gusynin VP, Sharapov SG, Unconventional integer quantum Hall effect in graphene. *Phys Rev Lett* 2005; 95.
283. Novoselov KS, McCann E, Morozov SV, Fal'ko VI, Katsnelson MI, Zeitler U, Jiang D, Schedin F, Geim AK, Unconventional quantum Hall effect and Berry's phase of 2 pi in bilayer graphene. *Nat Phys* 2006;2: 177–180.
284. Ghosh S, Calizo I, Teweldebrhan D, Pokatilov EP, Nika DL, Balandin AA, Bao W, Miao F, Lau CN, Extremely high thermal conductivity of graphene: prospects for thermal management applications in nanoelectronic circuits. *Appl Phys Lett* 2008; 92.
285. Balandin AA, Ghosh S, Bao WZ, Calizo I, Teweldebrhan D, Miao F, Lau CN, Superior thermal conductivity of single-layer graphene. *Nano Lett* 2008;8: 902–907.
286. Naeemi A, Meindl JD, Conductance modeling for graphene nanoribbon (GNR) interconnects. *IEEE Electron Device Lett* 2007;28: 428–431.
287. Shao Q, Liu G, Teweldebrhan D, Balandin AA, High-temperature quenching of electrical resistance in graphene interconnects. *Appl Phys Lett* 2008; 92.
288. Lan T, Pinnavaia TJ, Clay-reinforced epoxy nanocomposites. *Chem Mater* 1994;6: 2216–2219.
289. Moniruzzaman M, Winey KI, Polymer nanocomposites containing carbon nanotubes. *Macromolecules* 2006;39: 5194–5205.
290. Stankovich S, Dikin DA, Dommett GHB, Kohlhaas KM, Zimney EJ, Stach EA, Piner RD, Nguyen ST, Ruoff RS, Graphene-based composite materials. *Nature* 2006;442: 282–286.
291. Debelak B, Lafdi K, Use of exfoliated graphite filler to enhance polymer physical properties. *Carbon* 2007;45: 1727–1734.
292. Fukushima H, Drzal LT, Rook BP, Rich MJ, Thermal conductivity of exfoliated graphite nanocomposites. *J Therm Anal Calorim* 2006;85: 235–238.
293. Ganguli S, Roy AK, Anderson DP, Improved thermal conductivity for chemically functionalized exfoliated graphite/epoxy composites. *Carbon* 2008;46: 806–817.
294. Hung MT, Choi O, Ju YS, Hahn HT, Heat conduction in graphite-nanoplatelet-reinforced polymer nanocomposites. *Appl Phys Lett* 2006; 89.
295. Veca LM, Meziani MJ, Wang W, Wang X, Lu F, Zhang P, Lin Y, Fee R, Connell JW, Sun YP, Carbon nanosheets for polymeric nanocomposites with high thermal conductivity. *Adv Mater* 2009;21: 1–5.
296. Lin W, Zhang RW, Wong CP, High effective thermal conductivity of graphite nanosheet composites. *Journal of Electronic Materials* 2009; in press.
297. Prasher R, Thermal boundary resistance and thermal conductivity of multiwalled carbon nanotubes. *Phys Rev B* 2008; 77.
298. Sihn S, Ganguli S, Roy AK, Qu LT, Dai LM, Enhancement of through-thickness thermal conductivity in adhesively bonded joints using aligned carbon nanotubes. *Compos Sci Technol* 2008;68: 658–665.
299. Lin W, Moon KS, Wong CP, A combined process of in-situ functionalization and microwave treatment to achieve ultra-small thermal expansion of aligned carbon nanotube/polymer nanocomposites: toward applications as thermal interface materials. *Adv Mater* 2009; 21.

300. Schriemp JT, Laser flash technique for determining thermal diffusivity of liquid-metals at elevated-temperatures. *Rev Sci Instrum* 1972;43: 781–786.
301. Tada Y, Harada M, Tanigaki M, Eguchi W, Laser flash method for measuring thermal-conductivity of liquids – application to low thermal-conductivity liquids. *Rev Sci Instrum* 1978;49: 1305–1314.
302. Shaikh S, Li L, Lafdi K, Huie J, Thermal conductivity of an aligned carbon nanotube array. *Carbon* 2007;45: 2608–2613.
303. Zhang K, Chai Y, Yuen MMF, Xiao DGW, Chan PCH, Carbon nanotube thermal interface material for high-brightness light-emitting-diode cooling. *Nanotechnology* 2008; 19.
304. Wang SR, Liang ZY, Gonnet P, Liao YH, Wang B, Zhang C, Effect of nanotube functionalization on the coefficient of thermal expansion of nanocomposites. *Adv Funct Mater* 2007;17: 87–92.
305. Chen CI, Ni CY, Chang CM, Liu DS, Pan HY, Yuan TD, Thermal characterization of thermal interface materials. *Exp Tech* 2008;32: 48–52.
306. Chen CI, Ni CY, Pan HY, Chang CM, Liu DS, Practical evaluation for long-term stability of thermal interface material. *Exp Tech* 2009;33: 28–32.
307. Ohsone Y, Wu G, Dryden J, Zok F, Majumdar A, Optical measurement of thermal contact conductance between wafer-like thin solid samples. *J Heat Transf Trans ASME* 1999;121: 954–963.
308. Chu DC, Touzelbaev M, Goodson KE, Babin S, Pease RF, Thermal conductivity measurements of thin-film resist. *J Vac Sci Technol B Microelectron Nanometer Struct* 2001;19(6): 2874–2877.
309. Kading OW, Skurk H, Goodson KE, Thermal conduction in metallized silicon-dioxide layers on silicon. *Appl Phys Lett* 1994;65: 1629–1631.
310. Smith AN, Hostetler JL, Norris PM, Thermal boundary resistance measurements using a transient thermoreflectance technique. *Microscale Thermophys Eng* 2000;4: 51–60.
311. Miklos A, Lorincz A, Transient thermoreflectance of thin metal-films in the picosecond regime. *J Appl Phys* 1988;63: 2391–2395.
312. Paddock CA, Eesley GL, Transient thermoreflectance from thin metal-films. *J Appl Phys* 1986;60: 285–290.
313. Schoenlein RW, Lin WZ, Fujimoto JG, Eesley GL, Femtosecond studies of nonequilibrium electronic processes in metals. *Phys Rev Lett* 1987;58: 1680–1683.
314. Elsayedali HE, Norris TB, Pessot MA, Mourou GA, Time-resolved observation of electron-phonon relaxation in copper. *Phys Rev Lett* 1987;58: 1212–1215.
315. Hu HP, Wang XW, Xu XF, Generalized theory of the photoacoustic effect in a multilayer material. *J Appl Phys* 1999;86: 3953–3958.
316. Wang XW, Hu HP, Xu XF, Photo-acoustic measurement of thermal conductivity of thin films and bulk materials. *J Heat Transf Trans ASME* 2001;123: 138–144.
317. Wardle BL, Saito DS, Garcia EJ, Hart AJ, de Villoria RG, Verploegen EA, Fabrication and characterization of ultrahigh-volume-fraction aligned carbon nanotube-polymer composites. *Adv Mater* 2008;20: 2707–2714.

Nanomaterials via NanoSpray Combustion Chemical Vapor Condensation, and Their Electronic Applications

Andrew Hunt, Yongdong Jiang, Zhiyong Zhao, and Ganesh Venugopal

Abstract NanoSpray CombustionTM processing is a versatile and cost-effective manufacturing method for a wide range of materials, including nanopowders and nanostructured thin films. The NanoSpray process is used in combustion chemical vapor condensation (nCCVC) mode for making metal oxide, metal phosphate, and select metal nanopowders while combustion chemical vapor deposition (nCCVD) is used to make thin films. In this chapter, we will use NanoSpray Combustion process as a benchmark and review its capabilities to that of other nanomaterials' fabrication processes. Examples will mostly be of nanopowders synthesized using nCCVC. Applications of the nanomaterials in the electronic and energy sectors will be discussed.

Keywords nCCVC · nCCVD · Nanopowders · Nanomaterials · Nanocomposites · NanoSpray · Lithium-ion batteries · Conductive adhesives · Nonlinear optical materials · Capacitors

1 Introduction to Nanomaterials and Their Synthesis

The term nanomaterials is commonly used to refer to materials that have at least one dimension of less than 100 nm [1]. It is important to consider in the definition that there be some changes in properties due to this small dimension. In this manuscript, the term will be used interchangeably with recently coined terms like nanoparticles and nanopowders. Nanomaterials are not new; some have existed in nature forever, and man-made versions have been around for more than a 100 years. Volcanoes emit more nanoparticulates than nanomaterials' manufacturers might ever dream of making [2]. This has occurred before life even began and has continued to happen ever since. Carbon black, fumed silica, alumina and titania, and nano-silver (photographs) are a few examples of man-made nanomaterials that have been in use for several generations now. Tires are one of the oldest and most widely used

A. Hunt (✉)
nGimat Co., Atlanta, GA, USA
e-mail: ahunt@ngimat.com

nanomaterial enhanced products, with the wear and strength of the “rubber” coming predominantly via incorporation of nanosilica and carbon. Renewed interest in these materials over the last couple of decades stems from the correlation of the reduction in dimension to unique characteristics that were observed in the nanomaterial when compared to larger sized particles of the same composition [3]. Nanomaterials have also spawned advances in other multicomponent, multiphase materials like dispersions or composites, by enabling optimization of structure, property, and processing characteristics.

Nanomaterial fabrication methods can be broadly classified into three categories: mechanical, physical, and chemical means. For most of the physical and mechanical methods, the composition of the nanomaterial has already been determined via an upstream chemical synthesis process (by man or nature). The final fabrication is usually a conversion process whereby the larger particles are sized down to the nanometer scale. The mechanical approach is frequently referred to as the “top-down” approach [4]. For example, a mechanical process would involve mechanical milling either in a planetary or rotating ball mill.

Physical nano-methods typically involve a change in the material from one phase to another without a chemical change, or just a reduction in size from grinding or other physical process. For example, in physical precipitation, a material may be dissolved in a solvent and converted to a nanosized form via controlled physical precipitation of the material induced via temperature/pressure change or via addition of another component to reduce the solubility limit. Alternatively, drying of the solvent in a spray drying process could also lead to the formation of the nanomaterial. While it is difficult to achieve nanosized material by grinding or pulverization, such top-down (size reduction) processes are used but can result in hard agglomerated material and yield a wide size distribution. Physical methods also include vapor condensation methods such as thermal or plasma processing and laser ablation techniques [5, 6]. These techniques have been widely utilized for making nanomaterials of metals.

Chemical nano-methods are defined here as a processing method where the nanomaterial is synthesized from precursor materials directly in nanoscale form. Chemical approaches are frequently referred to as being “bottom-up” approaches [4]. Some post-processing could take place to further reduce or even increase the size of the particles to a desired value. Perhaps the most widely used chemical process for making nanomaterials is flame pyrolysis. Since the turn of the twentieth century, this method has been used to make carbon black by the pyrolysis of liquid or gaseous hydrocarbons [7]. Carbon black has since seen widespread application as a pigment and a reinforcement agent for plastics and elastomers. In addition to carbon black, flame pyrolysis can also be used to make fumed silica, fumed titania, and fumed alumina [8–10]. In spite of its popularity, flame pyrolysis is limited in terms of the compositional control and compound flexibility in the nanomaterials that it can produce. This is largely because limited number of high-volatility precursors are available that are suitable for flame pyrolysis. Care needs to be taken, or flame pyrolysis may promote formation of >100 nm secondary particles through aggregation [11]. Other chemical methods include variations of the flame pyrolysis

method, such as spray pyrolysis and wet-chemical methods such as sol–gel synthesis (chemical precipitation).

The line between physical and chemical processing can get blurred. There is physical vaporization combined with gas phase reaction to form compounds such as oxides, carbides, nitrides, and borides. This is the same physical vapor process used to make metal nanomaterials, but with a reaction gas metered in during processing to facilitate a chemical change. Surface adsorbed monolayers are another example of nanomaterials. These are mostly homogenous materials, except for where the monolayer is chemically or physicochemically bonded to the surface of a substrate. Although the primary material being deposited stays the same, the reaction at the substrate leads to modification of the properties of the surface and the adsorbed species.

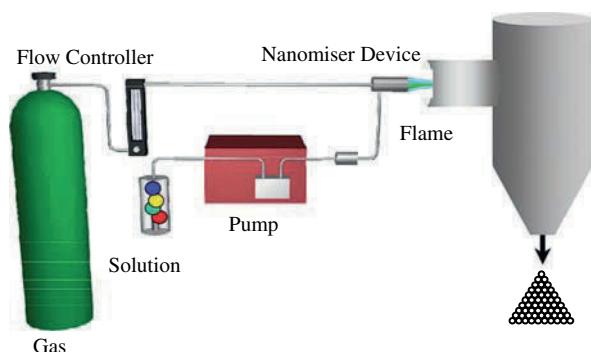
A large number of nanomaterial fabrication techniques currently exist and new ones continue to develop. While it is not possible to review each of these techniques in this chapter, it is important to make the distinction between the fabrication of nanoparticles or nanopowders, and other nanostructured materials that may exhibit nanoscale (<100 nm) features, but are not necessarily distinct particles. For example, lithographic patterning and vacuum deposition techniques like physical vapor deposition (PVD) and chemical vapor deposition (CVD) can produce nanostructured films, layers, or three-dimensional (3D) structures [12]. However, they are not very efficient for producing nanoparticles or nanopowders. Some of these processes have been referred to as “form-in-place” techniques for making nano-structured materials.

In the following section, we will focus on the combustion chemical vapor condensation method for making nanomaterials as a comparison to other processes discussed in this section. A brief mention of a related NanoSpray Combustion technique, called combustion chemical vapor deposition, will also be made. This method can be used to make nanostructured thin films and coatings.

2 NanoSpray Combustion Processing

NanoSpray Combustion processes to form nanomaterials have subsets depending on whether the material forms independently (homogeneous nucleation) or onto an existing solid (heterogeneous nucleation). Combustion chemical vapor condensation (nCCVC) is a homogeneous process for synthesizing powdered material by chemical change via combustion of precursor gases or solutions in droplet spray form, followed by the condensation of the material into ultrafine or nanopowder. nGimat Co. uses a proprietary atomizer technology, called the Nanomiser[®], to produce submicron liquid aerosol droplets (nanospray), which allows the vapor formation of the nanoparticles or nanopowder after combustion. Large volume commercial CCVC exists via gas precursor's combustion in forming carbon black, fumed silica and titania. The process of heterogeneously producing nanomaterial films via the combustion of a spray of submicron droplet has been patented by Georgia Tech [13] and licensed exclusively to nGimat, which formed the basis

Fig. 1 Schematic drawing of an nCCVC system comprising a Nanomiser device for NanoSpray Combustion processing



of the NanoSpray Combustion[®] technology umbrella. Thin-film coatings can be deposited onto substrates in the open atmosphere by way of Combustion Chemical Vapor Deposition (CCVD) process [14]. Figure 1 shows a schematic of the nCCVC system utilizing a Nanomiser device to produce nanopowders. In this chapter, the acronyms CCVC & nCCVC may be used interchangeably. Similarly, CCVD may be used instead of nCCVD and vice-versa.

The key advantage of the NanoSpray Combustion process is its ability to condense nanomaterials as nanoparticles or thin films in atmospheric pressure using low cost functional precursor chemicals in solution. This eliminates the need for costly high vapor pressure precursors, furnaces, vacuum equipment, and reaction chambers. Typical precursors, such as metal nitrates or metal carboxylates, are dissolved in a solvent, which typically also acts as the combustible fuel. Water-soluble precursors may be dissolved in water and then mixed with a suitable liquid or gas fuel. The solution is atomized to form submicron droplets, and these nano-droplets are then conveyed by an oxygen-containing stream to the flame where they combust in a manner similar to premixed gas fuels. In nCCVC of nanopowders, Fig. 1, the flame flash vaporizes the solvent and precursors, which decompose to yield metal and oxide vapors that consist of gaseous atoms, ions, and molecular-oxide species (e.g., M-O “monomers,” $<5\text{ \AA}$ in size). These vapor species then condense to form atomic clusters (e.g., $(\text{M-O})_x$, where $x \approx 10$ and the cluster is $<5\text{ nm}$ in size) that then may coalesce to form nanoparticles with a desired size distribution in any range from 2 to 100 nm. The synthesized nanoparticles are then captured in the dry form in bag house filters or via wet collections methods.

The number and the size of nuclei formed in the flame are a function of the flow rates, solution density, temperatures, time at temperature, and concentration. Depending on flame temperature and position in the flame where powder is collected, non-agglomerated particles can be readily produced. Atmospheric pressure flame temperature is in the range of 1,200–2,400°C, depending on oxygen ratio, flame size, solvent, precursor loading, and Nanomiser device atomization setting. In contrast to conventional spray pyrolysis from larger sized droplets, NanoSpray Combustion Processing technology is able to produce nCCVC nanopowders with

grain sizes less than 10 nm. By adjusting precursor solution concentrations and constituents, a wide range of nanopowder stoichiometries and compositions can be achieved.

In summary, the NanoSpray Combustion Processing technologies, nCCVC and nCCVD, offer the following capabilities for producing nanopowders or thin films:

- *Inexpensive precursor requirements:* Soluble precursors of any element are used instead of limiting, expensive, high vapor pressure metal-organics employed in traditional flame nanopowder technologies.
- *High degree of composition control:* Solution properties are adjustable to allow formation of a wide variety of complex, multicomponent compounds with precisely targeted stoichiometries.
- *Suitable for continuous production:* The equipment enables a robust production system shown to operate without failure around the clock for several days.
- *No need for specialized equipment:* CCVC nanopowders are formed at ambient conditions without the need for expensive, specialized equipment, thereby reducing down time and associated costs.
- *Accelerated development cycle for new applications:* Development of CCVC nanopowders for specific applications is achieved more rapidly than with nanopowders produced by traditional technologies; a large series of samples are quickly prepared for iterative testing and optimization.
- *Environmentally friendly:* Relatively safe chemical precursors (non-toxic and halogen-free) are used, resulting in benign by-products, thus reducing workplace hazards. By using biosolvents such as alcohols, the process is nearly carbon neutral.

3 Overview of Nanomaterials Capabilities

The NanoSpray Combustion process is ideally suited for producing oxides of metals. More than 150 compositions have been made, and these have been further varied over a range of stoichiometries. Most of the elements in the periodic table can be combined in different ratios, thereby enabling nanomaterials with novel properties. Phosphates and sulfates can also be produced, with the addition of stable precursors for phosphorus and sulfur. For sulfates, the process has to be controlled so that sulfur does not exit the nCCVC system as sulfur dioxide gas. By maintaining a reducing atmosphere in the combustion chamber downstream from the flame, metal nanopowders of moderate to low oxidation potential elements, like nickel, copper, tin, silver and palladium, can also be produced. A reducing atmosphere could also promote the formation a carbon coating on the surface of the nanoparticle. This surface treatment can be carried out both on reduced (metal) and oxidized (oxides, phosphates, etc.) nanoparticles. Unlike other competing processes, the surface modification takes place simultaneously with nanomaterial synthesis. This allows for low-cost production of complex nanomaterial compositions.

3.1 Single Metal Oxides and Phase Control

The CCVC process has been used to synthesize a wide variety of single metal oxides. Typical examples include MgO, Al₂O₃, CaO, TiO₂, Fe₂O₃, Fe₃O₄, ZnO, ZrO₂, SnO₂, and CeO₂. Single metal oxides are the simplest nanomaterials that can be made by nCCVC since metal stoichiometry is not an issue. Stable oxidation states such as the +2 state for MgO and CaO and the +4 state for ZrO₂ and SnO₂ are most easily achieved. Figure 2 shows the X-ray Diffraction (XRD) pattern and Transmission electron microscopy (TEM) micrograph for an SnO₂ nanopowder sample made using by NanoSpray Combustion of a Tin carboxylate precursor solution in an nCCVC system. Average particle sizes are less than 100 nm and the morphology of the particle is faceted.

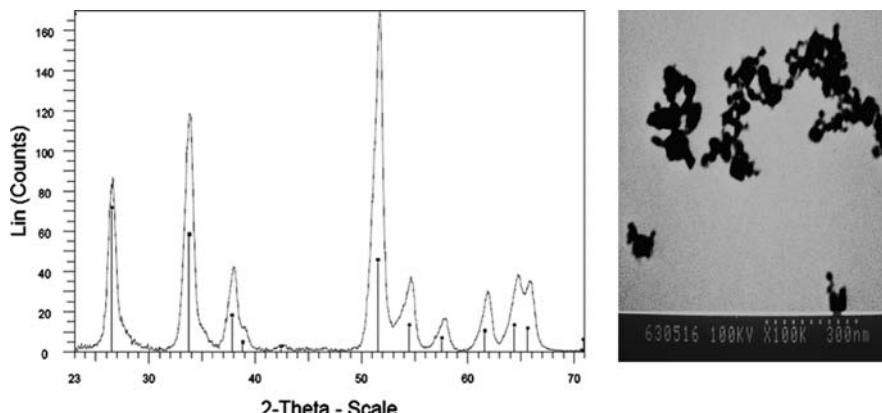


Fig. 2 X-ray diffraction pattern (*left*) and Transmission electron micrograph (*right*) for tin oxide nanopowder made by NanoSpray Combustion of a tin precursor solution in a CCVC system

The precursor material composition and the process parameters are controlled carefully to ensure that carbonates and hydroxides are not formed. Some elements, like Ca and Ba, are more susceptible to the formation of these additional by-products than others. Tight control of materials and process parameters are also required during synthesis of less stable oxidation states of specific metals. For example, iron oxides can be formed as FeO (iron (II) oxide), Fe₂O₃ (iron (III) oxide), or the mixed valence Fe₃O₄ (iron (II, III) oxide). While Fe₂O₃ is relatively simple to make by nCCVC, FeO and Fe₃O₄ require careful control of the combustion process, so that a slightly reducing atmosphere is maintained to promote formation of the lower oxidation states of the metal.

For some metals, the oxide may be capable of forming different phases, even though the oxidation state of the metal is unchanged. Titanium is a typical example, where the oxide (TiO₂) can form the Rutile, the Anatase, or the Brookite phase. Anatase and Brookite are known to change into the Rutile phase upon heating. Certain applications however require the Anatase phase preferentially. This is the case with dye-sensitized solar cells, where

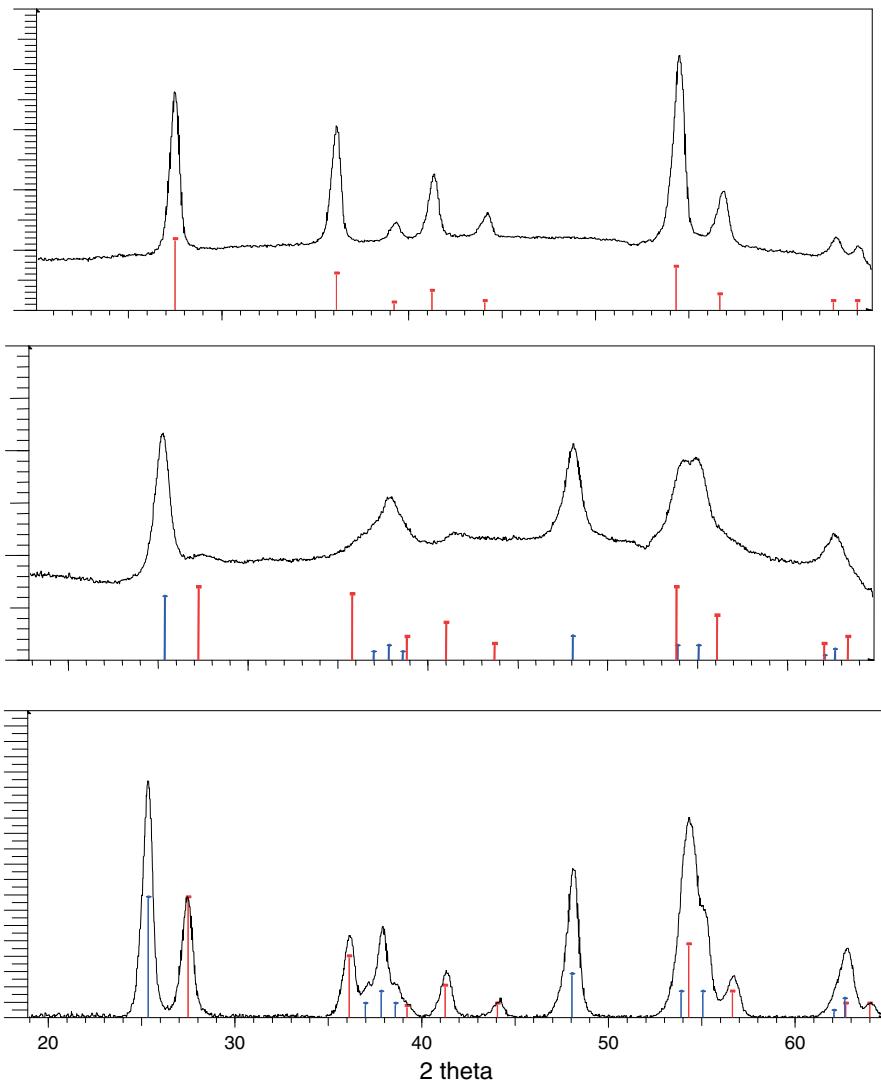


Fig. 3 X-ray diffraction patterns for Rutile (*top*), mostly Anatase (*center*), and Rutile + Anatase (*bottom*) forms of TiO_2 made using nCCVC

high efficiency is achieved when the TiO_2 is in the Anatase phase [15]. Both the Rutile and the Anatase phase of TiO_2 can be selectively produced using nCCVC. For Anatase, the precursor material solution and the process parameters are chosen so that the oxide forms under relatively cool conditions. Figure 3 shows the XRD curves for Rutile, mostly Anatase, and Rutile with Anatase forms of TiO_2 made by nCCVC operating under different materials and processing conditions.

3.2 Multi-metal Oxides

Synthesis of multi-metal oxides can be achieved by the combustion of precursor solutions containing multiple metal compounds in the desired stoichiometry. Precise formulation of stable precursor solutions is critical to achieving high-quality nanopowders. More than 100 multi-metal oxide nanopowders have been synthesized by nCCVC. Figure 4 shows the XRD and TEM (center) for as-synthesized Ni–YSZ nanopowder. In most cases, the material will form a single phase of similar stoichiometry to that of the feed solution, except when there are components that are not phase compatible as is the case with Ni and YSZ. The XRD pattern shows well-formed NiO and YSZ phases, even in as-synthesized form. The TEM shows mostly spherical particles with average particle size <50 nm. The TEM on the right-hand side of the same figure is for Cu/Sm/Ni/Gd-doped CeO₂ nanopowders, a complex metal oxide containing five metals synthesized by nCCVC.

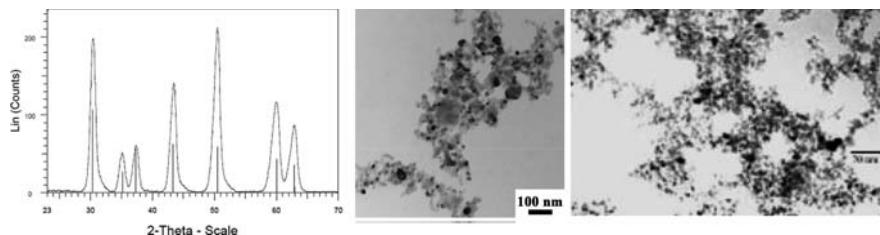


Fig. 4 XRD and TEM (center) for Ni–YSZ nanopowder made using CCVC; the TEM micrograph on the right is for Cu/Sm/Ni/Gd-doped CeO₂ (scale bar = 50 nm)

In most cases, nCCVC of a precursor solution containing multiple metals leads to the formation of either an amorphous mixture or a compound phase of the individual oxides of the component metal. Such is the case with some yttrium aluminum oxide materials, where an almost amorphous mixture is formed in the as-synthesized powder. Subsequent heating of the composite material at around 1,700°C leads to the formation of the crystalline garnet YAG phase. The XRDs for the nanopowder before and after heat treatment are shown in Fig. 5. Many elements can be combined

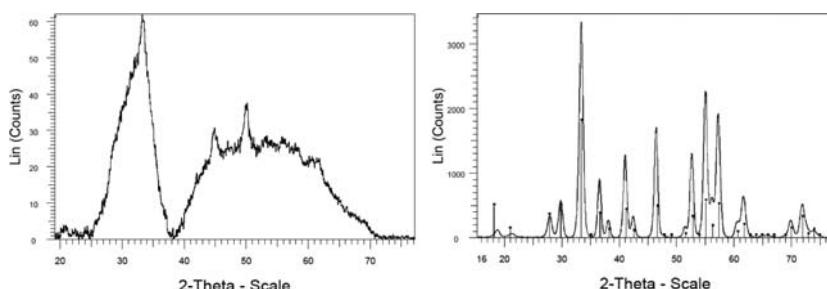


Fig. 5 XRD for yttrium aluminum oxide nanopowder before and after heat treatment at around 1,700°C, leading to the formation of the yttrium aluminum garnet (YAG) phase

in each grain, and this has been demonstrated in a multi-metal oxide, Cu/Sm/Ni/Gd-doped CeO₂, containing five metals. Both the CeO₂- and the YSZ-based materials discussed above play an important role in the electrolytes and electrodes that make up solid-oxide fuel cells [16] and as reforming catalysts for hydrogen generation [17]. By using nanopowders to fabricate the ceramic bodies in these applications, researchers have realized lower sintering temperatures, higher conductivities, and higher efficiencies in their respective end applications [16].

3.3 Metal Phosphates

Metal phosphate compositions are achieved by introducing phosphorus compounds into the metal precursor solution mix. nCCVC has been used to produce a variety of metal phosphates when a stable phase exists, including those of calcium, lanthanum, and cesium. Precursor selection and solution formulation are very important for stoichiometric control. Figure 6 shows the XRD and TEM for a synthetic nanopowder version of the bone mineral hydroxylapatite (HAP). Synthetic HAP is used for making bone cements for repairing fractured bone and as a sorbent for chromatographic separations [18, 19]. HAP has also been used as a cytocompatible growth promoter of bone tissue in tissue engineering applications [20]. For applications such as these, the higher surface area offered by nanosized particles could allow for faster growth and tissue regeneration.

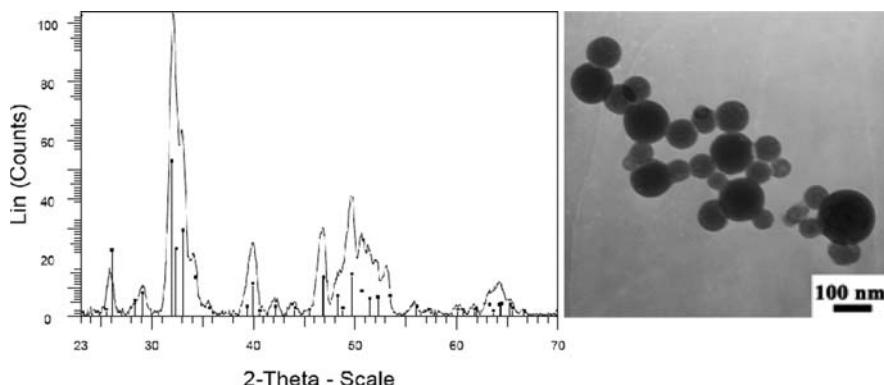


Fig. 6 XRD and TEM of synthetic hydroxylapatite nanopowder

Through careful control of precursor materials and process conditions, it has also been possible to produce other HAP-related calcium phosphate materials by CCVC. Examples of these are tri-calcium phosphate, amorphous calcium phosphate, and silicon-doped HAP. Synthetic versions of HAP with dopants have garnered a lot of attention recently for improving the biocompatibility of HAP in biomedical applications [21]. Nanosized phosphate materials have also attracted a lot of attention from

the battery industry. For example, lithium iron phosphate particles in the nanometer size range have enabled the fabrication of high power lithium-ion batteries that could be used for electric vehicle applications [22]. Other nanomaterials that serve the lithium-ion application are discussed in detail in a following section.

3.4 Metals

Since nCCVC is a combustion technology, particular attention has to be paid to the process conditions while using this method to make metal nanopowders. The list of metals that can be produced is limited to those that have moderate to low propensity to oxidize. More specifically, the metal has to have less oxidation potential than water or carbon dioxide. In combustion flames, water and carbon dioxide are the primary reactants. An element must have lower oxidation potential, be less competitive, to hydrogen and carbon to form as a metal in the high-energy flame environment. This list includes noble metals like, palladium, silver, and platinum. Other base metals such as iron, copper, and nickel have also been produced in nanopowder form by nCCVC, as has tin. Alloys of some of these metals such as silver–copper, iron–nickel, and tin–copper have also been produced. Figure 7 shows the XRD pattern and the TEM for silver metal nanopowder. Later in this chapter, we discuss in greater detail, the application of silver nanopowder in conductive adhesives for interconnect applications.

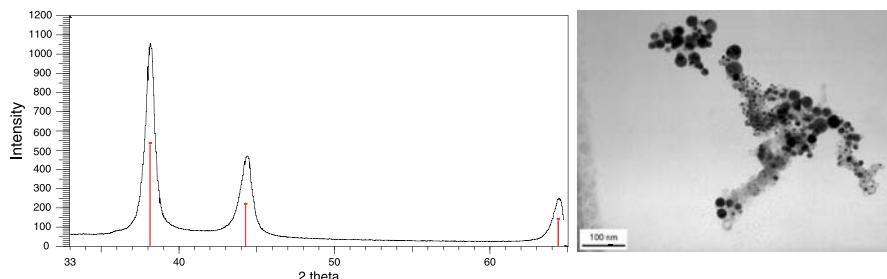


Fig. 7 XRD and TEM pattern for silver nanopowder

3.5 Nanocomposites

Nanocomposites are engineered multiphase materials that contain two or more distinct constituent materials, with at least one of the constituents having a dimension less than 100 nm. The resulting material has some physical, mechanical, or chemical property that is better than either of the constituents alone. The simplest nanocomposite consists of matrix material, the major phase, in which is uniformly dispersed

another material, the minor phase, which may have at least one dimension that is <100 nm.

Nanocomposites relevant to this discussion can be made in one of the several ways. The minor-phase material could be an nCCVC synthesized nanopowder that is dispersed in the matrix phase in a separate fabrication process, for example extrusion, casting, or spraying. Alternatively, the minor phase could be formed in situ by nCCVC during the nCCVD of the matrix inorganic material. Examples of each of these nanocomposites and their applications are discussed in the applications section.

3.6 Nanodispersions (Not Solutions)

In this context, the term nanodispersion refers to a suspension of solid nanomaterial in a liquid medium. Typical solids loadings for the dispersions are in the 5–30% range, although compositions outside this range are also feasible. The stability (length of settling time) of a nanodispersion depends on the surface characteristics of the nanoparticles, size, density, and also on the chemical composition and viscosity of the liquid medium. From this standpoint, nCCVC is capable of producing nanomaterials with specific surface characteristics that could compatibilize the nanoparticles with the liquid medium of choice. If surface modification does not address the desired compatibility between solid and liquid, a compatibilizer such as a surfactant or stabilizer might be added to achieve the desired result. Some people use the term “solution” incorrectly for dispersion. A true solution is molecular in scale which is smaller than solid nanoparticles. nCCVC-derived nanomaterials have been used to produce many dispersions including cerium oxide, zirconium oxide, silicon dioxide, hydroxylapatite, and cobalt aluminate.

4 Applications of Nanomaterials Made by CCVC

In addition to the dimensional reference, the term nanomaterial implies that the material in question also has a unique performance characteristic that is not exhibited by larger sized particles or bulk materials of the same composition [3]. The emergence of nanotechnology as an interdisciplinary science in recent years has mostly come about because of this realization. The performance characteristics that are being explored by research groups around the world are as varied as the materials themselves. Nanomaterials have been widely investigated for energy storage and energy conversion applications, where they are known to significantly improve power density in batteries or conversion efficiency in solar cells. For catalysis, nanomaterials can achieve similar efficiencies as micron-sized materials at lower activation temperatures, or higher efficiencies at similar temperatures. This behavior has a huge impact on myriad applications, from petroleum and petrochemical production to fuel cell electrodes. Nanomaterials also offer unique fabrication advantages, particularly for those related to ceramics processing. Small particle

sizes enable the fabrication of highly dense ceramic bodies at low processing temperatures that can have novel crystal structures. This is of particular importance to fuel cell electrolytes and laser host systems.

Whether in particle form, as made with nCCVC, or in thin-film form, as made by nCCVD, nanomaterials could be instrumental in the fabrication of miniaturized micro- or nanodevices. NanoSpray Combustion processing can enable a wide variety of the materials compositions for the components that make up these devices, in a form that can be easily integrated into the fabrication process. This opens the door for future application of these devices in electronics, electro-mechanical and biomedical areas. For example, in the area of biomedical device packaging, nanomaterials can enable novel biosensors and micro-fluidic devices that would have superior capability compared to what would otherwise be achieved with larger scale materials [23].

Nanocomposites have been used in structural applications to provide improved physical and mechanical properties. In addition to serving to improve the performance of conventional composites with larger sized reinforcing agents, nanocomposites are also being used in biomedical applications. A typical example is bone cement, made from synthetic hydroxylapatite nanopowders, which is used to repair broken or defective bone. Applications for nanodispersions have also emerged at a rapid pace. Typical examples include printing inks for printed electronics and pigments for coloring high temperature-stable ceramic tiles.

4.1 Conductive Adhesives as Electronic Interconnects

Continuous improvement in electronic interconnect technology has played a very important role in enabling sophisticated electronic devices that are reliable and environmentally friendly. Nanomaterials will play an important role in advancing future improvements in interconnect technology at several levels. Some examples have already started to emerge. Solder paste containing nanosized solder particles can be reflowed at lower temperatures than paste composed of conventional micron-sized solder particles [24]. This is of particular importance for lead-free solders, like tin–silver (Sn–Ag) and tin–silver–copper (Sn–Ag–Cu), where high melting points of the alloys require high reflow temperatures that potentially stress heat-sensitive components on the board. Similarly, nanomaterials with the right formulations can enhance performance in other areas of interconnect technology such as conductive adhesives [25], printing inks, underfill, thermal interface materials [26], and board technology.

Conductive adhesives can be used for connecting electronic components to rigid or flexible printed wiring (PWB) substrates instead of solder. They are particularly useful when one requires a low-temperature solder alternative for connecting temperature-sensitive components to conventional PWB substrates, conventional components to low-temperature stable boards, or a combination of both. When properly designed, conductive adhesives could also enable flexible electronic devices, since the adhesive composition can be tailored to impart the requisite mechanical properties in the interconnect material.

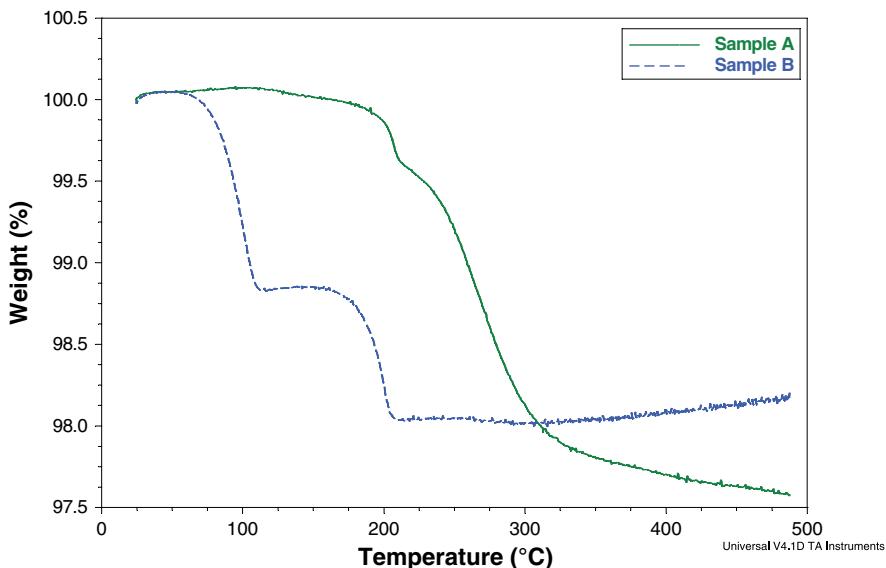


Fig. 8 Thermogravimetric analysis (TGA) curves for silver nanopowders made under relatively oxidizing (Sample B – dashed curve) and relatively reducing (Sample A – solid curve) conditions in the CCVC system

Silver-based nanopowders have been utilized to prepare conductive adhesive formulations [27]. In these formulations, the conductive component was composed of a mixture of micron-sized and nanosized silver particles. Silver nanopowders were made by nCCVC under oxidizing or reducing conditions. Figure 8 shows the thermogravimetric analysis (TGA) for two samples, representative of each of those production conditions, respectively. TGA experiments were done in air. Samples made under oxidizing conditions showed a weight loss of around 1%, leading up to approximately 100°C. This is believed to be due to the loss of adsorbed surface species. These samples also showed another pronounced weight loss of around 1%, leading up to approximately 210°C. This is attributed to decomposition of surficial silver oxide in the sample [28]. Since XRD shows only Ag metal peaks, the amount of oxide is nominal. Above 210°C, these samples are very stable and do not exhibit any further weight loss. Samples made under reducing conditions showed very little, if any, weight loss below 100°C and a relatively small loss (0.5%) at 210°C. These samples loose approximately 2% of their weight leading up to 300°C. This is indicative of the formation of carbonaceous surface coatings, as is to be expected of sample made under reducing conditions. The small loss leading up to 200°C is indicative of a relatively small amount of surface silver oxide being formed as a result of the reducing conditions used for this run. Furthermore, the negligible loss at temperatures below a 100°C suggests that the carbonaceous coating may be discouraging the adsorption of undesired chemical species on the surface of the nanomaterials.

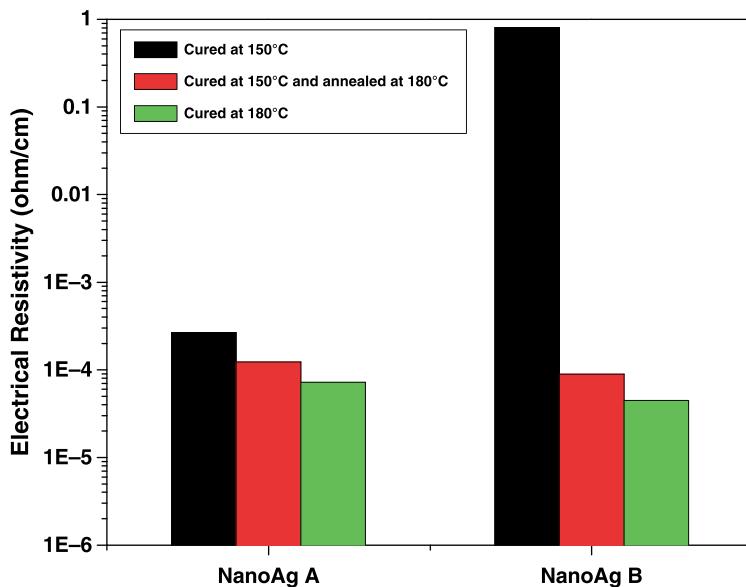


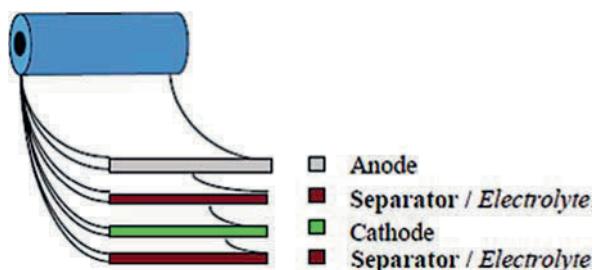
Fig. 9 Electrical resistivity measurements of conductive adhesives containing silver nanopowder made by CCVC under relatively oxidizing and relatively reducing conditions

Electrical resistivity measurements have been made on conductive adhesive formulations containing an epoxy matrix and a combination of micron-sized and nanosized silver particles. These measurements are shown graphically in Fig. 9. Samples with carbonaceous surface coatings yielded low resistivities ($\sim 2 \times 10^{-4} \Omega\text{cm}$) even at 150°C. Similar formulations consisting of silver nanopowders made under oxidizing conditions showed resistivity values $> 2 \times 10^{-2} \Omega\text{cm}$, indicating that the surficial silver oxide may be disrupting electron flow at the particle surface. Upon curing at 180°C, resistivities of all of the conductive adhesives dropped to lower than $1 \times 10^{-4} \Omega\text{cm}$, with the formulations containing the silver nanopowders made under oxidative conditions showing the lowest values. This indicates that at higher temperatures, the silver oxide layer is decomposed, thereby leading to lower resistivities. In summary, nanomaterial fabrication and the resulting particle characteristics play an important role in determining the performance of the material in an end-use application.

4.2 Lithium-Ion Battery Electrodes for Energy Storage

Lithium-ion (Li-ion) batteries have been the preferred choice for powering portable electronic devices such as laptop computers and cell phones for over a decade and a half. Figure 10 shows a schematic of a cylindrical lithium-ion cell in unwound

Fig. 10 Schematic of a cylindrical lithium-ion cell in unwound form showing anode, separator, and cathode layers



form, consisting of anode, cathode, and separator layers. Around a decade after its introduction into cellular phones, lithium-ion batteries completely replaced nickel-based chemistries. The selection of active (anode and cathode) materials for these energy-hungry applications has hence leaned toward high-capacity energy storage electrodes with layered structures such as LiCoO_2 and (for cathodes) and graphite (for anodes). These traditional materials are not the best in class when it comes to other characteristics such as power density, low temperature performance, thermal stability, calendar life, and cost. For this reason, Li-ion has found limited applications to date in areas such as hybrid electric vehicles (HEV), power tools, and back-up/storage, where traditional chemistries such as nickel metal hydride (NiMH), nickel cadmium (NiCd), and lead acid batteries are still preferred.

By sacrificing some electrode capacity, emerging Li-ion chemistries are capable of making significant improvements in the areas of power density (fast charge and discharge), cycle life, calendar life, thermal stability, and cost. This expands the utility of Li-ion batteries from portable electronic to aforementioned applications like HEV, power tools, and battery energy storage systems. Of these emerging materials classes, one that is considered a promising candidate is lithium titanate ($\text{Li}_4\text{Ti}_5\text{O}_{12}$, LTO). The 3D spinel structure of the LTO provides a more stable framework for the lithium ions to move back and forth during the charge discharge cycle, thus enhancing the cycle life and thermal stability [29]. The higher voltage of the LTO with respect to Li will eliminate the need to form a secondary interface (SEI) layer that is a prerequisite for stable graphite anodes [30]. Furthermore, due to the higher voltage of the LTO half-cell with respect to the Li/Li^+ electrode couple, the possibility of lithium plating is eliminated [30]. Both of these features render LTO-based batteries more thermally stable and safer compared to conventional Li-ion cells. The rate capability of conventional micron-sized LTO materials can be improved dramatically by synthesizing these electrode materials in the nanometer size range and/or coating with carbon. Nanosized particles allow shorter diffusion lengths for the lithium-ions in going from the bulk to the surface, and they allow high contact area between electrode and electrolyte, both requirements for high-power capability [31]. The benefits of coating LTO with carbon to improve the conductivity of the LTO and thereby the power capability of the resulting batteries has also been demonstrated [32]. While both of the aforementioned efforts, producing nanosized LTO and coating it with carbon, have been successful in synthesizing improved LTO

materials, they have relied on tedious multistep processing techniques. The versatility of the nGimat CCVC process simplifies what has otherwise been a laborious process for achieving carbon coating of nanometer-sized LTO, thereby enabling high power capable LTO at low cost. Figure 11 shows the XRD pattern for LTO nanopowder made using CCVC of a solution of lithium carboxylate and titanium alkoxide precursors. The pattern is in good agreement with the spinel phase of LTO ($\text{Li}_{1.33}\text{Ti}_{1.66}\text{O}_4$) published in the literature [33]. Lower cost lithium nitrate precursors can also be used to introduce lithium into the compound and form the spinel structure.

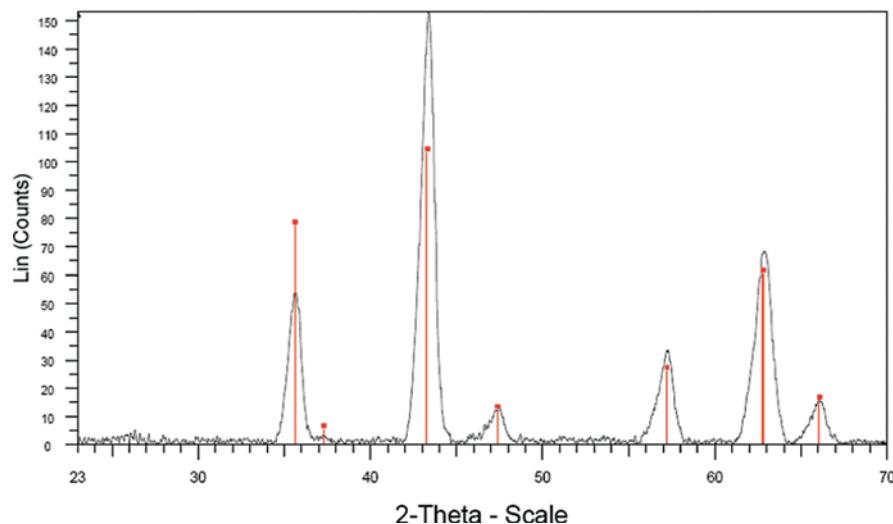


Fig. 11 XRD pattern for $\text{Li}_4\text{Ti}_5\text{O}_{12}$ made by nCCVC, the reference pattern from the library, corresponds to $\text{Li}_{1.33}\text{Ti}_{1.66}\text{O}_4$. (PDF number 26-1198)

Average particle size of the LTO was typically maintained to less than 100 nm. Particle morphology is determined by process conditions. For example, nanopowders formed at higher temperatures tended to form faceted morphologies, whereas particles that were formed at relatively low temperatures demonstrated spherical morphologies. This trend is shown in the TEM micrographs of the LTO samples in Fig. 12.

By adjusting the combustion environment and conditions, we have also been able to demonstrate the synthesis of carbon-coated LTO using a single-step process. This is an important process advantage offered by NanoSpray Combustion, since most commercial processes utilize at least two separate steps to manufacture carbon-coated electrode materials. The first step(s) in a commercial process typically involves the synthesis of the electrode powder (e.g., metal oxide), while the final step is dedicated to a carbon coating. Figure 13 shows a picture of three nanopowders that were made under oxidizing, slightly reducing and highly reducing conditions.

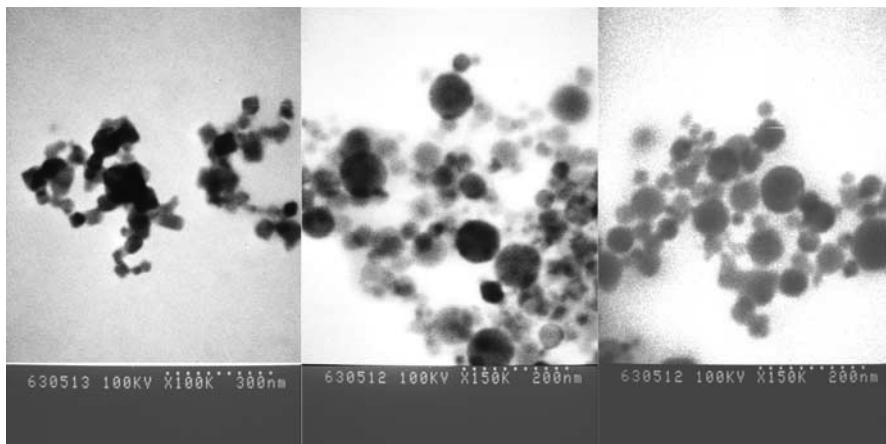
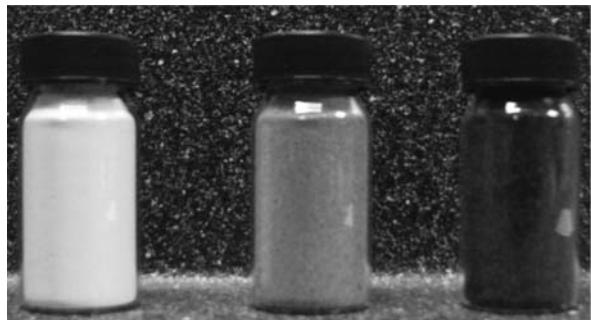


Fig. 12 TEM for LTO made using different processing conditions and different precursors: *left*: relatively high temperature, carboxylate precursor, *center*: relatively low temperature, carboxylate precursor, *right*: relatively low temperature, nitrate precursor

Fig. 13 Picture showing three nanopowders that were made under oxidizing (*left*), slightly reducing (*center*), and highly reducing (*right*) conditions



Thermogravimetric analysis (TGA) results for the carbon-coated materials indicated that the amount of carbon could be varied from 0.5 to 20%. Figure 14 compares the TGA scan for one typical carbon-coated LTO nanopowder sample with an LTO sample without any coating. Both scans were those of raw powder as removed from the nCCVC system prior to further treatment. The initial weight loss, up to 350°C, is expected to be due to surface adsorbed species that will need to be removed by heat treatment. This peak is visible even for uncoated LTO. The weight loss between 400 and 500°C is attributed to the decomposition of the carbon layer that is coated on the LTO. Around 6 wt% of carbon was coated in this case.

The best capacity for prototype cells containing nCCVC LTO nanopowder was approximately 143 mAh/g. Literature values have ranged anywhere between 120 and 155 mAh/g [32, 34]. The lower capacity, compared to the best-in-class, is

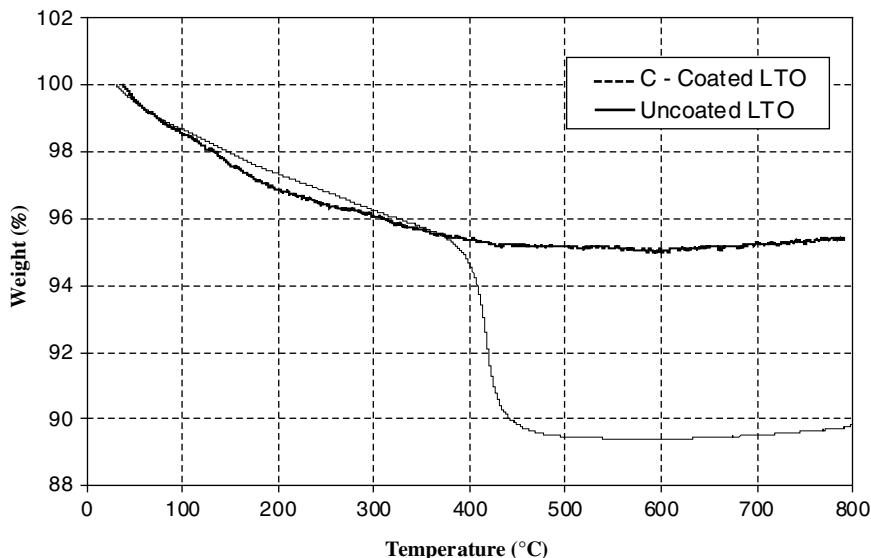


Fig. 14 TGA for uncoated (*solid*) and carbon-coated (*dashed*) LTO

attributed to surface-induced capacity losses due to high-surface area of the as-synthesized nanopowders that were used for the battery tests. Creating secondary particles that have larger particle size but at the same time retain some of nanostructure necessary to provide high rate capability could reduce these surface-induced losses. Nanostructured electrode powders of this kind have been shown by the DOE to provide optimum power and energy for LTO-based Li-ion batteries [35].

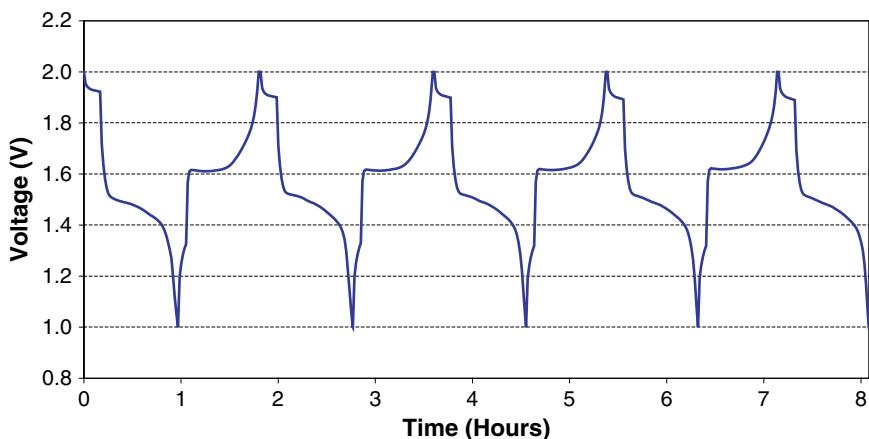


Fig. 15 1C–1C discharge–charge curves for an LTO | Li-metal T-cells over five cycles

Figure 15 shows the 1C–1C charge–discharge curves for the first five cycles of an LTO | Li half-cell. Even though the prototype cell is a half-cell that has not been optimized, it demonstrates relatively stable cycling behavior at 1C.

Rate capability studies were carried out at rates of 1C, 5C, and 10C. Figure 16 shows the discharge rate capability plot for a typical LTO | Li T-cell. Normalized to a 1C capacity of 100%, the 5C capacity was at around 80% and the 10C capacity was at around 60%. This performance is quite impressive, considering that the materials were tested in an unoptimized T-cell format. Most portable electronic applications can operate in the 1C–5C power range. However, higher rate capabilities are required for emerging lithium-ion batteries applications like power tools and electric vehicles.

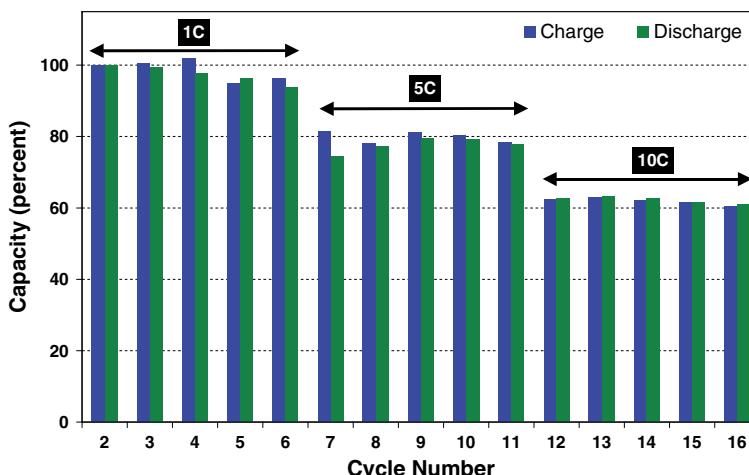


Fig. 16 Rate capability tests for nCCVC LTO | Li-metal T-cells at 1C, 5C, and 10C discharge rates. Charge rate was 1C for all. The left bar in each cycle corresponds to the charge capacity, while the right bar in each cycle corresponds to the discharge capacity

Some of the cells were cycled for an extended period to investigate preliminary cycle life capability. Figure 17 shows the chart for one such cell that was cycled to over 200 cycles at 1C charge and discharge rate. Prior to the test, this cell was cycled for five cycles in order to stabilize capacity. The sixth cycle capacity was set to 100% and the capacities of all subsequent cycles were normalized with respect to the sixth cycle. Around 75% of the capacity was retained after 100 cycles and around 60% after 200 cycles. Manufactured cells normally perform better than laboratory cells due to better elimination of oxygen and moisture.

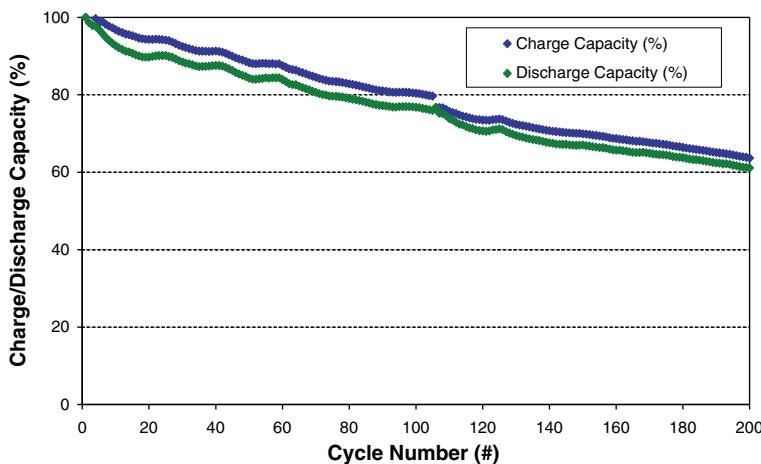


Fig. 17 Cycle life plot for a prototype Li-ion T-cell consisting of an LTO cathode and a lithium-metal anode. The upper trace represents the charge capacity and the lower trace represents the discharge capacity.

4.3 Polymer Nanocomposites for Capacitors

Development of electronic systems with reduced size, weight, cost, and improved functionality and performance will enable advances in applications such as telecommunications, network systems, automotive, and computer electronic devices. Electronic systems are composed of active components such as integrated circuits (ICs) and passive components including resistors, capacitors, and inductors. Currently more than 98% of passives used in the electronic industry are discrete passives. In a typical electronic system, discrete passives outnumber the active ICs by several times and occupy more than 70% real estate of the substrate. Discrete passives, especially capacitors, have already become the major barrier of the electronic systems miniaturization. Therefore, the development of integral passives is of much importance. Moreover, integral passives have the advantage of no separate interconnects to the substrate, reduced parasitic inductance, improved electrical performance, lower cost, and easier processing. Among passive components, special interest is focused on embedded capacitors because they are used for various functions, such as decoupling, by-passing, filtering, and timing capacitors. Important requirements for embedded capacitor materials include high dielectric constant, low capacitance tolerance, good processability, and low cost [36–41]. Polymer/ceramic nanocomposites have great potential to meet these practical requirements and are under extensive investigations [42–44]. However, to obtain a dielectric constant of 100 or higher, ceramic filler is generally required to be higher than 50 vol.%, which results in limited flexibility and processability. Meanwhile, polymer/metal nanocomposites have also attracted great attention for embedded capacitor applications [45–48]. Although polymer/metal composites have shown promise both

experimentally and theoretically, it is worthwhile to point out that the high dielectric constant of polymer/metal composites is usually achieved near the percolation threshold (the particle volume fraction at which particles have enough connectivity to create continuous paths through the structure) of metal particles. At these particle loadings, the composites have high leakage current and dissipation factors and low breakdown voltage, hence low performance and reliability.

In order to utilize the advantages of each materials system, nGimat has developed three-phase nanocomposites, in which polymer matrix dispersed with both ceramic and metal nanoparticles, which show promise to achieve high dielectric constant while retaining their flexibility and processability. In this study, the polymer matrix is cyclotene with a dielectric constant of about 2.6. All ceramic and metal nanoparticles were synthesized by the nCCVC technique. The nanocomposites were deposited by the NanoSpray technique onto platinized Si wafers. Capacitance and quality factor were measured by an HP 4285a Precision LCR meter at 1 MHz.

As an example, Fig. 18 shows SEM cross-section view of a typical cyclotene/(Ag+BTO) nanocomposite film with 5 vol.% Ag and 20 vol.% of barium strontium titanate (BTO) nanoparticles. The film is dense and uniform with nanometer-sized particulate features that are likely BTO particles. It is approximately 10 μm thick.

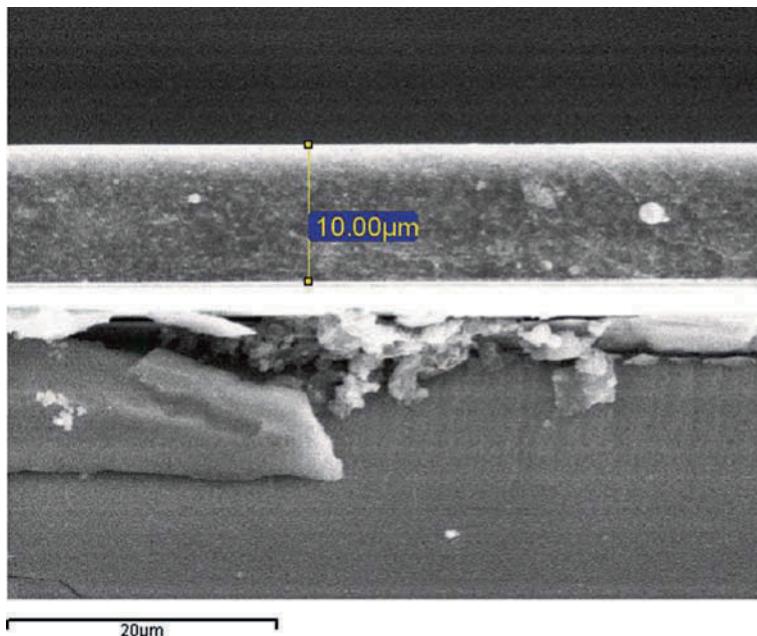


Fig. 18 SEM cross-section of cyclotene/(Ag+BTO) nanocomposites with 5 vol.% Ag and 20 vol.% BTO

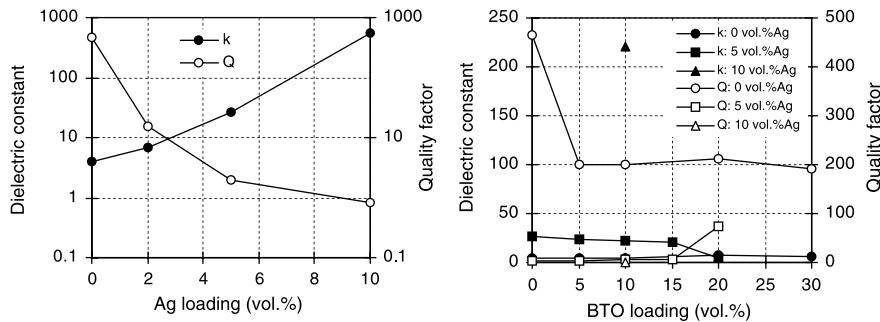
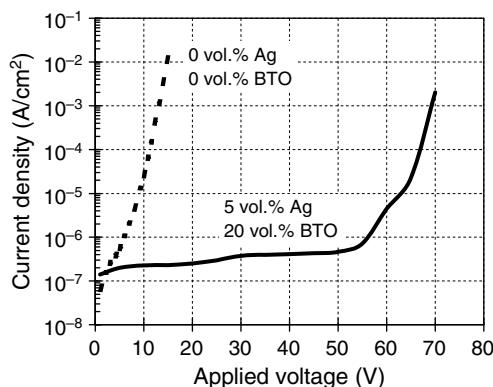


Fig. 19 Dielectric constant and quality factor as a function of (a) Ag loading and (b) BTO loading at different Ag loading levels

Capacitance and quality factor were measured at 1 MHz. Dielectric constant was calculated based on capacitance, film thickness, and electrode area. Figure 19 shows dielectric constant and quality factor of the nanocomposite films as a function of Ag loading and BTO loading at different Ag loading levels. For cyclotene/Ag nanocomposites (no BTO), dielectric constant increases and quality factor (Q) decreases with the increase of Ag loading, which suggests that Ag nanoparticles are partially percolated at relatively low loadings (2 vol.% or more). This is likely due to silver's propensity to diffuse at relatively low temperatures. Though the dielectric losses were large ($Q < 5$) for these partially percolated nanocomposites, impressive dielectric constants and capacitance densities of greater than 3,000 and 15 $\mu\text{F}/\text{in}^2$, respectively, were measured. Alternatively, for cyclotene/BTO nanocomposites (no Ag), the dielectric constant increases slightly as BTO loading is increased. Dielectric constant increases from 4 to 7 at 20 vol.% BTO. Quality factor decreases from 465 to 200 with increasing BTO loading from 0 to 5 vol.%, then remains approximately 200 at BTO loadings higher than 5 vol.%. For the cyclotene/(Ag+BTO) nanocomposites with 5 vol.% Ag, dielectric constant decreases and quality factor increases slightly with the increase of BTO loading. For example, the dielectric constant and quality factor are 27 and 2 at 0 vol.% BTO, and 21 and 5 at 15 vol.% BTO, respectively. A sharp decrease in dielectric constant and increase in quality factor were observed at 20 vol.% BTO, showing that Ag nanoparticles are not percolated at this point.

Figure 20 shows leakage current density as a function of applied voltage for pure cyclotene and cyclotene/(Ag+BTO) with 5 vol.% Ag and 20 vol.% BTO. The leakage current density of the pure cyclotene film increases rapidly with the increase of applied voltage. The leakage current density of the cyclotene/(Ag+BTO) film increases slightly with increasing applied voltage below 55 V. Above 55 V, it increases rapidly. At 5 V, leakage current density of the pure cyclotene and cyclotene/(Ag+BTO) with 5 vol.% Ag and 20 vol.% BTO is 5.3×10^{-7} and 2.0×10^{-7} A/cm^2 , respectively. The breakdown voltage of the nanocomposite film is about 55 V while that of pure cyclotene film is about 5 V, corresponding to breakdown strengths of 5.5 and 6.5 MV/m. Please note that the cyclotene/(Ag+BTO)

Fig. 20 Leakage current density as a function of applied voltage for pure cyclotene and cyclotene/(Ag+BTO)



nanocomposite is approximately ten times thicker than the pure cyclotene film.

In addition, nanocomposites of cyclotene embedded with Cu and BTO, cyclotene embedded with Cu and calcium copper titanate (CCT), and cyclotene embedded with Pt and BTO have also been developed. The nanocomposites were also scaled up onto large size copper foils for practical applications.

4.4 Inorganic Nanocomposites for Nonlinear Optical Materials

Nonlinear optical (NLO) effects play a major role in the development of high-speed data processing and communication systems. Materials with third-order optical nonlinearity ($\chi^{(3)}$) and fast response time are essential for light-controlled phase and refractive index modulation, which can replace electronic devices for next-generation applications in switching, routing, and signal processing. Future opportunities in photonic switching and information processing will depend critically on the development of improved photonic materials with enhanced third-order nonlinearity and reduced absorption coefficient (α).

It has been demonstrated that semiconductor or metal-dispersed dielectric thin films exhibited high third-order nonlinear properties which could be of great importance in future optical data processing. To obtain a high $\chi^{(3)}$ NLO material, a larger volume fraction of metallic particles with controllable shape and size in the dielectric matrix is desired [49]. Currently, techniques used for producing the nanocluster (NC) embedded composite materials include sol-gel, sputtering, ion implantation, and pulsed laser deposition (PLD). Most of these techniques are multistep processes, in which post-heat treatments are often needed to attain the desired properties. However, these treatments can also alter the average size, size dispersion, and spatial arrangement of the NCs, which could have unfavorable effects on the nonlinear properties. Furthermore, the vacuum-based methodologies are expensive, capital

intensive, batch-process limited, and suffer from low throughput. nGimat has successfully synthesized Pt nanoclusters embedded BaTiO₃ thin films, which exhibited high $\chi^{(3)}$ ($\sim 4.8 \times 10^{-10}$ esu) with a low α ($\sim 10^4$ cm⁻¹). This is the first time, to the best of our knowledge, that Pt was used as a metal dopant to improve NLO properties.

Since there is some mobility of Pt at high deposition temperatures, we deposited a Pt seed layer, then coated with BT using the CCVD at 600–1,000°C. It was predicted that the mobility of the Pt would enable it to diffuse into the dielectric matrix to form a nanocomposite, which it did.

Figure 21 shows the area detector XRD spectra of Pt seed layer and BT-coated Pt. The Pt seed layer is polycrystalline (slightly preferred). After BT deposition, the Pt diffraction peaks become sharper, which is contributed to possible Pt particle growth. This was found to be the case for Au seed layer as well. We theorize that the migrated Pt crystallites remain on the nanoscale after BT deposition (Fig. 21c).

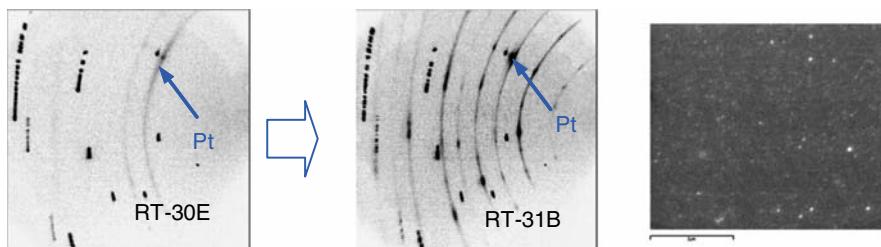


Fig. 21 XRD spectra of (a) Pt, (b) BT-coated Pt, and (c) plain-view SEM micrograph of Pt seed layer

Figure 22 shows the optical transmission spectra of Pt and BT/Pt. It is apparent that after BT deposition, the transmission increases, suggesting some Pt has migrated into small nanoparticles within the BT matrix. These Pt nanosized particles do not perturb light with wavelengths greater than 300 nm and therefore do not contribute to light scattering and absorption, resulting in higher transmission.

The z-scan technique [50] has been used to determine the sign and magnitude of nonlinear refractive index (n_2) of the nanocomposite films. The optical source uses a ~200-fs Ti-sapphire laser which permits peak power densities of ~1 GW/cm² requiring 50 mW average powers. The following parameters were given quantitatively or qualitatively through the z-scan measurement: the relative strength of the nonlinear index (n_2), the apparent two-photon absorption (TPA), which is an increased absorption with increasing intensity.

A pure n_2 response is indicated by a bipolar response. For example, a positive n_2 is indicated by a decrease in signal followed by an increase in signal strength. The signature is thus anti-symmetric around the z position corresponding to having the focus at the sample. Two-photon absorption is indicated by a decrease in signal strength and is symmetric about the focus location. Bleaching is indicated by an increase in signal strength and is also symmetric about the focus location.

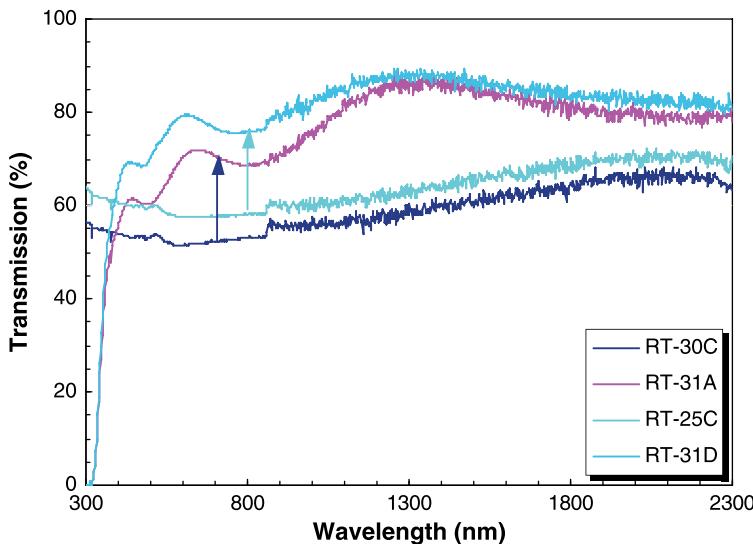


Fig. 22 Optical transmission spectra of Pt (RT-30C and RT-25C) and BT/Pt (RT-31A and RT-31D) layers. Arrows indicate the spectra of the same sample after BT deposition. In order of increasing transmission at 800nm, the curves are RT-30C, RT-25C, RT-31A and RT31-D

As a baseline, z-scan of a plain c-sapphire is shown in Fig. 23. Displayed characteristic nonlinear refraction coefficient of sapphire was on the order of $\sim 3 \times 10^{-16} \text{ cm}^2 \text{W}^{-1}$. Numerical estimation of z-scan response based upon measured experimental parameters corresponds very well to measured response. The

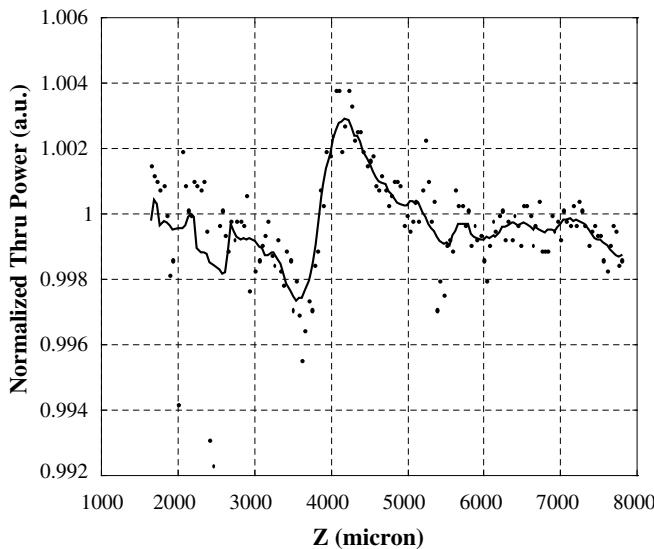


Fig. 23 z-Scan result of c-sapphire substrate

z-scan peak-to-peak amplitude was $\Delta II/I \sim 0.1\%$. z-Scan shows negligible nonlinear absorption.

Typical results of the z-scan measurement of BT and BT–Pt nanocomposite films are shown in Fig. 24 a (BT) and b (BT on Pt/sapphire with a thickness of 250 nm). The dots correspond to the far-field transmission data as a function of its distance (z) to the lens focus; the solid curve of Fig. 24b corresponds to the theoretical fit. The curve of Fig. 24b exhibits a valley-to-peak of $\sim 3.5\%$ and indicates a positive value of the nonlinear refractive index, n_2 . The sapphire substrate (0.47 mm thick) has a very small nonlinear optical response at 810 nm as shown above; thus the high nonlinear optical properties observed here result from the BaTiO₃/Pt films. Using the z-scan analysis, the nonlinear refractive index of BaTiO₃ on Pt layer is determined to be $\sim 4.8 \times 10^{-10}$ esu.

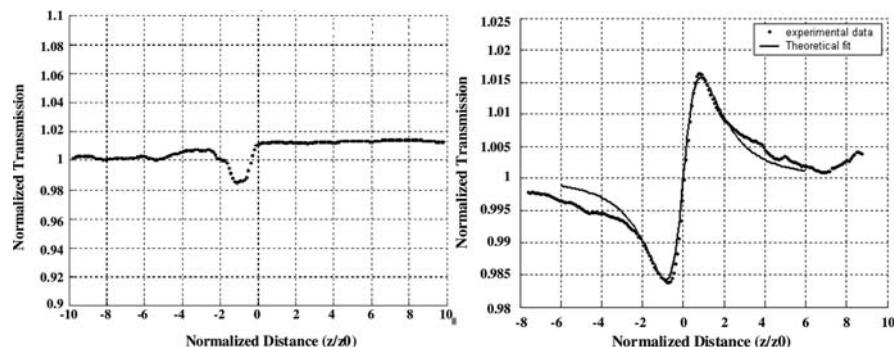


Fig. 24 z-Scan results of (a) BT film and (b) BT on Pt/sapphire

5 Conclusions

The versatility of NanoSpray Combustion technology is made evident by the variety of applications that are served by the nanomaterials made using both nCCVC and nCCVD. Nanopowders, nanocomposites, nanodispersions, and thin films of metals, metal oxides, mixed metal oxides, and metal phosphates have applications in a number of areas, including energy, electronics, biomedical, and printing. The ability to tailor the composition, particle size, particle morphology, and surface characteristics of these nanomaterials in a cost-effective process is critical for achieving superior applications' performance and therefore successful commercialization. NanoSpray CCVC and CCVD have demonstrated this ability to provide tailored materials and is poised to be the process of choice for the mass production of next-generation nanomaterials. As new applications for nanomaterials continue to emerge, there is an urgent need for a scalable and low-cost process to manufacture these materials in large volumes so that these technologies can be commercialized. The CCVC and CCVD processes offer the scaling efficiencies demanded by these applications

and do so for a broad range of chemistries and compositions. Compared to other nanomaterials manufacturing schemes, like those involving plasma, laser and vacuum processes, it can also claim to be significantly lower in power consumption characteristics. CCVC and CCVD can also be easily adapted to use biosolvents and chemicals derived from biomass sources, further reducing the carbon-footprint of the process.

Acknowledgments We thank all of the nGimat team members, past and present, who have contributed to the work discussed in this manuscript. In particular, we are grateful for the help provided by Dr. Elena Krumenaker and Mr. Michael Sapp. We would also like to thank Professor Steve Ralph at the Georgia Institute of Technology for performing z-scan measurements, and Professor C.P. Wong for his many discussions and collaborations. Last but not least, we are grateful for the funding provided by the US Department of Energy, Department of Defense, and the National Science Foundation that helped defray the cost of carrying out the research and development efforts described in this chapter.

References

1. ASTM Standards Document E 2456-06 (2006)
2. Buzea C., Pacheco I.I., and Robbie K., Nanomaterials and nanoparticles, sources and toxicity. *Biointerphases* 2007; 2(4): M17–M71.
3. Rempel A.A., Nanotechnologies. Properties and applications of nanostructured materials. *Russian Chemical Reviews* 2007; 76(5): 435–461.
4. Brechignac C., Houdy P., Lahmani M. (eds), *Nanomaterials and Nanochemistry*. Springer, Berlin (2008).
5. Vollath D., Plasma synthesis of nanopowders. *Journal of Nanoparticle Research* 2008; 10: 39–57.
6. Swihart M.T., Vapor-phase synthesis of nanoparticles. *Current Opinion in Colloid and Interface Science* 2003; 8: 127–133.
7. Carbon Black Users Guide, International Carbon Black Association, <http://www.carbon-black.org/carbonblackuserguide.pdf> (2004), Accessed 4 May 2009
8. Aerosil: Fumed Silica for Batteries, Technical Bulletin 2125, www.aerosil.com (2007), Accessed 4 May 2009
9. Aerioxide and Aeroperl: Titanium Dioxide as a Photocatalyst, Technical Bulletin 1243, www.aerosil.com (2005), Accessed 4 May 2009
10. Lin C., and Chung D.D.L., Nanstructured fumed metal oxides for thermal interface pastes. *Journal of Materials Science* 2007; 42: 9245–9255.
11. Pitkethly M.J., Nanomaterials – the driving force. *Nanotoday* 2004; 7(December issue): 20–28.
12. Rolland J.P., Hagberg E.C., Denison G.M., Carter K.R., and De Simone J.M., High-resolution soft lithography: enabling materials for nanotechnologies. *Angewandte Chemie* 2004; 43: 5796–5799.
13. Hunt A.T. et al., US Patents: 5652021, 5858465, 5863604 and 6013318
14. Hunt A.T. et al., *Applied Physics Letters* 1993; 63(2): 266.
15. Gratzel M., Photoelectrochemical cells. *Nature* 2001; 414: 338–344.
16. Hui S., Roller J., Sin Y., Zhang X., Deces-Petit C., Xie Y., Maric R., and Ghosh D., *Journal of Power Sources* 2007; 172: 493–502.
17. Cheekatamarla P.K., and Finnerty C.M., *Journal of Power Sources* 2006; 160: 490–599.
18. Fu Q., Zhou N., Huang W., and Wang D., Preparation and characterization of a novel bioactive bone cement: glass based nanoscale hydroxyapatite bone cement. *Journal of Materials Science: Materials in Medicine* 2004; 15: 133–1338.

19. Atkinson A. et al., Large-scale preparation of chromatographic grade hydroxylapatite and its application in protein separating procedures. *Journal of Applied Chemistry & Biotechnology* 1973; 23: 517–523.
20. Schnieder O.D., *Journal of Biomedical Materials Research Part B: Applied Biomaterials* 2008; 84B: 350–362.
21. Iliescu M. et al., Morphological and structural characterisation of osseointegrable Mn^{2+} and CO_3^{2-} doped hydroxylapatite thin films. *Materials Science and Engineering* 2007; 27(1): 105–109.
22. Chung S.-Y., Bloking J.T., and Chiang Y.M., *Nature Materials* 2002; 1: 123–128.
23. Ferrari M., (ed), *BioMEMS and Biomedical Nanotechnology*. Springer, New York (2006).
24. Wang X.D. et al., Nanomaterials and nanopackaging. In: Lu D., Wong C.P. (eds), *Materials for Advanced Packaging Reference*. pp 503–545, Springer, New York (2009).
25. Lu D., and Wong C.P., Electrically conductive adhesives. In: Lu D., Wong C.P. (eds), *Materials for Advanced Packaging Reference*. pp 365–405, Springer, New York (2009).
26. Prasher R., and Chiu C.-P., Thermal interface materials. In: Lu D., Wong C.P. (eds), *Materials for Advanced Packaging Reference*. pp 437–458, Springer, New York (2009).
27. Zhang R. et al., Georgia Institute of Technology, Manuscript in preparation
28. Lide D.R. (ed), *Handbook of Chemistry and Physics*, 87th Edition. CRC Press, Boca Raton, FL (2007)
29. Nazri G.A., Pistoia G., (eds), *Lithium Batteries: Science and Technology*. Kluwer Academic Publishers, Boston, Dordrecht, New York, London (2004).
30. Belharouak I. et al., *Journal of the Electrochemical Society* 2007; 154: A1083–A1087.
31. Manthiram A. et al., *Energy & Environmental Science* (2008). doi: 10.1039/b811802g
32. Amine K. et al., US Patent 2007/014845A1 (2007)
33. X-ray Diffraction Particle Data File # 26–1198
34. Thackeray M. et al., US Patent 7,452,630 B2 (2008)
35. Energy Storage R&D Annual Report, DOE Vehicle Technologies Program, (2007–2008)
36. Rao Y., and Wong C.P., *Journal of Applied Polymer Science* 2004; 92: 2228.
37. Rao Y., Ogitani S., Kohl P., and Wong C.P., *Journal of Applied Polymer Science* 2002; 83: 1084.
38. Kuo D.H., Chang C.C., Su T.Y., Wang W.K., and Lin B.Y., *Materials Chemistry and Physics* 2004; 85: 201.
39. Bhattacharya S.K., and Tummala R.R., *Microelectronics Journal* 2001; 32: 11.
40. Windlass H., Markondeya Raj P., Balaraman D., Bhattacharya S.K., and Tummala R.R., *IEEE Transactions on Advanced Packaging* 2003; 26: 10.
41. Pothukuchi S., Li Y., and Wong C.P., *Journal of Applied Polymer Science* 2004; 93: 1531.
42. Dang Z.M., Wang L., Wang H.Y., Nan C.W., Xie D., Yin Y., and Tjong S.C., *Applied Physics Letters* 2005; 86: 172905.
43. Kakimoto M., Takashi A., Tsurumi T., Hao J.J., Li L., Kikuchi R., Miwa T., Oono T., and Yamada S., *Materials Science and Engineering B* 2006; 132: 74.
44. Wu C.C., Chen Y.C., Su C.C., and Yang C.F., *European Polymer Journal* 2009; 45: 1442.
45. Li Y., Pothukuchi S., and Wong C.P., *Proceedings of the 9th Symposium on Advanced Packaging Materials* 2004; 175.
46. Sakzm D., Warjm N., Baalmann A., Simon, U., and Jaeger N., *Physical Chemistry Chemical Physics* 2002; 4: 2438.
47. Kelly J.M., Stenoien J.O., and Isbell D.E., *Journal of Applied Physics* 1953; 24: 258.
48. Wong C.P., Xu J.W., Moon K.S., and Tison C.K., *IEEE Transactions on Advanced Packaging* 2006; 29: 295.
49. Fukumi K., Chayahara A., Kadono K., Sakaguchi T., Horino Y., Miya M., Fujii K., Hayakawa J., and Satou M., Gold nanoparticles ion-implanted in glass with enhanced nonlinear-optical properties. *Journal of Applied Physics* 1994; 75: 3075.
50. Sheik-Bahae M., Said A.A., Wei T.H., Hagan D.J., and Van Stryland E.W., Sensitive measurement of optical nonlinearities using a single beam. *IEEE Journal of Quantum Electronics* 1990; 26: 760.

1D Nanowire Electrode Materials for Power Sources of Microelectronics

Jaephil Cho

Abstract Manipulation of nanostructured electrode material can provide versatile strategies toward improving the electrochemical properties. The nanostructure promotes fast Li-ion pathways for lithium ion transport and electronic conduction. Better rate capabilities are due to that the distance over which Li^+ must diffuse in the solid state dramatically decreases in the nanostructured electrode. Recently, there are growing applications of these materials in the area of micro-potable electronics, such as micro-robots, sensor nods, active radio frequency identification (RF-ID) tag, etc. Fast charging time and longer may be essential for these applications, and 1D-nanowire materials may enable to realize much more faster charging speeds than those achieved using conventional bulk materials, without significant degradation in storage capacity.

Keywords 1D Nanowire electrode · Li-ion · Fast charge · Microelectronics

1 Introduction

One-dimensional nanostructural materials have been intensively investigated for possible applications related to their magnetic properties and optical properties [1–4]. The superior characteristics of these materials over bulk counterparts originate mainly from their limited crystal dimensions with a very high surface-to-bulk ratio. Presently, the most popular power source for the mobile electronics is lithium-ion batteries for which the market is expected to grow to ~10 billion in 2009. On the other hand, lithium-ion batteries for possible applications in microelectronics (sensor nods (Fig. 1), active RF-ID (radio frequency identification)) need different electrode requirements. For instance, a micro-sensor nod may use the thin-film type Li-ion battery attached to the back panel for power supply, and electrode materials may need to use nanostructured morphology. Especially, these electronics

J. Cho (✉)

School of Energy Engineering, Ulsan National Institute of Science & Technology, Ulsan, Korea
e-mail: jpcho@unist.ac.kr

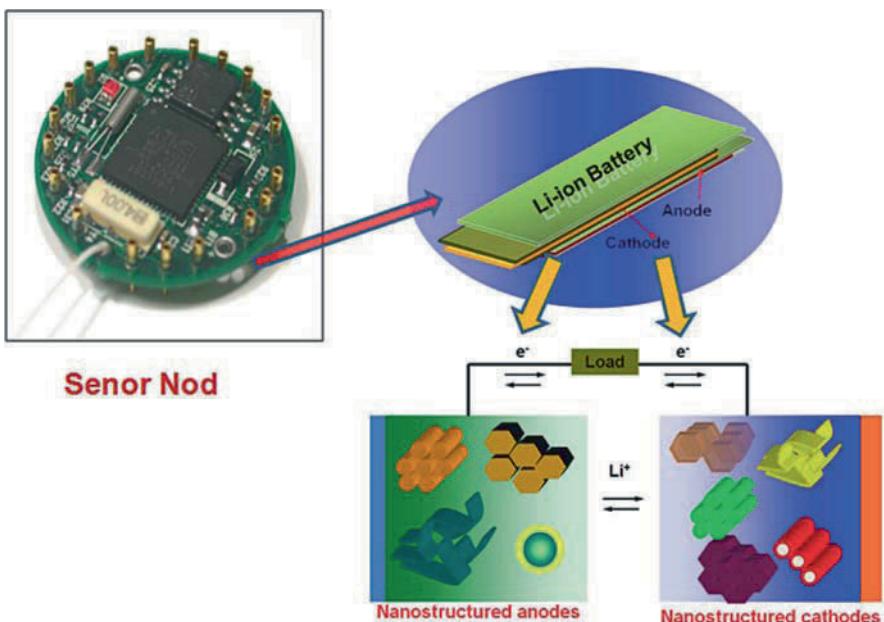


Fig. 1 Micro-sensor nod in which thin-film type Li-ion battery is implemented. The battery can be attached to the back panel of the nod and can use nanostructured cathode and anodes, such as nanowires and nanorods

require longer use time and fast charge time, and in this regard, electrode materials for power sources should have a high surface area, which can provide higher electrode/electrolyte contact areas, which can lead to shorter diffusion paths with the particles and more facile diffusion for electrons and Li ions. In addition, the reduced strain of intercalation and contributions from charge storage at the surface can also contribute to the Li capacity, compared with bulk counterparts.

In this study, 1D anode (SnO_2 , $\text{Sn}_{78}\text{Ge}_{22}$ @carbon) and cathode ($\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$) wires prepared by different synthetic methods were reviewed. These nanowires demonstrated superior capacity and rate capability (capacity retention at different current rates).

2 SnO_2 Nanowires

Due to the fact that SnO_2 is an n-type semiconductor with a wide band gap ($E_g = 3.6$ eV), it has been intensively investigated in the areas of gas sensors, solar cells, and catalysts [5–8]. Furthermore, it has been investigated for an anode material for lithium batteries due to its high capacity (>600 mAh/g) compared with graphite (372 mAh/g) [5–8]. However, similar to all other lithium-reactive materials,

it showed a volume change that was greater than 200% during the lithium alloy/dealloy process. This causes cracking and crumbling, resulting in electrical disconnection from the copper current collector. Eventually, rapid capacity fading occurs. In order to minimize such a drastic volume change, morphology controls of SnO₂ by a synthetic method, involving nanotubes [9], hollow nanotubes [10], nanowires [11], mesopores [12], and nanorods [13] have been reported.

To date, several preparative approaches utilizing a supramolecular templating mechanism have been reported for the preparation of mesoporous tin oxide [12, 14]. Upon removal of the surfactant, however, the mesoporous structure was destroyed or was somewhat disordered. Moreover, direct synthesis of these types of mesoporous materials using surfactants is often difficult; compared with silica, the surfactant/oxide composite precursors are often more susceptible to a lack of condensation, redox reactions, or phase transitions accompanied by thermal breakdown of the structural integrity [15]. Accordingly, an alternative method via a hard template method using a silica template with *p*6*mm* symmetry for preparing SnO₂ nanowire is used. It contains two-dimensional parallel cylindrical pores arranged with hexagonal symmetry.

The silica templates were prepared in an aqueous solution using Pluronic 123 (EO₂₀PO₇₀EO₂₀, MW = 5800, Aldrich) according to previously reported methods [15] that were modified. For preparing a silica template, 30 g of P123 and 30 g of *n*-butanol (Aldrich, 99.9%) were dissolved in 2000 g of distilled water and 60 g of 30 wt% HCl solution under stirring at 40°C; 30 g of tetraethoxysilane (TEOS) was added into the mixed solution at 40°C under stirring, and the remainder of the synthesis was identical to that of SBA-15.

Calcined silica templates at 550°C for 3 h were used to prepare mesostructural material and nanowire SnO₂; 18 g of SnCl₄5H₂O (99%, Aldrich) was dissolved in 20 g of distilled water. This solution was mixed with 20 g of the silica templates and stirred at room temperature until all of the solution was adsorbed. After drying in an oven at 120°C, this process (impregnation and drying) was repeated one additional time to obtain the SnO₂ precursor/SiO₂. The composite was further annealed at 400°C for 3 h and the silica template was removed from the SnO₂/SiO₂ composite by treating with dilute 1 M NaOH solution for 30 min and washing with distilled water and ethanol several times. Finally, the powers were vacuum dried at 120°C overnight.

Figure 2 shows high-angle XRD pattern of the SnO₂ nanowires and confirms the formation of nanocrystalline SnO₂, and the average crystal domain size was estimated as 6.5 ± 0.2 nm. ICP-MS results of the SnO₂ nanowire showed that residual Si from SiO₂ was 90 ppm, but residual Na from NaOH was not detected in both samples.

The nanowire morphology of SnO₂ prepared from the template is shown in Fig. 3a, b. An individual nanowire has the diameter of about 6 nm and >3 μ m in length, as evidenced by the TEM (Fig. 3a). In general, metal oxides prepared from the template led to aggregated nanowires, and our nanowire morphology is quite similar to that of nanowire CeO₂ [15]. A high-resolution TEM image (Fig. 3b) shows that the nanowires consist of nanocrystalline domains that are not structurally

Fig. 2 X-ray diffraction patterns of nanowire SnO₂

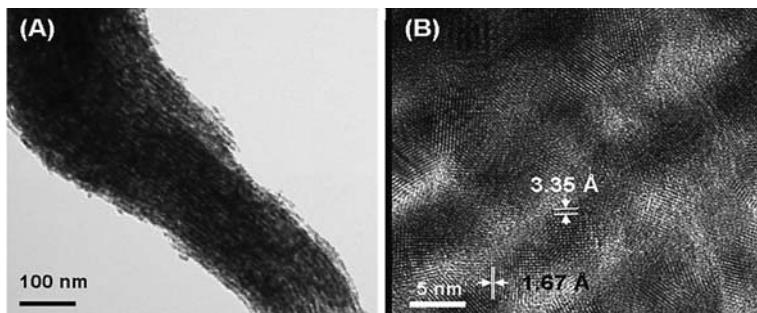
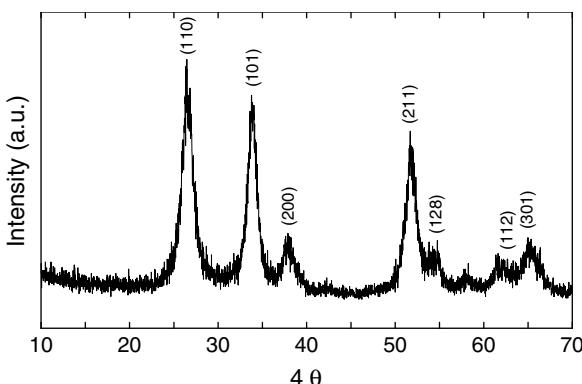


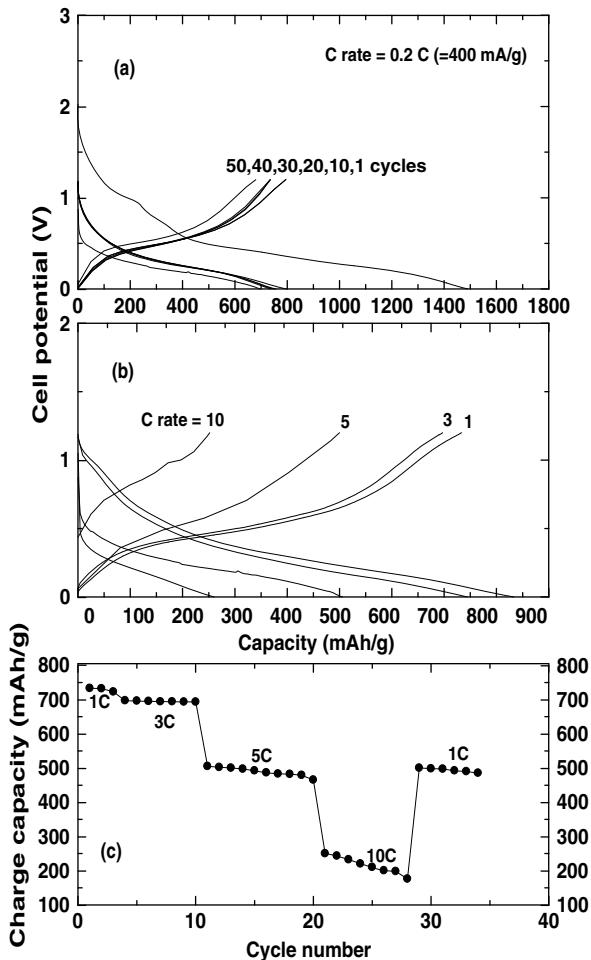
Fig. 3 TEM images of SnO₂ nanowires (**A**) along the (110) direction of the *p6mm* structure, (**B**) HREM image of (**A**). Copyright 2008, Royal Society of Chemistry: reproduced with permission from [16]

coherent, and that lattice fringes exist with d spacing values of 3.35 and 1.67 Å, corresponding to the (110) and (220) planes of SnO₂.

Figure 4 shows the cycle-life performance of the nanowire SnO₂ after 1, 10, 20, 30, 40, 50 cycles between 1.2 and 0 V at a rate of 0.2 C (= 80 mA/g). The first discharge and the charge capacity of the nanowires were 1400 and 795 mAh/g, respectively (the irreversible capacity ratio is 43%), and the capacity retention rate after 50 cycles is 85% at a rate of 0.2 C between 0 and 1.2 V. The high level of irreversible capacity during the first cycle is associated with the formation of an irreversible inactive Li₂O phase and possible side reactions with the electrolytes during the decomposition of SnO₂ ($\text{SnO}_2 + 4\text{Li}^+ + 4\text{e}^- \rightarrow \text{Li}_2\text{O} + \text{Sn}$) [17].

Compared with self-catalysis-grown SnO₂ nanowires, which has about 100 nm diameter that showed capacity decrease from 910 to 220 mAh/g at current density of 100 mA/g after 50 cycles [11], the capacity retention of the nanowires obtained from the template shows significantly improved value (680 mAh/g). This improvement may be related to that the spaces between the nanowires act as the buffer layers

Fig. 4 (a) Voltage profiles of the nanowire SnO₂ anode in coin-type half-cells during the 1st, 5th, 10th, 20th, 30th, and 50th cycles between 1.2 and 0 V. The charge C rate was maintained at 0.2 C (= 40 mA/g). (b) Voltage profiles of the nanowire SnO₂ anode in coin-type half-cells during each first cycle at different discharges of 1, 3, 5 and 10 C. The charge C rate was maintained at 0.2 C (= 40 mA/g).



during the lithium reactions. As the SnO₂ nanowires are aligned with different crystal orientations, an anisotropic volume change during the lithium alloy and dealloy processes can be expected. Although nanoscale materials can result in better capacity retention compared to micron-sized particles due to a more homogenous volume change and reduced absolute volume change between Sn and Li_xSn [18], nanoparticles would nonetheless aggregate into larger particles, resulting in pulverization and electrical disconnection with the current collector. Accordingly, a 50-cycled electrode shows some fragmentation of the nanowires into aggregated nanoparticles ~5 nm in size, inducing a capacity decay (Fig. 5). A TEM image shows the formation of tetragonal-Sn with different lattice fringes of (200) and (211), corresponding to 2.95 and 1.99 Å, respectively (Fig. 5).

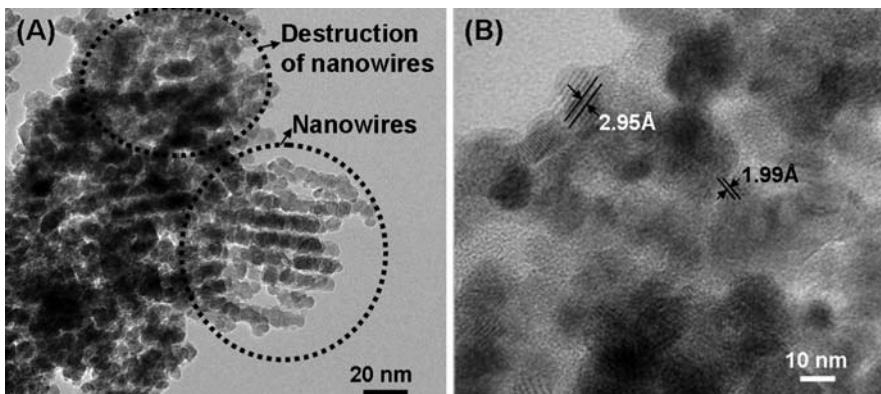


Fig. 5 TEM image of (A) the nanowire SnO_2 electrode after 50 cycles and (B) a high-resolution image of (A). Copyright 2008, Royal Society of Chemistry: reproduced with permission from [16]

When the cell is cycled at high rates from 1, 3, 5, and 10 C (with a discharge rate fixed at 0.2 C) (Fig. 4c), the charge capacity decreases slowly out to 3 C, after which it decreases rapidly, showing 697, 500, and 250 mAh/g at 3, 5, and 10 C rates, respectively; the capacity retention at a 10 C rate, compared with 1 C, is 31%. In addition, capacity recovery after 8 cycles at 10 C rate is 68% of the capacity at 1 C rate. This result indicates that nanowires may undergo a higher degree of nanowire pulverization than those cycled at lower C rate.

3 $\text{Sn}_{78}\text{Ge}_{22}$ @Carbon Nanowires

One of the most intensively studied anode materials for lithium batteries is Sn due to its relatively high theoretical capacity (991 mAh/g) and high electronic conductivities ($\sim 10^4$ s/cm) compared to other lithium-reactive metals [19]. In contrast to the metal hosts, lithium alloys Li_xM possess a highly ionic character; thus, are usually very brittle. Mechanical stresses related to volume changes in excess of 200% induce pulverization and aggregation as well as loss of the electrical interphase contact [20]. Although such a drastic volume change cannot be removed completely, the degree of the volume change can be reduced. In this regard, two methods have been proposed; the first involves forming an alloy with a ductile active metal so that it acts as a buffer for volume expansion [21]. The second method is similar to the first: bulk- or nanosized metals are coated with carbon or are dispersed in the carbon matrix [22,23]. Here, carbon and inactive conducting metal media act as electrical connector with the Cu current collector when the particles pulverized. These methods utilize zero-dimensional bulk particles or nanoparticles.

Syntheses and electrochemical studies of 1D lithium-reactive metal nanowires are rarely reported, and Si nanowires prepared by a laser ablation method have shown a reversible capacity < 500 mAh/g [24]. General synthetic methods for preparing metallic nanowires use vapor–liquid–solid (VLS) [25,26], solution–liquid–solid (SLS) or vapor–solid (VS) reactions [27,28].

In this method, for synthesizing the $\text{Sn}_{78}\text{Ge}_{22}$ @carbon nanowires, all the experimental procedures were carried out in purified Ar atmosphere except for the washing process. Sodium naphthalide solution was prepared from 5.4 g (0.23 mol) of sodium (Aldrich, 99.9%) and 19.38 g (0.15 mol) of naphthalene (99%, Aldrich), stirred in 100 ml 1,2-dimethoxyethane for 24 h; 3.978 ml of anhydrous SnCl_4 (Aldrich, 99.999%) and 0.684 ml of GeCl_2 (Aldrich, 99.99%) in a dried 50 ml of 1,2-dimethoxyethane (Aldrich, 99%) were thoroughly mixed in a flask under a dry argon atmosphere in a glove box for 3 h in order to get the atomic ratio of 85:15 (Sn:Ge) (atom%). This mixed solution was stirred for 24 h and then 60 ml of butyllithium ($n\text{-C}_4\text{H}_9\text{Li}$) (Aldrich, 99%) was added. This solution was stirred at room temperature for 4 h and naphthalene being removed by using a rotating evaporator at 110°C. The product was washed with distilled water seven times and finally vacuum-dried at 120°C for 48 h. The obtained product, which was a viscous liquid with a black color, was annealed at 600°C for 5 h under a vacuum atmosphere. The obtained nanoparticles were analyzed with an inductively coupled plasma-mass spectrometry (ICP-MS) to ensure the stoichiometry.

Our simple synthetic approach uses clusters of the butyl-capped $\text{Sn}_{78}\text{Ge}_{22}$ nanoparticles, and vacuum annealing of these clusters tuned into $\text{Sn}_{78}\text{Ge}_{22}$ @carbon core-shell nanowires (Fig. 6a). Figure 6b exhibits a TEM image of butyl-capped $\text{Sn}_{78}\text{Ge}_{22}$ nanoparticles before annealing. With these nanoparticles, particles of 1 μm in size are connected to each other. The electron diffraction (inset of b) pattern confirms the formation of the amorphous phase. Typically, Ge [29] or $\text{Sn}_{0.9}\text{Si}_{0.1}$ [30] nanoparticles obtained from similar synthetic methods are of a similar particle size while the particles are unconnected. Furthermore, the particles in the image consist of clusters of nanosized particles (~5–10 nm) capped with butyl groups. Hence, a similar occurrence can be expected with the sample before annealing. After thermal annealing at 600°C in a vacuum, the morphology before annealing changed completely, showing branched nanowires with diameters in a range of 50–100 nm and lengths that were several micrometers.

As $\text{Sn}_{78}\text{Ge}_{22}$ particles before annealing consist of many nanoparticles capped with butyl groups, these particles can be grown in different directions when the butyl groups are burnt out, as shown in Fig. 6c. Inductively coupled plasma-mass spectroscopy (ICP-MS) of the nanowires after annealing confirmed the formation of $\text{Sn}_{77}\text{Ge}_{23}$, and the carbon content determined by a CHS analyzer in the nanowires was 5 wt%. Figure 6d is an expanded TEM image of Region 2, in which an amorphous carbon layer with a thickness of ~2 nm is shown. This carbon layer was formed from the decomposition of C_4H_9 groups during annealing.

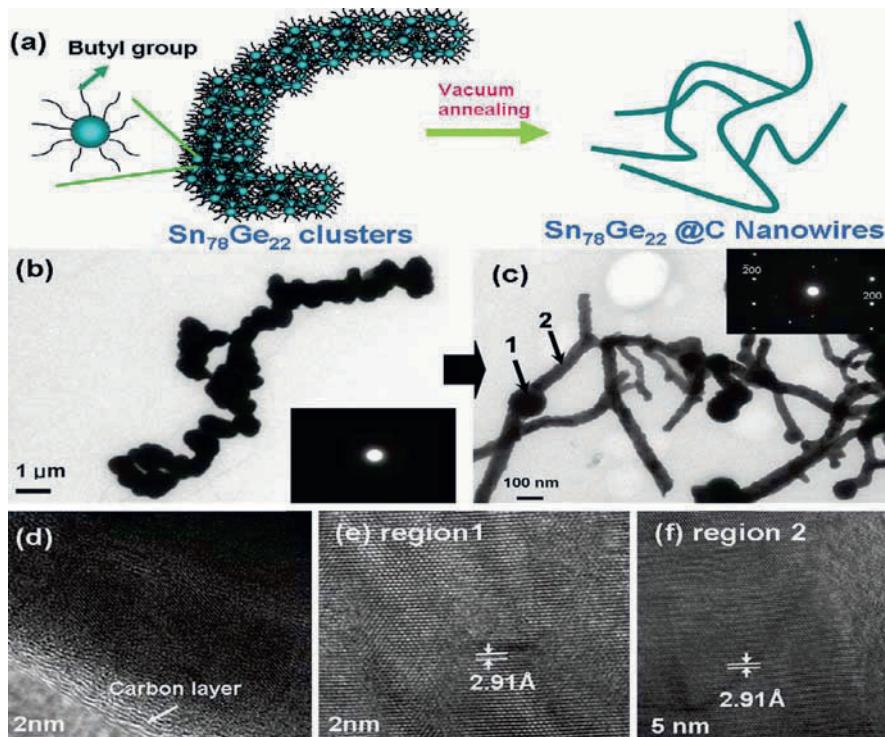
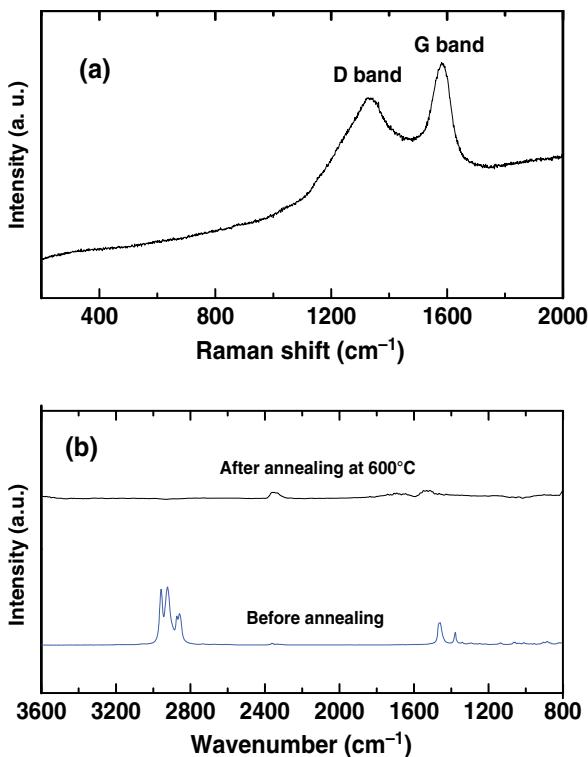


Fig. 6 (a) Schematic diagram for $\text{Sn}_{78}\text{Ge}_{22}$ @carbon nanowires growth from clusters of butyl-capped $\text{Sn}_{78}\text{Ge}_{22}$ nanoparticles and TEM images of the $\text{Sn}_{78}\text{Ge}_{22}$ @carbon nanowires (b) before thermal annealing and (c) after thermal annealing. Insets are electron diffraction patterns of (b) and (c). (d) TEM images of Region 1 in (c). (e) and (f) FE-SEM images of the nanowires. Copyright 2008, American Chemical Society: reproduced with permission from [31]

To investigate the ordering degree of amorphous carbon, Raman scattering of the annealed $\text{Ge}_{78}\text{Sn}_{22}$ nanowire was performed (Fig. 7a) and the dimensional ration of the D and G band was estimated, as the ratio of D/G and was found to be 2.13. The two peaks at ~ 1360 and $\sim 1580 \text{ cm}^{-1}$ are assigned to the (disordered band) D band and (graphene band) G band, respectively [32]. This is far larger than well-ordered graphite at 0.09 [6]. This result implies the formation of amorphous carbon. Figure 7b shows the FT-IR spectra of butyl-capped $\text{Sn}_{78}\text{Ge}_{22}$ particles before and after annealing at 600°C . The three peaks at 2955, 2924, and 2853 cm^{-1} fall where expected for C–H stretches of butyl groups, which is in good agreement with the results of an earlier study [33]. The peaks at 1376 and 1456 cm^{-1} are the positions expected for the symmetric and asymmetric bends of the butyl group, respectively [32]. The nanowire annealed at 600°C for 5 h showed no traces of butyl groups, indicating complete transformation into amorphous carbon. Typically, oxidized SnO_2 or GeO_2 are observed between 800 and 750 cm^{-1} ; these are attributed to metal–O

Fig. 7 Raman scattering of the Sn₇₈Ge₂₂@carbon nanowires after thermal analysis and FT-IR spectra of Sn₇₈Ge₂₂@carbon nanowires before and after thermal annealing



stretch [32]. In this case, no peaks in the range were observed, indicating no oxygen contamination.

Figure 8 shows the first discharge and charge capacities of the nanowires, exhibiting 1247 and 1107 mAh/g, respectively, corresponding to Coulombic efficiency of 88%. This value is an improved value compared to previous capped Sn [23] or Ge [24] nanoparticles that were measured below 85%. Moreover, such a Coulombic efficiency was superior to that of Si nanowires that shows the Coulombic efficiency below 50% [24]. Figure 8b shows that charge capacities at 1, 3, 6, and 8 C were 1107, 1075, 1060, and 1034 mAh/g with a capacity retention at 8 C of 93%.

Thus far, very few studies have reported a high-rate stability of nanoalloy particles. The excellent rate capability of the core-shell nanowires is attributed to the increased electrolyte contact area and the reduced lithium diffusion length that facilitates further Li reactions with the nanowire. For the 0D particles, repetitive particle aggregation and pulverization leads to the formation of a new surface, thus forming a new solid electrolyte interface (SEI) layer. Accordingly, part of the active metal becomes isolated by surface films formed on the fresh active surface by reactions with Li-metal compounds and electrolytes, resulting in rapid capacity fading

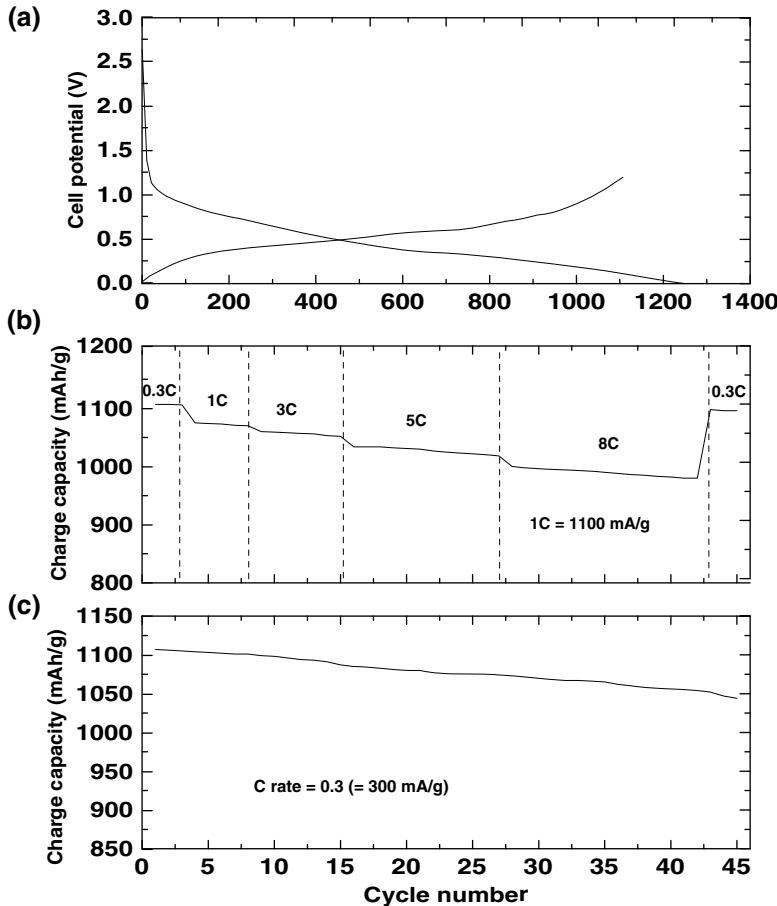


Fig. 8 (a) Rate capabilities of $\text{Sn}_{78}\text{Ge}_{22}$ @carbon nanowires at different C rates from 0.3, 1, 3, 5, 8 C (1 C = 800 mA/g) in a coin-type half-cell. Discharge rate was fixed at 0.3 C. (b) Plot of rate capabilities as a function of charge capacity during extended cycling. (c) Plot of cycle number as a function of the charge capacity of $\text{Sn}_{78}\text{Ge}_{22}$ @carbon nanowires

[34]. However, due to the nature of the nanowire/carbon with the branched morphology and carbon cell, direct contact among the nanotubes can be minimized, and less aggregation is expected. After finishing the rate capability test of the nanowire/carbon nanowire, a morphology change of the nanowire/carbon core shell occurred. As shown in Fig. 9a, the original branched nanowires were separated from the nanoparticles and nanowire. Furthermore, the separated nanowire has stacking faults, and microtwins are visible with their fault plane (101) parallel to the axis of the nanowire (Fig. 9b). These defects are not observed in the pristine nanowires, and may be associated with continuous volume change during the lithium dealloy/alloy

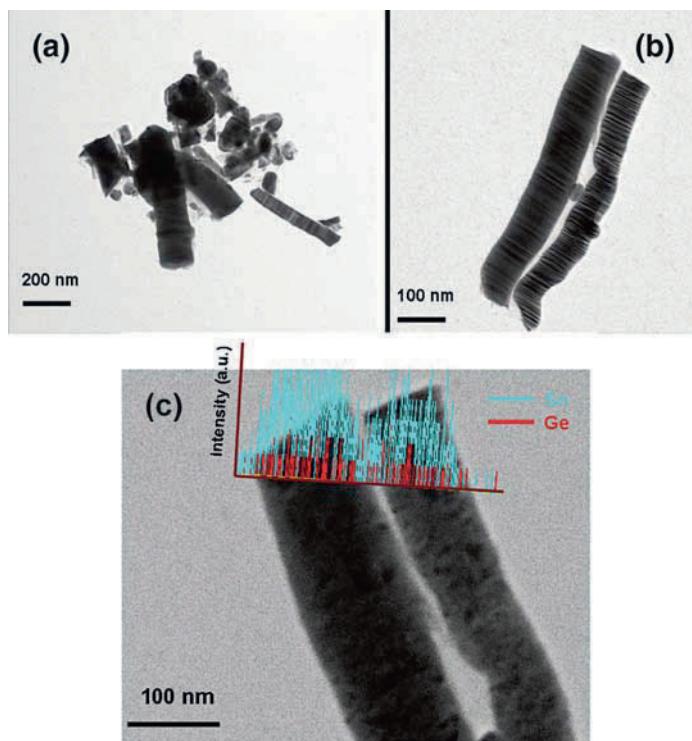


Fig. 9 (a and b) TEM images of $\text{Sn}_{78}\text{Ge}_{22}@\text{carbon}$ nanowires after the rate capability test in Fig. 8b. (c) The EDS (energy-dispersive spectroscopy) line scan across the particle. Copyright 2008, American Chemical Society: reproduced with permission from [31]

process. However, respective Sn and Ge atoms segregation was not observed during cycling, as evidenced by the line mapping of Sn and Ge in the nanowires (Fig. 9c).

4 Layered $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$ Nanowires

Recently, Mn-rich lithium metal oxides such as $\text{Li}[\text{Li}_{1/3-2x/3}\text{Mn}_{2/3-x/3}\text{M}_x]\text{O}_2$ cathodes are currently receiving significant interest as cathode materials for Li-ion batteries. These materials can provide capacities greater than 200 mAh/g if initially charged to 4.5 V or higher [35–41]. In spite of the capacity advantage, a rapid capacity fade occurs at higher C rates; for instance, over 50% capacity fade was observed when the current was increased from 20 to 200 mA/g in $\text{Li}[\text{Ni}_{0.2}\text{Li}_{0.2}\text{Mn}_{0.6}]\text{O}_2$ [42]. However, Cho et al. recently reported that $\text{Li}[\text{Ni}_{0.41}\text{Li}_{0.08}\text{Mn}_{0.51}]\text{O}_2$ nanoplates exhibited greatly improved rate capabilities at a 3 C rate compared with bulk counterpart prepared via a co-precipitation method [43].

For preparing the nanowire cathode, a solution containing 7.2 g of KMnO_4 and 200 ml of distilled water was slowly stirred for 30 min at 40°C and added to 2.1 g of fumaric acid which resulted in a rapid, exothermic reaction, forming a brown gel. This gel was further annealed at 400°C and 700°C for 6 and 12 h, respectively, and the resulting dark black powder was thoroughly washed with water six times, followed by vacuum drying at 200°C overnight. Inductively coupled plasma-mass spectrometry (ICP-MS) analysis of the sample confirmed $\text{K}_{0.3}\text{MnO}_2$. As-prepared K-birnessite was mixed with $\text{Co}(\text{NO}_3)_6 \cdot 6\text{H}_2\text{O}$: $\text{K}_{0.3}\text{MnO}_2$ with a weight ratio of 8:1 in 100 ml of distilled water, and kept in an autoclave for 5 days at 200°C, and finally washed with water six times in order to remove residues that did not participate in the reaction and dissolved K ions. ICP-MS analysis of the sample confirmed formation of $\text{Co}_{0.4}\text{Mn}_{0.6}\text{O}_2$. Finally, $\text{LiNO}_3 \cdot \text{H}_2\text{O}$ and $\text{Co}_{0.40}\text{Mn}_{0.60}\text{O}_2$ were combined with weight ratios of 4:1 were mixed in 100 ml of distilled water, and transferred in an autoclave, maintaining it at 200°C for 2 days. As-prepared powder was rinsed with water, and dried under vacuum at 120°C. Cathodes for battery test cells were made from the active material (~20 mg), super P carbon black (MMM, Belgium), and a polyvinylidene fluoride (PVdF) binder (Kureha Company, Japan) at a weight ratio of 90:5:5. A cathode-slurry was prepared by thoroughly mixing an *N*-methyl-2-pyrrolidene (NMP) solution with the PVdF, the carbon black and the powdery cathode-active material. The electrodes were prepared by coating the cathode-slurry onto an Al foil, followed by drying at 130°C for 20 min. Coin-type battery-test cells (size 2016) containing a cathode, a Li metal anode, and a micro-porous polyethylene separator were prepared in a helium-filled glove box. The electrolyte used was 1 M LiPF_6 with ethylene carbonate/diethylene carbonate/ethyl-methyl carbonate (EC/DEC/EMC) (30:30:40 vol.%) (Cheil Industries, Korea). After addition of the electrolyte, the test cells were aged at room temperature for 24 h before commencing the electrochemical tests.

Figure 10a shows an XRD pattern of the as-prepared $\text{Co}_{0.4}\text{Mn}_{0.6}\text{O}_2$ and $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$. The XRD pattern of the $\text{Co}_{0.40}\text{Mn}_{0.60}\text{O}_2$ is dominated by two major peaks at about 12° (110) and about 25° (200) scattering angles, which is similar to that of previously reported $\text{Ni}_{0.45}\text{Mn}_{0.55}\text{O}_2$ obtained from *K*-birnessite [43]. After a reaction with Li at 200°C, the $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$ sample clearly shows the formation of a hexagonal layered structure with a R-3m space group. The lattice constants of *a* and *c* were estimated as 2.834 Å and 14.208 Å, respectively. Very weak superlattice reflections at approximately 20° and 24° are known to correspond to the ordering of the Li and Co and Mn ions in transition metal sites of the layered lattice [37,38]. Moreover, preferred orientation along the (003) plane was observed. The ICP-MS results showed the formation of $\text{Li}_{1.3}\text{Co}_{0.4}\text{Mn}_{0.6}\text{O}_{4.9}$; when normalized to the LiMnO_2 , the compound had a formula $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$.

Figure 11 shows SEM and TEM images of the as-prepared $\text{Co}_{0.4}\text{Mn}_{0.6}\text{O}_2$ nanowires (a) and $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$ nanowires (b, d, and e), and $\text{Co}_{0.4}\text{Mn}_{0.6}\text{O}_2$ nanowires have a diameter of 40 nm and are >1 μm in length. After reaction with the Li nitrate, the dimension of the nanowires increases. The $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$ nanowires have a diameter of 100 nm and a length of

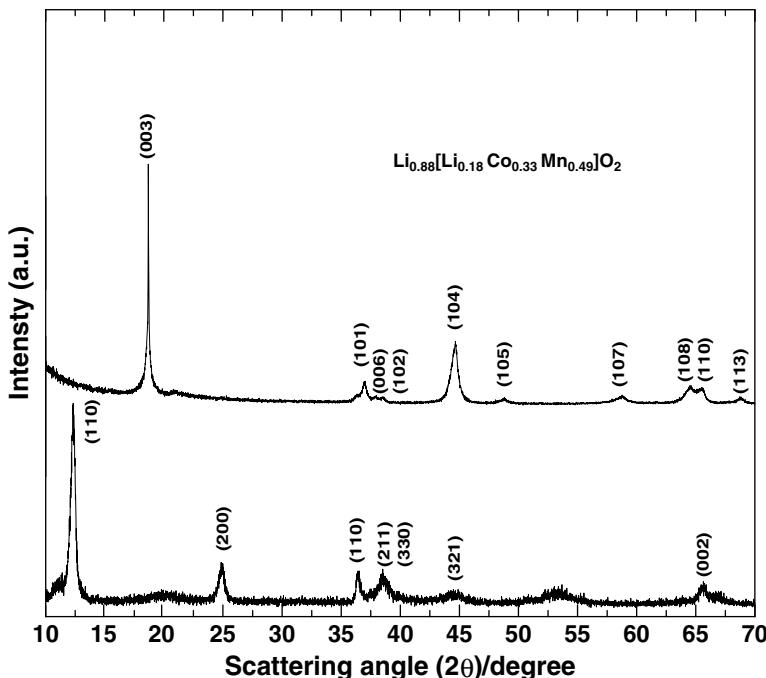


Fig. 10 Powder XRD patterns of $\text{Co}_{0.4}\text{Mn}_{0.6}\text{O}_2$ and $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}] \text{O}_2$ nanowires

$>3 \mu\text{m}$. An inset of Fig. 10b is an enlarged image of the edge parts of the nanowires, confirming that the particles consisted of aggregated individual nanowires. Energy-dispersive X-ray spectrometry (EDS) of the $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}] \text{O}_2$ nanowires shows both the presence of Co and Mn. The ICP-MS results confirmed stoichiometry of $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}] \text{O}_2$. The TEM image of Fig. 11d shows that the diameter of the nanowires is about 50 nm. An HREM image of the nanowires (image e) clearly shows the formation of the layered structure with the lattice fringe of the (003) plane corresponding to 0.46 nm. The Brunauer–Emmett–Teller (BET) surface area of the $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}] \text{O}_2$ nanowires was $50 \text{ m}^2/\text{g}$.

Figure 12 shows the charge and discharge curves between 4.8 and 2 V at rates of 0.2, 1, 3, 5, 10, and 15 C (1 C = 240 mA/g), corresponding to capacities of 245, 242, 238, 230, 225, and 220 mAh/g, respectively (charge rate was fixed at 1 C). In general, the loading level of the active cathode material critically affects the performance of the cathode. In this case, a cathode material loading level of 20 mg/cm^2 was used.

After the first charge process, the plateau vanished in the subsequent cycles. Although a similar irreversible voltage plateau during the first charge process has been reported with Li_2MnO_3 -based oxides, Li_2MnO_3 -based oxides showed a distinct flat plateau region at $\sim 4.5 \text{ V}$ [45]. In contrast to this,

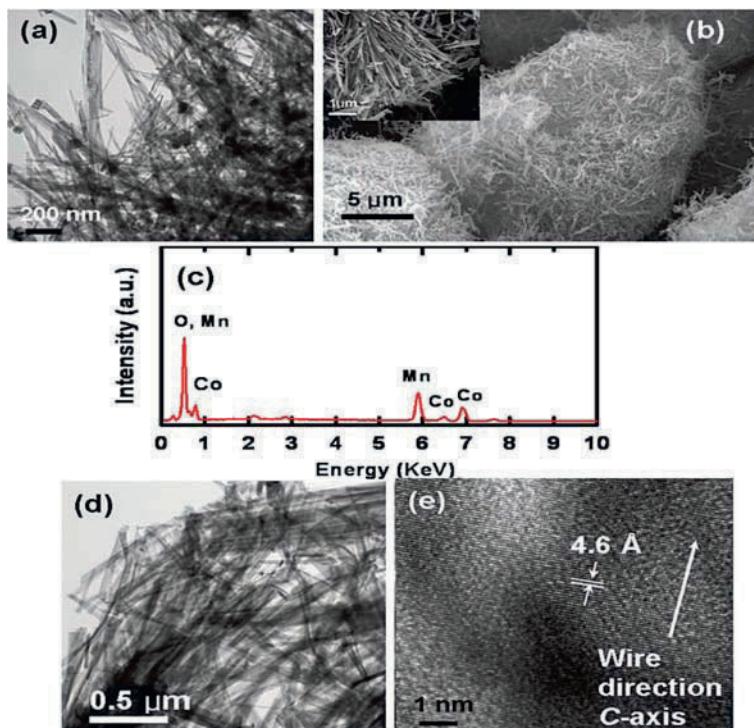


Fig. 11 TEM images of (a) $\text{Co}_{0.4}\text{Mn}_{0.6}\text{O}_2$ nanowires, and (b and d) SEM and TEM images of the $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$ nanowires, respectively. An inset of 2b is an expanded image of 2b. (c) Corresponding EDS spectrum of image b and (e) high-resolution TEM image of a single nanowire of (d). Copyright 2008, American Chemical Society: reproduced with permission from [44]

$\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$ has an inclined plateau at ~ 4.6 V. The presence of the inclined plateau at 4.5 V (as shown in Fig. 12a) instead of the flat plateau in $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$ nanowires is also observed in $\text{Li}[\text{Co}_{x}\text{Li}_{1/3-x/3}\text{Mn}_{2-x/3}]\text{O}_2$ and $\text{Li}[\text{Cr}_{x}\text{Li}_{1/3-x/3}\text{Mn}_{2-x/3}]\text{O}_2$, in which the flat plateau turn into a inclined plateau with increasing Co or Cr contents [39,40]. This is because more transition metals (Co, Cr) are available to be involved in the redox reaction, and oxygen loss will not be required for samples with higher transition metal content to remove all the Li from the Li layer [40].

During the first cycle, the charge and discharge capacities are 283 and 245 mAh/g, respectively, showing 87% columbic efficiency. After 50 cycles at a rate of 1 C, capacity retention was 92%. In particular, the capacity retention at a 15 C rate was 90%, compared to that at a 0.2 C rate. The capacity retention of the nanowires even after repetitive cycling at a 15 C rate was 98%. The rate capability of the cathode with the irregularly distributed particles was determined by the larger particles, and therefore it is hampered by the larger particles with longer Li ion diffusion distance. For instance, $\text{Li}[\text{Ni}_{0.20}\text{Li}_{0.2}\text{Mn}_{0.6}]\text{O}_2$ nanoparticles with

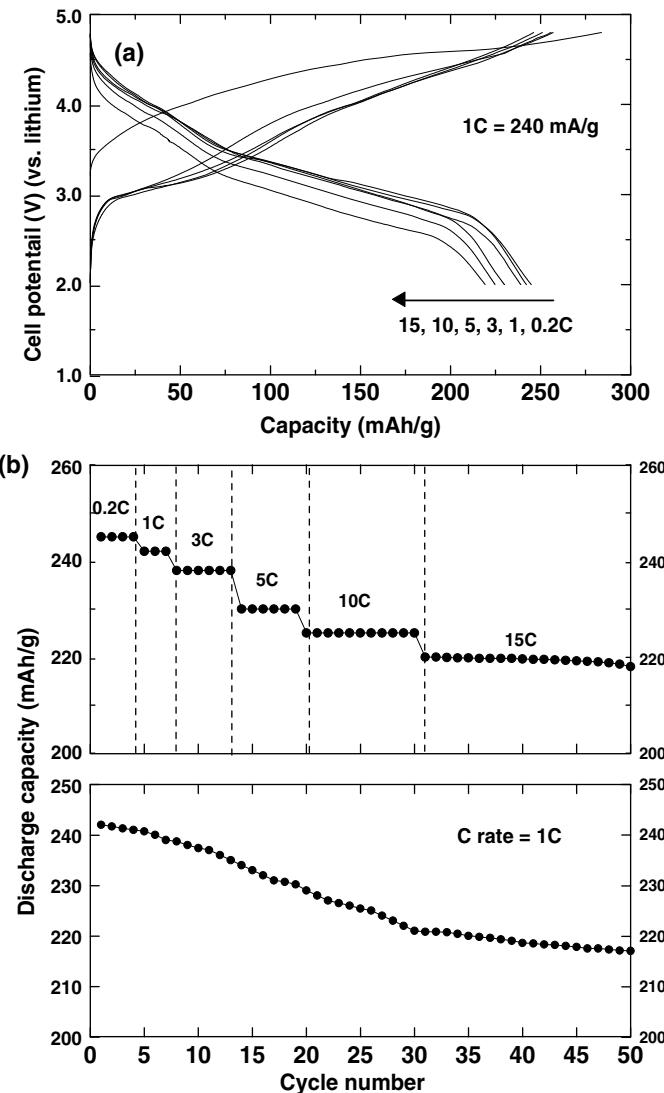


Fig. 12 (a) Voltage curves of $\text{Li}_{0.88}[\text{Li}_{0.18}\text{Co}_{0.33}\text{Mn}_{0.49}]\text{O}_2$ nanowires in a coin-type half-cell at different C rates of 0.2, 1, 3, 5, 10, and 15 C ($1\text{C} = 240 \text{ mA/g}$) between 2 and 4.8 V. The charge rate was fixed at 1 C, and the discharge capacities vs. cycle number were tested at different C rates. (b) Discharge capacity vs. cycle number of the $\text{Li}[\text{Li}_{0.08}\text{Co}_{0.33}\text{Mn}_{0.5}]\text{O}_2$ nanowires in a coin-type half-cell at a 1 C rate between 2 and 4.8 V

the primary particles were narrowly distributed in the range of 80–200 nm, and with large aggregates of 1–20 μm [42] showed the first discharge capacity of ~ 280 mAh/g at 20 mA/g but its discharge capacity significantly decreased to 210 mAh/g at 400 mA/g. However, our nanowires has a uniform diameter of about 50 nm

and shows first discharge capacity of 245 mAh/g at 48 mA/g and its capacity is 220 mAh/g at 3000 mA/g. The capacity retention of the Li[Ni_{0.41}Li_{0.08}Mn_{0.51}]O₂ nanoparticle with 20–200 nm sizes (BET surface area was 24 m²/g) at a current rate of 1200 mA/g rate is 89%, while that of the nanowire is 94% at the same current rate. Even though 6% difference is small at lower current rates, higher current rates can make a huge difference. Accordingly, the capacity retention of the nanowires is expected to be much higher than that of the nanoparticles at higher current rate than 1200 mA/g. This indicates that the nanowires with higher surface areas likely lead to higher electrode/electrolyte contact areas. This coupled with shorter diffusion paths into the lattices and reduced strain of intercalation may lead to improved cycling performance compared to nanoparticles [46].

5 Conclusions

We have reviewed 1D anode and cathode materials to obtain high surface area of the nanostructured electrode, which makes the current density more effective than for the conventional bulk electrode. The high polarization was believed to result from slow ionic and electronic diffusions in the active material or interfaces, but manipulation of the morphology provided versatile strategies toward improving the electrochemical properties. Better rate capabilities were obtained as there was a decrease in diffusion distance between Li⁺ and electrons within the nanostructured electrode.

Acknowledgments This research was supported by the converging Research center program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology.

References

1. X. Peng, L. Manna, W. Yang, J. Wickham, E. Scher, A. Kadavanich, *Nature* **2000**, 404, 59.
2. Y. Jun, S.-M. Lee, N.-J. Kang, J. Cheon, *J. Am. Chem. Soc.* **2001**, 123, 5150.
3. L. Manna, E. C. Scher, A. P. Alivisatos, *J. Am. Chem. Soc.* **2000**, 122, 12700.
4. Y.-W. Jun, J.-S. Choi, J. Cheon, *Angew. Chem. Int. Ed.* **2006**, 45, 3414.
5. Z. Liu, D. Zhang, S. Han, C. Li, T. Tang, W. Jin, X. Liu, B. Lei, C. Zhou, *Adv. Mater.* **2003**, 15, 1754.
6. Y. Liu, J. Dong, M. Liu, *Adv. Mater.* **2004**, 16, 353.
7. Z. R. Dai, Z. W. Pan, Z. L. Wang, *J. Am. Chem. Soc.* **2002**, 124, 8673.
8. J. Q. Hu, X. L. Ma, N. G. Shang, Z. Y. Xie, N. B. Wong, C. S. Lee, S. T. Lee, *J. Phys. Chem. B* **2002**, 106, 3823.
9. Y. Wang, J. Y. Lee, *J. Phys. Chem. B* **2004**, 108, 17832; Y. Wang, H. C. Zeng, J. Y. Lee, *Adv. Mater.* **2006**, 18, 645; Y. Wang, J. Y. Lee, H. C. Zeng, *Chem. Mater.* **2005**, 17, 3899.
10. Y. Wang, F. Su, J. Y. Lee, X. S. Zhao, *Chem. Mater.* **2006**, 18, 1347; X. W. Lou, Y. Wang, C. Yuan, J. Y. Lee, L. A. Archer, *Adv. Mater.* **2006**, 18, 2325; S. Han, B. Jang, T. Kim, S. M. Oh, T. Hyeon, *Adv. Funct. Mater.* **2005**, 15, 1845.
11. M.-S. Park, G.-X. Wang, Y.-M. Kang, D. Wexler, S.-X. Dou, H.-K. Liu, *Angew. Chem. Int. Ed.* **2007**, 46, 750.
12. F. Chen, M. Liu, *Chem. Commun.* **1999**, 1829.
13. D.-F. Zhang, L.-D. Sun, J.-L. Yin, C.-H. Yan, *Adv. Mater.* **2003**, 12, 1022.

14. N. Ulagappan, C. N. R. Rao, *Chem. Commun.* **1996**, 1685; K. G. Severin, T. M. Abdel-Fattah, T. J. Pinnavaia, *Chem. Commun.* **1998**, 1471.
15. S. C. Laha, R. Ryoo, *Chem. Commun.* **2003**, 2138; K. P. Gierszal, T.-W. Kim, R. Ryoo, M. Jaroniec, *J. Phys. Chem. B* **2005**, 109, 8774.
16. H. Kim, J. Cho, *J. Mater. Chem.* **2008**, 18, 771.
17. M.-S. Park, G.-X. Wang, Y.-M. Kang, D. Wexler, S.-X. Dou, H.-K. Liu, *Angew. Chem. Int. Ed.* **2007**, 46, 750.
18. I. A. Courtney, J. R. Dahn, *J. Electrochem. Soc.* **1997**, 144, 2045; I. A. Courtney, J. S. Tse, O. Mao, J. Hafner, J. R. Dahn, *Phys. Rev. B* **1998**, 58, 15583.
19. G.-A. Nazri, G. Pistoia, *Lithium Batteries Science and Technology*, Kluwer, Boston, **2004**.
20. J. O. Besenhard, J. Yang, M. Winter, *J. Power Sources*, **1997**, 68, 87.
21. H. C. Shin, M. Liu, *Adv. Funct. Mater.* **2005**, 15, 582; Y. Xia, T. Sakai, M. Wada, H. Yoshinaga, *J. Electrochem. Soc.* **2001**, 148, 471; S. D. Beattie, J. R. Dahn, *J. Electrochem. Soc.* **2003**, 150, A894; O. Mao, J. R. Dahn, *J. Electrochem. Soc.* **1999**, 146, 414.
22. I.-S. Kim, G. E. Blomgren, P. N. Kumta, *Electrochem. Solid-State Lett.* **2004**, 7, A44; G. X. Wang, J. Yao, H. K. Liu, *Electrochem. Solid-State Lett.* **2004**, 7, A250; L. K. Lee, Y. S. Jung, S. M. Oh, *J. Am. Chem. Soc.* **2003**, 125, 5652; M. Noh, Y. Kim, M. G. Kim, H. Lee, Y. Kwon, Y. Lee, J. Cho, *Chem. Mater.* **2005**, 17, 3320.
23. Y. Kwon, M. G. Kim, Y. Kim, Y. Lee, J. Cho, *J. Electrochem. Solid-State Lett.* **2006**, 9, A34; M. Noh, Y. Kwon, H. Lee, J. Cho, Y. Kim, M. G. Kim, *Chem. Mater.* **2005**, 17, 1926; H. Kim, J. Cho, *J. Electrochem. Soc.* **2007**, 154, A462; H. Kim, J. Cho, *Electrochim. Acta* **2007**, 52, 4197.
24. B. Gao, S. Sinha, L. Fleming, O. Zhou, *Adv. Mater.* **2001**, 13, 816.
25. F. J. Himpsel, T. Jung, A. Kirakosian, J. L. Lin, D. Y. Petrovykh, H. Rausher, J. Viernow, *MRS Bull.* **1999**, 24, 20.
26. R. S. Wagner, W. C. Ellis, *Appl. Phys. Lett.* **1964**, 4, 89; G. Cao, *Nanostructure and Nanomaterials*, Imperial College Press, London, **2004**.
27. T. J. Trentler, K. M. Kickman, S. C. Goel, S. M. Viano, P. C. Gibbons, W. E. Buhro, *Science*, **1995**, 270, 1791.
28. P. Yang, C. M. Lieber, *Science*, **1996**, 273, 1836; P. Yang, C. M. Lieber, *J. Mater. Res.* **1997**, 12, 2981.
29. H. Lee, H. Kim, S.-G. Doo, J. Cho, *J. Electrochem. Soc.* **2007**, 154, A343.
30. Y. Kwon, H. Kim, S.-G. Doo, J. Cho, *Chem. Mater.* **2007**, 19, 982.
31. H. Lee, J. Cho, *Nano Lett.* **2007**, 7, 2638.
32. C.-S. Yang, Q. Li, S. M. Kauzlarich, B. Phillips, *Chem. Mater.* **2000**, 12, 983.
33. A. C. Ferrai, R. Roberston, *J. Phys. Chem. B* **2000**, 61, 14095.
34. D. Aurbach, A. Nimberger, B. Markosky, E. Levi, E. Sominski, A. Gedanken, *Chem. Mater.* **2002**, 14, 4155.
35. Z. Lu, J. R. Dahn, *J. Electrochem. Soc.* **2002**, 149, A815.
36. A. R. Armstrong, M. Holzapfel, P. Novak, C. S. Johnson, S.-H. Kang, M. M. Thackeray, P. G. Bruce, *J. Am. Chem. Soc.* **2006**, 128, 8694.
37. S.-H. Park, S.-H. Kang, C. S. Johnson, K. Amine, M. M. Thackeray, *Electrochem. Commun.* **2007**, 9, 262.
38. M. M. Thackeray, C. S. Johnson, J. T. Vaughey, N. Li, S. A. Hackney, *J. Mater. Chem.* **2005**, 15, 2257.
39. C. S. Johnson, N. Li, C. Lefief, M. M. Thackeray, *Electrochem. Commun.* **2007**, 9, 787.
40. Y.-S. Hong, Y. J. Park, X. Wu, K. S. Ryu, S. H. Chang, *Electrochem. Solid State Lett.* **2003**, 6, A166.
41. B. Ammundsen, J. Paulsen, *J. Adv. Mater.* **2001**, 13, 943.
42. Y.-S. Hong, Y. J. Park, K. S. Ryu, S. H. Chang, M. G. Kim, *J. Mater. Chem.* **2004**, 14, 1424.
43. J. Cho, Y. Kim, M. G. Kim, *J. Phys. Chem. C* **2007**, 111, 1186.
44. Y. Lee, M. G. Kim, J. Cho, *Nano Lett.* **2008**, 8, 957.
45. Y. J. Park, Y.-S. Hong, X. Wu, M. G. Kim, K. S. Ryu, S. H. Chang, *J. Electrochem. Soc.* **2004**, 151, A720.
46. C. Kim, M. Noh, M. Choi, J. Cho, B. Park, *Chem. Mater.* **2005**, 17, 3297.

Mechanical Energy Harvesting Using Wurtzite Nanowires

Xudong Wang and Zhong Lin Wang

Abstract This chapter reviews a novel technology that uses piezoelectric nanowires to harvest nanoscale mechanical energy. Wurtzite ZnO nanowires were first applied for this purpose. Due to the bending of nanowires, the piezoelectric effect induces positive and negative potentials on the stretched surface and compressed surface, respectively. This effect has been proved by using a conductive AFM tip to deflect individual wurtzite ZnO and CdS nanowires as well as vertically aligned ZnO nanowire arrays. On the basis of this phenomenon, a prototype nanodevice (nanogenerator) has been successfully developed to convert ultrasonic wave energy into electricity. After optimization, the nanogenerator's output has reached ~ 83 nW/cm², which shows a great potential to power nanosensors. A flexible substrate and textile fibers were also integrated with the ZnO nanowire array for harvesting energy from low-frequency perturbations. At the end of this chapter, several pathways are suggested for further performance improvements.

Keywords ZnO · Nanowires · Piezoelectricity · Nanopiezotronics · Nanogenerator

1 Introduction

Facing current global energy shortage, seeking for renewable energies is essential for releasing the heavy reliance on mineral based energy and improving the cleanliness of our environment. Solar, mechanical (mostly wind), and bioenergies are common substitutions and have been more and more widely used. Meanwhile, scientific research is intensively focused on these areas for improving the energy efficiency and lowering the cost. Based on the nature of those energy sources, advantages and limitations are obvious for each of them. The type of energy source to be

X. Wang (✉)

Department of Materials Science and Engineering, University of Wisconsin-Madison, Madison, WI, USA

e-mail: xudong@engr.wisc.edu

chosen has to be decided on the basis of specific applications. For instance, the supply of solar energy is endless, while it does not apply to dark areas or under-water/underground. Mechanical energy more universally exists on the earth but mostly in a lower energy level, such as friction, body movement, wave, and acoustic. Considering small electronics or advanced sensing systems that do not have high energy requirement, the mechanical energy would be a perfect power source for them to operate in a “self-sufficient” mode. Small-level mechanical energy harvesting was typically realized using piezoelectric materials, which generates electric potential when subject to mechanical deformation. However, due to the limitation of micro-fabrication, its power density is rather low and a big mass is usually required for improving the sensitivity to low-level vibrations. Using piezoelectric wurtzite ZnO nanowire (NW) arrays, we have developed a novel technology for converting nanoscale mechanical energy into electric energy. The operation mechanism of the electric generator relies on the unique coupling of piezoelectric and semiconducting dual properties of ZnO as well as the elegant rectifying function of the Schottky barrier formed between the metal electrode and the NW. The nanodevice developed under this concept is named as *nanogenerator* (NG). The details of this development and corresponding mechanisms will be addressed in this chapter.

2 Mechanical to Electrical Energy Conversion by Wurtzite Nanowires

2.1 Signal ZnO Wires

The mechanical energy to electrical energy conversion process was first revealed on a ZnO wire under atomic force microscope (AFM) by directly correlating the manipulation process of the wire and the electric output voltage/current signal [1]. In the experiment, a long ZnO wire that was large enough to be seen under an optical microscope was used. One end of the ZnO wire was affixed on a silicon substrate by silver paste, while the other end was left free. The substrate was intrinsic silicon, so its conductivity was rather poor. The wire was laid on the substrate but kept at a small distance from the substrate to eliminate the friction with the substrate except at the affixed side (Fig. 1a). The measurements were performed by AFM using a Si tip coated with Pt film coating, which had a tetrahedral shape with an apex angle of 70° and height of 14 μm and a spring constant of 1.42 N/m. The measurements were done in AFM contact mode under a constant normal force of 5 nN between the tip and sample surface, and the scan area was $70 \times 70 \mu\text{m}^2$.

Both the topography (feedback signal from the scanner) and the corresponding output voltage (V) images across a load were recorded simultaneously when the AFM tip was scanned across a wire/belt. The topography image reflects the change in normal force perpendicular to the substrate, which shows a bump only when the tip scans over the wire. The output voltage between the conductive tip and the ground was continuously monitored as the tip scans over the wire. No external

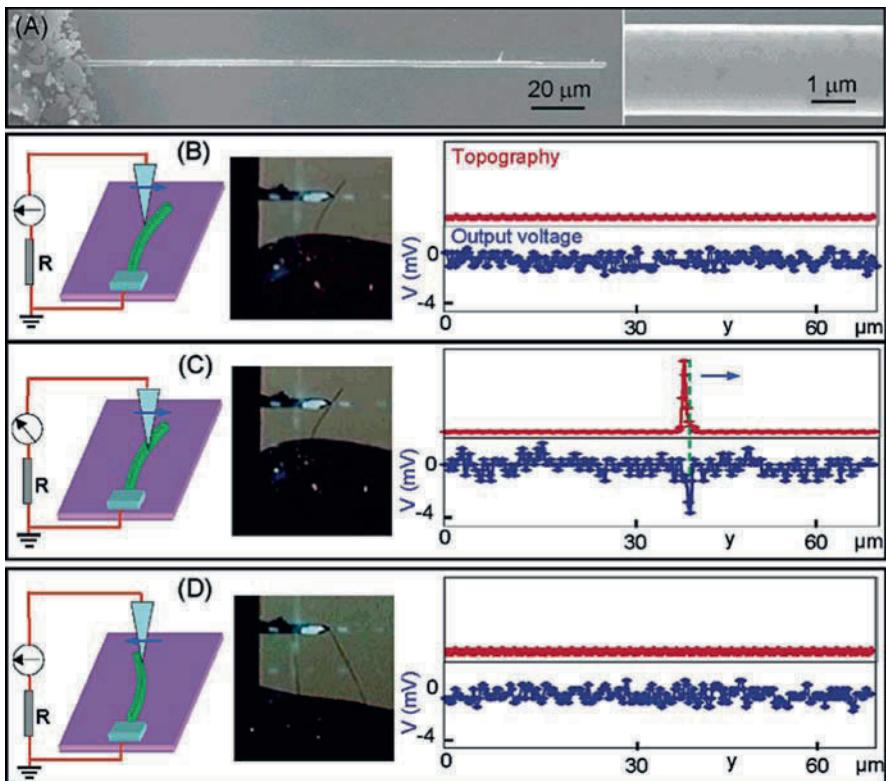


Fig. 1 In situ observation of the process for converting mechanical energy into electric energy by a piezoelectric ZnO belt. **(a)** SEM images of a ZnO belt with one end affixed by silver paste onto a silicon substrate and the other end free. **(b, c, and d)** Three characteristic snapshots and the corresponding topography (*top curve*) and output voltage (*bottom curve*) images when the tip scanned across the middle section of the belt. The schematic illustration of the experimental condition is shown at the left-hand side, with the scanning direction of the tip indicated by an *arrowhead*

voltage was applied in any stage of the experiment. The output data to be presented were obtained from snapshots of the characteristic events observed during the experiments.

The AFM tip scanned line by line at a speed of 105.57 $\mu\text{m/s}$ perpendicular to the wire and pushed the wire at its middle section. When subjected to a displacement force, one side of the wire is stretched (left-hand side) and the other side (the right-hand side) is compressed. The objective of this experiment is to observe if the piezoelectric discharge occurs when the tip touches either the stretched side or the compressed side, or both sides. The topography image directly captured whether the tip passed over the belt or not because it was a representation of the normal height received by the cantilever. When the tip pushed the wire but did not go over and across it, as judged by the flat output signal in the topography image (Fig. 1b), no voltage output was produced, indicating the stretched side produced no piezoelectric

discharge event. Once the tip goes over the belt and touches the compressed side, as indicated by a peak in the topography image, a sharp voltage output peak is observed (Fig. 1c). By analyzing the positions of the peaks observed in the topography image and the output voltage image, we noticed that the discharge occurred after the tip nearly finished crossing the wire. This clearly indicates that the compressed side was responsible for producing the negative piezoelectric discharge voltage. For another wire, when the tip retracted from the right-hand side to the left-hand side (Fig. 1d), no output voltage was detected because the tip just touched the stretched side of the belt without crossing it.

As a summary of the above experimental observations, there are three key results. First, piezoelectric discharge is observed only when the AFM tip touches the end of the bent wire. Second, the piezoelectric discharge occurs only when the AFM tip touches the compressed side of the wire, and there is no voltage output if the tip touches the stretched side of the wire. Third, the piezoelectric discharge gives negative output voltage as measured from the load R_L . Finally, in reference to the topography image, the voltage output event occurs when the AFM tip has almost finished crossing the width of the wire/belt, which means that the discharge event is delayed to the last half of the contact between the tip and the wire.

2.2 Aligned ZnO Nanowire Arrays

Vertically aligned ZnO NW array is a perfect self-assembled nanostructure that allows series of macroscopic manipulation and operation. Aligned ZnO NWs used in the experiments were grown on $a\text{-Al}_2\text{O}_3$ substrate using Au particles as a catalyst, by the vapor–liquid–solid (VLS) process [2]. An epitaxial relation between ZnO and $a\text{-Al}_2\text{O}_3$ allows a thin, continuous layer of ZnO to form at the substrate, which serves as a large electrode connecting the NWs with a metal electrode for transport measurement (Fig. 2a) [3]. The NW grows along the [0001] direction and has side surfaces of {0110} (Fig. 2b). Most of the Au particles at the tips of the NWs either evaporate during the growth or fall off during scanning by the AFM tip. For most of the NWs, the growth front is free of Au particles or has a small hemispherical Au particle that covers only a fraction of the top (inset, Fig. 2b). For the purpose of our measurements, we have grown NW arrays that have relatively less density and shorter length (0.2–0.5 mm), so that the AFM tip can exclusively reach one NW without touching another.

Similar measurements were performed by AFM using a Si tip coated with Pt film [4]. The rectangular cantilever had a calibrated normal spring constant of 0.76 N/m (Fig. 2c). In the AFM contact mode, a constant normal force of 5 nN was maintained between the tip and the sample surface. The tip scanned over the top of the ZnO NW, and the tip's height was adjusted according to the surface morphology and local contacting force. The thermal vibration of the NWs at room temperature was negligible. For the electric contact at the bottom of the nanowires, silver paste (uncured) was applied to connect the (large) ZnO film on the substrate surface with the measurement circuit. The output voltage across an outside load of resistance

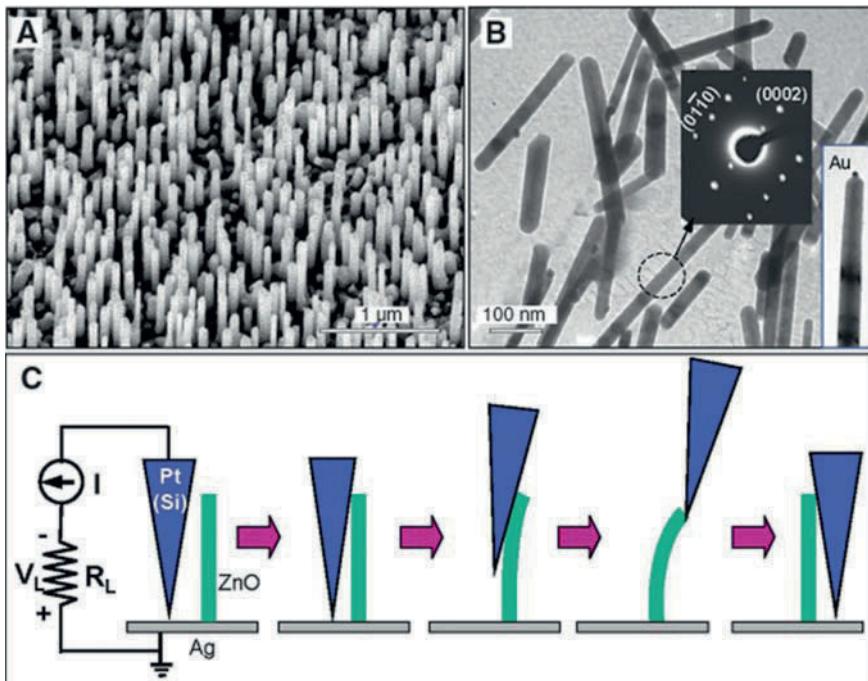


Fig. 2 Experimental design for converting nanoscale mechanical energy into electrical energy by a vertical ZnO NW. (a) SEM images of aligned ZnO NWs grown on $\alpha\text{-Al}_2\text{O}_3$ substrate. (b) TEM images of ZnO NWs, showing the typical structure of the NW without an Au particle or with a small Au particle at the top. Inset at center: an electron diffraction pattern from a NW. Inset at right: image of a NW with an Au particle. (c) Experimental setup and procedures for generating electricity by deforming a PZ NW with a conductive AFM tip

$R_L = 500$ megohms was continuously monitored as the tip scanned over the nanowires (note the defined polarity of the voltage signal). No external voltage was applied in any stage of the experiment.

Experimentally, both the topography (feedback signal from the scanner) (Fig. 2a) and the corresponding output voltage (V_L) images across the load (Fig. 2b) were recorded simultaneously when the AFM tip scanned over the aligned NW arrays. In contact mode, as the tip scanned over the vertically aligned NWs, the NWs were bent consecutively. The bending distance was directly recorded in the topography image, from which the maximum bending deflection distance and the elastic modulus of the NW as well as the density of NWs that have been scanned by the tip were directly derived [5].

In the V_L image, many sharp output peaks (like discharge peaks) were observed. These peaks, typically about 4–50 times the noise level, are rather sharp and narrow, and sometimes one or two pixels represent one voltage peak because of the limited scanning speed of the AFM tip, so that the color distribution in the plot is not easy to display. By reducing the scan range and increasing the scan frequency, more

complete profiles of the discharge peaks were captured. Most of the voltage peaks are $\sim 6\text{--}9\text{ mV}$ in height. The density of NWs contacted by the tip is counted from Fig. 3a to be $\sim 20\text{ }\mu\text{m}^{-2}$, and the average density of NWs whose voltage output events had been captured by the tip in Fig. 3b is $\sim 8\text{ }\mu\text{m}^{-2}$, thus $\sim 40\%$ of the NWs were contacted.

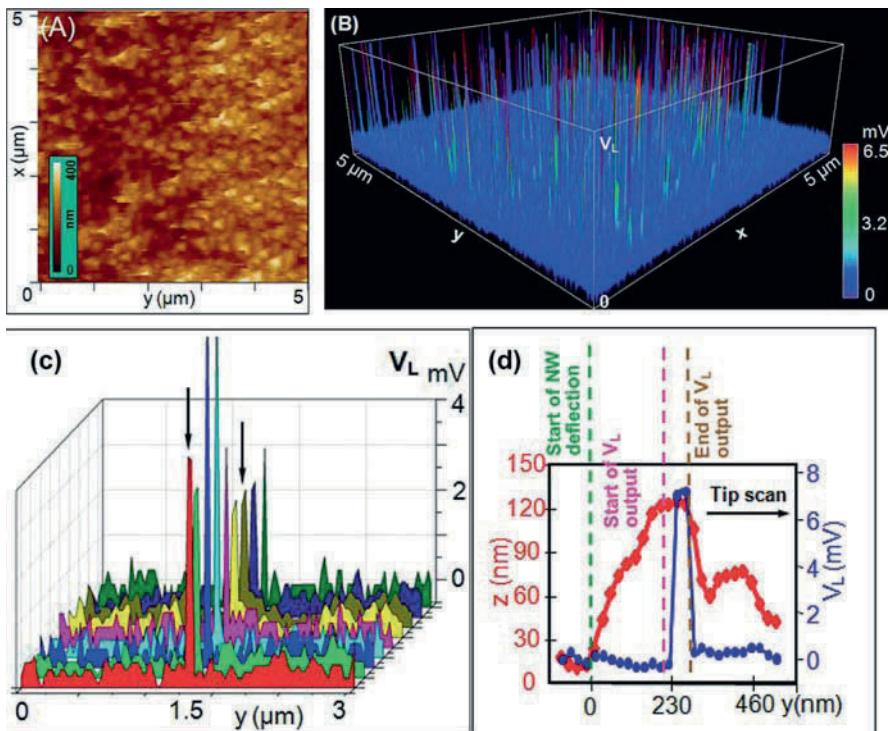


Fig. 3 Electromechanically coupled discharging process of aligned piezoelectric ZnO NWs observed in contact mode. Topography image (a) and corresponding output voltage image (b) of the NW arrays. (c) A series of line profiles of the voltage output signal when the AFM tip scanned across a vertical NW at a time interval of 1 min. (d) Line profiles from the topography (diamond marks) and output voltage (circle marks) images across a NW

The location of the voltage peak is directly registered at the site of the NW. A time series of the voltage output line profiles across one NW acquired at a time interval of 1 min is shown in Fig. 3c. Because the dwell time for each data point (or pixel) is 2 ms, which is longer than the average lifetime of the voltage peak of $\sim 0.6\text{ ms}$ (Fig. 3d), the peak at which V_L reached the maximum was possibly missed by the “slow” scanning tip, so that V_L shows a chopped top (arrows in Fig. 3c). A sharp peak can be identified continuously at the location of the NW and in the NW output voltage in each scan of the tip. When the tip started to deflect the NW, no voltage output was observed (Fig. 3d); V_L was detected when the deflection of the NW approached its maximum. When the NW was released by the AFM tip, V_L

dropped to zero, indicating that the output of piezoelectricity was detected toward the end of the AFM scan over the NW.

2.3 CdS Nanowires

As another semiconductor material with wurtzite crystal structure, CdS also exhibited the ability for mechanical energy conversion [6]. Vertically CdS NW arrays were grown through a hydrothermal process following the procedures reported in literature with minor modifications [7]. The morphology of the hydrothermally grown CdS NWs is shown by a SEM image (Fig. 4a). These NWs were 150 nm in diameter and several micrometers in length. The x-ray diffraction (XRD) pattern of the as-grown CdS NWs is shown in Fig. 4b. Also included in Fig. 4b are the XRD patterns of the reference cubic zinc blende (ZB) and hexagonal wurtzite (WZ) phases of CdS crystals for comparison purposes. All of the diffraction peaks, except the one located at 26.4° , can be indexed to the WZ phase. The diffraction peak at 2θ of 26.4° is a combined contribution from WZ (0002) and ZB (111) planes. The coexistence of the two phases can be clearly seen from the high-resolution transmission electron microscopy (HRTEM) image shown in the inset of Fig. 4a. The NW is composed of alternating ZB and WZ phases, with alternating atomic layer stacking sequence from ABC to AB along the growth direction of the NWs. The growth direction was identified to be along ZB (111) and WZ <0001>.

The energy conversion experiments were conducted through the same manner as the measurements on ZnO NW arrays presented in Section 2.2. Similarly, the process for generating the electric current can be derived from a detailed analysis of the output voltage peak and the topological profile received by the tip when scanned across a NW. During the tip scans, no voltage output signals were observed if the tip touched only the stretched side of the NW and did not lift up to go beyond the central line of the NW to reach the compressed side. A negative voltage output signal was detected when the tip went beyond the NW to reach the compressed side of the NW, as shown in the topography and output voltage images of Fig. 4c. This is clearly indicated by a delayed output in voltage signal in reference to the surface profile image of the NW. The typical magnitude of the output voltage was $\simeq(0.5\text{--}1)$ mV, where the negative sign means that the generated current flowed from the tip to the NW. The low-voltage output is likely to be resulting from the phase transition between ZB and WZ phases along the growth direction of the NW, because WZ phase is piezoactive, while ZB phase is not. However, the presence of the ZB phase along the growth direction is disadvantageous to the piezoelectronic performance of the CdS NWs.

3 Energy Conversion Mechanisms

The mechanical to electrical energy conversion phenomenon relies on the piezoelectric and semiconducting coupled property of wurtzite crystals, which controls the charge creating, preserving, and releasing processes [8]. For a vertical, straight

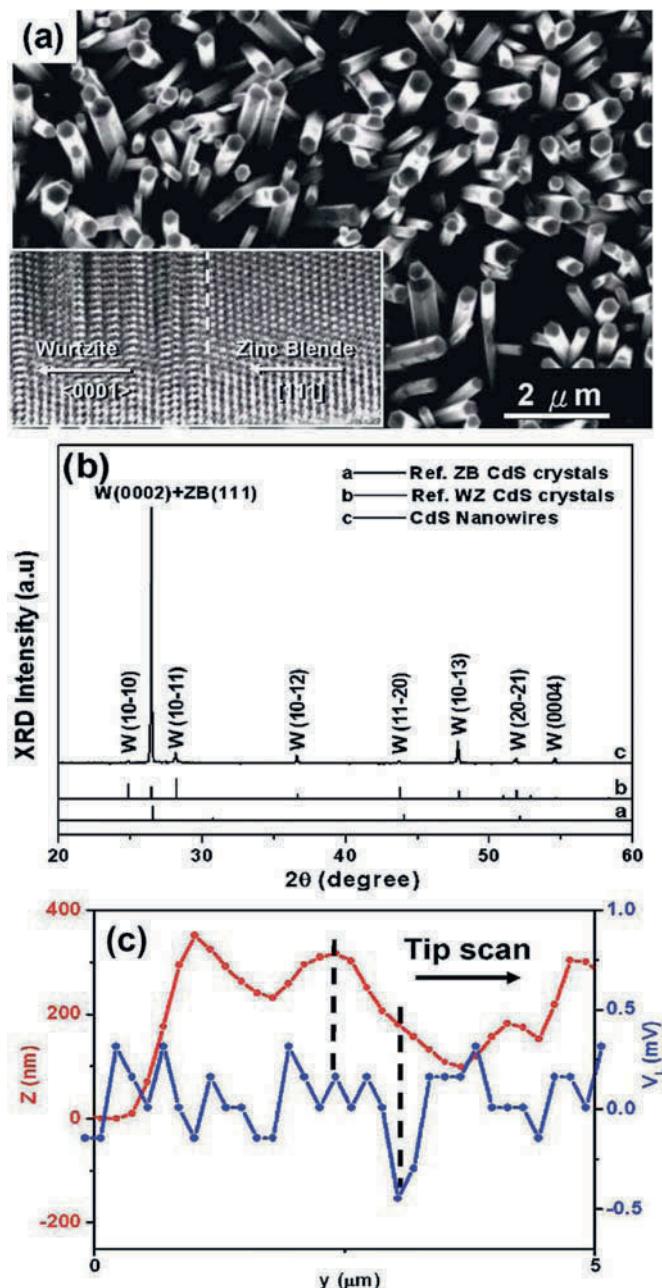


Fig. 4 (a) Top view SEM image of the CdS NWs prepared from the hydrothermal process, *inset* is an HRTEM image of a single CdS nanowire. (b) An XRD spectrum of the CdS NW arrays. (c) AFM line profiles of topography (*top*) and output voltage (*bottom*) scanned across CdS NWs

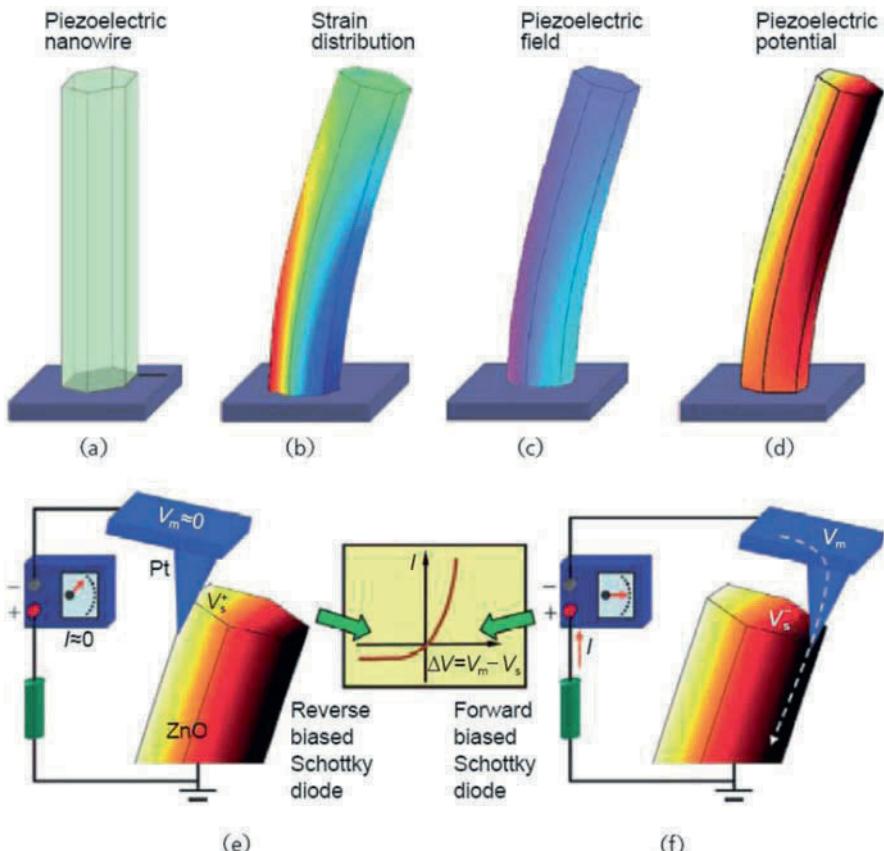


Fig. 5 (a) Schematic definition of a NW and the coordination system. (b) Longitudinal strain ε_z distribution in the NW after being deflected by an AFM tip from the side. (c) The corresponding longitudinal piezoelectric-induced electric field E_z distribution in the NW. (d) Potential distribution in the NW as a result of the piezoelectric effect. (e and f) Contacts between the AFM tip and the semiconductor ZnO NW [boxed area in (d)] at two reversed local contact potentials (positive and negative), showing reverse- and forward-biased Schottky rectifying behavior, respectively. The *inset* shows a typical current–voltage (I – V) relation characteristic of a metal–semiconductor (n-type) Schottky barrier

ZnO NW (Fig. 5a), the deflection of the NW by AFM tip creates a strain field, with the outer surface being tensile (positive strain ε) and inner surface compressive (negative strain ε) (Fig. 5b). A piezoelectric field E_z along the NW is then created inside the NW volume by the piezoelectric effect, with the piezoelectric field closely parallel to the NW at the outer surface and antiparallel to the NW at the inner surface (Fig. 5c). As a result, across the width of the NW at the top end, the electric potential distribution from the compressed to the stretched side surface is approximately between V_s^- (negative) to V_s^+ (positive).

In order for releasing the piezoelectric charge, a Schottky barrier formed between the NW and AFM tip is the key controlling factor. In the first step, the AFM conductive tip that induces the deformation is in contact with the tensile surface of positive potential V_s^+ (Fig. 5d, e). The Pt metal tip has a potential of nearly zero, $V_m = 0$. So the metal tip ZnO interface is negatively biased for $\Delta V = V_m - V_s^- < 0$. With consideration of the n-type semiconductor characteristic of the as-synthesized ZnO NWs, the Pt metal ZnO semiconductor (M-S) interface in this case is a reverse-biased Schottky diode (Fig. 5e) (a Schottky contact is a metal-semiconductor contact at which a potential barrier is formed, so that it behaves like a diode that only allows the current to flow from metal to semiconductor), and little current flows across the interface. This is the process of creating, separating, and accumulating the charges. In the second step, when the AFM tip is in contact with the compressed side of the NW (Fig. 5f), the metal tip ZnO interface is positively biased for $\Delta V = V_m - V_s^- > 0$. The M-S interface in this case is a positively biased Schottky diode and it produces a sudden increase in the output electric current. The current is the result of ΔV -driven flow of electrons from the semiconductor ZnO NW to the metal tip. This is the charge releasing process.

Therefore, the theoretical background for such an energy conversion process is based on a voltage drop created across the cross section of the NW when it is laterally deflected, with the tensile side surface in positive voltage and compressive side in negative voltage. To quantify the piezoelectric potential generated at the two side surfaces of the NW is essential to predict the energy conversion efficiency and the output power. We have applied the perturbation theory for calculating the piezoelectric-potential distribution in a nanowire as pushed by a lateral force at the tip [9]. The analytical solution given under the first-order approximation produces a result that is within 6% from the fully numerically calculated result using a finite element method. Under a simplified condition, where we assume that the nanowire has a cylindrical shape with a uniform cross section of diameter $2a$ and length l , the maximum potential at the surface of the NW is given by equation

$$\varphi_{\max}^{(T,C)} = \pm \frac{3}{4(\kappa_0 + \kappa_{\perp})} [e_{33} - 2(1 + \nu)e_{15} - 2\nu e_{31}] \frac{a^3}{l^3} v_{\max} \quad (1)$$

where φ is the electric potential; κ_0 and κ_{\perp} are the dielectric constants of vacuum and ZnO crystal along its c -plane, respectively; e_{33} , e_{15} , and e_{31} are the linear piezoelectric coefficients; ν is Poisson ratio; and v_{\max} is the maximum deflection at the NW's tip. This equation clearly shows that the electrostatic potential is directly related to the aspect ratio of the NW instead of its dimensionality.

For a typical NW that was grown through the vapor-liquid-solid (VLS) process [10], the diameter is $d = 50$ nm, length is $l = 600$ nm. When it is bent 145 nm to the right by an 80 nm lateral force, which is a common scanning situation in AFM, ± 0.3 V piezoelectric potential can be induced on the two side surfaces, as shown in Fig. 6b, the calculated potential distribution across the NW cross section. Figure 6a is the potential distribution along the bent NW generated by finite element calculation using the above equation. Calculation also shows that the piezoelectric

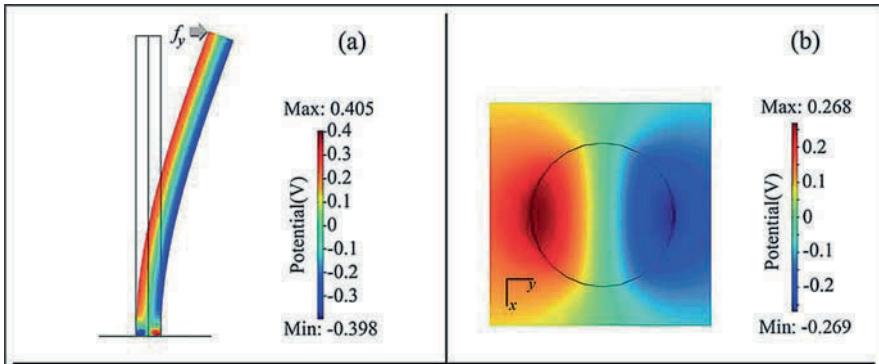


Fig. 6 Potential distribution for a ZnO nanowire with $d = 50$ nm, $l = 600$ nm at a lateral bending force of 80 nN. (a) and (b) are side and cross-sectional (at $z = 300$ nm) output of the piezoelectric potential in the NW given by finite element calculation

potential in the NW almost does not depend on the z -coordination along the NW. Therefore the potential is uniform along z -direction except for regions very close to the ends of the NW. This means that the NW is approximately like a “parallel plate capacitor.” The maximum potential at the surface of the NW is directly proportional to the lateral displacement of the NW and inversely proportional to the length-to-diameter aspect ratio of the NW. For a larger size NW with $d = 300$ nm and length $l = 2 \mu\text{m}$, the surface piezoelectric potential can reach ± 0.6 V, when it is deflected by a 1000 nN force. From this calculation, it can be concluded that the maximum potential at the surface of the NW is directly proportional to the lateral displacement of the NW and inversely proportional to the cube of length-to-diameter ratio of the NW. The voltage that can be generated in a conventional ZnO NW is in the range of several tenths of volts, which is high enough to drive the M-S Schottky diode at the interface between the AFM tip and the ZnO NW, as the key for nanogenerator operation.

4 Direct Current Nanogenerator

AFM-based experiments successfully showed the potential in applying piezoelectric and semiconducting wurtzite NWs for mechanical energy harvesting. Supported by theoretical calculation, the voltage that can be generated by a NW can be sufficiently high for realistic applications. However, in order to convert this concept into an applicable device for power generation, a few challenges need to be addressed. First, sophisticated AFM needs to be eliminated in deflecting the NWs so that the power generation can be achieved by an adaptable, mobile, and cost-effective approach over a larger scale. Second, all of the NWs are required to generate electricity simultaneously and continuously, and all the electricity can be effectively collected and outputted. Finally, the energy to be converted into electricity has to be provided in

a form of wave/vibration from the environment, so the NG can operate “independently” and wirelessly. In addressing those challenges, an aligned NW array-based NG model has been developed, which is operated by ultrasonic waves and outputs a direct current (DC) electric signal – known as DC NG [11].

4.1 The DC Nanogenerator Model

The DC NG is built on vertically aligned ZnO NW arrays and the structure is schematically shown in Fig. 7a. The ZnO NW array was covered by a zigzag Si electrode coated with Pt. The Pt coating not only enhanced the conductivity of the electrode, but also created a Schottky contact at the interface with ZnO. The NWs were grown on either GaN substrates (Fig. 7b) or sapphire substrates that were covered by a thin layer of ZnO film [12], which served as a common electrode for directly connecting the NWs with an external circuit. The density of the NWs was $\sim 10/\mu\text{m}^2$, and the height and diameter were $\sim 1.0 \mu\text{m}$ and $\sim 40 \text{ nm}$, respectively. The top electrode was composed of parallel zigzag trenches fabricated on a (001)-orientated Si wafer [13] and coated with a thin layer of Pt (200 nm in thickness) (Fig. 6c). The electrode was placed above the NW arrays and manipulated by a probe station under an optical microscope to achieve precise positioning; the spacing was controlled by soft-polymer stripes between the electrode and the NWs at the four sides. The assembled device was sealed at the edges to prevent the penetration of liquid. A cross-sectional image of the packaged NW arrays is shown in Fig. 7d; it displays a “lip/teeth” relationship between the NWs and the electrode. Some NWs are in direct contact with the top electrode, but some are located between the teeth of the electrode. The inclined NWs in the SEM image were primarily caused by the cross sectioning of the packaged device.

In this design, the zigzag trenches on the top electrode act as an array of aligned AFM tips. Figure 8a–c, shows four possible configurations of contact between a NW and the zigzag electrode. When subject to the excitation of an ultrasonic wave, the zigzag electrode can move down and push the NW, which leads to lateral deflection of NW I. This, in turn, creates a strain field across the width of NW I, with the NW’s outer surface being in tensile strain and its inner surface in compressive strain. The inversion of strain across the NW results in an inversion of piezoelectric field E_z along the NW, which produces a piezoelectric-potential inversion from V^- (negative) to V^+ (positive) across the NW (Fig. 8b). When the electrode makes contact with the stretched surface of the NW, which has a positive piezoelectric potential, the Pt metal–ZnO semiconductor interface is a reversely biased Schottky barrier, resulting in little current flowing across the interface. This is the process of creating, separating, preserving, and accumulating charges. With further pushing by the electrode, the bent NW I will reach the other side of the adjacent tooth of the zigzag electrode (Fig. 8c). In such a case, the electrode is also in contact with the compressed side of the NW, where the metal/semiconductor interface is a forward-biased Schottky barrier, resulting in a sudden increase in the output electric current flowing from the top electrode into the NW. This is the discharge process.

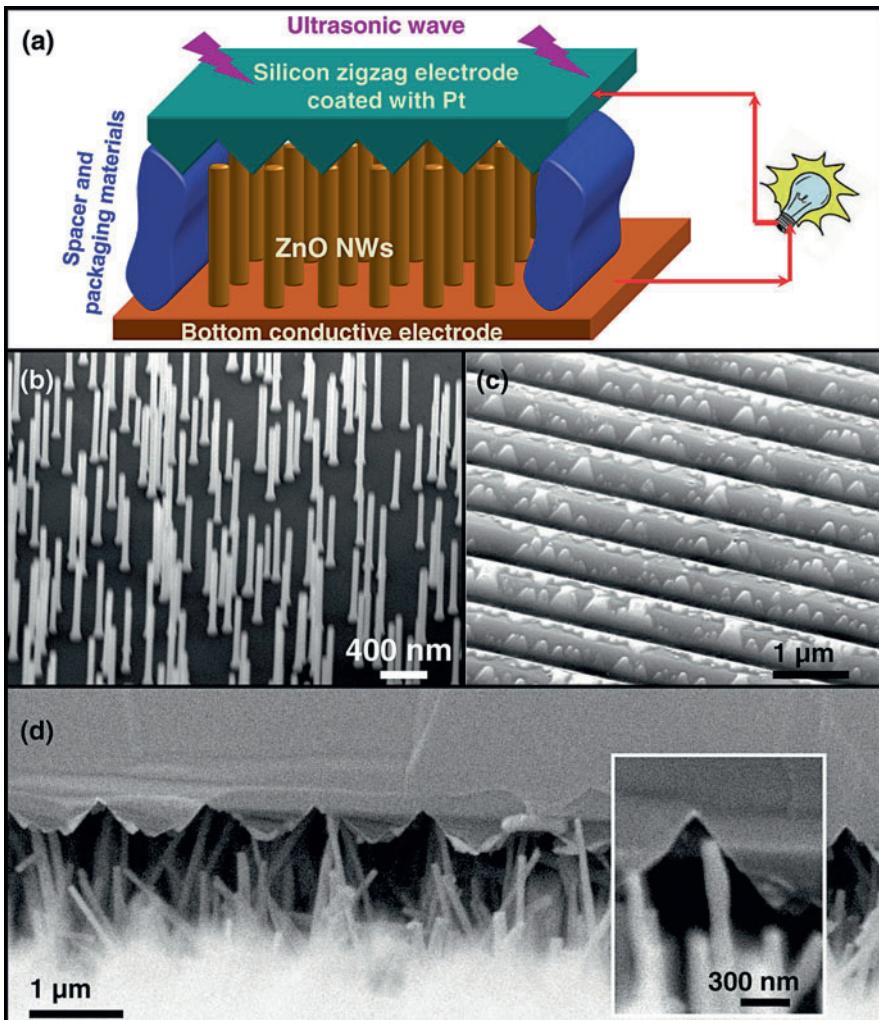


Fig. 7 Nanogenerators driven by an ultrasonic wave. (a) Schematic diagram showing the design and structure of the nanogenerator. (b) Low-density aligned ZnO NWs grown on a GaN substrate. (c) Zigzag trenched electrode coated with Pt. (d) Cross-sectional SEM image of the nanogenerator. *Inset:* A typical NW that is forced by the electrode to bend

Analogous to the situation described for NW I, the same processes apply to the charge output from NW II. NW III is chosen to elaborate on the vibration/resonance induced by an ultrasonic wave. When the compressive side of NW III is in contact with the electrode, the same discharge process as that for NW I occurs, resulting in the flow of current from the electrode into the NW (Fig. 8c). NW IV, which is short in height, is forced (without bending) into compressive strain by the electrode. In

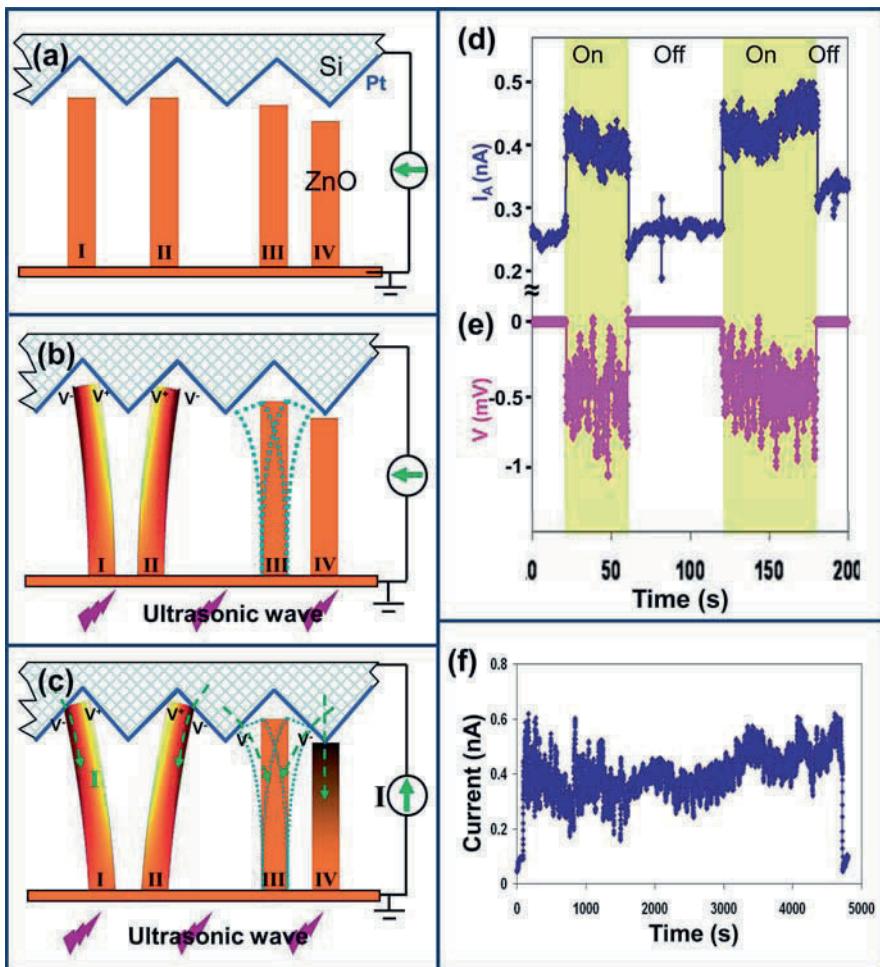


Fig. 8 (a–c) The mechanism of the nanogenerator driven by an ultrasonic wave. (d) and (e) Current and voltage measured on the nanogenerator, respectively, when the ultrasonic wave was turned on and off. (f) Continuous current output of the nanogenerator for an extended period of time

such a case, the piezoelectric voltage can also be created at the top of the NW, thus contributing to the electricity output.

The current and voltage outputs of the NG are shown in Fig. 8d, e, respectively, with the ultrasonic wave being turned on and off regularly. A jump of ~ 0.15 nA was observed when the ultrasonic wave was turned on, and the current immediately fell back to the baseline once the ultrasonic wave was turned off. Correspondingly, the voltage signal exhibited a similar on and off trend but with a negative output of $\simeq 0.5$ mV. The size of the NG is ~ 2 mm² in effective substrate surface area. The number of NWs that were actively contributing to the observed output current

is estimated to be 250–1000 in the current experimental design. The NG worked continuously for an extended period of time of beyond 1 h (Fig. 8f).

4.2 Nanogenerator in Fluid and Three-Dimensional Integration

The first prototype of the NG has been demonstrated by indirect driving using ultrasonic waves through a metal frame suspended in air. An improved design has been thereafter developed for the operation in biocompatible fluid as driven by ultrasonic waves [14].

The new NG was modified from our original model composed by vertically aligned ZnO NWs and a Pt-coated zigzag top electrode, as described in Section 4.1. The NG core was completely packaged by a polymer to prevent the infiltration of biofluid into the NG. The polymer also has certain flexibility to remain free of relative movement between ZnO NWs and the top of the electrode. As shown in Fig. 9a, the NG was placed inside a container filled with 0.9% NaCl solution, which is a typical biocompatible solution. The container had a diameter of 11 cm and a height of 9.5 cm. The substrate and the top electrode were connected by waterproof extension cord to the outside of the container and marked as the positive and the negative electrodes, respectively. The definition of the two electrodes is according to the flow of current in the external circuit, which is always from the bottom substrate to the top zigzag electrode owing to the rectifying effect of the Schottky barrier between ZnO and Pt in the core of the NG. An ultrasonic stimulation was applied through the bottom of the container with a frequency of 41 kHz. The ultrasonic wave transported through the fluid and triggered the vibration of the electrode and NWs to generate electricity. The size of the NG was about 2 mm^2 , on which there are more than 1 million NWs.

Inside the fluid container, the output of NG exhibited a strong correspondence to the local intensity of the ultrasonic wave. The ultrasonic wave source was placed at the center beneath the container. Once the ultrasonic wave was excited inside the container, it was reflected by the container's wall and the water surface. As a result, the wave intensity was enhanced in a certain region inside the fluid. A NG was placed inside the fluid and held by a clamp that can be freely moved in any direction to trace the enhanced ultrasonic wave, while the output current was continuously monitored. First, the NG was placed at the center region above the water surface and the ultrasonic wave was kept on. The corresponding current signal is shown in Fig. 9b. Then the NG was slowly moved into the fluid along the z-direction (depth direction) and a jump in current for $\sim 1\text{ nA}$ was immediately detected once the NG touched the water surface. When the NG reached $\sim 3.3\text{ cm}$ below the water surface, the current quickly jumped to $\sim 20\text{ nA}$. The output can be kept at such a high level as long as the NG stayed at this depth. After 15 s of steady high output, the NG was moved further down and the current dropped back to the 1–2 nA level again. When the bottom of the container was reached, the NG was pulled upward to the water surface. The same 20–25 nA high current output was observed again once the NG reached 3.3 cm depth. The current signal dropped back to its baseline after the NG was pulled out of water.

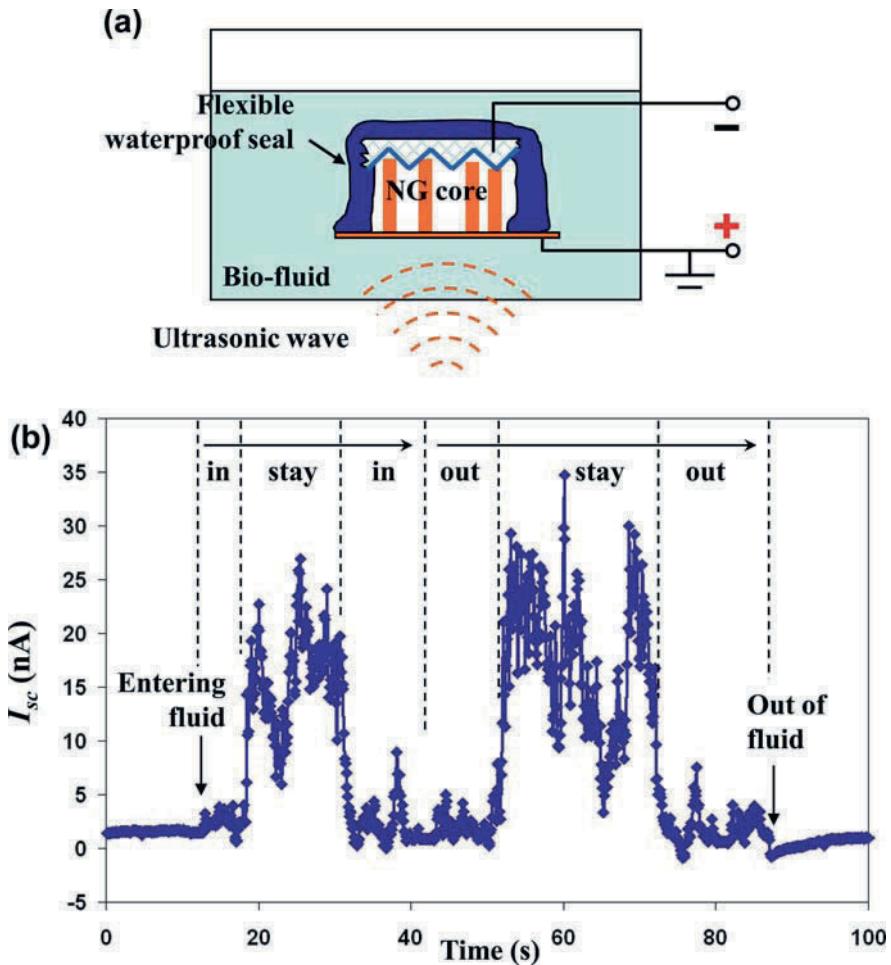


Fig. 9 (a) Schematic of a NG that operates in biofluid and the definition of its two electrodes. (b) Short circuit current signal measured during the movement of NG along z -direction (from water surface to the bottom and then back to surface)

Raising the output of the NG is essential to advance this technology into an applicable power source. If we take each NG as a battery, the most straightforward way to increase the current/voltage is to put them in parallel/serial. This has been demonstrated experimentally. As shown in parts a and b of Fig. 10, two NGs I and II have been tested under the same experimental conditions. NG I exhibited an average I_{sc} of ~ 0.7 nA, and NG II showed a lower noisy signal of 1 nA. These two NGs were then connected in parallel (inset of Fig. 10c) and tested under the same condition again. The resultant output current reached an average of ~ 1.8 nA, which is the sum of the two individual outputs (Fig. 10c). This concept was further proved by antiparallel-connecting NG I and NG II (inset of Fig. 10d). Since the magnitudes

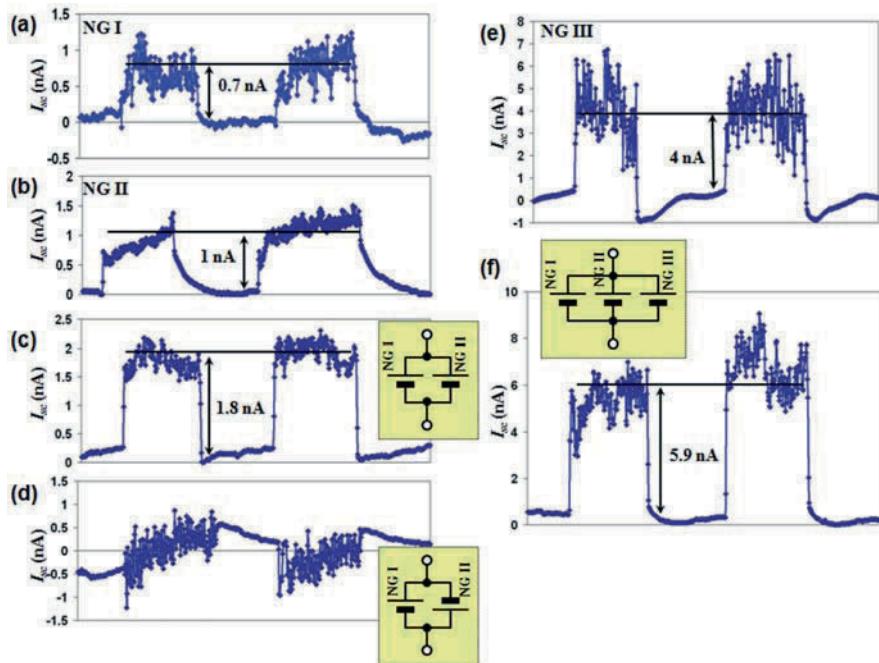


Fig. 10 Short circuit current measured from an integrated NG system. (a) and (b) Current signal measured from two individual NG I and II. (c) and (d) Current signal measured from parallel- and antiparallel-connected NG I and NG II, respectively; the connection configurations are schematically shown in the *insets*. (e) Current signal measured from NG III with a better performance. (f) Current signal measured from parallel-connected NG I, II, and III; *inset* shows the connection configuration

of the I_{sc} of the two NGs were very close, the total current was mostly cancelled out by the “head-to-tail” connection in parallel, and the received signal was around the baseline (Fig. 10d), just as expected.

If more NWs were located in the activate position when packaging the top zigzag electrode, the output current can be improved by several times. As shown in Fig. 10e, with a better assembly, NG III exhibited an average I_{sc} of ~ 4 nA. Once NG III was added to the parallel connection with NG I and NG II (inset of Fig. 10f), their total output current was the sum of the three NGs (5.9 nA $\approx 0.7 + 1 + 4$ nA), as shown in Fig. 10f. This series of experiments demonstrated that the short circuit current can be effectively improved by parallel-connecting multiple NGs. This sets the platform for developing three-dimensionally stacked NGs.

4.3 Performance Analysis – The Carrier Density

The operation mechanism of NG requires that the conductivity and carrier density of the ZnO NW are adequately low in the first step to preserve the piezoelectric-potential distribution in the NW from being “neutralized” by the freely flowing

charge carriers, which are electrons for n-type ZnO, but they need to be high enough to transport the current under the driving of the piezoelectric potential in the charge releasing process [15]. To evaluate the performance of NG in response to carrier density, two experiments have been carried out. Figure 11a shows the UV-induced carrier density change of the vertically aligned ZnO NWs and the underlying ZnO thin film by connecting two electrodes at the two ends of the film, as shown in the inset. By applying a constant voltage V_0 , the transport current was monitored when the UV was turned on or off. The conductance of the film responded rapidly to UV and reached saturation after ~ 300 s. However, the current decayed very slowly after the UV was turned off, possibly due to the following reason. The UV light can effectively create electron–hole pairs in ZnO and greatly increase the carrier density. After turning off the UV, the electrons and holes would recombine and thus decrease the conductivity. However, some carriers may be trapped in the surface/vacancy states in ZnO, which greatly delayed the electron–hole recombination rate, possibly resulting in a long decay in conductivity.

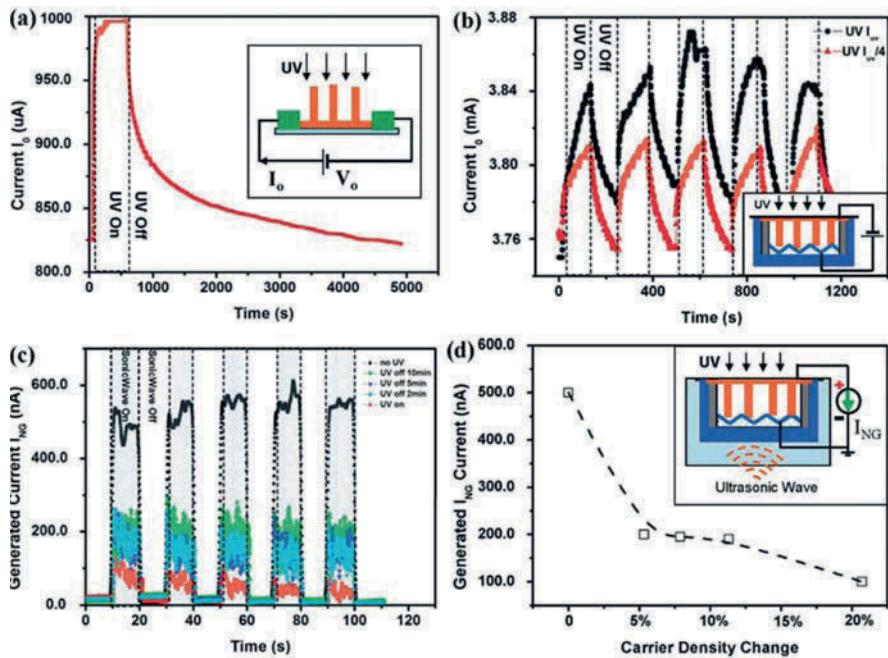


Fig. 11 Tuning the transport property of ZnO and an assembled NG by UV excitation. (a) Response of the current transported through a thin ZnO film with NWs on top when subjected to UV illumination. The *inset* is the measurement setup. (b) Current transported through an assembled NG when the UV light was tuned on and off. The *black* and *gray* curves represent the response curve at full UV intensity I_{UV} and $I_{UV}/4$, respectively. The *inset* is the measurement setup. (c) Current generated by the NG when the ultrasonic wave was turned on and off under different illumination conditions of UV light. (d) Plot of the NG-generated current as a function of the increased percentage of carrier density in the NWs. The *inset* is the measurement setup

The long-lasting tail in the recovery of conductivity in Fig. 11a may possibly indicate that ZnO can hold the charges for an extended period of time before they are completely and entirely neutralized/screened by external electrons. One may suggest that the piezoelectric charges and potential created in a NW may hold for an equivalent length of time before being neutralized/screened by external electrons, so that it can effectively gate the charge carriers as proposed for nanopiezotronics [16], giving sufficient time for experimental observation of the effect of the piezoelectric potential across a NW. Such length of time may be even longer than tens to hundreds of seconds.

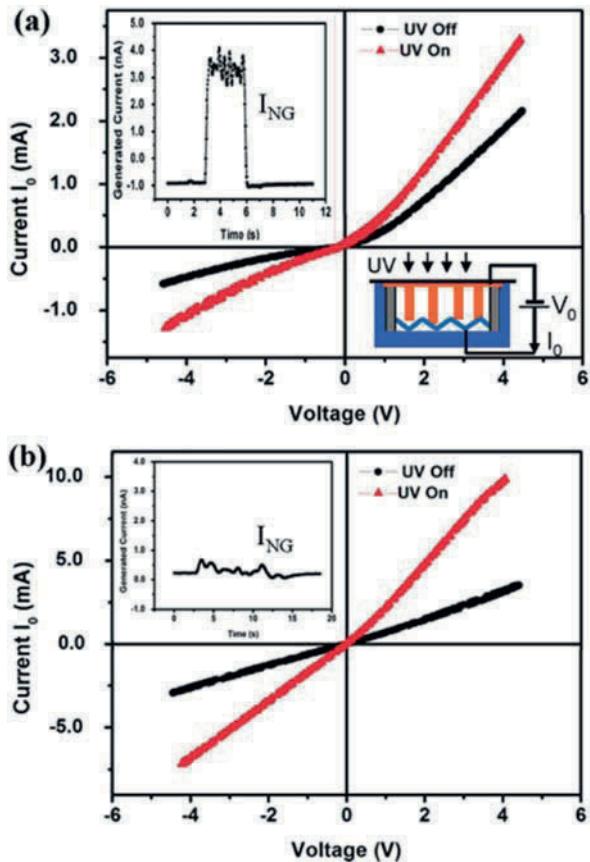
The second experiment was to characterize the UV-induced carrier density change of an assembled NG, as shown in the inset in Fig. 11b. By reducing the relative UV intensity from I_{UV} to $\sim I_{UV}/4$, the conductance of the NG changed by $\sim 2\%$. The conductivity also sharply responded to the on and off of the UV, clearly showing that the UV can effectively increase the carrier density in the NG.

After observing the increased conductivity in NG under UV irradiation, we investigated the performance of the NG for generating electricity. Figure 11c is the generated current (I_{NG}) of a NG when subject to ultrasonic wave excitation by changing the UV irradiating condition. When the UV light was off for an extended period of time, the generated current was ~ 500 nA for a NG with 6 mm^2 in size, corresponding to an output current density of $8.3 \mu\text{A/cm}^2$. The key to obtaining a high output current density is to have NWs that are uniform in size and with patterned distribution on substrate that matches the design of the electrode. As soon as the UV light was turned on, the generated current dropped to as low as 30–80 nA. At the same time, we monitored the I_{NG} after the UV was turned off for 2, 5, and 10 min, and the generated current increased to ~ 180 , 190, and 205 nA, respectively. Such a gradual increase in the generated current corresponds to the slow decay in conductivity as shown in Fig. 11b. This set of experiments apparently shows that an increase in conductivity, e.g., carrier density, reduces the output current of the NG. By extracting the relative change in carrier density from Fig. 11a and the correspondingly measured NG output I_{NG} from Fig. 11c, Fig. 11d illustrates the monotone decrease relationship between the increased percentage in carrier density and the output current I_{NG} . This result indicates that the carrier density in ZnO is one of the key characteristics that dictate the output power of a NG. A too high carrier density would screen the piezoelectric charges, resulting in a lower or even vanishing output current. However, a too low carrier density would increase the inner resistance and greatly decrease the output current. Therefore, there should be an optimum choice of NW carrier density or conductivity to maximize the NG output.

4.4 Performance Analysis – The Schottky Barrier

The Schottky contact between the metal contact and ZnO NW is another key factor to the current generation process [15]. Investigating how it affects the performance of NG will provide effective guidance in designing and fabricating high-output NGs.

Fig. 12 I - V characteristic of an assembled NG for identifying its performance for producing current. (a) For a NG that is actively generating current, as shown in the inset, its I - V curve when the ultrasonic wave was off clearly shows a Schottky diode behavior. UV light not only increased the carrier density (or conductivity) but also might reduce the barrier height. (b) For a “defective” NG that did not produce current, the I - V curve shows a clearly ohmic behavior. The UV light clearly increased the conductivity



We characterized the I - V characteristic of a set of NGs to illustrate the effects of Schottky barrier for current generation. Figure 12a shows the I - V characteristics of a NG that did generate output current (top inset of Fig. 12a). The experimental setup for the measurement is shown in the inset at the lower right corner in Fig. 12a. The NG exhibits a Schottky-like I - V characteristic, which means that the NG has different responses for forward and reverse biases. When being illuminated by UV, the NG showed an increased conductance, and its Schottky-like rectifying effect was largely reduced due to the increased conductance. Alternatively, for a “defective” NG that did not produce current in responding to ultrasonic wave, its I - V characteristic was clearly ohmic even prior to UV illumination (Fig. 12b). By examining over 10 NGs, their I - V characteristics reliably indicated whether the NGs were effective for producing current. This study establishes a criterion for identifying the working NG from the defective ones.

To examine the role played by the Schottky barrier in the NG, we used the AFM-based manipulation and measurement system as used in our first study for demonstrating the piezoelectric NG3 (see Fig. 13a). When a 100 nm thick Pt-coated

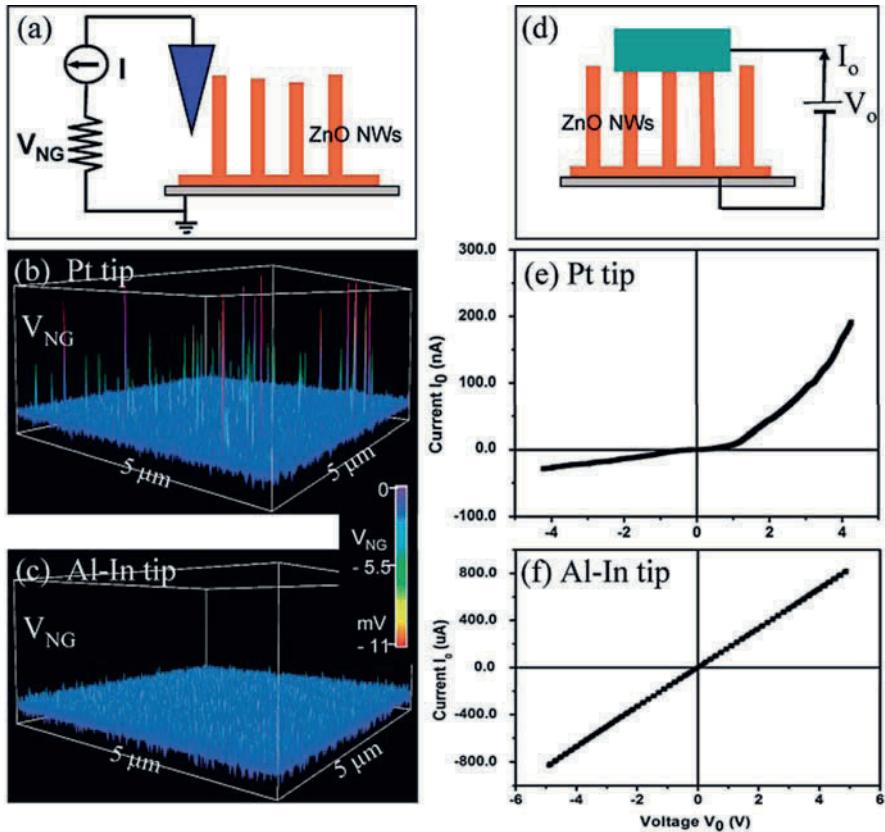


Fig. 13 (a) AFM-based measurement setup for correlating the relationship between the metal–ZnO contact and the NG output. (b) Output potential generated by ZnO NW array when scanned by a Pt-coated Si tip. (c) No output potential is generated by ZnO NW array when scanned by an Al–In coated Si tip. (d) Experimental setup for characterizing the I – V transport property of metal–ZnO NW contact. (e) I – V curve of a Pt–ZnO NW contact, showing Schottky diode effect. (f) I – V curve of an alloyed Al/In–ZnO NW contact, showing ohmic behavior

Si tip was used for scanning the NWs in contact mode, voltage peaks were observed (Fig. 13b), and the output voltages were in the order of ≈ 11 mV (the negative sign means that the current flowed from the grounded end through the external load). By changing the tip to an Al–In (30 nm/30 nm) alloy coated Si tip, the ZnO NWs showed no piezoelectric output (Fig. 13c). In order to understand the two distinct performances of the two types of scanning tips, we have measured their corresponding I – V characteristics with the ZnO NWs. To ensure the stability of the contact, we used a large electrode that was in contact with a group of NWs, as shown in Fig. 13d. The Pt–ZnO contact clearly presented a Schottky diode (Fig. 13e), whereas the Al–In alloy–ZnO was an ohmic contact. In reference to the piezoelectric output presented in Fig. 13b, c, we conclude that a Schottky contact between the metal electrode and ZnO is a must for a working NG.

4.5 Output Improvement

The amplitude of the NG's output is largely dependent on the number of NWs that involve in the power conversion. In order to maximize the number of NWs that would contribute to the power generation process while minimizing the parasitic capacitance that could bring down the voltage output, the NW arrays were carefully chosen so that the variance of the average length of the NW was minimal [17].

The NG was then placed in the water bath to measure the closed circuit current and open circuit voltage. The ultrasonic wave generator was periodically turned on every other 15 s. Measurements were taken under both connection polarities to rule out any possible artifact caused by the measurement setup. Figure 14a shows the closed circuit current when the ultrasonic wave was turned on and off. The data clearly indicate that the current output was originated from the NG as a result of ultrasonic wave excitation, as the output coincided with the working cycle of the ultrasonic wave generator. Furthermore, the output signal switched in sign from positive to negative when the measurement polarity was switched from forward to reversed connection. Finally, there is no destabilization of the output amplitude over time in the graph. A similar pattern in the open circuit voltage output was also observed, as shown in Fig. 14b. Both the current and the voltage outputs exhibit high levels for this type of NG, with a current of ~ 500 nA and voltage of ~ 10 mV. Considering the effective area of the NG (6 mm^2), it is equivalent to a current generation density of $\sim 8.3 \mu\text{A}/\text{cm}^2$, which is ~ 20 times higher than previously reported [14]. A power generation density of $\sim 83 \text{ nW}/\text{cm}^2$ is reported, which shows a great potential to power nanosensors [18].

However, even the highest output signal was still more than 10 times smaller than the output voltage calculated by the bending theory of a single NW. To understand the reason we need to keep in mind that a realistic NG cannot be simply treated as a single voltage source, but has to be treated as multiple sources along with a capacitance and an internal resistance. As Fig. 15 shows, each voltage source $V_1, V_2 \dots V_n$ corresponds to one of the NWs that generates electric power. For each of the NWs, the output consists of multiple pulse outputs:

$$V_i(t) = \sum_{t_{p,i}} V^*(t - t_{p,i}) \quad (2)$$

where $V^*(t)$ is a pulse generated by one instance of power release process between the NW and the top electrode. $t_{p,i}$ is the time when such a pulse p happens for the i th NW. $V^*(t)$ can be approximated as a Gauss package:

$$V^*(t) = V_0 \exp \left(-\frac{t^2}{2\tau^2} \right) \quad (3)$$

where $V_0 = 0.3$ V is the output of a single pulse, as calculated with the bending theory of single NWs, and τ is the width of each output pulse. When driven by ultrasonic waves at ~ 40 kHz, τ can be as small as $0.1 \mu\text{s}$.

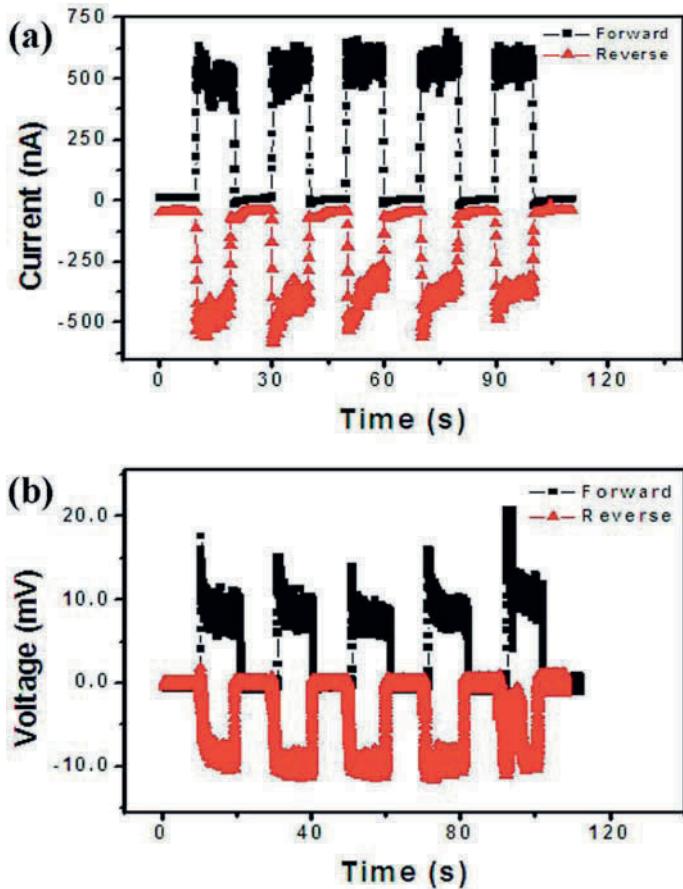


Fig. 14 Performance of a high-output NG when periodically excited by ultrasonic wave. (a) Closed circuit current output and (b) open circuit voltage output measured at forward polarity (black line) and reversed polarity (gray line) connection with the measurement system

The internal resistance is due to the leakage of electric current through the NWs from the substrate to the top electrode. Suppose a single NW has an average resistance of $R_1 = R_2 = \dots = R_n = R_0$, experiment shows that $R_0 \sim 10 \text{ M}\Omega$. r is the resistance due to the nanowires that do not generate power. Those NWs that touch the electrode thus gives a bypass to the output. Experiment shows that $r \sim 100 \text{ k}\Omega$.

The capacitance between the top electrode and the substrate is

$$C = \varepsilon_{\text{Eff}} \frac{A}{d} \quad (4)$$

where A is the area of the electrode and the d is the spacing between the top electrode and substrate. ε_{Eff} is the effective permittivity, which can be approximated to be the

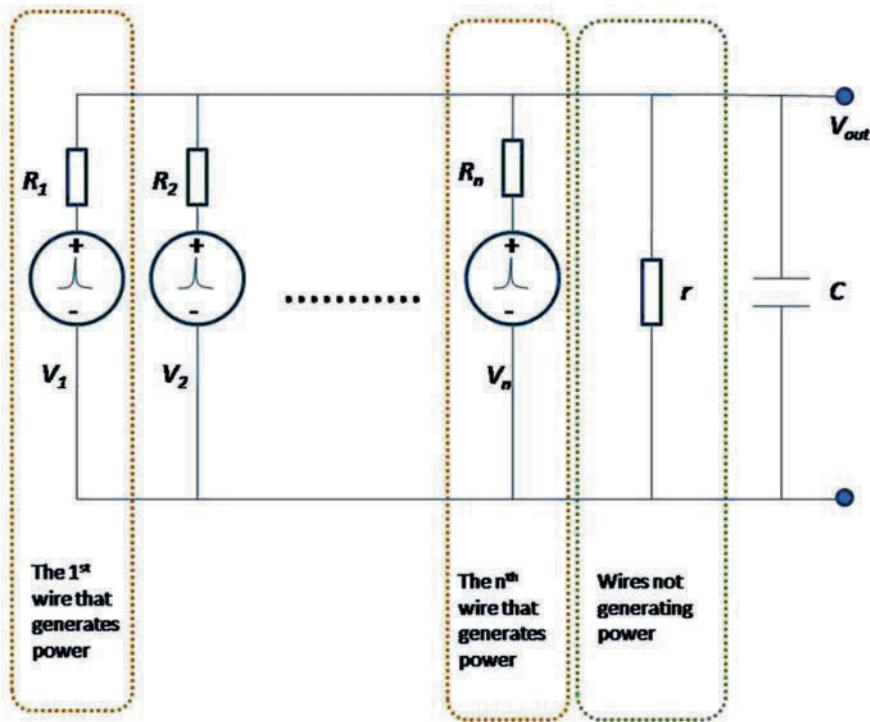


Fig. 15 Equivalent circuit of a realistic nanogenerator

permittivity of the vacuum as an estimation. Plugging $\epsilon_{\text{Eff}} = \epsilon_r \epsilon_0 = 7.77 \times 8.85 \times 10^{-12} \text{ F/m}$, $A = 2 \text{ mm}^2$, $d = 1 \mu\text{m}$, we have $C = 0.14 \text{ nF}$.

Now the timescale of the R - C impedance circuit is

$$\tau_{RC} \sim rC = 100 \text{ k}\Omega \times 0.14 \text{ nF} = 14 \mu\text{S} \quad (5)$$

In solving the system in Fig. 15, we make three assumptions:

- (A) The resistance r is a constant that does not change with time. Experimentally, the fluctuation of r is observed very small, meaning that the number of NWs touching the top electrode does not change very much when sonicating.
- (B) The change of the capacitance C can be neglected. In the realistic situation, C does change with time; but such a change does not give rise to a significant effect in a NG application. As a simple theory illustrating how the output voltage is decreased, we do not consider the variation of the capacitance C .
- (C) The NWs do not interfere with each other. That is, the i th wire and j th wire work independently. Therefore, we can first calculate the output V_{out} due to each NW, then linearly sum up the contributions of all the NWs.

By Kirchoff's law and the above assumptions, we integrate the output from a single pulse of a single NW, the output voltage is found to be

$$V_{\text{out}}^{\text{total}}(t) = ri_r^{\text{total}}(t) = \sum_i \sum_{t_{p,i} < t} \frac{\sqrt{2\pi} V_0 \tau}{R_0 C} \exp\left(-\frac{t - t_{p,i}}{\tau_{\text{RC}}}\right) \quad (6)$$

The power generation process of each NW can be assumed as a Poisson-type random process. Instead of solving the random process rigorously, we are only interested in the averaged DC output of the system. Now let us assume that there are n NWs generating power. For each of the NWs, it takes an average time of T between every two instances of power release, namely, the period of each generation process is T . Taking a long time $T_1 \gg T$ to calculate the average, it is found that the average voltage output is

$$\bar{V}_{\text{out}}^{\text{total}} = \sqrt{2\pi} \frac{\tau}{T} \frac{n}{n'} V_0 \quad (7)$$

where n' is the total number of NWs in a NG.

Equation (7) shows that we need to improve the ratio n/n' in order to improve the output voltage. Increasing τ/T would also improve the performance, which means that we should increase the time during which the power release happens or shorten the generation period. Both of these two ways to improve the performance require better control over the device morphology.

5 Flexible Nanogenerator

5.1 ZnO Nanowires on Flexible Substrate

NGs presented previously in this chapter were all based on ZnO NWs that were grown on ceramic and semiconducting substrates, which are hard and brittle and cannot be used in applications that require a foldable or flexible power source. By growing ZnO NW arrays on a flexible plastic substrate, we demonstrate the first successful flexible power source built on conducting polymer films [19]. This approach has two specific advantages: it uses a cost-effective, large-scale, wet-chemistry strategy to grow ZnO NW arrays at temperatures lower than 80°C, and the growth of aligned ZnO NW arrays can occur on a large assortment of flexible plastic substrates.

The aligned ZnO NWs arrays were grown on a plastic film using a solution-based method [20]. Serving as one of the electrodes for a later electrical connection and also functioning as uniform nucleation sites for NW growth, a thin, ca. 100 nm, layer of Au was deposited on the plastic substrate before the hydrothermal growth of ZnO NW arrays. Through this method, aligned ZnO NWs can be grown uniformly covering an area of 2 inch plastic wafer (Fig. 16a) or even larger. In order for the manipulation of AFM, the density of ZnO NWs was controlled to be

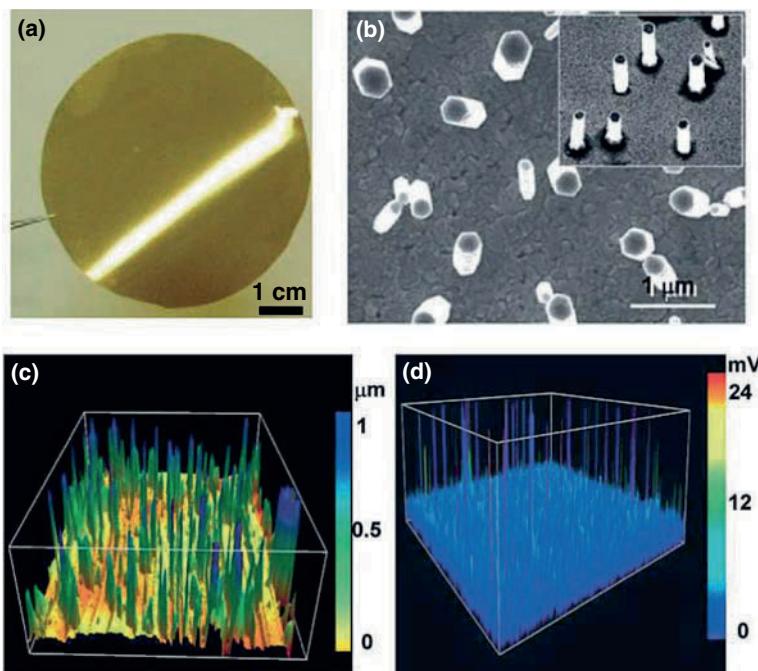


Fig. 16 (a) Optical image of the polymer substrate, on which aligned ZnO nanowires had been grown. (b) A SEM image showing a sparsely grown ZnO NW array on a plastic substrate, *inset* is a tilted image. (c) A 3D plot of the AFM topography image and (d) a voltage output profile obtained by scanning the AFM tip over a $40 \times 40 \mu\text{m}$ area of aligned ZnO NWs

low by adjusting the reaction temperature, solution concentration, and pH value. Figure 16b shows a SEM image of sparsely grown ZnO NW arrays on a plastic substrate. The NWs have a density of ca. $1 \mu\text{m}^{-2}$.

A similar AFM scanning process was applied on the NW arrays and the voltage signal was recorded simultaneously. Figure 16c is a 3D plot of a topographic AFM scan of a $40 \times 40 \mu\text{m}$ area of the ZnO NW array. The scanning direction is from the front to the back, which can be seen from the raised linear traces on the flexible plastic substrate. Because of the flexibility of the plastic substrate and in spite of the firm adhesion to the Ag paste, the substrate surface profile under tip contact was somewhat wavy. Relative to the plastic substrate, the NW array was revealed to have a height distribution of 0.5–2.0 μm . The corresponding 3D plot of the voltage output image is shown in Fig. 16d. There are a number of sharp peaks ranging from 15 to 25 mV that represent voltage outputs. By counting the pulse numbers shown here with respect to the number of NWs in the topography profile in Fig. 16c, the ratio of voltage peaks to the number of available NWs is 90:150 or about 60%, which suggests that the discharge events of the NWs captured by the AFM tip correspond to at least 50% of the NWs. In fact, because of the limitation in the data-collecting speed of the AFM, which has a data step size of ca. 156.2 nm

in this measurement, the discharge peaks of some NWs were missed, possibly due to poor contact between the tip and the NWs and/or due to multiple contacts with neighboring NWs. It is suggested that both suitable bonding strength between the ZnO NWs and the polymer substrate and a uniform density distribution of NWs in the array might be very important in terms of improving the piezoelectric discharge efficiency.

5.2 Power Fiber

Fabric is another type of widely used soft material, which can be an excellent media for harvesting mechanical energy. By growing ZnO NWs around Kevlar 129 fibers, which have high strength, modulus, toughness, and thermal stability, we demonstrated the power generation ability of a textile fiber system [21]. The preliminary research was based on a double-fiber model as shown in Fig. 17a. Two fibers, one coated with a 300-nm-thick gold layer and the other as-grown, were entangled to form the core for power generation. The relative brushing between the two fibers was simulated by pulling/releasing the string using an external rotor with a controlled frequency. In this design, the gold-coated ZnO NWs acted as an array of scanning metal tips that deflected the ZnO NWs rooted at the other fiber; a coupled piezoelectric and semiconducting property resulted in a process of charge creation and accumulation and charge release. The gold coating completely covered the ZnO NWs and formed a continuous layer along the entire fiber. Once the two fibers were

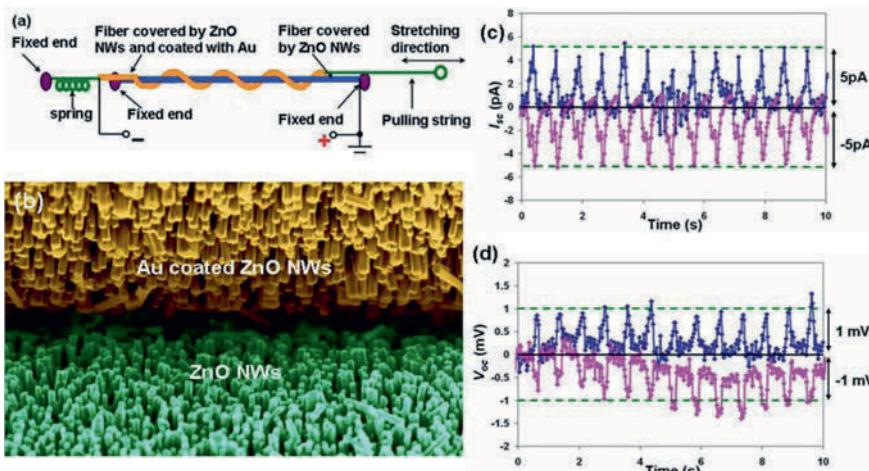


Fig. 17 (a) Schematic experimental setup of the fiber-based NG. (b) SEM image at the “teeth-to-teeth” interface of two fibers covered by NWs with/without gold. (c) and (d) I_{sc} and V_{oc} of a double-fiber NG

firmly entangled together, some of the gold-coated NWs penetrated slightly into the spaces between the uncoated NWs rooted at the other fiber, as shown by the interface image in Fig. 17b. Thus, when there was a relative sliding/deflection between them, the bending of the uncoated ZnO NWs produced a piezoelectric potential across their width, and the Au-coated NWs acted as the “zigzag” electrode (as for the DC NG) for collecting and transporting the charges.

The short circuit current (I_{sc}) and open circuit voltage (V_{oc}) were measured to characterize the performance of the fiber NGs. The pulling and releasing of the gold-coated fiber was accomplished by a motor at a controlled frequency. The resulting electric signals detected at 80 r.p.m. are shown in Fig. 17c, d. A “switching polarity” testing method was applied during the entire measurement to rule out system artifacts. When the current meter was forward-connected to the NG, which means that the positive and negative probes were connected to the positive and negative electrodes of the NG as defined in Fig. 17a, respectively, ~ 5 pA positive current pulses were detected at each pulling–releasing cycle (blue curve in Fig. 17c). Negative current pulses with the same amplitude were received (pink curve in Fig. 17c) when the current meter was reverse-connected, with its positive and negative probes connected to the negative and positive electrodes of the NG, respectively. The small output current (~ 5 pA) is attributed mainly to the large loss in the fiber due to an extremely large inner resistance ($R_i \approx 250$ M Ω). Open circuit voltage was also measured by the switching polarity method. Corresponding positive and negative voltage signals were received when the voltage meter was forward- and reverse-connected to the fiber NG, respectively (Fig. 17d). The amplitude of the voltage signal in this case was $\sim 1\text{--}3$ mV.

After demonstrating the electricity generation principle, a few approaches were investigated to increase the output power and prototype integration. To simulate a practical fabric made of yarn, a single yarn made of six fibers was tested; three fibers were covered with NWs and coated with Au, and three were only covered with NWs. All of the gold-coated fibers were movable in the testing (Fig. 18b). At a motor speed of 80 r.p.m., an average current of ~ 0.2 nA was achieved (Fig. 18a), which is $\sim 30\text{--}50$ times larger than the output signal from a single-fiber NG: this is due to the substantially increased surface contact area among the fibers. The width of each pulse was broadened, apparently due to the unsynchronized movement and a relative delay in outputting current among the fibers.

Reducing the inner resistance of the fiber and the NWs was found to be effective for enhancing the output current. By depositing a conductive layer directly onto the fiber before depositing the ZnO seeds, the inner resistance of the NG was reduced from ~ 1 G Ω to ~ 1 k Ω ; thus the output current I_{sc} of a double-fiber NG was increased from ~ 4 pA to ~ 4 nA (Fig. 18c). The current I_{sc} is approximately inversely proportional to the inner resistance of the NG (inset in Fig. 18c). This study shows an effective approach for increasing the output current.

The textile-fiber-based NG has demonstrated the following innovative advances in comparison to the DC NGs. First, using ZnO NWs grown on fibers, it is possible to fabricate flexible, foldable, wearable, and robust power sources in any shape (such as a “power shirt”). Second, the output electricity can be dramatically enhanced

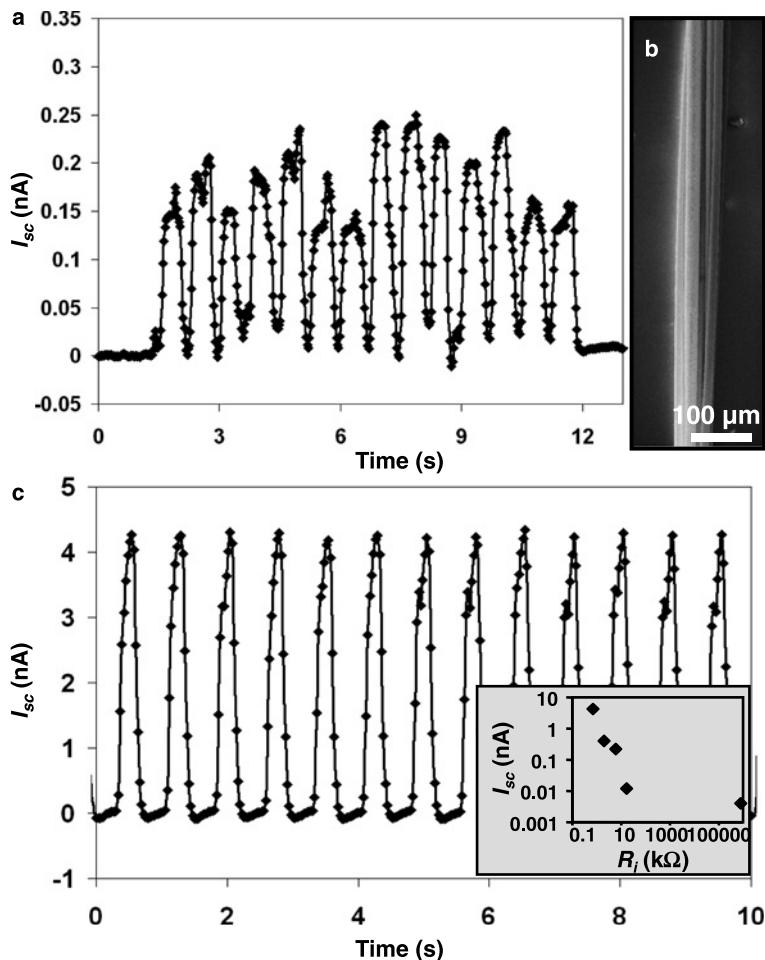


Fig. 18 (a) I_{sc} of a multi-fiber NG. (b) SEM image of the multi-fiber NG composed of six ZnO NW-coated fibers, three of which were gold-coated. (c) Enhancement of I_{sc} by reducing the inner resistance of the NG

using a bundle of fibers as a yarn, which is the basic unit for fabrics. The optimum output power density from textile fabrics can be estimated on the basis of the data we have reported, and an output density of 20–80 mW per square meter of fabric is expected. Third, the NG operates at low frequency, in the range of conventional mechanical vibration, footsteps, and heartbeats, greatly expanding the application range of NGs. Last, as the ZnO NW arrays were grown using chemical synthesis at 80°C on a curved substrate, we believe that our method should be applicable to growth on any substrate, so the fields in which the NGs can be applied and integrated may be greatly expanded. This work establishes a methodology for scavenging light-wind energy and body-movement energy using fabrics.

6 Prospect

Relaying on the unique piezoelectric and semiconducting coupled property, ZnO NWs have been successfully applied for harvesting nanoscale mechanical energy and converting it into electricity. The NGs have been made on solid substrates for harvesting high-frequency acoustic wave or on flexible soft materials that are able to harvest low-frequency perturbations, such as footsteps, wind blows, or even heartbeats. The development of NGs exhibits unique advantages: First, the NW-based NGs can be subjected to extremely large deformation, so they can be used for flexible electronics as a flexible/foldable power source. Second, the large degree of deformation that can be withstood by the NWs is likely to result in a larger volume density of power output. Third, ZnO NW NGs can directly produce current as a result of their enhanced conductivity with the presence of oxygen vacancies. Fourth, the flexibility of the polymer substrate used for growing ZnO NWs/NBs makes it feasible to accommodate the flexibility of human muscles so that we can use the mechanical energy (body movement, muscle stretching) in the human body to generate electricity. Finally, ZnO is a biocompatible and biosafe material [22]; it has great potential as an implantable power source within the human body.

The following development will be focused on improving the output power of a current NG model, while a full article addressing some concerns about NG's performance [23] can be found elsewhere [24]. Combining with the existing knowledge on the relationship between the ZnO NW arrays and the performance of NG, we summarize the key factors that are important to boost up the voltage output as follows:

- (A) Increase the number of active NWs for participating in electricity generation. This is crucial in determining the voltage output performance. There are two possible approaches. One is to utilize ZnO NW arrays with uniform size, especially uniform in length. The other approach is to pattern the NW arrays according to the dimension and shape of the top electrode. An increase in output voltage will be accompanied with a simultaneous increase in output current.
- (B) Increase the charges generated by individual NW during the deflection process. This may be possible by increasing the magnitude of the external excitation, because the magnitude of the generated voltage is proportional to the deflection of the NW.
- (C) Increase the total charge to be the output by a NW to the external load. This requires a great decrease in the contact resistance between the metal electrode and the NW. In a general NG device, the sum of the contact resistance and inner resistance of the wire was estimated to be $\sim 35\text{ M}\Omega$. This large resistance dissipates a large voltage at the contact. It is thus essential to reduce the contact resistance to receive larger output voltage.
- (D) Optimize the electric conductivity of ZnO NW. The relationship between the carrier density and the performance of the NG indicates that a too high conductivity destroys the Schottky contact at the interface; while a too low conductivity consumes too much voltage. It is expected that by tuning the carrier density to

an optimal value, both the current and the voltage outputs can be significantly increased. However, the optimum value of the conductivity needs to be modeled with consideration of the charge transport dynamics.

- (E) Decrease the capacitance between individual NW and the top electrode, as well as the system capacitance. The size of the electrodes can make large contribution to the system capacitance. The other factor is the capacitance of the measurement circuit.

By addressing the above challenges, we expect to have >80% NWs involved in electricity generating. Under this ideal situation, the power density per unit substrate area is $\sim 1\text{--}2 \text{ nW}/\mu\text{m}^2$, i.e., $0.1\text{--}0.2 \text{ W/cm}^2$, which is large enough for powering regular small electronic devices.

Acknowledgments This work was supported by DOE BES, NSF, DARPA, and NASA. Thanks to the contributions from our group members: Dr. Jinhui Song, Dr. Yong Qin, Dr. Jin Liu, Dr. Jun Zhou, Dr. Rusan Yang, Yifan Gao, Peng Fei, Dr. Puxian Gao, Dr. Jr-Hau He, Dr. Changshi Lao, Dr. Yi-Feng Lin, and Wenjie Mai.

References

1. Song J.H., J. Zhou, Z.L. Wang. Piezoelectric and Semiconducting Coupled Power Generating Process of a Single ZnO Belt/Wire. A Technology for Harvesting Electricity from the Environment. *Nano Letters* 2006; 6: 1656–1662.
2. Wang X.D., C.J. Summers, Z.L. Wang. Large-scale Hexagonal-Patterned Growth of Aligned ZnO Nanorods for Nano-optoelectronics and Nanosensor Arrays. *Nano Letters* 2004; 4: 423–426.
3. Wang X.D., J.H. Song, C.J. Summers, J.H. Ryou, P. Li, R.D. Dupuis, Z.L. Wang. Density-Controlled Growth of Aligned ZnO Nanowires Sharing a Common Contact: A Simple, Low-Cost, and Mask-Free Technique for Large-Scale Applications. *Journal of Physical Chemistry B* 2006; 110: 7720–7724.
4. Wang Z.L., J.H. Song. Piezoelectric Nanogenerators Based on Zinc Oxide Nanowire Arrays. *Science* 2006; 312: 242–246.
5. Song J.H., X.D. Wang, E. Riedo, Z.L. Wang. Elastic Property of Vertically Aligned Nanowires. *Nano Letters* 2005; 5: 1954–1958.
6. Lin Y.F., J.H. Song, Y. Ding, S.Y. Liu, Z.L. Wang. Piezoelectric Nanogenerator Using CdS Nanowires. *Applied Physics Letters* 2008; 92: 022105.
7. Cao B., Y. Jiang, C. Wang, W. Wang, L. Wang, M. Niu, W. Zhang, Y. Li, S.-T. Lee. Synthesis and Lasing Properties of Highly Ordered CdS Nanowire Arrays. *Advanced Functional Materials* 2007; 17: 1501–1506.
8. Wang Z.L., X.D. Wang, J.H. Song, J. Liu, Y.F. Gao. Piezoelectric Nanogenerators for Self-Powered Nanodevices. *IEEE Pervasive Computing* 2008; 7: 49–55.
9. Gao Y.F., Z.L. Wang. Electrostatic Potential in a Bent Piezoelectric Nanowire. The Fundamental Theory of Nanogenerator and Nanopiezotronics. *Nano Letters* 2007; 7: 2499–2505.
10. Wang X.D., J. Song, P. Li, J.H. Ryou, R.D. Dupuis, C.J. Summers, Z.L. Wang. Growth of Uniformly Aligned ZnO Nanowire Heterojunction Arrays on GaN, AlN, and Al_{0.5}Ga_{0.5}N Substrates. *Journal of the American Chemical Society* 2005; 127: 7920–7923.
11. Wang X.D., J.H. Song, J. Liu, Z.L. Wang. Direct-Current Nanogenerator Driven by Ultrasonic Waves. *Science* 2007; 316: 102–105.

12. Wang X.D., J.H. Song, C.J. Summers, J.H. Ryou, P. Li, R.D. Dupuis, Z.L. Wang. Density-Controlled Growth of Aligned ZnO Nanowires Sharing a Common Contact: A Simple, Low-Cost, and Mask-Free Technique for Large-Scale Applications. *Journal of Physical Chemistry B* 2006; 110: 7720–7724.
13. Fröhlauf J., S. Krönert. Wet Etching of Silicon Gratings with Triangular Profiles. *Microsystem Technologies* 2005; 11: 1287–1291.
14. Wang X.D., J. Liu, J.H. Song, Z.L. Wang. Integrated Nanogenerators in Biofluid. *Nano Letters* 2007; 7: 2475–2479.
15. Liu J., P. Fei, J.H. Song, X.D. Wang, C.S. Lao, R. Tummala, Z.L. Wang. Carrier Density and Schottky Barrier on the Performance of DC Nanogenerator. *Nano Letters* 2008; 8: 328–332.
16. Wang Z.L. Nanopiezotronics. *Advanced Materials* 2007; 19: 889–892.
17. Liu J., P. Fei, J. Zhou, R. Tummala, Z.L. Wang. Toward High Output-Power Nanogenerator. *Applied Physics Letters* 2008; 92: 173105.
18. Yu C., Q. Hao, S. Saha, L. Shi, X. Kong, Z.L. Wang. Integration of Metal Oxide Nanobelts with Microsystems for Nerve Agent Detection. *Applied Physics Letters* 2005; 86: 063101.
19. Gao P.G., J.H. Song, J. Liu, Z.L. Wang. Nanowire Piezoelectric Nanogenerators on Plastic Substrates as Flexible Power Sources for Nanodevices. *Advanced Materials* 2007; 19: 67–72.
20. Vayssières L. Growth of Arrayed Nanorods and Nanowires of ZnO from Aqueous Solutions. *Advanced Materials* 2003; 15: 464–466.
21. Qin Y., X.D. Wang, Z.L. Wang. Microfibre-Nanowire Hybrid Structure for Energy Scavengings. *Nature* 2008; 451: 809–813.
22. Zhou J., N.S. Xu, Z.L. Wang. Dissolving Behavior and Stability of ZnO Wires in Biofluids: A Study on Biodegradability and Biocompatibility of ZnO Nanostructures. *Advanced Materials* 2006; 18: 2432–2435.
23. Alexe M., S. Senz, M.A. Schubert, D. Hesse, U. Gosele, Energy Harvesting Using Nanowires? *Advanced Materials* 2008; 20: 4021.
24. Wang Z.L. Energy Harvesting Using Piezoelectric Nanowires – A Correspondence on ‘Energy Harvesting Using Nanowires?’ by Alexe et al. *Advanced Materials* 2009; 21: 1311–1315.

Nanolead-Free Solder Pastes for Low Processing Temperature Interconnect Applications in Microelectronic Packaging

Hongjin Jiang, Kyoung-sik (Jack) Moon, and C.P. Wong

Abstract Lead-free solder, tin, tin/silver (SnAg) and tin/silver/copper (SnAgCu) alloy nanoparticles with various sizes were synthesized via a low temperature chemical reduction method and their thermal properties were studied by differential scanning calorimetry. The particle size dependency of the melting temperature and the latent heat of fusion was observed. The wetting test for the as-prepared SnAg and SnAgCu alloy nanoparticle pastes on a Cu surface showed the typical Cu₆Sn₅ intermetallic compound (IMC) formation. These low melting point SnAg or SnAgCu alloy nanoparticles could be used for low reflow temperature lead-free interconnect applications.

Keywords Lead-free solders · Low reflow temperature · Nanoparticles · Chemical reduction · Melting point depression

Solder has long played an important role in the assembly and interconnection of integrated circuit(IC) components on substrates, i.e., ceramic or organic printed circuit boards. This type of alloy can provide electrical, thermal, and mechanical connections in electronic assemblies. Although the electronics industry has made considerable advances over the past few decades, the essential requirements of communications among all types of components in all electronic systems remain unchanged. Components need to be electrically connected for power and ground as signal transmissions. Tin/lead (Sn/Pb) solder alloy has been the de facto interconnect material in most areas of electronic packaging, including interconnection technologies such as pin through hole (PTH), surface mount techniques (SMT), ball grid array (BGA) (Fig. 1), chip scale packaging (CSP), flip chip.

There are increasing concerns nowadays about the use of tin/lead alloy solders. Lead, a major component in solder, has long been recognized as a health threat to human beings. The major concern is that lead from discarded consumer electronics

H. Jiang (✉)
Intel Corporation, Chandler AZ, USA
e-mail: hongjinjiang@gmail.com

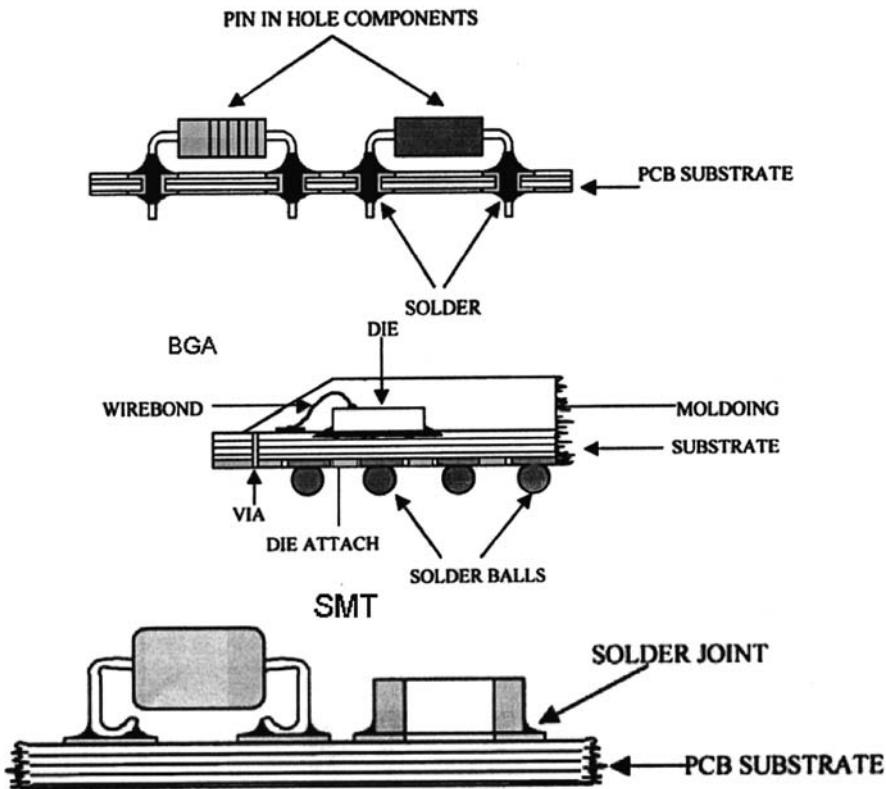


Fig. 1 Schematic structures of pin through hole (PTH), ball grid array (BGA), and surface-mounted technology (SMT) packages using solder interconnects

products in landfills could leach into underground water and eventually into drinking water system [1]. According to the latest report (US Geological Survey, Jan 2009), the total lead consumption by the US industries was 1,620,000 tons in 2008, a significant amount of which was used to produce alloy solders. Worst of all, most of the electronic products have a very short service life (e.g., cell phones, pagers, electronic toys, PDA), which often end up in landfills in just a few months or years. Recycling of lead-containing consumer electronic products has been proven to be very difficult, compared to that of lead-containing batteries or cathode ray tubes (CRT) displays.

The attempt to ban lead from electronic solder was initiated in the US Congress in 1990; however, the lead-free movement has advanced much more rapidly in Japan and Europe. The leading Japanese equipment manufacturers (OEMs) have already introduced products that contain no lead in the interconnection. Although the WEEE (Waste from Electrical and Electronic Equipment) Directive has put the European Union's lead-free legislation back to 2006, many companies such as Ericsson, Nokia, and Philips have already implemented "green" marketing strategies

into their new products. These differences in lead-free progress have triggered great concerns about maintaining business opportunity among users of lead-containing solders, as such; it further expedites the advancement of lead-free-soldering programs. Hence, lead-free electronics are not only perceived as a health issue, but their development is also driven by governmental and commercial interests. Many IC chip and board manufacturers including Intel, the largest microprocessor manufacturer, claim that their lead-containing products would be phased out around 2006.

A variety of lead-free solder alloys have been investigated as potential replacements for tin/lead solders. The main requirements for a lead-free solder alloy are

- (1) Low melting point: The melting point should be low enough to avoid thermal damage to the packages and high enough for the solder joint to bear the operating temperature. The solder should retain adequate mechanical properties at these temperatures.
- (2) Wettability: The solder materials should wet the base metal properly to provide the electrical, thermal, and mechanical connections between chips and substrates.
- (3) Availability and cost: There should be adequate supplies and at the same time the cost should be reasonable. The microelectronics industry is extremely cost conscious. The electronic manufactures are unlikely to change to an alternative solder material with an increased cost unless it has demonstrated better properties or there is legislative pressure to do so.

Although several commercial and experimental Sn-based lead-free solder alloy exist, none can meet all the above requirements, especially if the melting point of the candidate alloys is required to be very close to that of the SnPb. Lead-free candidates and their respective melting points are listed in Table 1. Two alloy families, tin/silver/copper (SnAgCu) and tin/copper (SnCu), seem to be generating the most interest. SnAgCu alloy composition (with or without the addition of a fourth element) appears to be the most popular replacement and has been chosen to be the

Table 1 Lead-free alloys

Alloy	Melting point
Sn96.5Ag3.5	221°C
Sn96Ag3.5Cu0.5	217°C
Sn20Au80	280°C (mainly used in interconnects for optoelectronic packaging)
Sn99.3/Cu0.7	227°C
SnAgCuX(Sb, In)	Ranging according to compositions, usually above 210°C
SnAgBi	Ranging according to compositions, usually above 200°C
Sn95Sb5	232–240°C
Sn91Zn9	199°C
SnZnAgAlGa	189°C
Sn42Bi58	138°C

benchmark, with SnPb being the baseline, that all other potential alloys for the industry to be tested against. Concerns with this alloy family include higher processing temperatures, poorer wettability due to their higher surface tension, and their compatibility with lead bearing finishes. The SnCu alloy composition is a low-cost alternative for wave soldering and is compatible with most lead bearing finishes. Process considerations must be addressed with this alloy due to its higher melting temperature than most SnAgCu alloys.

As shown in Table 1, the melting points of the SnAgCu and SnCu alloys are more than 30–40°C higher than that of the eutectic SnPb alloy requiring the reflow temperature for soldering components on assembly boards should be raised by 30–40°C. The higher reflow temperature leads to a number of undesirable consequences such as higher residual stress of components and substrates, which adversely affects their reliability. It is also considered likely to have an increased tendency of the “pop-corning” found in the plastic molded/encapsulated ICs during the reflow process. The high required temperature can potentially create serious warpage in organic board substrates. Furthermore, heat-sensitive components such as electrolytic capacitors might not survive the high process temperatures of lead-free assembly. Solders tend to be re-oxidized at high reflow temperatures and the wettability becomes worsened unless expensive nitrogen reflow is used. Therefore, industry’s attentions have been paid on lowering the processing temperature of the lead-free metals.

The melting point of many materials can be dramatically reduced by decreasing the size of the materials. The melting and freezing behavior of finite systems have been of considerable theoretical and experimental interests for many years. As early as 1888, J.J. Thomson suggested that the freezing temperature of a finite particle depends on the physical and chemical properties of the surface. It was not until 1909, however, that an explicit expression for a size-dependent solid–liquid coexistence temperature first appeared. By considering a system consisting of small solid and liquid spheres of equal mass in equilibrium with their common vapor, it was shown that the temperature of the triple point was inversely proportion to the particle size. A similar conclusion was later reached based on the conditions for equilibrium between a solid spherical core and a thin surrounding liquid shell. Systematic experimental studies of the melting and freezing behavior of small particles began to appear in the late 1940s and early 1950s: first in a series of experiments on the freezing behavior of isolated micrometer-sized metallic droplets and later in an electron diffraction study of the melting and freezing temperatures of vapor-deposited discontinuous films consisting of nano-sized islands of Pb, Sn, and Bi. These studies demonstrated that small molten particles could often be dramatically undercooled and that solid particles melted significantly below their bulk melting temperature. To date, the melting point of substances can be dramatically decreased when their size is reduced to nanometer size [2–12]. This is due to the high surface area to volume ratio for nanoparticles, which as a consequence of the higher surface energy substantially affects the interior “bulk” properties of the material, resulting in a decrease of the melting point. The outer surface of the nanoparticles plays a relevant role in the melting point depression. And the surface premelting

process has been suggested as one of sources of the melting point depression of the nanoparticles [13].

Transmission electron microscopy (TEM) and nanocalorimeter have been used to study the melting behavior of single Sn nanoparticle or cluster. At the melting point, the diffraction pattern of the crystal structure in TEM exhibited the order-disorder transition [14]. However, unlike calorimeters, TEM cannot measure the latent heat of fusion (ΔH_m), which is very useful to understand the thermodynamics of a finite material system. Lai et al. investigated the melting process of supported Sn clusters by nanocalorimeter [3] and found that the melting point depended nonlinearly on the inverse of the cluster radius R , which was in contrast to the traditional description of the melting behavior of small particles (Fig. 2a). They first reported a particle-size-dependent reduction of ΔH_m for Sn nanoparticles (Fig. 2b). Bachels et al. studied the melting behavior of isolated Sn nanoparticles or clusters by nanocalorimeter as well [4]. The melting point of the investigated Sn clusters was found to be lowered by 125 K and the latent heat of fusion per atom was reduced by 35% compared to bulk Sn. However, for both the TEM and nanocalorimetric measurements, the single element of Sn nanoparticle or the cluster was synthesized by the deposition method inside the measurement equipments in order to prevent the particles from oxidation.

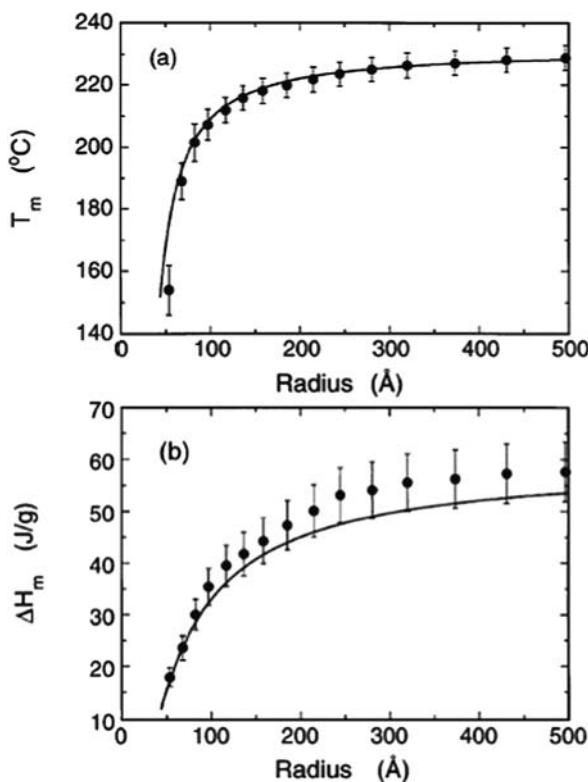


Fig. 2 (a) Size dependence of the melting points of Sn nanoparticles. The *solid line* is calculated in terms of equation (1). (b) Size dependence of the normalized heat of fusion. The *solid line* is calculated in terms of equation (2)

Although this synthesis method was sufficiently sophisticated for the pure observation, the real world synthesis and the characterization possess more experimental parameters to consider such as oxidation during moving samples and differences in melting/wetting behavior on different substrate materials.

$$T_m = 232 - 782 \left[\frac{\sigma_s}{15.8(r - t_0)} - \frac{1}{r} \right] \quad (1)$$

$$\Delta H_m = \Delta H_0 \left(1 - \frac{t_0}{r} \right)^3 \quad (2)$$

Various approaches to synthesize the single element nanoparticles have been reported [42–44], which can be largely categorized into “bottom-up” (chemical reduction) and “top-down” method (physical method), respectively. The chemical reduction methods include the inert gas condensation, sol-gel, aerosol, micelle/reverse micelle, and irradiation by UV, γ -ray, microwave, etc [15–23]. For bimetallic or multicomponents nanoparticles, chemical “bottom-up” and physical “top-down” methods have also been used as well. The chemical methods use the co-reduction of dissimilar metal precursors or successive reduction of two-metal salts, which is usually carried out to prepare a core–shell structure of bimetallic nanoparticles. The binary alloys reported in the form of nanoalloys or core–shell structures such as Ag–Au are an alloy system that forms solid solutions and their structures which can be controlled by the reduction order. On the other hand, in the case of alloys that do not favor the formation of solid solutions such as eutectic alloys (Sn-based alloys), little research has been reported because more sophisticated synthetic methods are required due to their oxidative nature. The physical method can be used to synthesize monometallic [24] and bimetallic nanoparticles [25]. Using this technique, nanoparticles can be synthesized in gram quantities directly from the bulk materials without complex reaction procedures. This is suitable for the low melting point metal precursors and their alloys.

Duh et al has reported Sn–3.5Ag– x Cu ($x = 0.2, 0.5, 1.0$) nanoparticles synthesis for the lead-free solder application [26], where the differential scanning calorimetry (DSC) profile showed the melting (endothermic) peak of their SnAgCu alloy nanoparticles at $\sim 216^\circ\text{C}$, which is the melting point of the micron meter-sized SnAgCu alloy powders. No obvious melting point depression from their particles might be due to the surface oxidation or heavy agglomeration of the nanoparticles.

Tin and its alloys are easily oxidized due to their low chemical potential. For nano-sized tin and its alloys, oxidation more easily happens due to the higher surface area to volume ratio of nanoparticles. However, the oxides of Sn and its alloys cannot form interconnection between chips and substrates. Therefore, trapping each nanoparticle is critical for prevention oxidation and capping agents/surfactants can cover the particle surfaces to serve as an effective barrier against the penetration of oxygen [45,46]. Our research has found that some surfactant, such as 1,10-phenanthroline, is an effective capping agent on preventing tin and its alloy nanoparticles from oxidation. When Sn, SnAg, or SnAgCu alloy nanoparticles are formed,

they are instantly coordinated to 1,10-phenanthroline through the two lone pair of chelating nitrogen donor sites adjoining to the two heterocyclic aromatic rings.

1 Size-Dependent Melting Point of Tin Nanoparticles

Figure 3a shows the TEM and inserted HRTEM images of Sn nanoparticles synthesized by using 2.1×10^{-4} mol tin (II) acetate as a precursor in the presence of 0.045 mol surfactants. The average particle size calculated from the TEM picture was around 61 nm and the obvious lattice fringes in HRTEM image implies the crystalline structure. The interplanar spacing was about 0.29 nm which corresponds to the orientation of (200) atomic planes of the tetragonal structure of Sn. In the XRD pattern (Fig. 3b), all the peaks can be indexed to a tetragonal cell of Sn with

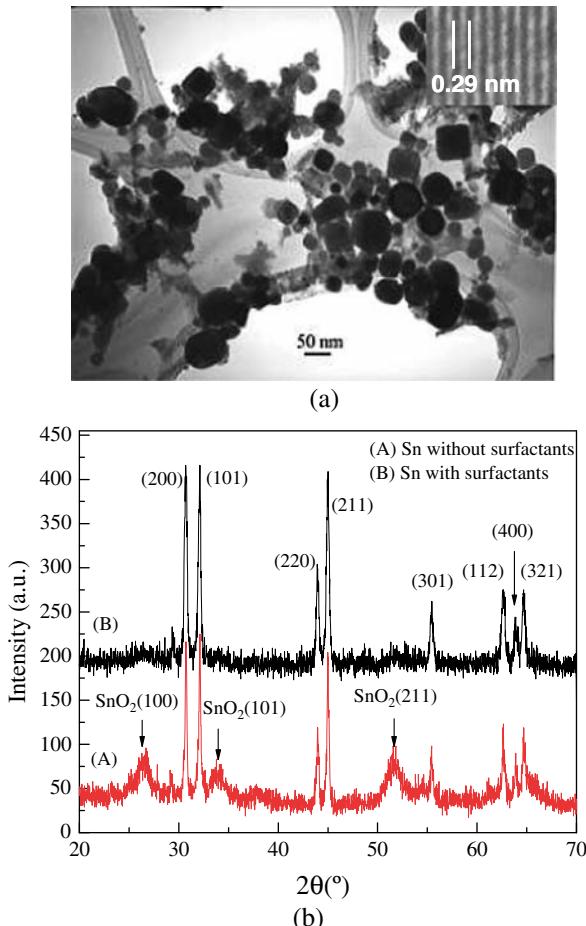


Fig. 3 (a) TEM and inserted HRTEM images; (b) XRD of Sn nanoparticles which were synthesized by using 2.1×10^{-4} mol tin (II) acetate as a precursor in the presence of 0.045 mol surfactants

$a = 0.582$ and $c = 0.317$ nm. The relative intensity of the peaks was consistent with that of the Sn nanoparticles reported elsewhere [27]. No obvious peaks at 20.6° , 33.8° , and 51.8° were found, which match the crystal planes of SnO_2 : (100), (101), and (211) [28]. However, the obvious oxidation peaks of SnO_2 were observed from the x-ray diffraction (XRD) pattern of the Sn nanoparticles without surfactants. The XRD and high-resolution transmission electron microgram (HRTEM) results showed that the Sn nanoparticles synthesized with surfactants in this study were nearly oxide free.

Figure 4a shows the TGA curve of the as-synthesized Sn nanoparticles. The weight loss at below 275°C may be due to the evaporation of absorbed moisture and the decomposition of the surfactants. Above 275°C , the weight gain was

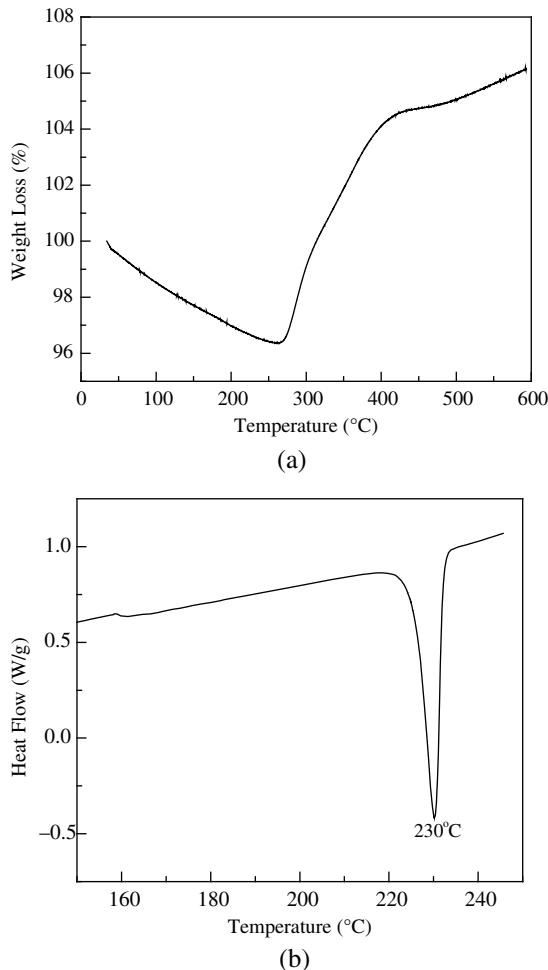


Fig. 4 TGA (a) and DSC (b) curves of the as-synthesized Sn nanoparticles

observed. This may be attributed to thermal oxidation of the pure Sn nanoparticles. The TGA results showed that the Sn nanoparticles were covered with the surfactants by ~ 4 wt% to the particle weight. And these surfactants can protect the Sn nanoparticles from oxidation. Figure 4b shows the thermal profiles of the as-synthesized Sn nanoparticles obtained from DSC where the melting point (T_m) was observed at 230°C . Compared to the T_m of micron-sized Sn particles, the Sn nanoparticles exhibited the T_m depression by $2\text{--}3^\circ\text{C}$.

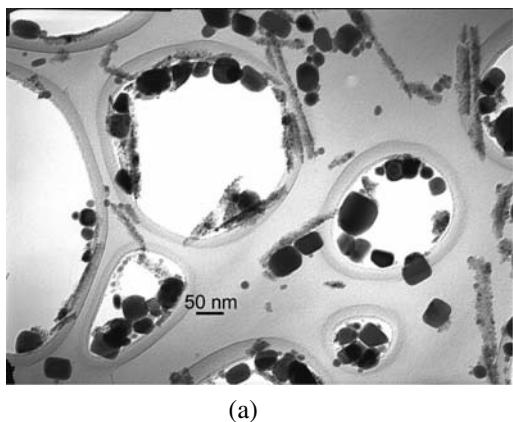
The oxide-free Sn nanoparticles were obtained by using the different molar ratios between a precursor and a surfactant. Thus, different sized Sn nanoparticles can be obtained.

Figure 5a, b shows the TEM image and DSC profile of the Sn nanoparticles which were synthesized by using 4.2×10^{-4} mol tin (II) acetate as a precursor in the presence of 0.045 mol surfactants. The average particle size was around 52 nm. The melting point of the Sn nanoparticles was around 228.0°C , which was 4°C lower than that of micron-sized Sn particles. Figure 5c, d shows the TEM image and DSC profile of the Sn nanoparticles which were synthesized by using 1.1×10^{-3} mol tin (II) acetate as a precursor in the presence of 0.045 mol surfactants. The average particle size was around 85 nm. The melting point of the as-synthesized Sn nanoparticles was 231.8°C , which was still lower than the melting point of micron-sized Sn particles. The TEM image of Sn nanoparticles which were synthesized by using 1.75×10^{-4} mol tin (II) acetate as a precursor in the presence of 0.045 mol surfactants is shown in Fig. 5e. The average particle size was around 26 nm. The melting point of these particles was around 214.9°C , which was 17.7°C lower than that of micron-sized Sn particles. The melting transition of this sample took place over a temperature range of about 60°C , which was much wider than the other three samples. This phenomenon can be attributed to a broadening of the phase transition due to the finite size effect [29] and wide size distribution of the Sn nanoparticles. Schmidt et al. also found a melting temperature range of 60 K for cluster size with the number of atoms between 70 and 200 [8].

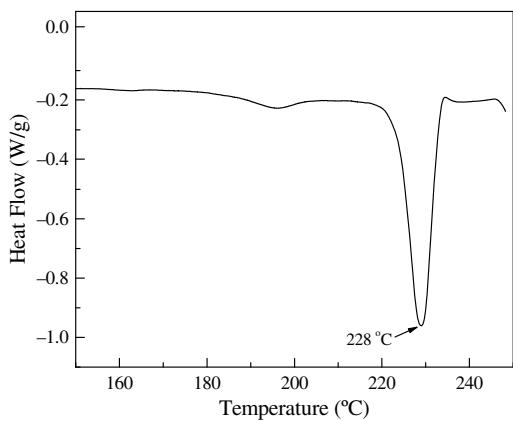
From Table 2, it can be seen that the larger molar ratios between the surfactants and the precursors were used, the smaller the particle size was. This is because a larger amount of surfactants can restrict the growth of Sn nanoparticles. The surfactant molecules coordinate with the nanoclusters, resulting in the capping effect to restrict the particle growth. This was also found by Pal et al. on their gold nanoparticle synthesis that increasing the concentrations of surfactants would limit the particle size through the restriction of particle growth [30]. The DSC results showed size-dependent melting depression behavior and size-dependent latent heat of fusion.

Figure 6 shows the size dependence of the melting points of the synthesized Sn nanoparticle powders, which was compared with Lai et al.'s model [3]. The solid line was calculated from equation (1). Lai et al. obtained this equation based on the model of Hanszen [31] in which it was assumed that the solid particle was embedded in a thin liquid overlayer and the melting temperature was taken to be the temperature of equilibrium between the solid sphere core and the liquid overlayer of a given critical thickness t_0 .

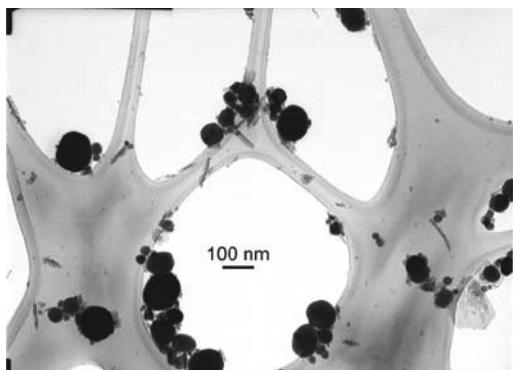
Fig. 5 TEM image and thermal behavior of Sn nanoparticles which were synthesized by using 4.2×10^{-4} mol (a) and (b), 1.1×10^{-3} mol (c) and (d), 1.75×10^{-4} mol (e) and (f) tin (II) acetate as a precursor in the presence of 0.045 mol surfactants



(a)



(b)



(c)

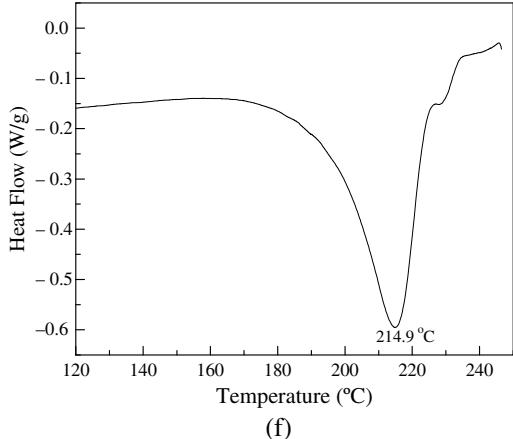
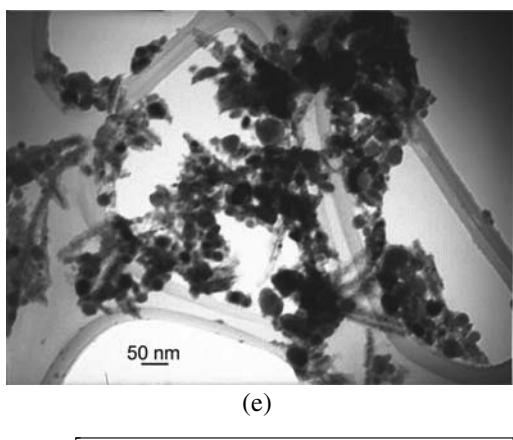
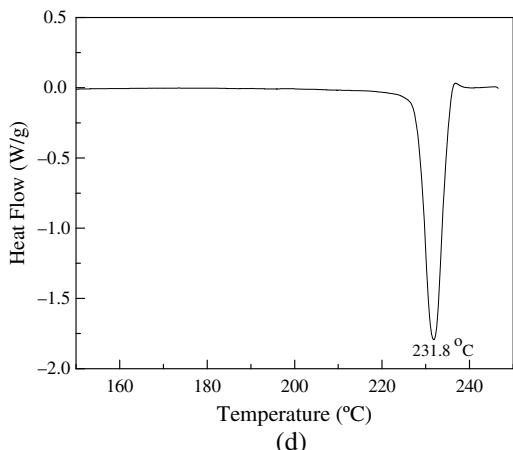
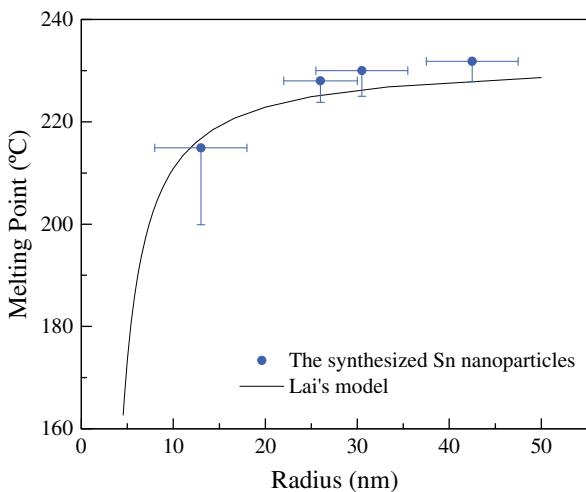
Fig. 5 (continued)

Table 2 Melting points and heat of fusion of different sized Sn nanoparticles

Samples	Surfactants/ precursor	Average size (nm)	Melting point (°C)	ΔH (J/g)
1	N/A	Micron size	232.6	60.0
2	$0.045/1.1 \times 10^{-3}$	85 ± 10	231.8	32.1
3	$0.045/2.1 \times 10^{-4}$	61 ± 10	230.0	24.5
4	$0.045/4.2 \times 10^{-4}$	52 ± 8	228.0	28.7
5	$0.045/1.75 \times 10^{-4}$	26 ± 10	214.9	27.4

Fig. 6 Size dependence of the melting points of Sn nanoparticles (Y error bars stand for the melting temperature from the onset point to the peak point of DSC curves). The solid line is calculated from equation (1) by Lai *et al* [3]



$$T_r = 232 - 782 \left[\frac{\sigma_{sl}}{15.8(r - t_0)} - \frac{1}{r} \right] \quad (3)$$

where T_r (°C) is the melting points of particles with a radius of r (Å). t_0 (Å) is the critical thickness of the liquid layer. σ_{sl} is the interfacial surface tension between the solid and the liquid, the liquid and its vapor. From their experiment, σ_{sl} is determined to be 48 ± 8 mN/m and the best fit for t_0 is 18 Å.

It is found that our experimental results are in reasonable agreement with the Lai *et al.*'s model and the melting point depends nonlinearly on the cluster radius.

From Table 1, it could also be found that the latent heat of fusion of the different sized Sn nanoparticles was smaller than that of micron-sized Sn powders. Ercolelli *et al.* have found that ΔH_m of gold nanoparticles decreases steadily from 114 meV/atom (bulk) to 23 meV/atom ($N = 879$) and 10 meV/atom ($N = 477$) by molecular dynamic simulation [32]. It is also found by experiments that the normalized heat of fusion of Sn nanoparticles decreases markedly from the bulk

value (58.9 J/g) by as much as 70% when the particle size is reduced, which can be interpreted as a solid core melting following the gradual surface melting for small particles [3]. We tried to compare the heat of fusion (ΔH_m) of our synthesized Sn particles with Lai et al.'s model as well [3]. It was found that the ΔH_m values of our Sn particles were smaller than those of their models. This might be due to the existence of the surfactants, solvents, and some small amounts of oxides which might develop from the heating process of DSC test.

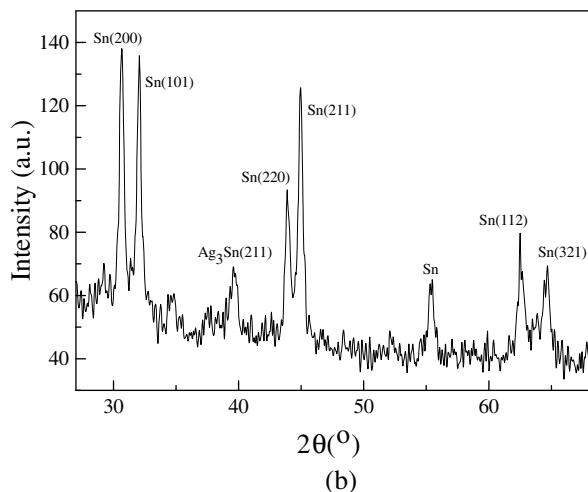
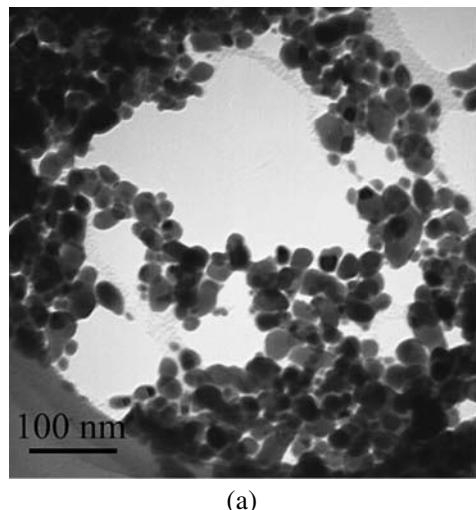
2 Size-Dependent Melting of Tin/Silver Alloy Nanoparticles

Figure 7a shows the TEM image of the SnAg alloy nanoparticles synthesized by using 7.4×10^{-4} mol tin (II) 2-ethylhexanoate and 3.0×10^{-5} mol silver nitrate as precursors in the presence of 5.6×10^{-4} mol surfactants at 0°C. The average diameter of the particles was ~ 24.0 nm. The XRD patterns of the as-synthesized SnAg alloy nanoparticles are shown in Fig. 7b. In addition to the peaks indexed to a tetragonal cell of Sn with $a = 0.582$ and $c = 0.317$ nm, the Ag_3Sn phase ($\sim 39.6^\circ$) was found in the XRD patterns, indicating the successful alloying of Sn and Ag after the reduction process [26,33]. No prominent oxide peak was observed from the XRD patterns. This indicates that the surfactants could help to protect the synthesized SnAg alloy nanoparticles from oxidation [34,35].

Figure 8 shows the HRTEM image of the as-synthesized SnAg alloy nanoparticles, where the particles showed core–shell structures. The dark core and the brighter shell correspond to the crystalline metal and the amorphous organic surfactants, respectively. The surfactant shells on the particle surface protected the SnAg alloy nanoparticles from oxidation.

Figure 9a shows the TGA curve of the dried SnAg alloy nanoparticles in nitrogen atmosphere. The weight loss at below 180°C might be due to the evaporation of small amount of absorbed moisture and surfactants. Above 180°C, the weight gain was observed, which was attributed to thermal oxidation of the SnAg alloy nanoparticles. Figure 9b displays the thermal profile of the dried SnAg alloy nanoparticles. The melting point of the SnAg nanoparticles was found at $\sim 209.5^\circ\text{C}$, about 13°C lower than that of the micron-sized 96.5Sn3.5Ag particles (222.6°C). This was an obvious size-dependent melting point depression. The latent heat of fusion (ΔH_m) of the SnAg alloy nanoparticles (24.2 J/g) was smaller than that of micron-sized 96.5Sn3.5Ag powders (68.6 J/g). It has already been observed by experiments that the normalized heat of fusion of Sn nanoparticles decreases markedly from the bulk value by as much as 70% when the particle size is reduced. This can be interpreted as a solid core melting following the gradual surface melting for small particles [3]. The recrystallization peak of the as-synthesized SnAg alloy nanoparticles was found at 109.2°C, which was 91.2°C lower than that of the micron-sized 96.5Sn3.5Ag particles (200.4°C). Such a supercooling effect in the recrystallization of the melted Sn nanoparticles has already been observed [36], which can be explained by the critical-sized stable grain that has to form for solidification to take place [37]. Solidification

Fig. 7 The TEM image (**a**) and XRD patterns (**b**) of the SnAg alloy nanoparticles which were synthesized by using 7.4×10^{-4} mol tin (II) 2-ethylhexanoate and 3.0×10^{-5} mol silver nitrate as precursors in the presence of 5.6×10^{-4} mol surfactants



of the melted nanoparticles can only occur once the temperature is low enough so that the critical-size solidification grain can be accommodated in the small volume.

The TEM image and the corresponding DSC curves of SnAg alloy nanoparticles which were synthesized by using 7.4×10^{-4} mol tin (II) 2-ethylhexanoate and 3.0×10^{-5} mol silver nitrate as precursors in the presence of 1.1×10^{-3} mol surfactants at 20°C are shown in Fig. 10a, b. The average diameter of the as-synthesized nanoparticles was ~ 10 nm (Fig. 10a). From the DSC studies, the first peak melting temperature was 194.3°C , which was about 28.3°C lower than that of micron-sized $96.5\text{Sn}3.5\text{Ag}$ particles. The onset peak temperature was 172.6°C , which was 46.9°C lower than that of micron-sized particles (219.5°C). The melting transition of this sample took place over a temperature range of about 21.7°C , much wider than the

Fig. 8 HRTEM image of SnAg alloy nanoparticles which were synthesized by using 7.4×10^{-4} mol tin (II) 2-ethylhexanoate and 3.0×10^{-5} mol silver nitrate as precursors in the presence of 5.6×10^{-4} mol surfactants

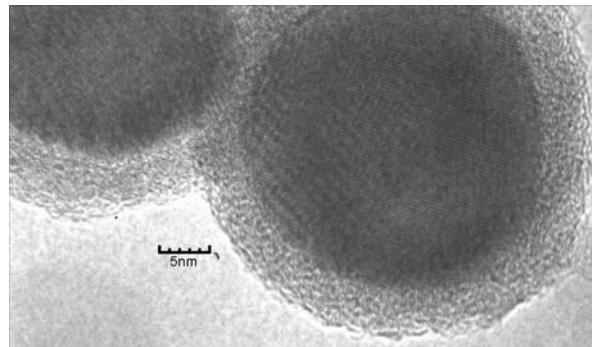


Fig. 9 The TGA (a) and DSC (b) curves of the SnAg alloy nanoparticles which were synthesized by using 7.4×10^{-4} mol tin (II) 2-ethylhexanoate and 3.0×10^{-5} mol silver nitrate as precursors in the presence of 5.6×10^{-4} mol surfactants

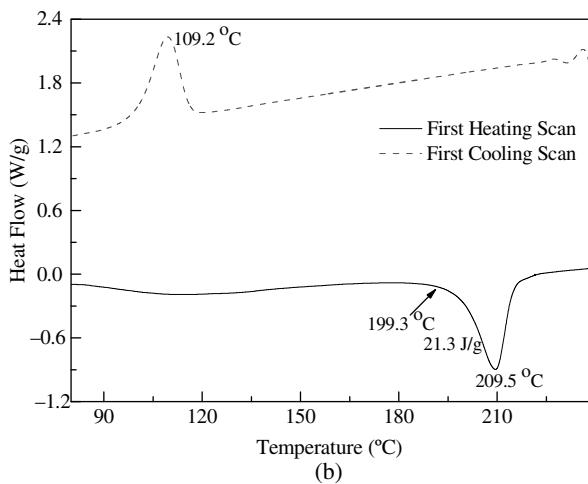
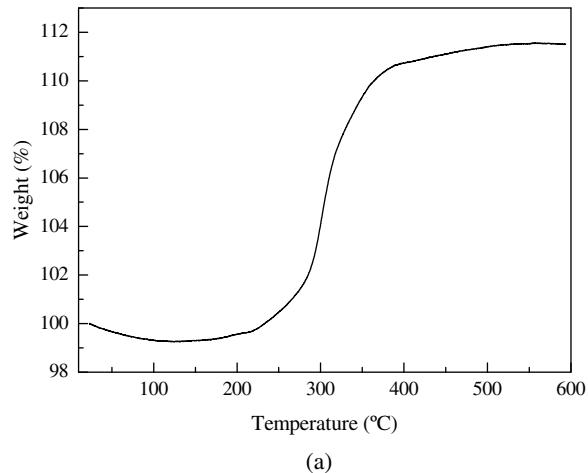
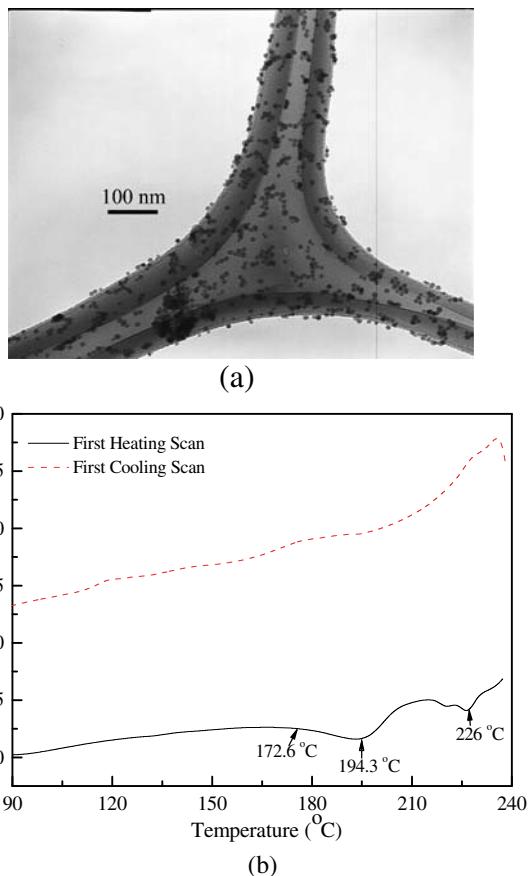


Fig. 10 The TEM image (a) and DSC curves (b) of the SnAg alloy nanoparticles which were synthesized by using 7.4×10^{-4} mol tin (II) 2-ethylhexanoate and 3.0×10^{-5} mol silver nitrate as precursors in the presence of 1.1×10^{-3} mol surfactants



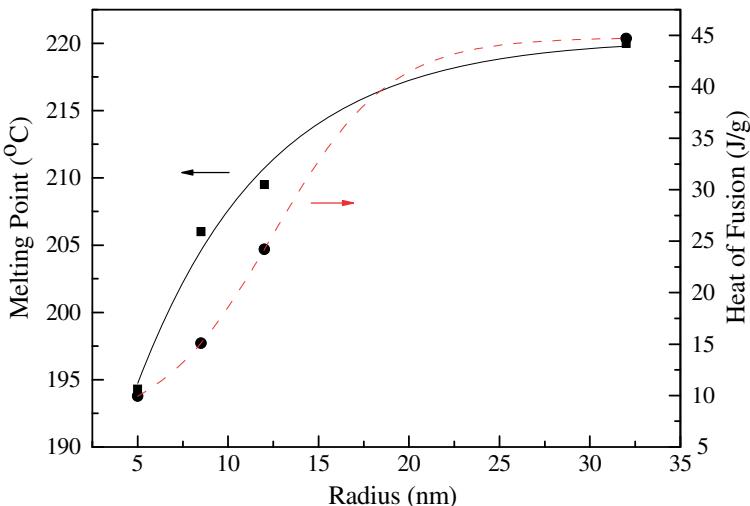
micron-sized particles (3.1°C). This phenomenon can be attributed to a broadening of the phase transition due to the finite size effect [38]. A small peak at 226.0°C was also observed in the first heating scan, which might have come from the melting of a very small amount of larger sized particles existing in this sample.

The melting point and latent heat of fusion for micron-sized 96.5Sn3.5Ag are 222.6°C and 68.5 J/g , respectively.

Table 3 shows the melting point and latent heat of fusion of different sized SnAg alloy nanoparticles. We plotted the melting point and ΔH_m vs. the corresponding particle radius in Fig. 11. Both the particle size-dependent melting point depression and latent heat of fusion have been observed. This is due to the surface premelting of nanoparticles. It has already been found that surface melting of small particles occurs in a continuous manner over a broad temperature range, whereas the homogeneous melting of the solid core occurs abruptly at the critical temperature T_m [39,40]. For smaller size metal nanoparticles, the surface melting is strongly enhanced by curvature effects. Therefore with the decreasing of particle size, both the melting point and latent heat of fusion will decrease too. Among

Table 3 The melting points and heat of fusion of the as-synthesized different sized SnAg alloy nanoparticles

No.	Surfactants (mol)	Reaction temperature (°C)	Average diameter (nm)	Melting point (°C)	ΔH_m (J/g)
1	5.6×10^{-4}	25	64	220.0	44.7
2	5.6×10^{-4}	0	24	209.5	24.2
3	8.0×10^{-4}	≈ 10	17	206.0	15.1
4	1.1×10^{-3}	≈ 20	10	194.3	9.95

**Fig. 11** The relationship between the radius of the as-synthesized SnAg alloy nanoparticles and their corresponding melting points and heat of fusion, respectively

all the synthesized particles in Table 3, the 10 nm (average diameter) SnAg alloy nanoparticles have a melting point at $\sim 194.3^\circ\text{C}$, which will be a good candidate for the low melting temperature lead-free solders and can solve the issues from the high temperature reflow for the micron-sized lead-free particles.

3 Size-Dependent Melting of Tin/Silver/Copper Alloy Nanoparticles

SnAgCu alloy nanoparticles were synthesized by a chemical reduction method. Figure 12 shows the TEM image of the as-synthesized SnAgCu alloy nanoparticles. It can be calculated that the average diameter of the particles is around 22 nm.

The XRD patterns of the as-synthesized SnAgCu alloy nanoparticles are shown in Fig. 13. In addition to the peaks indexed to a tetragonal cell of Sn with $a = 0.582$

Fig. 12 The TEM image of ~22 nm SnAgCu alloy nanoparticles

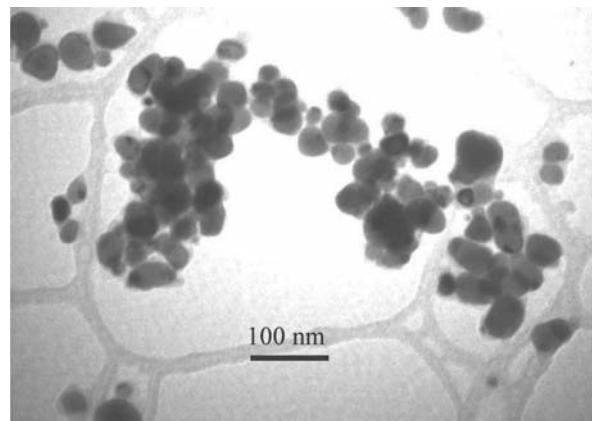
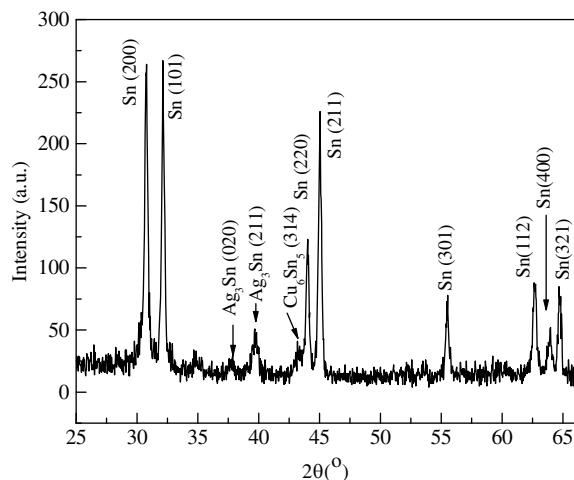


Fig. 13 The XRD patterns of ~22 nm SnAgCu alloy nanoparticles



and $c = 0.317$ nm, the Ag_3Sn phase (39.6°) was found in the XRD patterns, indicating the successful alloying of Sn and Ag after the reduction process [26,33,41]. At the same time, Cu_6Sn_5 was formed which was due to the alloying of Sn and Cu [26]. No prominent oxide peaks were observed from the XRD patterns. This indicates that surfactants can help protect the synthesized SnAgCu alloy nanoparticles from oxidation [34,41]. The HRTEM characterizations already showed that the surfactants covered the particle surface and formed a core–shell structure [41]. The core was from the crystalline metal particles and the shell was from the amorphous surfactants. The amorphous surfactant shells on the particle surface helped prevent the diffusion of oxygen to SnAgCu alloy nanoparticles.

Figure 14 shows the TGA curve of the as-synthesized dried SnAgCu alloy nanoparticles in a nitrogen atmosphere. The weight loss below 180°C might be due to the evaporation or decomposition of a small amount of absorbed moisture and

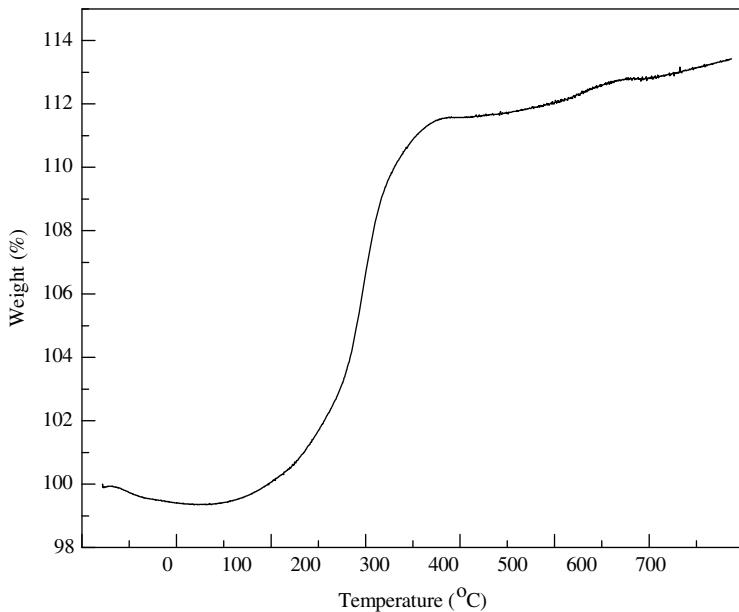


Fig. 14 The TGA curve of ~ 22 nm SnAgCu alloy nanoparticles

surfactants. Above 180°C , the weight gain was observed, which can be attributed to thermal oxidation of the SnAgCu alloy nanoparticles.

The thermal properties of the as-synthesized SnAgCu alloy nanoparticles were studied by a differential scanning calorimeter (Fig. 15). In the first heating scan of the DSC curve, an endothermic peak point at $\sim 207^\circ\text{C}$ was obtained, which is around $10\text{--}12^\circ\text{C}$ lower than the melting point of micron-sized SnAgCu ($217 \sim 219^\circ\text{C}$) alloy particles. In the first cooling scan, the supercooling of the nanoparticles with a crystallization peak at 113.3°C was observed, lower than that of micron-sized SnAgCu alloy nanoparticles (145°C). Such a supercooling effect in the recrystallization of the melted Sn and SnAg alloy nanoparticles has already been observed [36,41], which can be explained by the critical-sized stable grain that needs to form in order for solidification to take place [37]. Solidification of the melted nanoparticles can only occur once the temperature is low enough so that the critical-size solidification grain can be accommodated in the small volume.

Figure 16a shows the TEM image of $10\text{--}13$ nm SnAgCu alloy nanoparticles which were synthesized by the chemical reduction method at a different reaction temperature. From DSC studies (Fig. 16b), the peak melting temperature was 199°C , which was around 20°C lower than that of micron-sized $96.5\text{Sn}3.0\text{Ag}0.5\text{Cu}$ particles. The onset peak temperature is 177°C , 37.5°C lower than micron-sized particles (214.5°C). The melting transition of this sample took place over a temperature range of about 22°C . This phenomenon can be attributed to broadening of the phase transition due to the finite size effect [38].

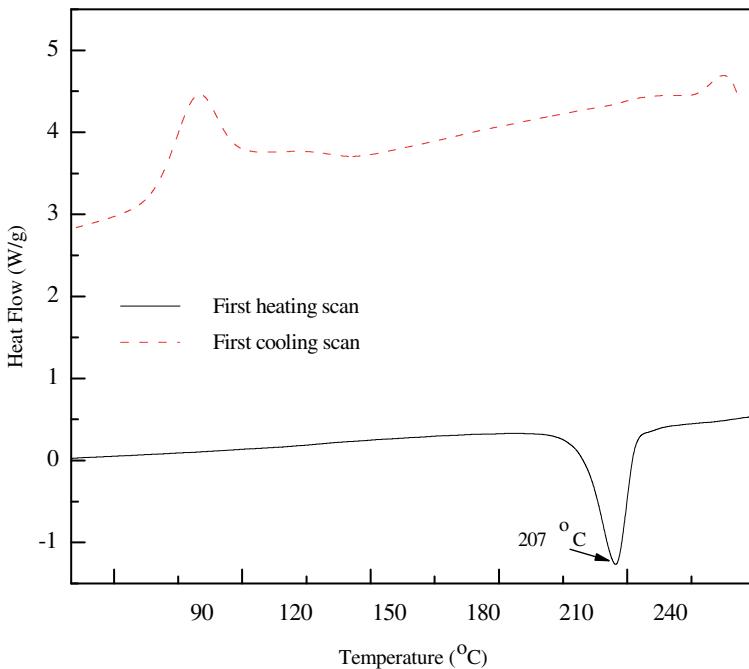


Fig. 15 The DSC curves of ~ 22 nm SnAgCu alloy nanoparticles

Table 4 shows the melting point, latent heat of fusion, and recrystallization temperature of different sized SnAgCu alloy nanoparticles. Both size-dependent melting point depression and latent heat of fusion have been observed. It has already been found that surface melting of small particles occurs in a continuous manner over a broad temperature range, whereas homogeneous melting of the solid core occurs abruptly at the critical temperature T_m [40]. For smaller sized metal nanoparticles, surface melting is strongly enhanced by the curvature effects. Therefore, with the decrease of particle size, both the melting point and latent heat of fusion decrease as well. Among all the synthesized particles in Table 4, the 10–13 nm (average diameter) SnAgCu alloy nanoparticles have the lowest melting point at $\sim 199^\circ\text{C}$, indicating it to be a good candidate for lead-free solders as its melting point is comparable to the reflow temperature of the conventional eutectic micron-sized Sn/Pb alloy particles.

4 Wetting Properties of Tin/Silver and Tin/Silver/Copper Alloy Nanoparticle Pastes

The solder paste is a viable interconnecting material, providing electrical, thermal, and mechanical properties applicable to electronics assemblies. It is a homogeneous

Fig. 16 The TEM image (a) and DSC curves (b) of 10–13 nm SnAgCu alloy nanoparticles

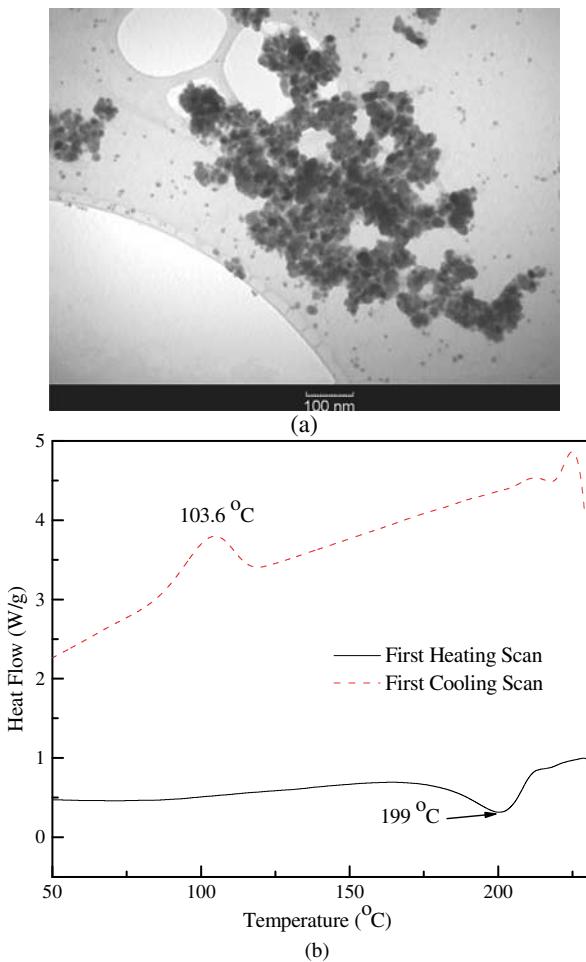
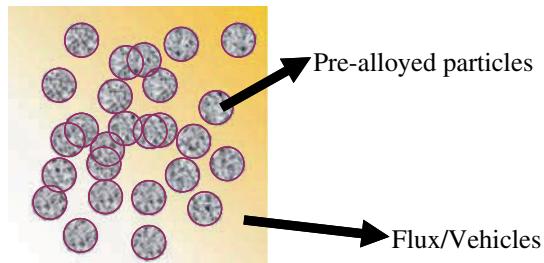


Table 4 The melting and recrystallization points, heat of fusion of different sized SnAgCu alloy nanoparticles

Size	Melting point (°C)	ΔH (J/g)	Recrystallization (°C)
Micron	217	72.2	145
~ 28 nm	210.4	42.7	112.3
~ 22 nm	207.3	31.0	109.9
~ 18 nm	206.1	24.7	108.7
~ 10–13 nm	199	15.9	103.6

Fig. 17 The composition of solder pastes



and kinetically stable mixture of solder alloy powder, flux, and vehicle, which is capable of forming metallurgical bonds at given soldering conditions (Fig. 17).

Solder alloy powders are usually tin/lead (SnPb), lead-free solder materials, such as tin/silver (SnAg), tin/silver/copper (SnAgCu), and tin/copper (SnCu). The flux can cause the alloy powders and the surfaces to be joined to maintain a clean and metallic state. The vehicle is a carrier for the solder powder, provides a desirable rheology, and protects the molten solder and the cleaned substrates from re-oxidation.

The elements used in solder alloys are tin, lead, silver, copper, bismuth, indium, antimony, and cadmium. Phase diagram can be used to determine the eutectic point and the corresponding compositions for the solid solution. Several methods can be used to make the solder alloy powders, such as chemical reduction, decomposition, mechanical processing of solid metal, and atomization of liquid metals. Chemical reduction method and decomposition can be used to prepare fine solder alloy particles. Mechanical processing of solder metal can be used to prepare flake-like particles, while powders to be used in a solder paste are mostly produced by atomization because of desirable inherent morphology and shape of resulting particles.

Melting range, flow rate, and particle morphologies are important parameters for solder powders. The melting range of solder powders are usually tested by differential scanning calorimeter (DSC). Table 1 shows the melting point of SnPb solders and lead-free solders. The particle size, size distribution, and shapes are usually tested by scanning electron microscopy (SEM), light scattering, and so on.

Some of the other physical properties, like viscosity, surface tension, thermal properties, electrical properties, and thermal expansion coefficient are also very important for the solder materials. For the SnPb alloys, the surface tension, thermal and electrical conductivity decrease with increasing lead content, while viscosity and thermal expansion coefficient increase with increasing lead content.

The function of fluxes in a solder paste is to chemically clean the surfaces to be joined, to clean the surface of solder powder, and to maintain the cleanliness of both substrate surface and solder powder surface during reflow so that a metallic continuity at the interface and a complete coalescence of the solder powder during reflow can be achieved (Fig. 18).

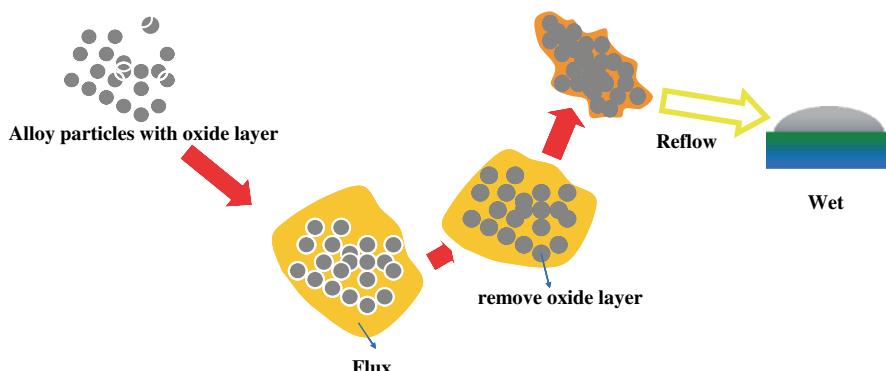


Fig. 18 The function of flux

There are several kinds of fluxes:

- Type R-rosin flux, the weakest, containing only rosin without the presence of activator.
- Type RMA (mildly activated rosin) is a system containing both rosin and activator.
- Type RA is a fully activated rosin and resin system, having a higher flux strength than the RMA type.
- Type OA is an organic acid flux and possessing high fluxing activity and is generally considered corrosive.
- Type SA (synthetic activator) is designed to improve fluxing activity on hard-to-solder surfaces. The SA flux displays better wetting ability than RA flux and is equivalent to OA flux.

The major components of unmodified rosin are abietic acid, isopimaric acid, neoabietic acid, pimaric acid, dihydroabietic acid, and dehydroabietic acid. The chemical structures of these acids are shown in Fig. 19.

The main compositions of vehicles we used in this section are tackifier, solvent, antioxidation agent, and surfactants.

The tackifier is typically a medium-to-high viscosity, high surface tension liquid serving to wet the printed circuit board and the component, and retain the component in position during handling and reflow soldering. The tackifier usually comprises one or more alcohols, aromatic hydrocarbon solvents, aliphatic hydrocarbon solvents, or polymers.

A suitable solvent vehicle includes any solvent which is chemically inert with the other components in the flux. Three important parameters for solvents are boiling point, viscosity, and polarity. Alcohols are usually used as solvents. The more hydroxyl groups the solvents have, the better activity the flux will have.

Antioxidation agents are used to protect the Sn and its alloy nanoparticles from oxidation. Butylated hydroxytoluene (BHT) and hydroquinone (Fig. 20) are

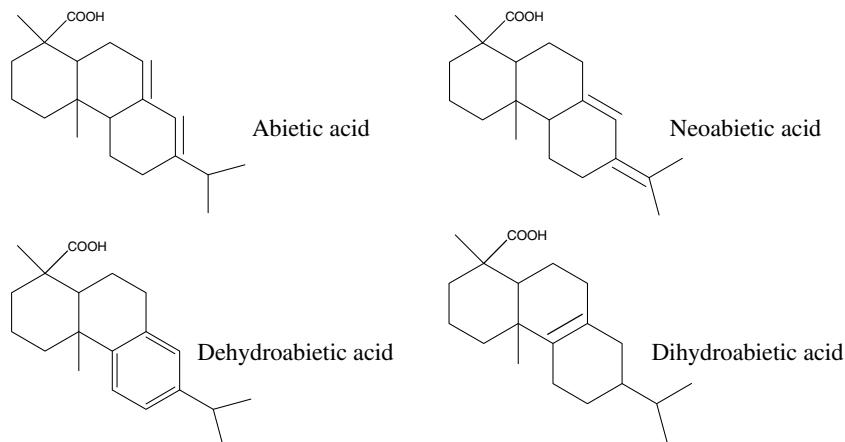


Fig. 19 The chemical structures of the acids in rosin

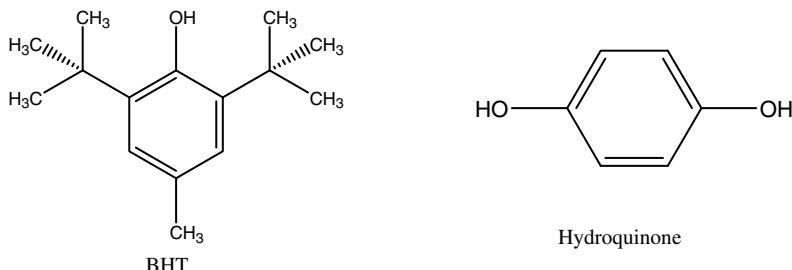


Fig. 20 The chemical structures of BHT and hydroquinone

the typically used antioxidant agents. Oxygen reacts preferentially with BHT or hydroquinone rather than oxidizing Sn alloy, thereby protecting them from oxidation.

The solder surfactant is a compound which improves the solder wetting rate of a surface and enables better and more uniform spreading of molten solder across the surface to be soldered. Suitable surfactants include polybasic acids, e.g., polycarboxylic acids such as dicarboxylic and tricarboxylic acids. The dibasic acids typically have 4–10 carbon atoms. Suitable tricarboxylic acids typically comprise acids having 6–7 carbon atoms. Other suitable surfactants include hydroxyl-substituted polybasic acids, such as tartaric acid and citric acid. The selected surfactant is present in the flux mixture in an amount of at least 1 wt.% of the resultant flux mixture.

The solderability, flow properties, wetting properties, and solidification are important physical properties for the solder pastes. Solderability is the ability to achieve a clean, metallic surface on solder powder and on substrates during the dynamic heating process so that a complete coalescence of solder powder particles

and good wetting of molten solder on the surface of the substrates can be formed. Solderability depends on fluxing efficiency by the solder pastes and the quality of surface of substrates.

When heat is applied to the pastes through any means, the paste tends to spread or slump due to gravity and thermal energy generated. The surface energies of the liquid and the solid substrates are key factors in determining the spreading and wetting properties. For a system with liquid to wet the solid substrate, the spreading occurs only if the surface energy of the substrate to be wetted is higher than that of the liquid to be spread.

The solidification of liquid metal occurs by nucleation and crystal growth. Dendritic growth during solidification is a common phenomenon in pure metals and alloys. The heat flow during the solidification and the crystal structure of the alloy are crucial factors to the properties and structure of the solidified alloy.

Application techniques for the solder pastes usually include printing and paste dispensing. Printing is a viable method to economically produce accurate and reproducible transfer of paste onto the designed pattern, but only suitable for flat surface. However, paste dispensing can be used for irregular surface or hard-to-reach area.

Conduction reflow is most suitable for the assemblies with flat surfaces, composed of thermal conductive materials as the substrates and with single-side component/device populations (fast heating rate and operational simplicity). Infrared reflow is a dynamic process. The precise temperature profile depends on solder alloy composition, properties of the pastes, and assembly involved. Some other reflow methods, such as vapor phase reflow, convection reflow, hot gas reflow, resistance reflow, laser reflow, and induction reflow, are also being used.

Residue is composed of polar organics, nonpolar organics, ionic salts, and metal salts of organics. The cleaning solvents include trichlorotrifluoroethane, 1,1,1-trichloroethane, acetone, methylene chloride, low carbon-chain alcohols, and water. Basic techniques: vapor degreasing, liquid spray, liquid immersion, high pressure spray, and liquid immersion with ultrasonic acid.

A certain amount of fluxing agent, solvent, antioxidation agent, surfactants, and tackifier were mixed together to prepare the flux and vehicle for the SnAg or SnAgCu alloy nanoparticles. Table 5 shows the composition of the flux and vehicle we made and their corresponding chemical structures.

The SnAg alloy nanoparticles (sample 1 in Table 3) were mixed with an acidic-type flux to form the nanosolder pastes at room temperature. A copper foil was cleaned by hydrochloric acid to get rid of the oxide layer and then rinsed with DI water for four times. Thereafter, the nanosolder pastes placed on top of the copper foil were put into a 230°C oven in an air atmosphere for 5 min. The cross-section of the sample after reflow was shown in Fig. 21. It was observed that the SnAg alloy nanoparticles completely melted and wetted on the cleaned copper foil surface. The energy-dispersive spectroscopy (EDS) results revealed the formation of the intermetallic compound (IMC) of Cu₆Sn₅, which showed scallop-like morphologies in Fig. 21. The thickness of the IMC was approximately 4.0 μm. Further studies

Table 5 The composition of flux and vehicle made by our group

Fluxing agent	
Tackifier	
Surfactant	
Solvent	
Antioxidation	

Fig. 21 SEM image of the cross-section of the wetted SnAg alloy nanoparticles (sample 1 in Table 3) on the cleaned copper foil [47]

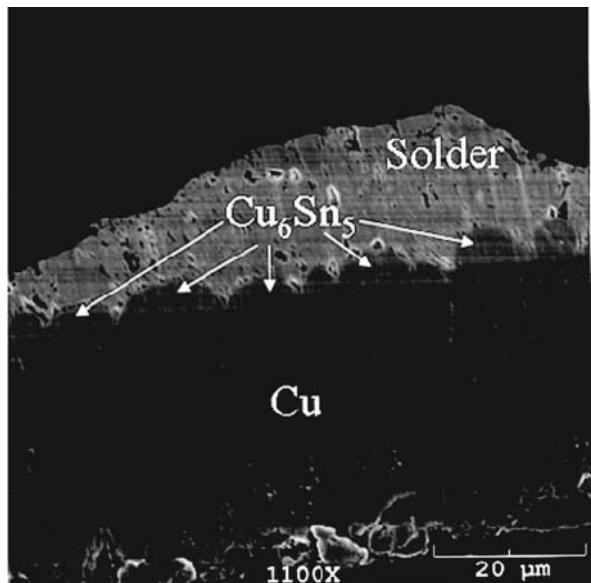
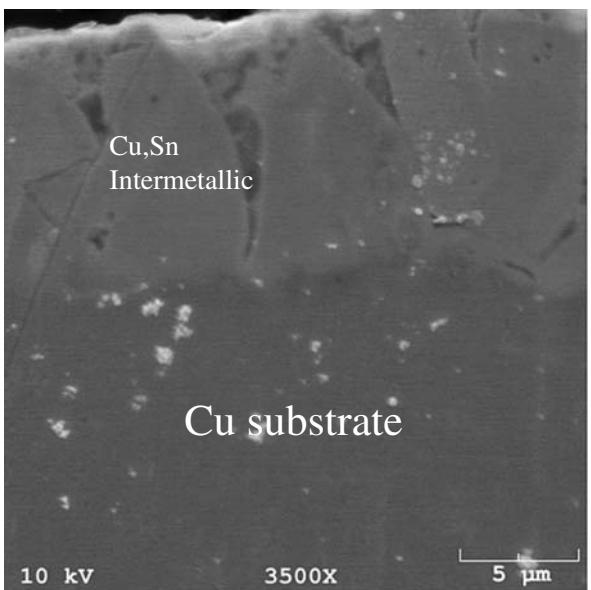


Fig. 22 SEM image of the cross-section of the wetted 50 nm SnAgCu alloy nanoparticles on copper foil



energy-dispersive spectroscopy (EDS) results revealed the formation of the intermetallic compounds (Cu_6Sn_5), which showed scallop-like morphologies in Fig. 22. The thickness of intermetallic compounds was approximately 10.0 μm . The wetting properties of different sized SnAgCu alloy nanoparticles at different reflow temperatures need to be further studied.

5 Conclusion

This chapter addresses synthesis routes for lead-free nanoparticles including a tin single element and alloys and their structures and thermal behavior based on melting point depression. During this study, we observed obvious melting point depression behavior of synthesized nanoparticles and learned, most of all, that the surface oxidation of the non-noble nanoparticle is a critical obstacle in the case where melting of the particles is needed. Therefore we found the surface capping by functionalization was imperative and when and how the surface can be capped were very important factors to obtain oxide-free non-noble metal nanoparticles, together with the careful selection of the surfactant. Finally in order to make nanosolder pastes, the appropriate flux vehicle and the reflow profile optimization were very important, which are also the lessons we learned.

The aim of this study is the reduction of the reflow temperature in microelectronic packaging processes which may be a representative case study to employ cutting-edge nanoscience and technology into the real world research and development arena.

References

1. M. Abtew, G. Selvaduray “Lead-free solders in microelectronics”. *Materials Science & Engineering: R: Reports* 2000; 27, 95.
2. C. R. M. Wronski “Size dependence of melting point of small particles of tin”. *Br. J. Appl. Phys.* 1967; 18, 1731.
3. S. L. Lai, J. Y. Guo, V. Petrova, G. Ramanath, L. H. Allen “Size-dependent melting properties of small tin particles: nanocalorimetric measurements”. *Phys. Rev. Lett.* 1996; 77(1): 99–102.
4. T. Bachels, H. J. Guntherodt, R. Schafer “Melting of isolated tin nanoparticles”. *Phys. Rev. Lett.* 2000; 85(6): 1250–1253.
5. S. J. Zhao, S. Q. Wang, D. Y. Cheng, H. Q. Ye “Three distinctive melting mechanisms in isolated nanoparticles”. *J. Phys. Chem. B* 2001; 105(51): 2857–12860.
6. A. A. Shvartsburg, M. F. Jarrold “Solid clusters above the bulk melting point”. *Phys. Rev. Lett.* 2000; 85, 2530.
7. C. L. Cleveland, W. D. Luedtke, U. Landman “Melting of gold clusters: Icosahedral precursors”. *Phys. Rev. Lett.* 1998; 81, 2036.
8. M. Schmidt, R. Kusche, B. Issendroff, H. Haberland “Irregular variations in the melting point of size-selected atomic clusters”. *Nature* 1998; 393(6682): 238–240.
9. L. J. Lewis, P. Jensen, J. L. Barrat “Melting, freezing, and coalescence of gold nanoclusters”. *Phys. Rev. B* 1997; 56, 2248.
10. C. L. Cleveland, U. Landman, W. D. Luedtke “Phase coexistence in clusters”. *J. Phys. Chem.* 1994; 98, 6272.
11. F. G. Shi “Size-dependent thermal vibrations and melting in nanocrystals”. *J. Mater. Res.* 1994; 9, 1307.
12. Q. Jiang, F. G. Shi “Entropy for solid-liquid transition in nanocrystals”. *Mater. Lett.* 1998; 37, 79.
13. L. Allen, R. A. Bayles, W. W. Gile, W. A. Jesser “Small particle melting of pure metals”. *Thin Solid Film* 1986; 144, 297.
14. P. Buffat, J. P. Borel “Size effect on melting temperature of gold particles”. *Phys. Rev. A* 1976; 13, 2287.

15. R. Birringer, H. Gleiter, H. P. Klein, P. Marquart "Nanocrystalline materials an approach to a novel solid structure with gas-like disorder?". *Phys. Lett.* 1984; 102A, 365.
16. B. I. Lee, E. J. A. Pope, "Chemical Processing of Ceramics", Marcel Dekker, New York, 1994.
17. O. G. Raabe, in "Fine Particles", ed. by B. Y. H. Liu, Academic Press, Inc., New York, 1975; p. 60.
18. J. P. Wilcoxon, A. Martino, R. L. Baughmann, E. Klavetter, A. P. Sylvester, "Synthesis of transition metal clusters and their catalytic and optical properties", in "Nanophase and Nanocomposite materials" ed. by S. Komarneni, J. C. Parker and G. J. Thomas, MRS, Pittsburgh, PA, 1993; p. 131.
19. J. Thomas "Preparation and magnetic properties of colloidal cobalt particles". *J. Appl. Phys.* 1966; 37, 2914.
20. G. L. Rochfort, R. D. Rieke "Preparation, characterization, and chemistry of activated cobalt". *Inorg. Chem.* 1986; 25, 348.
21. C. C. Koch "Materials synthesis by mechanical alloying". *Ann. Rev. Mater. Sci.* 1989; 19, 121.
22. K. Klabunde, Y. Li, B. Tan "Solvated metal atom dispersed catalysts". *Chem. Mater.* 1991; 3, 30.
23. F. Mafune, J. Y. Kohno, Y. Takeda, T. Kondow "Dissociation and aggregation of gold nanoparticles under laser irradiation". *J. Phys. Chem. B* 2001; 105, 9050.
24. Y. B. Zhao, Z. J. Zhang, H. X. Dang "Preparation of tin nanoparticles by solution dispersion". *Mater. Sci. Eng.* 2003; A359, 405.
25. Y. B. Zhao, Z. J. Zhang, H. X. Dang "Synthesis of In-Sn alloy nanoparticles by a solution dispersion method". *J. Mater. Chem.* 2004; 14, 299.
26. L. Y. Hsiao, J. G. Duh "Synthesis and characterization of lead-free solders with Sn-3.5Ag-xCu ($x=0.2, 0.5, 1.0$) alloy nanoparticles by the chemical reduction method". *J. Electrochem. Soc.* 2005; 152(9): J105–J109.
27. Y. Kwon, M. G. Kim, Y. Kim, Y. Lee, J. Cho "Effect of capping agents in tin nanoparticles on electrochemical cycling". *Electrochem. Solid-State Lett.* 2006; 9, A34.
28. Y. Wang, J. Y. Lee, T. C. Deivaraj "Controlled synthesis of V-shaped SnO₂ nanorods". *J. Phys. Chem. B* 2004; 108, 13589.
29. Y. Imry, D. Bergman "Critical points and scaling laws for finite systems". *Phys. Rev. A* 1971; 3, 1416.
30. M. Mandal, S. K. Ghosh, S. Kundu, K. Esumi, T. Pal "UV photoactivation for size and shape controlled synthesis and coalescence of gold nanoparticles in micelles". *Langmuir* 2002; 18, 7792.
31. K. J. Hanszen "Theoretische untersuchungen über den schmelzpunkt kleiner kugelchen – Ein Beitrag zur thermodynamik der grenzflächen". *Z. Phys.* 1960; 157, 523–553.
32. F. Ercolelli, W. Andreoni, E. Tosatti "Melting of small gold particles – mechanism and size effects". *Phys. Rev. Lett.* 1991; 66(7): 911–914.
33. H. L. Lai, J. G. Duh "Lead-free Sn-Ag and Sn-Ag-Bi solder powders prepared by mechanical alloying". *J. Electron. Mater.* 2003; 32(4): 215–220.
34. H. J. Jiang, K. Moon, H. Dong, F. Hua, C. P. Wong "Size-dependent melting properties of tin nanoparticles". *Chem. Phys. Lett.* 2006; 429, 492–496.
35. L. Balan, R. Schneider, D. Billaud, J. Ghanbaja "A new organometallic synthesis of size-controlled tin(0) nanoparticles". *Nanotechnol.* 2005; 16(8): 1153–1158.
36. F. Banhart, E. Hernandez, M. Terrones "Extreme superheating and supercooling of encapsulated metals in fullerenelike shells". *Phys. Rev. Lett.* 2003; 20(18): 185502.
37. H. K. Christenson "Confinement effects on freezing and melting". *J. Phys. Condens. Matter.* 2001; 13(11): R95–R133.
38. Y. Imry, D. Bergman "Critical points and scaling laws for finite systems". *Phys. Rev. A* 1971; 3(4): 1416.

39. R. Garrigos, P. Cheyssac, R. Kofman “Melting of lead particles of small sizes – influence of surface phenomena”. *Z. Phys. D* 1989; 12, 497.
40. W. Y. Hu, S. G. Xiao, J. Y. Yang, Z. Zhang “Melting evolution and diffusion behavior of vanadium nanoparticles”. *Eur. Phys. J. B* 2005; 45(4): 547–554.
41. H. Jiang, K. Moon, F. Hua, C. P. Wong “Synthesis and thermal and wetting properties of tin/silver alloy nanoparticles for low-melting point lead-free solders”, *Chem. Mater* 2007; 9(8): 4482–4485.
42. H. Jiang, K. Moon, Y. Sun, C. P. Wong, F. Hua, T. Pal, A. Pal “Tin/indium nanobundle formation from aggregation or growth of nanoparticles”. *J. Nanopart. Res.* 2008; 10(1): 41–46.
43. H. Jiang, L. Zhu, K. Moon, C. P. Wong “The preparation of stable metal nanoparticles on carbon nanotubes whose surface were modified during production”. *Carbon* 2007; 45(3): 655–661.
44. H. Jiang, K. Moon, Z. Zhang, S. Pothukuchi, C. P. Wong “Variable frequency microwave synthesis of silver nanoparticles”. *Journal of Nanoparticle Research* 2006; 8(1): 117–124.
45. H. Jiang, K. Moon, Y. Li, C. P. Wong “Surface functionalized silver nanoparticles for ultrahighly conductive polymer composites”. *Chemistry of Materials* 2006; 18(13): 2969–2973.
46. H. Jiang, K. Moon, J. Lu, C. P. Wong “Conductivity enhancement of nano Ag filled conductive adhesives by particle surface functionalization”. *Journal of Electronic Materials* 2005; 34(11): 1432–1439.
47. H. Jiang, K. Moon, C. P. Wong, “Tin/silver/copper alloy nanoparticles for low temperature solder pastes interconnect”, 2008 IEEE 58th Electronic Components & Technology Conference, May 27–30, 2008; 1400–1404.

Introduction to Nanoparticle-Based Integrated Passives

Ranjith John and Ajay P. Malshe

Abstract This chapter discusses state of the art in nanoengineered integrated passives. In particular, the chapter elaborates nanoparticle-based integrated passives that are being developed using various thin and thick films processes, such as printing. The discussion provides need for next generation of integrated passives and opportunities offered by the nanotechnology to meet those needs. This is a novel area of research and the chapter shares results for integrated capacitors, inductors, and resistors for materials selection, deposition processes, and testing for integrated applications.

Keywords Integrated passives · Nanotechnology · Nanoparticles · Electronic packaging · Capacitor · Resistor · Inductor

1 Introduction and Background

Faster, denser, cheaper: this has been the battle cry of the electronic age and nanotechnology even though in its early stage has shown considerable promise in taking us in further along the road. The dawn of the electronic age was marked by the invention of vacuum tubes. Vacuum tubes (VTs) slowly but surely drove the commercialization of radios, television, radar, and digital computers. But with the invention of transistors, VTs lost their appeal as transistors offered a greater reduction in size which made electronic devices more appealing to consumers. Figure 1 shows the difference in size between a VT and a transistor. Additionally, transistors offered increased reliability in comparison to VT.

Transistors started replacing VTs effectively in electronic devices, but when engineers tried building complex devices such as computers they quickly realized that transistors had their limitations. For instance, assembly workers had to manually

A.P. Malshe (✉)

College of Engineering, University of Arkansas, Fayetteville, AR, USA

e-mail: apm2@uark.edu



Fig. 1 Vacuum tube vs. transistor [1]

solder each component to the circuit using long metal wires. This resulted in reducing the effectiveness of the computer and other electronic devices. So they ran into the problem of building complex circuitry using transistors.

In the summer of 1958 Jack Kilby a Texas Instrument engineer invented the integrated circuit (also known as IC or silicon chip) which revolutionized the electronic industry. He built and tested the first IC which had a transistor and other components built together [2]. Figure 2 shows the first working IC built by Jack Kilby. After some skeptical analysis ICs have successfully replaced discrete transistors in radios, television sets, and many other applications. In fact we can even say that without the invention of ICs none of the modern technology could be possible.

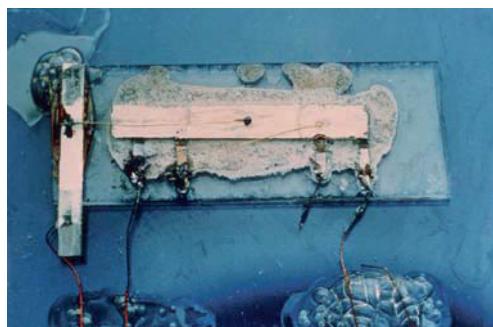


Fig. 2 Jack Kilby's first working IC [2] (Photo courtesy of Texas Instruments)

An IC is a miniaturized electronic circuit consisting of both active components (which have a gain, such as transistors, logic gates) and passive components (which do not have gain, such as resistors, capacitors, inductors). Due to IC's mass production capability, reliability, and building block approach to circuit design, it has made a rapid transition in replacing discrete transistors. Even though an early attraction of the IC technology was a dramatic reduction in size due to the reduction in the number of components in electronic systems, it never came to be a reality. Engineers have worked on incorporating more features and functionality into ICs. The past four decades have seen tremendous progress in miniaturization and integration of ICs in the field of electronics. For instance, we have moved from few hundred transistors to a billion transistors on an IC [3]. This progress has brought electronic packaging technology (the craft of miniaturizing and interconnecting complex circuitry while maximizing the functionality of the circuit) to the forefront of the semiconductor industry. The evolution of the electronic packaging technology started during the early 1980s when through-hole packages such as dual-in-line packages (DIPs) were being replaced by surface mount packages such as quad flat packages (QFP), thin quad flat packages (TQFP). The 1990s saw the emergence of packaging technology called chip scale packaging (CSP) technology (packages are of the size $1.2 \times$ a die area) which offered smaller and lighter package. The most recent technology in packaging is the wafer-level chip scale package (WLCSP) which involves the blending of both the front-end and the back-end processes of packaging. There is strong research in academia and industry toward the development of a system-level integrated package. For instance, NSF Packaging Research Center (PRC) of Georgia Institute of Technology has proposed a system-on-package which is shown in Fig. 3

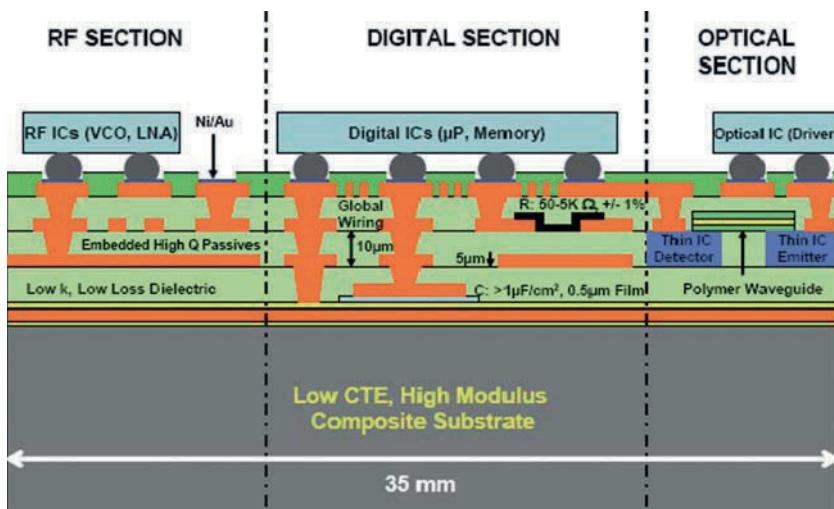


Fig. 3 Schematic of SOP structure [4] (Copyright © 2004 by IEEE)

Even though the development of ICs has advanced by leaps and bounds, passive components are still lagging and have evolved at a slower pace with respect to their size and density. This continuous need for miniaturization of electronic devices is the driving force behind integrated passives (IPs). The rest of this chapter will address the background of passives, their classification; the need for nanoparticle (NP)-based IPs, the advantages and challenges of NP-based IPs, and future work.

2 History of Passive Technology

“Passives” usually refer to resistors, capacitors, and inductors, but could include any device which does not have a gain. The “passive” industry has grown into a multi-billion dollar business which supports electronic products in automotive, telecommunications, computer, consumer, and aerospace industries, for digital and analog–digital applications. Passives are generally used for connecting the complex network of ICs. They primarily serve as decoupling capacitors, bypass filters, and other purposes by which the functionality of the ICs are preserved.

Passives can be classified into two basic types, namely discrete passives (DPs) and integrated passives (IPs). DPs encompass single passive devices which have their own leaded or surface mount technology (SMT) package (technology in which electronic components are mounted directly on the surface of printed circuit boards (PCB)). Today’s handheld devices have greater functionality and the complexities of these devices have greatly increased the number of discrete components. Figure 4 shows a circuit with discrete passives occupying a majority of the space on a mother board.

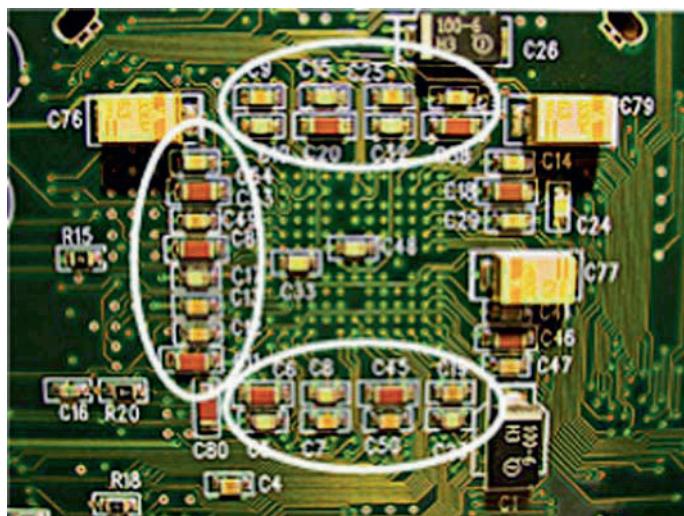


Fig. 4 Surface mount DPs [5] (Photo courtesy of nGimat)

Analysts believed that DPs would be integrated away into ICs; however, exactly the opposite has occurred [6]. In 1984, passives accounted for 25% of all components on a PCB; by 1998 this grew to about 90% [7], and with the incessant growth toward complexity this figure is expected to go up further. Additionally a typical handheld device today averages about 15–20 ICs and 300–400 passives. This results in a 20:1 ratio between passive and active devices [8]. Even though DPs are readily available and cheap, the challenges posed by them indirectly to cost and reliability are vital to the designer. For example, studies show that even though the cost of individual components averages only a few cents the cost of assembly involving DPs is great [9]. Additionally, the reliability cost is another component of indirect cost. For instance, published studies show that the solder joint failure is one of the most important reasons for product failure [10] which leads to hidden cost. So relative to their integrated circuit (IC) counterparts which cannot give up the space (due to increasing functionality of products), passives are being closely investigated to provide the necessary reduction in size, cost while maintaining performance.

The idea of “integrated passives” includes the manufacturing of passives in groups or on a common substrate rather than in their own individual packages. They can be further classified as “embedded passives” and “passive arrays” depending on the packaging technology provided. Embedded passive (EP) technology involves packaging of passive devices inside the primary interconnect substrate rather than being on the surface. Passive arrays are those in which an array of passive devices are packaged in a single SMT case and mounted on the primary substrate. Figure 5 shows the hierarchy of IPs [12].

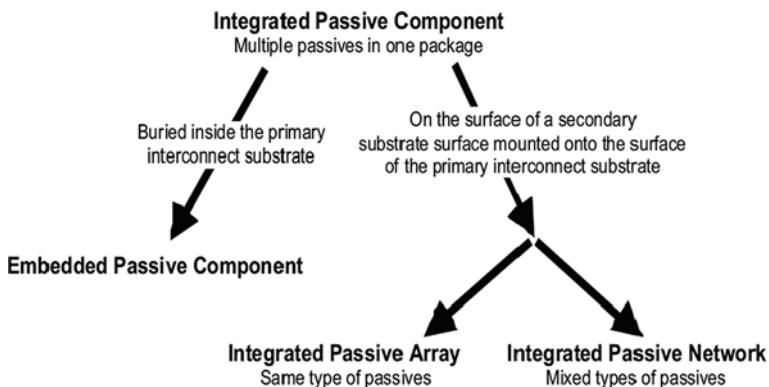


Fig. 5 Hierarchy of IPs [11] (Copyright © 1996 by IMAPS)

IPs are acknowledged to be the most likely alternative to DPs. Since, they promise to offer better reliability and more area saving on a PCB, there is widespread research being done to successfully incorporate them as a part of

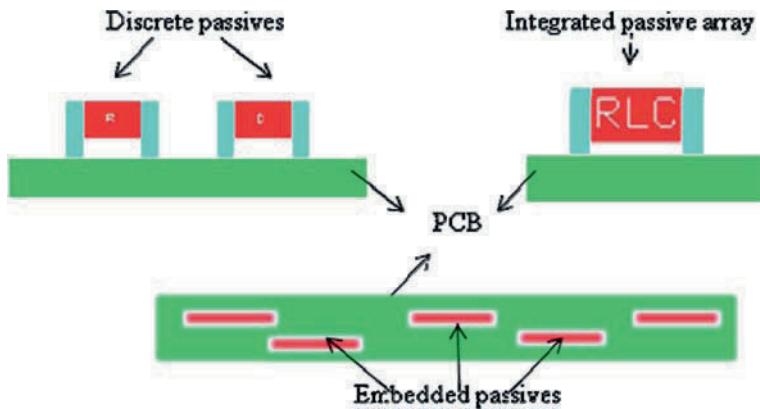


Fig. 6 Simple schematic of size reduction offered by IPs in comparison with DPs

everyday semiconductor technology. Here is a simple figure (Fig. 6) that shows the size reduction of IPs in comparison to DPs.

2.1 Next Generation IP Needs

In 2000 the US market share for passives was estimated to be \$18 billion. This huge multi-billion dollar industry is expected to grow by about 3.7% per year [13]. Table 1 shows the trend in parameter of semiconductor technology which forecasts a paradigm shift in the passive industry.

Table 1 Projected needs for the portable emulator [11]

First year of significant production	Metric	2003	2005	2007	2013
Board assembly cost	c per I/O	0.5	0.45	0.4	0.3
Package I/O pitch	mm	0.5	0.5	0.5	0.5
Substrate lines and spaces	Microns	75	65	65	35
Substrate pad diameter	Microns	225	200	175	125
Max components per cm ²	#/cm ²	50	55	60	25
Frequency on board	MHz	150	250	300	400
Solder	Composition	Lead/lead-free	Lead/lead-free	Lead-free	Lead-free
RF component thickness	mm	2.5	1.5	1.5, MEMS	1.5, MEMS
Passives		0201	0201	EPs	EPs
Product introduction cycle time, platform	Months	10	9	8	6
Product introduction cycle time, spin	Months	4	3	2	1

As per the forecast of NEMI to meet the ever increasing need for miniaturization, reduction of price, while maintaining or increasing the performance of commercial electronic devices (such as cell phones, IPODS, handheld cameras) the passive industry will have to adapt and make the transition from DPs to IPs.

IPs have several benefits over DPs, namely

1. Overall system mass reduction
2. Better electrical performance and reliability
3. Design flexibility and
4. Reduced unit cost.

As indicated in Introduction, DPs which occupy a major portion of the substrate due to their bulky structure and can be replaced by EPs and IP arrays which can either be placed inside a layer of the substrate or can be arrayed into multiple passives which are then mounted on the surface of the substrate thereby freeing up the board space and providing a major reduction in footprint. Also, IPs provide better electrical performance than their discrete counter parts because of the simplicity of design. The issues of parasitic capacitance and resistance which are present in DPs can be reduced greatly with the use of IPs. For instance, compared to their discrete counterparts IP arrays do not require as many leads for electrical connection thereby reducing the effect of parasitic and they can be reduced further by the application of EPs which are fabricated under the surface of the board. This goes a long way in reducing the failure of product by reducing the most common failure seen in the industry – failure of solder joints. IPs also provide the designer with the design flexibility by giving options as to where and how to address the placement of passives. Additionally, since IPs can be formed simultaneously with active devices the extra steps in manufacturing and assembly are eliminated.

2.2 Advantages of EPs

By definition EP technology involves the processes of burying passive devices inside the substrate (such as ceramic, FR4, or silicon). EPs have been recognized as the most promising technology to replace DPs as they provide several advantages over DPs. As seen in Fig. 7 EPs offer a sizable advantage in size reduction but apart from that they also completely dissolve the need for solder joints thereby eradicating failure of solder joints [14]. Additionally, due to the fact that they are embedded in the substrate they provide better electrical performance as they have shorter interconnects paths which reduce the signal delay and parasitic effects. Also, they provide increased silicon device efficiency and eliminate the need for separate packaging process and provide efficient circuit design [15–22]. Finally, placement time for discrete components is one of the major factors that determine the cost of passive components; by using EPs which are mass produced during the fabrication of substrate there is a substantial reduction in assembly time and cost.

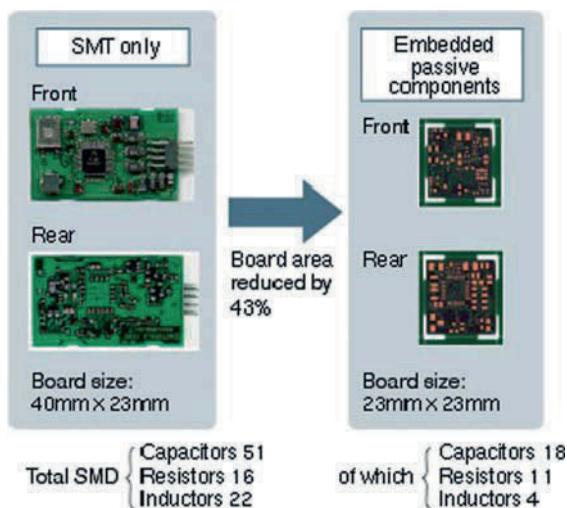


Fig. 7 Real estate saving and size reduction using EPs [64] (Photo courtesy of TechOn)

So if EPs offer all these features which will take us to our destination of smaller, cheaper, and better products then why are they only used for niche applications? The answer to this question deals with multiple issues such as materials, processes, design tools, tolerance issues, and lack of a clear understanding of cost models. There are scores of research projects which have proven that EPs are attainable in almost any substrate, but there is still the need for identifying new materials and processes before EPs are standardized. The call for product scale-up and qualification is still in the developmental stages. For example, manufacturers who previously made only boards have to expand their business if they are to incorporate EPs. They will need to develop the background not just in processing but in issues such as performance, sizing, tolerance, parasitic, and reliability which were previously dealt with by a separate entity. Additionally the placement of DPs is one of the last steps to be performed on a PCB prior to its inclusion in the system but with EPs which are fabricated during the process of board manufacturing and portions of board formed earlier should be able to withstand the thermal, chemical, and mechanical stress during the fabrication process. Also, the issues with rework due to yield problems will be compounded and might result in the scrapping of an entire board due to problem with one bad component. Finally, there is great demand for design tools to be established to effectively incorporate IP design into a system. All these aspects of the IP technology are being inspected in order to keep up with the demands of modern electronics.

2.3 Applications of IPs

Passives have different roles in the three main sectors of electronics – RF and microwave, digital, and analog-mixed signal applications. The emergence of the

wireless telecommunications network and the demand for high-frequency devices required by the military have led to the demand for integration of passives and interestingly enough, the passive integration technology in this field (RF and microwave) is one which has been in practice for several decades. The main proponent of IPs in RF applications is the necessity to reduce the effects of parasitics. So there is a great demand to produce high operating frequency components, which have good electrical characteristics. Table 2 shown below addresses the application range of passives in RF.

Table 2 Passive component requirements for RF and microwave applications [9]

Application	Value range	Tolerance (%)	Requirement
Signal inductors	1–20 nH	1–10	High Q and self-resonant frequency
Signal capacitors	1–20 pF	5–10	High Q and self-resonant frequency
Decoupling capacitors	0.01–0.1 μF	10–20	Low series inductance
Choke inductors	1–10 μH	10–20	High Q and self-resonant frequency
Signal resistors	10–100 Ω	1–10	Tightly matched ratios
Terminating resistors	20–100 Ω	1–10	

In digital applications ICs are accompanied by usually one and sometimes more than one decoupling capacitors whose main requirement is to be large in value so their tolerance is usually not an issue. Also, resistors are used as pull-up and pull-down devices between a signal line and either a power or ground supply terminal. Additionally, resistors are employed in terminating interconnections in high-speed digital circuits. This is not a comprehensive list of application for passive but does state some of the most important applications of passives in digital circuits. Table 3 given below lists the applications of passives in digital circuits.

Table 3 Passive component requirement for digital circuits [9]

Application	Value range	Tolerance (%)	Requirement
Decoupling capacitor	0.01–0.1 μF	10–20	Low-series inductance
Pull-up/down resistors	1–30 KΩ	10–20	
Terminating resistors	2–100 Ω	1–10	
Timing capacitors	10–100 pF	10–20	
Filter resistors	1–10 MΩ	20	

In analog and mixed-signal applications, decoupling capacitors are found to accompany ICs. Since these circuits are not responsive to high frequencies, it is not important to have low series inductance capacitors. Capacitors and inductors are used to isolate analog parts of a circuit from digital parts. Table 4 shows the requirements of passives in analog and mixed signal applications.

To summarize, passives find a wide range of applications and have wide range of values in each area of application, placing severe demands on the need to inspect new materials, processes, design tools for successfully integrating passives

Table 4 Passive component requirements for analog and mixed-signal circuits [9]

Application	Value range	Tolerance (%)	Requirement
Resistors	10–100 MΩ	1–10	Tightly matched
Signal capacitors	10 pF–10 nF	5–10	Tightly matched
Decoupling capacitors	0.01–0.1 μF	10–20	
EMI filter capacitors	1–10 nF	10–20	
Choke inductors	1–10 μH	10–20	

for future needs, and the general expectation is that nanotechnology will provide solutions for these needs making EPs the staple of next generation devices.

3 Nanotechnology and Nanoparticles

Nanotechnology can be defined as the science and engineering of designing physical, chemical, and related functional properties of nanomaterials and nanostructures (<100 nm) and their fabrication, integration, and manufacturing to realize advanced integrated systems, many times across the scale boundaries. It is a highly diverse field which shows promise in the development of new materials and devices which are functional in the nanometer scale. It has wide ranging applications in the field of medicine, electronics, and energy. Even though nanotechnology is new, history shows us that research in the nanometer (10^{-9} m) scale is not new. For instance, gold nanoparticles have been used as an inorganic dye by the Chinese for more than a thousand years [23, 24]. The use of colloidal gold is seen in treating arthritis and the diagnosis of a number of diseases by their interaction with spinal fluids of patients [25]. The semiconductor industry remaining true to Moore's law [26] combined with the recent explosion in our ability to image, engineer, and manipulate in the nanometer scale has brought about the drive toward the era of nanotechnology.

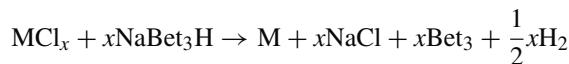
Over the last decade nanoparticles have been of great scientific interest. They range from several nanometers to several tens of nanometers and act as a bridge between the bulk materials and atomic structures. Properties of materials change as they approach the nanometer scale and these nanoparticles exhibit special properties relative to bulk materials [27–31]. Nanoparticles can be used as agents of change for different processes or as building blocks of new materials or devices. They show a strong dependence on size, shape, and surface structure. Over the last couple of years, various commercial suppliers (NNI, Nanoscale, NanoSonic, etc.) have started making unique nanomaterials. These nanomaterials are unique in size, shape, and functionality due to surface chemistry. They have exceptional functional properties which are not possible in microsize particles. Further, they provide scientists and engineers the opportunity to assemble these particles in 1D, 2D, and 3D configurations through self-assembly and directed assembly processes. Due to these remarkable qualities NPs have already found widespread use in application such as coatings, dispersions, biological, and environmental systems. In the semiconductor

industry scientists and engineers aim to take advantage of the new physical, chemical, and functional properties caused by this size scaling to reach new heights in the passive technology.

3.1 Synthesis of Nanoparticles

The science of synthesis of nanoparticles has gained a lot of attention over the past two decades and has seen enormous improvement. In fact the synthesis of NPs is a vast topic which cannot be thoroughly explained as a section of a chapter. There are countless articles and several books which deal exclusively with the different practices involved in the fabrication of nanoparticles. So in this section we will briefly discuss a few common methods used in the synthesis of nanoparticles. The most common methods used in the synthesis of nanoparticles are (i) ball milling, (ii) chemical reduction, (iii) laser ablation, and (iv) sol–gel method.

- Ball milling is a top-down approach where nanoparticles are produced from microsized particles. In this process microsized particles are loaded into vials and pounded by steel balls for a long duration. This grinding of microsized particles results in fine-sized nanoparticles whose size depends on the initial particle size and the number of hours spent in the ball mill. The possible drawback of this process is the presence of impurities in the nanoparticles contributed by the steel balls [32].
- The chemical reduction process is generally used in the reduction of metal compounds into elementary nanosized metal particles. Generally, a solution of the metallic compound is mixed with an alcohol (such as ethanol). Once it is well mixed an alkaline solution such as NaOH/NaHCO₃/NaBethyl solution is added. The mixture is heated to a given temperature under nitrogen gas and the resulting liquid contains the reduced metal nanoparticles. For example, a reaction equation can be written as (M:metal, Et:Ethyl) [33]



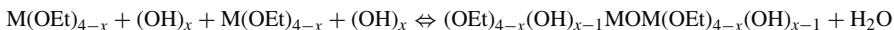
- Laser ablation is the process where solid material is reduced to nanoparticles by a laser beam. This is very similar to the cutting of either metal or semiconductor materials. The raw material is the bulk form of the nanomaterial which is subjected to a focused laser beam resulting in the production of fine nanoparticles. For instance, Lee et al. [34] used Lumonics KrF excimer laser of wavelength 248 nm and pulse-width of 12 ns in the production of well dispersed 60–100 nm permalloy nanoparticles. The size of the nanoparticles can be controlled by process parameters, such as laser fluence. Nanoparticles produced in this method have the distinction of being very pure, spherical in shape, and not agglomerated.

- Sol-gel method is a wet chemical process which results in the synthesis of colloidal dispersions of organic, inorganic, and hybrid materials. Generally, sol-gel processing involves hydrolysis and condensation of both organic or inorganic salts and metal alkoxides. Organic or aqueous solvents are used to dissolve the above materials and some catalysts are usually added to promote the reaction. One of the best known examples of sol-gel process is probably the production of SiO₂. The hydrolysis and subsequent condensation reaction of Si–O methyl compound results in the formation of SiO₂ gel or SiO₂ nanocomposite (which is gel-like material with nanoparticles). Below is an example of the chemical reaction in a sol-gel process [35].

Hydrolysis:



Condensation:



3.2 Nanocomposites

Over the last decade nanocomposites have garnered enormous interest due to the potential in building large area, flexible electronic devices [35–39]. Scientists and engineers have been intrigued by the remarkable physical and electrical properties of this unique material. For instance, upon optical pumping the addition of ZnO nanoparticles into a polymer matrix results in laser-like behavior [34]. Addition of silicate nanoparticles into gel electrolytes results in high conductance [37]. Additionally, the application of nanocomposites to meet the continuous demand for miniaturization, high density, high speed, and flexible microelectronic devices is being investigated in the area of embedded passives.

3.2.1 Nanocomposite-Embedded Capacitors

In today's electronic devices capacitors occupy about 50–60% of overall passive components [40]. So, the replacement of discrete capacitors with embedded capacitors offers the best potential in the drive toward miniaturization and high-density electronic packaging. They can be integrated into the substrate to provide decoupling, bypass, termination, and frequency-determining functions [40]. The packaging technologies available for integrating capacitors into a substrate are of three kinds: (i) MCM-C (packaging technology in which the substrate is ceramic), (ii) MCM-D (packaging technology in which the substrate is glass or silicon), and (iii) MCM-L (which uses PWBs as the substrate). The pros and cons of each packaging technology have been discussed in detail in several books and publications

[19, 41]. Of the three, MCM-L technology is considered to be the best option for embedded capacitors due to the fact that PWBs offer the possibility of processing large number of MCMs on a single substrate. The fabrication of embedded capacitors has been demonstrated through several approaches, namely sputtering, chemical vapor deposition, anodization, ceramic green-sheet processes for the above packaging technologies. Nevertheless, these processes are too expensive or not process compatible with PCB fabrication. Nanoparticles dispersed in a polymer offer the best of both worlds, they offer the low-temperature processing which is compatible with PCB manufacturing as well as the high dielectric properties of the nanoparticle fillers which are essential for integration of capacitors into substrate. Some of the advantages offered by these nanocomposites are (i) good process compatibility with PCB, (ii) good CTE match with PWB, (iii) high temperature stability, (iv) low moisture absorption, (v) good adhesion to copper, interlayer-dielectrics, and substrates, (vi) ability to withstand high pH electroless plating bath, and (vii) good mechanical properties [19]. There are several different factors that are critical to the integration of nanoparticle-based embedded capacitors, namely (i) Materials, (ii) Fabrication, and (iii) Characterization.

Materials

There is wide range of nanoparticles which are being currently investigated to increase the dielectric constant; capacitance density, and streamlining of NP-based embedded capacitors. The most common nanoparticles which are being investigated for the dielectric properties fall under two categories: (i) ceramic-in-polymer nanocomposites and (ii) metal-in-polymer nanocomposites. For instance, Biswas et al. [42] have investigated the use of a ceramic-in-polymer approach where barium titanate nanoparticles were deposited in a polar polyaniline–polyurethane polymer composite to achieve a capacitance density of 15 nF/cm^2 for radio frequency (RF) applications. Table 5 is a list of material choices being explored by various academic research groups. Bhattacharya et al. [48] have investigated the properties of lead magnesium niobate–lead titanate (PMN-PT) ceramic powder in epoxy composites for the fabrication of parallel plate capacitors. A capacitance density of 10 nF/cm^2 was achieved in epoxy nanocomposites. Windlass et al. [49] have investigated the effects of both lead magnesium niobate–lead titanate and barium titanate as ceramic fillers in epoxy composites with capacitance densities of 35 nF/cm^2

Table 5 Material choices in nanoparticle-based embedded passives

Material	References
Polymer/Al nanocomposite	Xu et al. [43]
Silver/carbon black/epoxy nanocomposite	Lu et al. [44, 45]
Barium titanate/epoxy nanocomposites	Das et al. [46]
Resin-coated copper nanocomposites	Das et al. [47]

with film thickness of 2–3 μm . Additionally, Das et al. [46] have also investigated the effects of barium titanate nanoparticles in epoxy polymer composites with resulting capacitance densities of 10–100 nF/in.² at 1 MHz. There are others who have employed metal nanoparticles to enhance the dielectric properties of polymer composites in achieving high capacitance densities for embedded capacitors. For example, Lu et al. [44, 45] achieved high dielectric constant and low dissipation factor for nanoparticle-based embedded capacitors by using in situ formed silver nanoparticles in carbon black/polymer composites. Additionally, Xu et al. [43] have investigated the effects of aluminum nanoparticles in polymer composites for the fabrication of embedded capacitors.

Fabrication

There are several approaches involved in the fabrication of nanoparticle-based composites for embedded capacitors. One of the basic steps involved in the fabrication is the dispersion of nanoparticles. The dispersion and the size of nanoparticles play a very important role in determining the quality and capacitance density of nanoparticle-based embedded capacitors. According to the dispersant first rule of Cannon et al. [50] the efficient de-agglomeration which is essential for the fabrication quality of capacitors can be achieved in shorter times by dispersing powder in the solvent before adding the polymer. The most prevalent technique in dispersion is to mix the nanoparticles in a solvent such as 95% ethanol or propylene glycol methyl ether and then mix the epoxy polymer to the above solvent with the help of an ultrasonicator or a high-shear homogenizer to achieve a homogeneous mixture of the nanocomposite. The nanocomposite is then oven dried at the chemical specific temperature and then either laminated or deposited on the substrate by either spin coating or meniscus coating. For example, Bhattacharya et al. [48] blended epoxy and ceramic powder using Pro Scientific high-shear homogenizer (Model PRO200) and roll milled the mixture 3–5 days before fabrication to prevent agglomeration and settling. The mixture was then spin coated for small area copper clad FR4 substrate deposition and meniscus coated for substrates of size 300 × 300 mm glass and copper clad PCB panels. The samples then underwent photolithography processes before the top copper electrode was deposited and processed. Additionally, Xu et al. have demonstrated the importance of surface treatment of nanoparticles before dispersion in epoxy. Windlass et al. [49] have stated in his paper that good dispersion of filler results in a homogeneous packaging resulting in high reliable, uniform properties of film which result in high dielectric constants. Das et al. have demonstrated the fabrication of multilayer embedded capacitors on large coatings using the resin-coated copper capacitive (RC3) lamination process [47] where barium titanate nanoparticles and epoxy were dispersed in organic solvents. A thin layer of this nanocomposite was then deposited on copper substrate and dried to produce free standing RC3 which can be directly laminated onto a substrate. Das et al. [51] have also demonstrated the use of laser micromachining in the

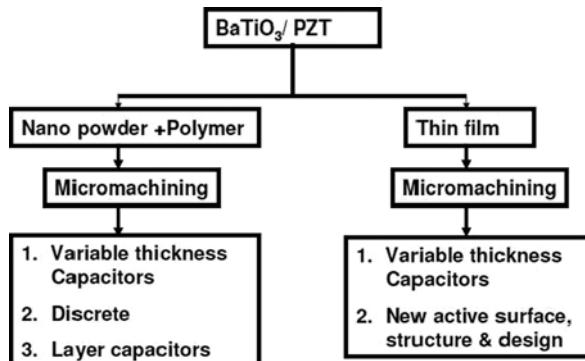


Fig. 8 Process flow chart [51] (Copyright © 2007 by IEEE)

fabrication of barium titanate nanocomposite-based flexible embedded capacitors. In the laser micromachining process, the barium titanate/epoxy nanocomposite is deposited onto a substrate and dried to remove the organic solvent and then micro-machined using Nd:Yag laser. This process resulted in high performance, large area, and flexible capacitors. The schematic and conditions of the laser micromachining process are as shown in Fig. 8 and Fig. 9, respectively.

Additionally, Lu et al. [44, 45] have investigated the fabrication of embedded capacitors using an *in situ* processing of silver/epoxy nanocomposite, where silver nanoparticles were *in situ* synthesized in epoxy resin through the chemical reduction of silver nitrate. This was mixed with the carbon black/epoxy composite via ultrasonication to have a uniform, homogeneous mixture which resulted in a dielectric constant of ~ 2200 at 10 kHz.

In summation, the general steps in fabrication of nanoparticle-based embedded capacitors are as follows:

1. Dispersion of nanoparticles in chemical specific solvent via ultrasonication.
2. Mixing of epoxy polymer in chemical specific solvent via ultrasonication.
3. Mixing of nanoparticle/solvent mixture and epoxy/solvent mixture via mixing and ultrasonicator.
4. Roll milling or ultrasonication of the nanoparticle/epoxy/solvent mixture to reduce settling and agglomeration.
5. The nanoparticle/epoxy/solvent mixture is deposited on substrate via spin coating or meniscus coating.
6. Then in case of photo-definable nanocomposite mixture is patterned and oven baked whereas non-photo-definable nanocomposites are oven baked to remove all organic solvents.
7. The top electrode is then deposited, patterned, and etched to form capacitors which can be laminated onto a substrate.

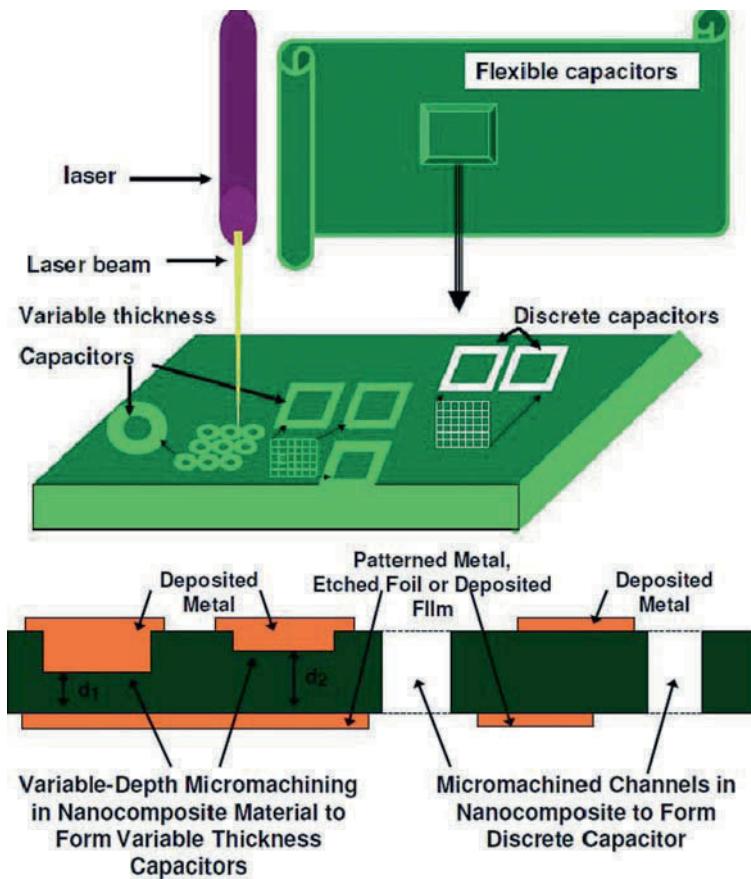


Fig. 9 Top (a) and cross-sectional (b) views of capacitors [51] (Copyright © 2007 by IEEE)

Characterization

Capacitance values are determined by the thickness, quality of the dielectric film, and the feature size. The characterization of embedded capacitors can be classified broadly into three different categories, namely: (i) characterization of nanocomposite, (ii) electrical characterization of the dielectric, and (iii) the reliability testing of the fabricated capacitors. So the first step in the characterization process is to ascertain the quality of the nanocomposite. This is done using techniques such as transmission electron microscopy (TEM), scanning electron microscopy (SEM), energy dispersive X-ray spectroscopy (EDS), Fourier transformed infrared spectroscopy (FTIR), X-ray diffraction (XRD), and stress rheometer. Once the quality of the nanocomposite has been characterized the next step is to complete the electrical characterization of the dielectric which involves analyzing the thickness, dielectric constant, dissipation factor, frequency response,

and capacitance. The final step in the characterization process is the reliability testing which is done by thermal shock, pressure cooker test (PCT), and IR-reflow.

The challenge with nanoparticle/epoxy composites is the incompatibility between the hydrophilic nature of nanoparticles and the hydrophobic nature of epoxy. This opposing character of nanoparticles and epoxy results in agglomeration which results in poor quality of capacitors due to the non-uniformity of the dielectric thickness. So in the first step, characterization of the nanocomposite is performed using transmission electron microscopy (TEM) to verify the size, shape of the nanoparticle, and the dispersion of the nanoparticles in epoxy polymer matrix. For example, Xu et al. [43] claim that silane functionalization of aluminum nanoparticles helped in forming a uniform mixture of nanoaluminum/epoxy composite which was verified by the TEM images, which are shown in Fig. 10.

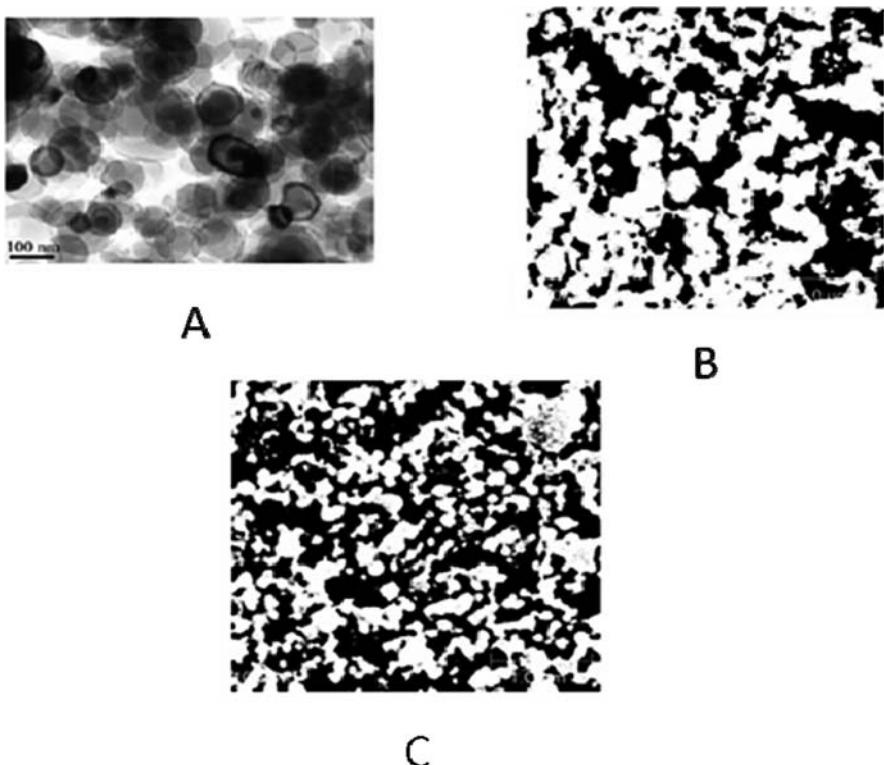


Fig. 10 (a) Nanoaluminum, (b) untreated nanoaluminum composite, and (c) silane-functionalized nanoaluminum composite [43] (Copyright © (2005) by IEEE)

Second, it is important to identify the interlinking of the coupling agent or epoxy to the surface-treated nanoparticle. In the work done by Xu et al. [43] this verification was done via Fourier transformed infrared spectroscopy (FTIR). The surface chemistry of the aluminum nanoparticles was verified by oven drying the aluminum

nanocomposite and analyzing the dried KBr pellets prepared with the dried aluminum under the FTIR. The results obtained indicated the presence of Si and the absence of hydroxyl groups indicating that good interaction between nanoaluminum and silane-coupling agent. There are other techniques such as X-ray diffraction (XRD) which can be used to determine the size of the nanoparticle in a given nanocomposite and SEM and EDS have been utilized for several years to identify the presence of different chemicals in a given compound. For example, Lu et al. [44, 45] have employed TEM, SEM, EDS, and XRD to determine the size, size distribution, and verification of silver nanoparticles in the nanocomposite. The TEM pictures indicated the presence of capping agent and the effect of capping agent in inhibiting the growth and agglomeration of particles. SEM images verified the size and dispersion of the nanoparticles. EDS results confirmed the presence of silver nanoparticles in the polymer matrix. Finally XRD was used to verify the effects of the mixture on the size of silver nanoparticles in the epoxy polymer matrix. Additional tests are conducted to determine the rheology properties of the nanocomposite which is essential for the uniform coating of the substrate during the spin cycle (when the nanocomposite is deposited on a substrate at some predetermined conditions via spin coater or meniscus coater). The rheology properties of nanocomposites have been reported in several papers. For instance, Xu et al. [43] report that untreated nanoaluminum composites were more viscous than silane-treated composites. This indicates the importance of functionalization or presence of a coupling agent. Additionally, it also shows the effects of agglomeration of nanoparticles and how it adversely affects the dielectric properties. The next step involves the electrical characterization of the dielectric which accounts for the quality and capacitance density of the fabricated devices. The thickness of the dielectric plays a vital role in determining the capacitance. Capacitance is given as

$$C = \frac{\varepsilon_0 \varepsilon_r A}{d}$$

where C is the capacitance in farads, ε_0 is the relative permittivity of free space, ε_r is the relative permittivity of the dielectric material, A is the area of each plate in square meters, and d is the distance between top and bottom plates.

The thickness of the nanocomposite film plays a very vital role in determining the capacitance and can be determined using a DEKTAK profilometer. Bhattacharya et al. [52] have reported the use of a DEKTAK profilometer as an electrical characterization tool by measuring the dielectric thickness. Additionally, the quality of the dielectric is imperative to the quality and capacitance density. In a two-phase composite system the dielectric constant of the base polymer is vital in achieving higher permittivity, thereby higher capacitance density in the composite system [53]. The dielectric constant of the nanocomposite film can be measured via a precision LCZ meter, which measures the capacitance density and the dielectric constant as a function of temperature. There is plethora of information that is provided by several authors who have used the LCZ meter to examine the capacitance and dielectric constant of capacitors. These measurements are at a certain

frequency which also provides information upon the quality of the capacitors at different frequencies. For instance, Das et al. [47] report that the capacitance density of the RC3 nanocomposite-embedded capacitors decreased as the frequency was increased from 10 kHz to 1 MHz. Authors also indicate that the capacitance density depends on the loading of the nanomaterial by giving us evidence that the measured capacitance density of 33% loading nanomaterial have shown high capacitance value of 515 nF, which is six times of nanomaterial at 25% loading. Further electrical characterization can be done by obtaining the capacitance and dissipation factor of the nanocomposite film using an impedance/gain analyzer and the frequency response of the nanocomposite film is obtained using a DEA test. The final stage in the characterization of embedded capacitors is the reliability testing. Embedded capacitors eliminate the failure due to solder joints, but they enhance the potential for defects due to material mismatch which leads to failure due to cracks and delamination. The reliability of nanoparticle-based embedded capacitors is examined by the tests such as IR-reflow, pressure cooker test (PCT), and solder shock tests. These tests will accentuate the defects due to delamination or cracking. Additionally, thermal aging tests are run on these new generation passives to make sure that they are at least as reliable as their surface mount counterparts.

In conclusion, the steps necessary to ascertain the quality of passives are in place and have been developed and implemented in universities and research labs. It is a matter of time before we figure how to efficiently scale up the process and take advantage of this new line of passives which are on track to revolutionize the electronic industry.

3.2.2 Nanocomposite-Embedded Inductors

There is a great need for research and development in the field of embedded magnetic components which exhibit good performance characteristics. Magnetic components need compatible fabrication sequence in integrating them with embedded capacitors. A logical approach is to integrate these passive components using the appropriate MCM technology, the challenge is to make sure there is process compatibility while the quality is maintained.

In general, embedded inductors suffer from poor performance due to various core losses and low Q factor (which is the ratio of energy stored to energy lost). The losses in an inductor can be divided into three main categories: (i) eddy current loss, (ii) hysteresis loss, and (iii) residual loss. The ratios of these losses vary depending on the test and operating conditions such as frequency, magnetic flux density, and temperature. For instance, at high frequencies we see that the eddy current losses increase, but this loss can be controlled by increasing the core resistance and by the effect of grain boundaries which result in an increase in the electrical resistance at the interface. So it is important to have high core resistivity and narrow hysteresis to limit losses and improve performance. The integration of embedded inductors falls in the same categories as capacitors: (i) materials, (ii) fabrication, and (iii) characterization.

Materials

The first step in the integration of embedded inductors is the availability of materials. Due to the fact that high electrical resistivity is a need in producing inductors, ferrite materials play a vital role in the fabrication of inductors. The electromagnetic characteristics of these inductors depend on their composition and their structure. Since their discovery ferrites have been involved in the fabrication of inductors but have been based on ceramic processing which involves high temperature fabrication steps such as firing of ferrite-based paste. This cannot be integrated into mainstream electronics due to the fact that most main stream electronic components use low-cost polymer substrates which cannot tolerate high temperatures. So, researchers have been working on developing ferrite-based polymer composites which can be used in the temperature compatible fabrication of inductors. In fact there are several researchers who have developed ferrite particle filled polymer composite inductors which can be used with organic substrates [54–60]. Additionally, there are several ferrite-filled polymer composites which are available commercially and are used in the fabrication of embedded capacitors. The shortcomings of these micron-sized ferrite particles are size limitations in the fabrication of thin film inductors and the presence of wide hysteresis loops which result in high hysteresis losses. These can be overcome by the use of nanoparticle-sized ferrite fillers which exhibit properties not seen in their bulk counterparts.

Nanoparticles of ferrites can be synthesized using anyone of the different techniques explained in earlier section but one of the most common methods used is the method of chemical reduction. This method has been used significantly by researchers in the low-temperature synthesis. In this method the water-based reagents such as $\text{Co}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ and FeCl_2 are mixed along with SDS which acts as a catalyst and forms a micellar solution. This solution is then heated to some predetermined temperature. This process results in the formation of nanophase CoFe_2O_4 which can be separated by centrifuge and ferrite nanoparticles are obtained after oven drying for a predetermined time. Detailed procedure for the synthesis is documented elsewhere [58]. The general procedure followed during the synthesis of ferrite nanoparticles are as follows. First a salt solution that contains Fe compounds is prepared by dissolving Fe-linked chemicals in a predetermined ratio, and then adding a hydroxyl group such as NH_4OH solution into the ferrite solution. The hydroxyl group is added for pH adjustment of the salt solution. This results in the formation of Fe–O powder which can be reduced to nanophase ferrite solution at low temperature. The ferrite nanoparticles are then obtained by centrifuge and oven drying at certain predetermined level.

Fabrication

The fabrication of nanoparticle-based embedded inductors comprises the use of the binding properties of the polymer matrix to provide uniform adhesion to the magnetic nanoparticles. Commercially available epoxy resins are used as the base resin to which the nanoparticles are added. Then a commercial hardener is added

to the mixture and the composite is then cured for a predetermined time at a set temperature. The time and temperature for curing of the nanocomposite are dependent on the commercial epoxy and hardener that are used. For an instance, Dong et al. [59] used EPON 8281 bisphenol A-type epoxy as the base resin and 4-methylhexahydrophthalic anhydride as a hardener and a mythylimidazole as catalyst and added the ferrite nanoparticles to form the magnetic nanocomposite which was cured at 120°C for 30 min and 150°C for 2 h. In his work Raj et al. [60] synthesized a nickel–zinc–ferrite nanoparticle of size 5–15 nm and dispersed them in a commercially available epoxy to form a thick paste. The paste was then screen printed onto an FR4 board and cured to a set temperature and fabricated the desired nanoparticle-based inductor to specific dimensions. Figure 11 shows the fabrication steps of a silicon-based inductor and Fig. 12 is the spiral inductor designed at the University of Toulouse. In conclusion the most common method of producing nanocomposites is to disperse synthesized nanoparticles into an epoxy/hardener mixture

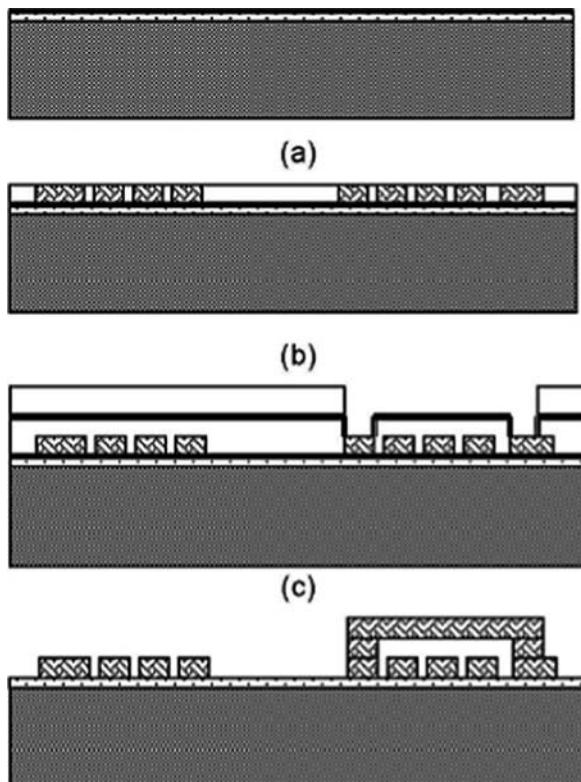
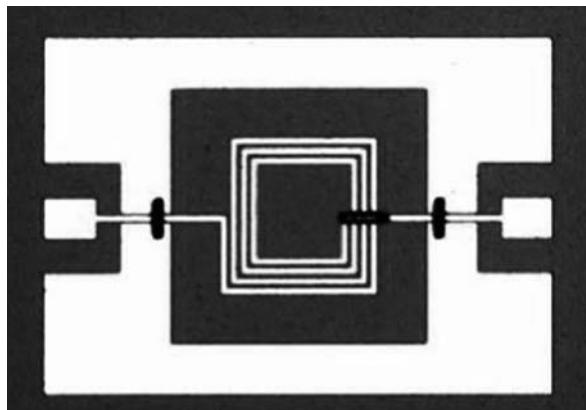


Fig. 11 Scheme of the inductor fabrication processes (a) 0.7 gm SiO_2 and $\text{Ti}/\text{Cu}/\text{Ti}$ layer deposition. (b) First copper plating for the coil part of the inductor. (c) Air-bridge pattern and Cu/Ti layer deposition. (d) Air-bridge plating and seed layer removal [61] (Copyright © (2006) by IEEE)

Fig. 12 Silicon-integrated inductors [62] (Copyright © (2007) by IEEE)



and cured to manufacturer specifications which are then characterized to determine their quality.

Characterization

The characterization process can be separated into different phases similar to the characterization of nanoparticle-based capacitors. The first step is to characterize the nanoparticles to verify the size and crystalline properties. The XRD is generally used to verify the crystalline structure of the nanoparticles. The nanoparticles are examined to see if their XRD curves match that of their bulk samples. Additionally, TEM or HRTEM is used to verify the size of the nanoparticles. The final and most important characterization of the nanoparticles is to analyze their magnetic property. This measurement is done using a SQUID (superconducting quantum interference device) magnetometer. The hysteresis response of the ferrite nanoparticles is plotted and if the resulting loop is narrow then the quality of the nanoparticle is considered to be very desirable for inductor application as it shows that there is almost no hysteresis loss. An example of a desirable hysteresis loop which will result in low loss and good quality inductor is shown in Fig. 13.

Second, the electrical properties of the nanocomposite are tested to verify that the quality of the nanocomposite is desirable for the fabrication of inductors. The dielectric properties of the nanocomposite play an important role in the performance of the inductor. The dielectric constant and loss at low frequencies are measured using a precision LCR meter. The high-frequency measurement of the permittivity of the nanocomposites is conducted using an s-parameter Network Analyzer. The impedance of the fabricated inductors can be obtained from these measurements.

In conclusion, there is a lot of room for research and development in the area of nanoparticle-based embedded inductors. We still need to develop universally standard synthesis and fabrication techniques which when followed will result in good quality inductors. Additionally, there is room for exploration in the area of materials available for the fabrication of inductors. Finally there is a great need to

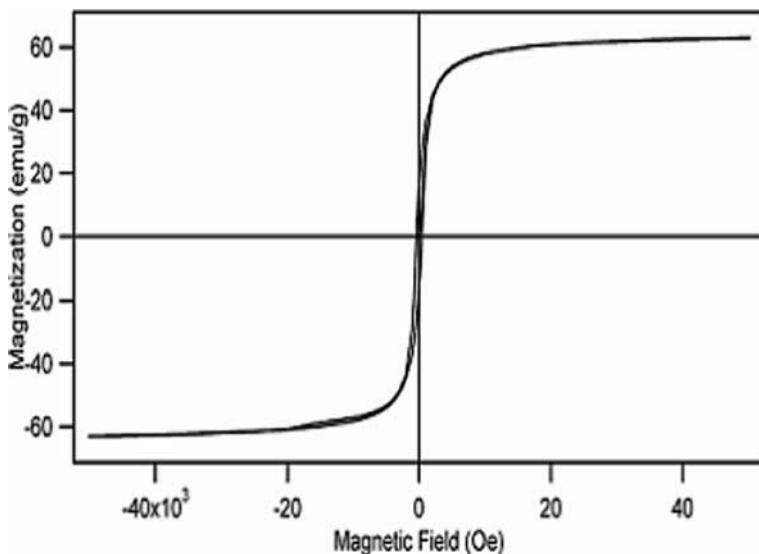


Fig. 13 Hysteresis curve of magnetic nanocomposite [56] (Copyright © (2004) by IEEE)

develop standard methodology before nanoparticle-based embedded passives can be an integral part of main stream electronics.

3.2.3 Nanoparticle-Based Embedded Resistors

The behavior of elemental metals is fairly predictable and has been well documented over the years. In order for nanoparticle-based embedded resistors to gain acceptance into main stream electronics it has to overcome several challenges. A few issues that need to be solved are tight tolerance, high yield, and reliability and availability of wide range of resistances. In fact iNEMI (International electronic Manufacturing Initiative) predicts that maximum resistance of 100 kΩ/sq is essential for embedded resistors to become a part of main stream electronics. Another challenge is that even though solder joints are eliminated, better optimization of process flow is necessary as rework is not an option.

In embedded resistor technology, resistors are patterned as strips of the resistive material. The ends of the strip have pads for interconnection to other components on the board. The resistance is therefore dependent on the property of the material and the dimensions of the strip. The general formula for resistance is given below

$$R = \frac{\rho L}{A}$$

where R is the resistance (ohms), ρ is the resistivity of the material (Ω cm), L is the length of the strip (cm), and A is the cross-sectional area (cm^2). For majority of the applications, resistance ranges from 10Ω to $1 \text{ M}\Omega$ is a requirement [58]. The

challenge in integrating resistors into the board deals with finding materials which are process compatible while giving the reliable performance.

Material Choices

The industry has used metal paste in the fabrication of resistors for a very long time. The paste is a mixture of micron-sized metal particles and some organic resin. This is applied onto a substrate and allowed to cure, after curing the paste is fired and the metal particles melt and fuse forming a conductive film. This technology requires the substrate to withstand temperatures ranging between 550°C and higher. For embedded technology to become an everyday practice in industry the fabrication of embedded resistors needs to be done at low temperatures. Organic boards can generally withstand a temperature in the order of few hundred degrees Celsius.

The key to nanoparticle-based embedded resistors is to use the natural property of nanoparticles. Nanoparticles have a tendency to agglomerate due to their surface energy, and this property of nanoparticles can be used effectively in the fabrication of nanoparticle-based embedded resistors. Metal nanoparticles can be mass produced by solid state thermal decomposition process which is done at relatively low cost and at low temperatures. For instance, silver nanoparticles can be manufactured by heated decomposition of solid fatty acid silver at 200°C for 5 h [58]. The synthesis of metal alloys for use as nanoparticles in the fabrication of embedded resistors is being actively pursued by several researchers in the industry and at educational institutions. Silver palladium alloys, alloys of copper nanoparticles are being investigated in the fabrication of embedded resistors. For example, Nakamoto [58] has investigated the synthesis of silver palladium alloys as nanoparticles for the fabrication of resistors. He claimed that silver–palladium nanoparticles which were synthesized have superior migration resistance which is a very important issue in the electronic industry. The fabrication of resistors using these silver–palladium alloys have been done at 300°C for 30 min. The resistivity of these nanoparticle-based components was calculated to be 10 ohm-cm which is favorable for use with polyimide substrates. Water drop method evaluation showed that the migration resistance was higher than in pure nanoparticles. This results in the fabrication of narrow electronic structures with a very high reliability. Thus low-cost mass production of high-quality metal nanoparticles is feasible for volume production of nanoparticle-based embedded resistors.

Fabrication

Resistors are an integral part of every electronic circuit, so the fabrication of nanoparticle-based resistors is a vital component in realizing embedded circuits. There are several methods of fabricating embedded resistors which have been demonstrated over the years (embedded resistor fabrication). The most promising method for the fabrication of nanoparticle-based embedded resistors is the technology of ink-jet printing.

Ink-jet printing is the method of dispensing liquids and pastes at certain specific regions of a circuit which can be fired to form the required component. In one of the designs, this technology is based on the volumetric change in the fluid by displacement of piezoelectric material that is coupled to the fluid (a method of characteristics model of a DOD ink-jet device using an integral method drop formation model). The volumetric change causes the conductive region to be fabricated without the need for photolithography and other practices which are common in the fabrication of passives in the semiconductor industry. Ink-jet printing technology is considered to exhibit considerable promise in making embedded passive technology a reality and the backbone to meet the demands of next generation electronic devices. There are two specific types of modes used in the dispensing of conductive solutions, namely demand mode and continuous mode. In the demand mode the conductive paste or fluid is dispensed at only a particular region whereas in continuous mode the paste is dispensed in a continuous fashion to form passives over a specified area and can be used in the fabrication of capacitors and other devices. Shah et al. [61] reported the use of ink-jet technology which has been employed in the fabrication of passives using droplets of polymers, solder, organo metallic, and metal nanoparticle solutions of sizes 25–125 μm in diameter at rates up to 4000 droplets/s. Figures 14 and 15 show the embedded resistors printed using ink-jet technology by Shah et al. at MicroFab Technologies

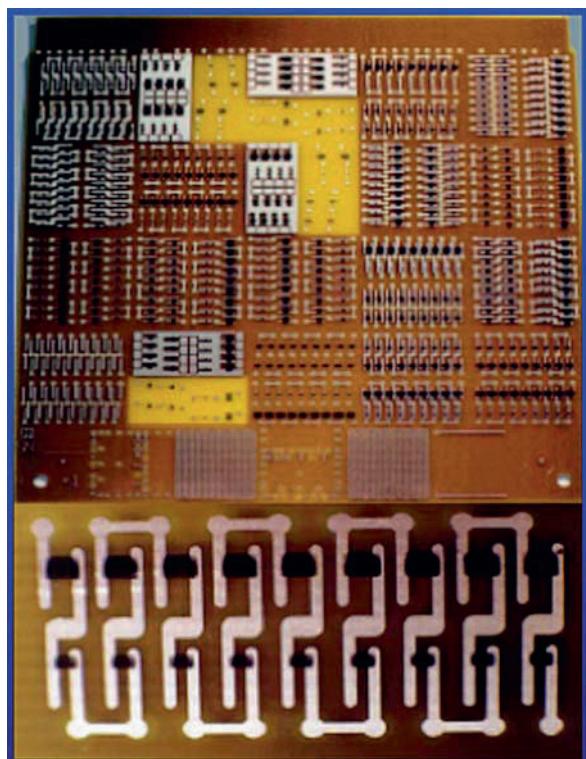


Fig. 14 Ink-jet printed embedded resistor using conductive polymer ($100 \Omega/\text{sq.}$ resistivity) [65]
Copyright (c) 1999 by IMAPS

The fabrication of embedded resistor technology has been demonstrated successfully by researchers at the National Institute of Standards and Technology (NIST) [59]. They have demonstrated the fabrication of embedded resistor technology by fabricating resistors ranging from $100\ \Omega$ to several mega ohms. Printed resistors have ranged from a few micrometers to several millimeters long. Thus there is ongoing research which has given us glimpses that the embedded resistor technology can become a reality and that nanoparticles will become an integral part in the drive toward miniaturized, reliable, electronic devices.

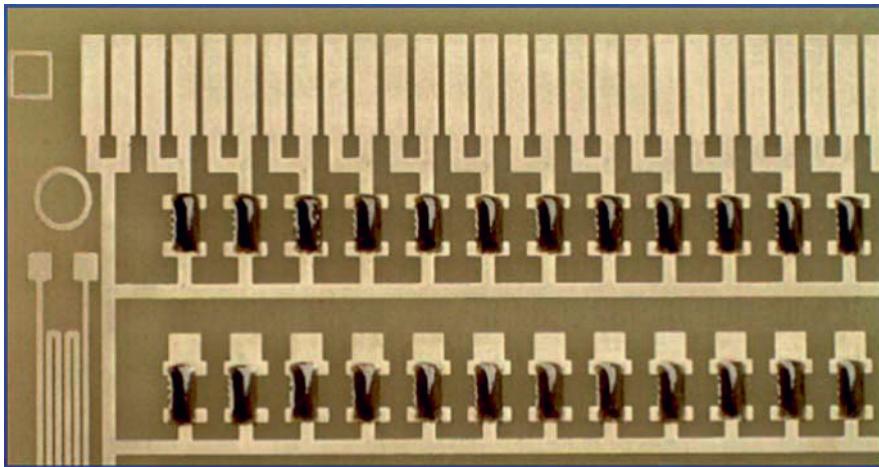


Fig. 15 Ink-jet printed carbon nanotube filled epoxy resistor [61]. Copyright (c) 2002 by IMAPS

Characterization

The characterization techniques for nanoparticle-embedded resistors involve the analysis of the quality of the nanoparticles and quality and reliability of the fabricated resistors. The first step is to characterize the metal nanoparticles and this can be done using TEM, SEM, and other analysis techniques described in earlier sections. These analysis tools will give a good understanding of the size, shape, and other qualities of the nanoparticles. The second step is to verify that the fabricated resistors give the correct resistance values for which they were designed. This can be accomplished using a two-probe method or four-probe methods which are well-known measurement techniques.

Additionally, reliability tests need to be completed on embedded resistors as it becomes difficult to rework or replace embedded resistors after they are integrated into the substrate. So it is imperative to complete reliability test before the embedded resistor technology can be called a success. The reliability of nanoparticle-based embedded resistors should be subjected to the same tests that embedded resistors are subjected to. For instance, embedded passives are subjected to temperature shock test, high accelerated stress test (HAST), electrostatic charge discharge (ESD),

thermal cycling, etc. Min et al. [32] describe in his work that embedded resistors fabricated during his collaborative effort with NIST have passed the reliability tests conducted.

In conclusion the synthesis, fabrication, and characterization of nanoparticle-based embedded resistors have been explored at universities' labs and are being currently investigated by the R&D departments at industries. The success of nanoparticle-based embedded passives depends on integrating all the different synthesis, fabrication, and characterization techniques into scalable production. Additionally, the rapidly evolving microelectronic technology coupled with the swift ascent of nanotechnology is leading us toward the goal of miniaturization. The unique properties found in nanoparticle which cannot be realized in microparticles and bulk material give us the opportunity to design devices with functional density which cannot be realized using other technologies. Thus it is not a question of whether nanoparticle-based embedded resistors will be realized but how soon will they be realized.

4 Summary

The advent of nanotechnology has provided glimpses of potential in solving some of the problems associated with embedded passives. The nanosize and the unique properties associated with this small size have brought into focus a new area for research and development. Literature, conferences, and workshops have started to focus on using the potential shown by nanoparticles in making nanoparticle-based embedded passives. Over the years, NEMI has clearly shown that the lack of infrastructure to be one of the major problems associated with embedded passives, which also is needed for nanoengineered passives. In order for nanoparticle-based embedded passives to become a reality there is a need for more research and development in materials, processes, designs, and proven scale-up and cost models along with environmental and health safety (EHS) data. This area is in the early stages of fundamental understanding and development. Depending upon the breakthroughs and systematic developments, next decade could see integration of nanoparticle-based integrated passives.

Acknowledgments Authors would like to acknowledge financial support of the Office of Naval Research (ONR). Also, one of the authors (APM) acknowledges support of National Science Foundation (NSF; grant # 0501597).

References

1. http://nobelprize.org/educational_games/physics/integrated_circuit/history/ (downloaded 10/3/08) Nobelprize.org. The history of integrated circuits 2005
2. http://www.ti.com/corp/graphics/press/image/on_line/co1034.jpg(downloaded 10/9/08) The Chip that Jack Built, (c. 2008), (HTML), Texas Instruments.
3. Clarke P. Intel enters the billion-transistor processor era. EE Times 14 Nov 2005.

4. Sundaram V., Tummala R., Bhattacharya S., and Pulugurtha R.M. System-on-a-Package (SOP) Substrate and Module with Digital, RF and Optical Integration. *Electronic Components and Technology Conference* 54th Proceedings 2004; 1: 17–23.
5. <http://www.ngimat.com/electronics/passives.html> (downloaded 11/12/08) nGimat.com
6. Sandborn P., Etienne B., and Subramanian G. Application-specific economic analysis of integral passives. *IEEE Transactions on Electronic Packaging Manufacturing* 2001; 24: 203–213.
7. Lasky R. Growth continues for passive components. *Electronic Packaging and Production* 1998; 38: 77–78.
8. Rector J. Jr., Doughtery J., Brown V., Galvagni J., and Pyrmak J. Integrated and Integral Passive Components: A Technology Roadmap. *Electronic Components & Technology Conference* 1997; 713–723.
9. Frye R. Passive components in electronic applications: requirements and prospects for integration. *The International Journal of Microcircuits and Electronic Packaging* 1996; 19: 483–490.
10. Oberschmidt J. and Humenik J.A. Low Inductance Capacitor Technology. *Electronic Components & Technology Conference* 40th Proceedings 1990; 1: 284–288.
11. Dougherty J.P., Galvani J., Marcanti L., Sheffield R., Sandborn P., and Ulrich R.K. The NEMI roadmap: Integrated Passives Technology and Economics. *Proceeding of the Capacitor and Resistor Technology* 2003; 1–11.
12. Ulrich R.K. and Schaper L.W. *Integrated Passive Component Technology*. Wiley-IEEE Press, New York, 2003; 19–22.
13. Electronic Industry Market Research and Knowledge Networks. *Passive Components – Private Companies Report April 2001*.
14. <http://techon.nikkeibp.co.jp/english/img2/nea0305manu1fig1.jpg> (downloaded 7/07/09) TechOn. Cover Story: Motorola Ships Passive-Embedded PCB for Mobile Phones May 2003.
15. Bhattacharya S.K. and Tummala R.R. Epoxy nanocomposite capacitors for applications as MCM-L compatible integral passives. *Journal of Electronic Packaging* 2002; 124: 1–6.
16. Tummala R., Chahal P., and Bhattacharya S.K. Recent Advances in Integral Passives at PRC. *Proceedings of the IMAPS 35th Nordic Conference* 1998.
17. Chahal P., Tummala R.R., Allen M., and Swaminathan M. A novel integrated decoupling capacitor for MCM-L technology. *IEEE Transactions on Components, Packaging, and Manufacturing Technology* 1998; 21: 184–193.
18. Power C., Realff M., and Bhattacharya S.K. A Decision Tool for Design of Manufacturing Systems for Integrated Passive Substrates. *IMAPS Advanced Technology Workshop* 1999.
19. Bhattacharya S.K. and Tummala R.R. Next generation integral passives: materials, processes, and integration of resistors and capacitors on PWB substrates. *Journal of Materials Science* 2000; 11: 253–268.
20. Yoon H., Hou J., Bhattacharya S.K., Chatterjee A., and Swaminathan M. Fault detection and automated fault diagnosis for embedded integrated passives. *Journal of VLSI Signal Processing* 1999; 21: 265–276.
21. Chahal P., Tummala R., Allen M., and White G. Electro-less Ni–P–W–P thin film resistors for MCM-L based technologies. *Electronic Components and Technology Conference* 48th Proceedings 1998; 232–239.
22. Agarwal V., Chahal P., Tummala R., and Allen M. Improvements and Recent Advances in Nanocomposites Capacitors using a Colloidal Technique. *Proceedings of 48th Electronic Components and Technology Conference* 1998; 165–170.
23. Camusso L. and Bortone S. *Ceramics of the World: From 4000BC to the Present*. Harry N. Abrams, New York, 1992; 284.
24. Zhao H. and Ning Y. Techniques used for Preparation and Application of Gold powder in Ancient China. *Gold Bull* 2000; 33: 103.
25. Turkevich J. *Colloidal Gold: Part I - Historical and Preparative aspects Morphology and Structure*. *Gold Bull* 1985; 18: 86.

26. Moore G.E. Cramming more components onto integrated circuits. *Proceedings of the IEEE* 1998; 86: 82–85.
27. Carpenter E.E. Iron nanoparticles as potential magnetic carriers. *Science Direct* 2001; 255: 17–20.
28. Papp S., Szucs A., and Dekany I. Preparation of PdO nanoparticles stabilized by polymers and layered silicate. *Applied Clay Science* 2001; 19: 155–172.
29. Rataboul F., Nayral C., Casanove M.J., and Maisonnat A. Synthesis and characterization of monodisperse zinc and zinc oxide nanoparticles from the organo-metallic precursor [Zn(C₆H₁₁)₂]. *Journal of Organometallic Chemistry* 2002; 643–644: 307–312.
30. Kruis F.E., Fissan H., and Peled A. Synthesis of nanoparticles in the gas phase for electronic, optical and magnetic applications—a review. *Journal of Aerosol Science* 1998; 29: 511–535.
31. Sundaram V., Tummala R., Bhattacharya S., and Pulugurtha R.M. System-on-a-Package (SOP) Substrate and Module with Digital, RF and Optical Integration. *Electronic Components and Technology Conference* 54th Proceedings 2004; 1: 17–23.
32. Min G. Embedded Passive Resistors: Challenges and Opportunities for Conducting Polymers. *Proceedings of the International Conference on Science and Technology of Synthetic Metals* 2005; 49–52.
33. Xu J. Dielectric Nanocomposite for High Performance Embedded Capacitors in Organic Printed Circuit Boards. PhD Dissertation Georgia Institute of Technology 2006; 29.
34. Lee J., Becker F.M., Brock R.J., Keto W.J., and Walser M.R. Permalloy nanoparticles generated by laser ablation. *IEEE Transaction on Magnetics* 1996; 32: 4484–4486.
35. Anglos D., Stassinopoulos A., Das R.N., Zacharakis G., Psyllaki M., Anastasiadis S.H., Vaia R.A., and Giannelis E.P. Random laser action in organic/inorganic nanocomposites. *Journal of Optical Society of America B* 2004; 21: 208–213.
36. Lappas A., Zorko A., Wortham E., Das R.N., Giannelis E.P., Cevc P., and Arcon D. Low-energy magnetic excitations and morphology in layered hybrid perovskite-poly(dimethylsiloxane) nanocomposites. *Chemistry of Materials* 2005; 17: 1199–1207.
37. Scully S.R., Lloyd M.T., Herrera R., Giannelis E.P., and Maliaras G.G. Dye-sensitized solar cells employing a highly conducting and mechanically robust nanocomposite gel-electrolyte. *Synthetic Metals* 2004; 144: 291–296.
38. Schmidt D., Saha D., and Giannelis E.P. New advances in polymer/layered silicate nanocomposites. *Current Opinion in Solid State and Materials Science* 2002; 6: 205–212.
39. Vaia R.A. and Giannelis E.P. Polymer nanocomposites. Status and Opportunities *MRS Bulletin* 2001; 26: 394–401.
40. Herndon V.A. Passive Components Technology Roadmap, National Electronics Manufacturing Technology Roadmaps, International Electronic Manufacturing Initiative (www.inemi.org), December 1998.
41. Brown W.D. Advanced Electronic Packaging: With Emphasis on Multichip Modules. *IEEE Press*, Piscataway, NJ, 1999; 24–28.
42. Biswas A. Nanostructured Barium titanate composites for embedded radio frequency applications. *Applied Physics Letters* 2007; 91: 1–3.
43. Xu J. and Wong C.P. High-K Nanocomposites with Core-Shell Structured Nanoparticles for Decoupling Applications. *Electronic Components and Technology Conference* 55th Proceedings 2005; 1234–1240.
44. Lu J., Moon K.S., Xu J., and Wong C.P. Synthesis and dielectric properties of novel high-K polymer composites containing in-situ formed silver nanoparticles for embedded capacitor applications. *Journal of Material Chemistry* 2006; 16: 1543–1548.
45. Lu J. and Wong C.P. Dielectric Performance Enhancement of Nanoparticle-Based Materials for Embedded Passive Applications. *Electronic Components and Technology Conference* 58th Proceedings 2008; 736–741.
46. Das R.N., Poliks M.D., Lauffer J.M., and Markovich V.R. High Capacitance, Large Area, Thin Film, and Nanocomposite Based Embedded Capacitors. *Electronic Components and Technology Conference* 56th Proceedings 2006; 1510–1515.

47. Das R.N. Resin Coated Copper Capacitive (RC3) Nanocomposites for Multilayer Embedded Capacitors. *Electronic Components and Technology Conference 58th Proceedings 2008*; 729–735.
48. Bhattacharya S.K. Epoxy nanocomposite capacitors for applications as MCM-L compatible integral passives. *Journal of Electronic Packaging 2002*; 124: 1–6.
49. Windlass H., Raj M., Balaraman D., Bhattacharya S.K., and Tummala R.R. Colloidal processing of polymer ceramic nanocomposite integral capacitors. *IEEE Transactions on Electronic Packaging Manufacturing 2003*; 26: 100–105.
50. Cannon W.R. Dispersants for non-aqueous tape casting. *Advanced Ceramics 1984*; 9: 164–183.
51. Das R.N. Laser Micromachining of Nanocomposite-Based Flexible Embedded Capacitors. *Electronics Components and Technology Conference 57th Proceedings 2007*; 435–441.
52. Bhattacharya S.K. and Tummala R.R. Integral passives for next generation of electronic packaging: application of epoxy/ceramic nanocomposites as integral capacitors. *Microelectronics Journal 2001*; 32: 11–19.
53. Jayasundere N. and Smith B. Dielectric constant for binary piezoelectric 0-3 composites. *Journal of Applied Physics 1993*; 73: 2462–2466.
54. Tang S.C., Hui S.Y.R., and Chung H.S. A low-profile power converter using printed-circuit board (PCB) power transformer with ferrite polymer composite. *IEEE Transactions on Power Electronics 2001*; 16: 493–498.
55. Brandon E.J., Wesseling E.E., Vincent C., and Kuhn W.B. Printed micro-inductors on flexible substrates for power applications. *IEEE Transactions on Components and Packaging Technologies 2003*; 26: 517–523.
56. Park J.Y. and Allen M.G. Low Temperature Fabrication and Characterization of Integrated Packaging-Compatible, Ferrite Core Magnetic Devices. *Twelfth Annual APEC '97 Conference Proceedings 1997*; 1: 361–367.
57. Park J.Y. and Allen M.G. Packaging-composite high Q microinductors and micro filters for wireless applications. *IEEE Transactions on Advanced Packaging 1999*; 22: 207–213.
58. Rondinone A.J., Samia A.C.S., and Zhang Z.J. A chemo metric approach for predicting the size of magnetic spinel ferrite nanoparticles from the synthesis conditions. *Journal of Physical Chemistry B 2000*; 104: 7919–7922.
59. Dong H., Liu F., and Wong C.P. Magnetic Nanocomposite for High Q Embedded Inductor. *9th International Symposium on Advanced Packaging Materials 2004*; 171–174.
60. Raj M., Muthana P., Xiao T.D., Wan L., Balaraman D., Aboitu I.R., Bhattacharya S.K., Swaminathan M., and Tummala R.R. Magnetic Nanocomposites for Organic Compatible Miniaturized Antennas and Inductors. *10th International Symposium on Advanced Packaging Materials: Processes, Properties and Interfaces 2005*; 272–275.
61. Shah V.G. and Hayes D.J. Fabrication of passive elements using ink-jet technology. *Printed Circuit Fabrication 2002*; 25: 36–38.
62. Nakamoto M. Fine electronic circuit pattern formation on plastic substrates by metal nano particle pastes. *Welding International 2007*; 21: 831–835.
63. Chao T.Y. and Cheng Y.T. Synthesis and Characterization of Cu/CoFe₂O₄ Magnetic Nanocomposite for RFIC Application. *IEEE Transactions 2006*; 810–813.
64. <http://techon.nikkeibp.co.jp/english/img2/nea0305manu1fig1.jpg> (downloaded 7/07/09) TechOn. Cover Story: Motorola Ships Passive-Embedded PCB for Mobile Phones May 2003.
65. Hayes D.J., Wallace D.B., and Cox W.R. Microjet Printing of Solders and Polymers for Multichip Modules and Chip Scale Packages. *Proceedings, IMAPS International Conference on High Density Packaging and MCM's 1999*; 1–6.

Thermally Conductive Nanocomposites

Jan Felba

Abstract The heat dissipation problem is becoming a crucial barrier in the continuous process of electronic devices and systems miniaturization, and thermal interface materials play a key role in transport at all level microelectronics packaging. In this chapter, thermally conductive composites, as one of the main types of thermal interface materials are discussed. Such composites consist of the polymer base material matrix and thermally conducting filler. In modern bio- and microelectronics, the thermal conductivity of the filler material ought to be as high as possible and its size—in the nanometer range. For this purpose, the nanosized particles of metals (mostly silver) or carbon allotropes (mostly carbon nanotubes) are used. Unfortunately, even when the fillers with thermal conductivity in the range of hundreds or more W/m×K are applied, the conductivity of nanocomposites elaborated currently are not higher than a few W/m×K. The analysis of heat transport in filled composites shows that the reason of such limitation of the heat transport is related with thermal contact resistance between filler's particles. The way of the thermal conductivity improvement is discussed; however, in order to evaluate this progress correctly, the proper measurement methods have to be used. Currently there are many technical solutions to measure thermal conductivity which are based on either steady-state or transient techniques, but for relatively low thermal conductivity of nanocomposites, only few methods can be applied. Here the analysis of the thermal conductivity measurement is presented.

Keywords Microelectronics · Thermal conductivity · Thermally conductive composites · Nanomaterials · Nanosilver · Carbon nanotubes · Nanocomposites · Thermal conductivity measurement

The miniaturization is a steady process and the real progress of microelectronic circuits is influenced by practical state of technological level as scaling parameters,

J. Felba (✉)

Faculty of Microsystem Electronics and Photonic, Wrocław University of Technology,
Wrocław, Poland

e-mail: jan.felba@pwr.wroc.pl

impurity, or contamination of the materials and processes, as well as the technical level of technological equipment. The heat dissipation problem is becoming a crucial barrier in miniaturization process and new materials for this purpose are elaborated. The problem of heat transport and heat dissipation, especially in the first and second packaging level, is solved usually by conductive adhesives. At least, there are two main requirements for such material, namely a sufficient mechanical strength of joined components and high thermal conductivity.

Generally, the thermally conductive adhesives can be treated as thermal interface materials (TIMs). Materials of this type play a key role for the heat dissipation at all levels within microsystems and are used for filling the gaps between surfaces to enhance the heat conduction in the package.

It means that the good adhesion to the surfaces cannot be always needed, while the requirement of high thermal conductivity remains crucial to its efficiency. Thermal interface materials can be classified into a few general categories [1, 2] such as thermal greases, phase changed materials, or filled polymers. It is worth to note that low melting point metal solder alloys commonly used for microelectronic packaging are also thermal interface materials, can be classified as phase changed materials.

In this chapter, only filled polymers will be discussed. These composites consist of the polymer base material matrix and thermally conducting filler. All polymeric materials, epoxy or other types, thermoset or thermoplastic, have the thermal conductivity in the range of 0.2–0.3 W/m×K. Practically, only the filler material is responsible for heat transport. Conduction is provided by conductive additives, as high conductivity requires high filler content, considerably above the percolation threshold. It is believed that at this concentration, all conductive particles contact each other and form a three-dimensional network. The size, shape, and content of the filler particles change significantly the viscosity and rheology of the composite in comparison with the pure polymer. The expected conductivity value is obtained after the curing process due to better contacts between filler particles resulting from shrinkage of the polymer matrix. The thermal conductivity is limited by the so-called thermal contact resistance between filler particles. The contact thermal resistance depends on both properties of the materials and geometric parameters of the contact areas between the particles. The geometric parameters are related to the contact pressure within the contact area.

In modern microelectronics, gaps for thermal interface materials are usually not higher than tens of micrometers. This needs filler particles with much smaller dimensions—with size in micro- or even nanoscale. Today's technology makes it possible to produce metal particles with dimension of a few nanometers, which theoretically can be used as the conductive fillers (mostly silver and gold). Also one of the most promising thermally conductive filler—carbon nanotubes have diameters range of single nanometers or less [3]. This allows us to give such a material a name, *composites filled with nanosized particles* or simply *nanocomposites*. The important question which arises is the “nanoeffect” of nanoparticle or fiber inclusion relative to their larger scale counterparts. In fact there are changes of glass transition temperatures for some polymer matrix with nanofiller incorporation [4]. Nevertheless,

many nanocomposite systems can be modeled using continuum models where the absolute size is not important since only shape and volume fraction loading are necessary to predict their thermal properties.

1 Model of Heat Transport

The general transient heat transfer equation can be written as follows:

$$\lambda \nabla^2 T - c \rho \frac{\partial T}{\partial t} + q_v = 0 \quad (1)$$

where T is the temperature, λ the coefficient of thermal conductivity, c the specific heat, ρ the material density, while q_v stands for other inside and outside heat sources or phenomena. The above equation consists of three components of which:

- the first one is responsible for heat conductance,
- the second one for heat accumulation,
- while the third one for the internal heat generation or outside heat dissipation.

The basic principle of heat conductance through the thermally conductive adhesive layer dx (Fig. 1) is that heat is transferred in proportion to the gradient of temperature, which can be rewritten as follows:

$$\frac{dQ}{dt} = -\lambda A \frac{dT}{dx} \quad (2)$$

where Q is the quantity of heat energy (J), t the time (s), λ the thermal conductivity ($\text{W}/(\text{m} \times \text{K})$), dT/dx is temperature gradient in the heat flow direction (K/m), x the direction of heat flow (m), and A the area of the cross-section (m^2). Thermal conductivity is given by

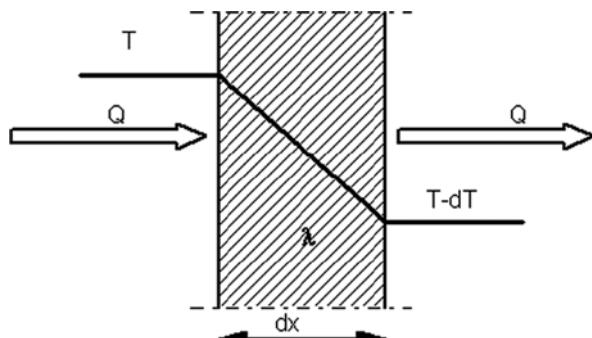


Fig. 1 Heat transfer through a layer

$$\lambda = \frac{1}{3} (c_e v_e L_e + c_{ph} v_{ph} L_{ph}) = \lambda_e + \lambda_{ph} \quad (3)$$

where c_e and c_{ph} are the heat capacities per unit volume (J/m^3K) of electrons and phonons, respectively, v_e and v_{ph} are their root-mean-square velocities and L_e , L_{ph} are their mean free paths.

For metals, the λ_e is the dominant part of the thermal conductivity. Thermal conductivity depends on temperature and for temperatures higher than room temperature, λ of pure metals remains essentially independent on temperature or decreases with increasing temperature in the case of metals of high conductivity (Fig. 2) [5].

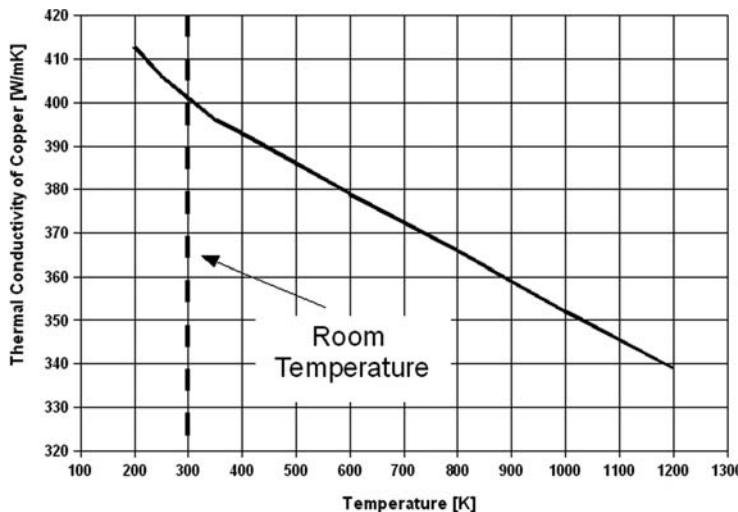
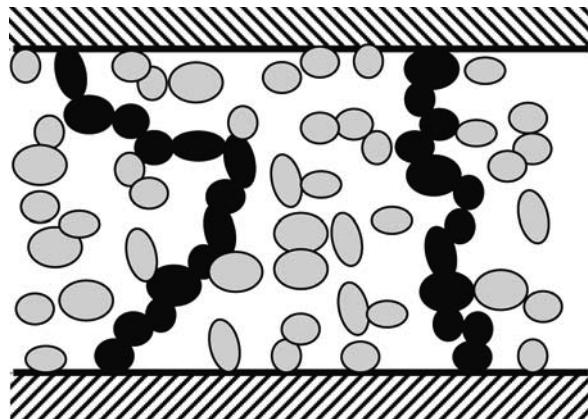


Fig. 2 Dependence of thermal conductivity on temperature for copper [5]

The transport of heat in nonmetals occurs mainly by phonons. The thermal conductivity is restricted by various types of phonon scattering process. In order to maximize the thermal conductivity, these phonon scattering processes have to be minimized. For insulators thermal conductivity linearly increases with temperature.

It is assumed that in macroscale the formulation of thermally conductive composites, which consist of the polymer base material matrix, filler, and some special additives are uniform and isotropic, although slight heterogeneity was also measured [6]. With enough high content of the filler when, the conductive paths appeared (Fig. 3), the heat conduction may depend on bulk thermal conductivity of a filler material. For the single particle of filler its bulk conductivity is determined by the material used. From this point of view, the diamond with conductivity of $2000 \text{ W/m}\times\text{K}$ is a better material than alumina (slightly better than $30 \text{ W/m}\times\text{K}$). In the case of composites formulated such as electrically conductive adhesives, which contain silver flakes with average particle dimensions of several micrometers, and thermosetting epoxy resin as polymer matrix, the thermal conductivity value does not exceed $3 \text{ W/m}\times\text{K}$ [7]. This enormous difference between thermal conductivity

Fig. 3 The phenomenon of percolation occurring with the formation of a continuous chain (*dark balls*) between the two surfaces by the conducting particles



of pure silver (about $420 \text{ W/m} \times \text{K}$) and thermally conductive adhesive with this filler is due to the constraint by the existing so-called thermal contact resistance between filler particles. It is the main “bottleneck” for heat transport inside composite formulation.

There are many analytical models to predict the heat transport between particles, which may be used as filler in thermally conductive composite. According to literature review concerning thermal interface materials [8, 9] the most common and important analytical models mostly taking into account spherical particles can be listed as follows:

- The Maxwell–Garnett effective medium model [9–12] can be considered as a good model for low-volume fraction, up to 40%. Therefore it is not used in modeling of the heat transport for most common composites in which conductive particle content exceeds 60% [8].
- The Bruggeman symmetric model for high-volume fraction. One of the biggest drawbacks of this model is that it does not include the interface resistance between the particle and the matrix [8].
- The Bruggeman asymmetric model. After modification it includes thermal contact resistance between the particle and the matrix. This model is capable of predicting thermal conductivity of spherical particles for larger volume fractions.
- Percolation models. Percolation is a geometrical phenomenon, which means that above a particular volume fraction (called the percolation threshold) there is a continuous path for heat conduction through the particles because the conducting particles start to touch each other as shown in Fig. 3.

The classic percolation theory assumes a statistical distribution of the filler particles in a matrix. Nevertheless, the distribution of nanoparticles in a polymer matrix cannot be completely random because of an aggregation of particles. In the case of nanosilver with epoxy resin, it was observed that a conductive network can be formed even if the content of particles is lower than the percolation

threshold estimated conventionally [13]. With increasing temperature, self-organization of particles helps in the formation of the conducting path throughout the matrix.

2 Thermal Contact Resistance

There are two basic materials formulating thermally conductive composites—polymer matrix and thermally conductive filler. As polymeric materials are about 2000 times less conductive than metals with high thermal conductivity, such as silver or copper and several thousand times less than carbon allotropes (e.g., carbon nanotubes or graphenes), the heat transport is possible by conducting pathways (Fig. 3). Theoretically, the effectiveness of heat transfer depends on conductivity of a filler bulk material and thermal contact resistance between filler particles.

Generally, the thermal resistance of both bulk material and contacts can be expressed as thermal conductors. This is allowed because of analogy between thermal and electrical resistance defined by Ohm's law, where instead of the potential difference across the object and the current passing through the object, the temperature difference occurs provoking a heat flow. Due to this analogy the concept of a thermal resistance can be introduced

$$\Theta_\lambda = \frac{l}{\lambda A} \quad (4)$$

where Θ_λ is the thermal resistance of a thermal conductor (K/W), λ the thermal conductivity, l the length of thermal conductor (m), and A the area of the conductor cross-section (m^2). The total thermal resistance in “a three-dimensional network” of conducting filler particles can be summed using the concepts of serial and parallel resistances as in an electrical circuit. The thermal resistance Θ_p of the singular contact path in the space between two surfaces (filled by composite) in the steady-state conditions can be expressed in a simple equation:

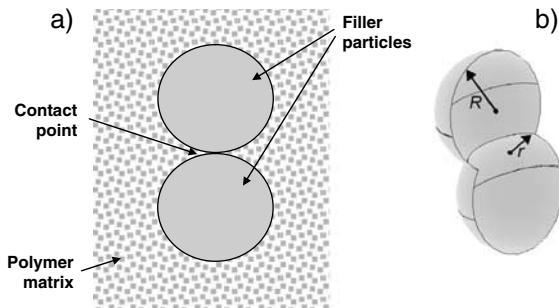
$$\Theta_p = \Theta_{\lambda 1} + \Theta_{\lambda 2} + \dots + \Theta_{\lambda n} + \Theta_{TC1} + \Theta_{TC2} + \dots + \Theta_{TC(n-1)} \quad (5)$$

where Θ_λ is the thermal resistance of the filler particle (bulk resistance) and Θ_{TC} the thermal contact resistance between particles forming a chain in the direction of the temperature gradient. The total thermal resistance of composite Θ_{TA} is

$$\Theta_{TA} = \left(\frac{1}{\Theta_{p1}} + \frac{1}{\Theta_{p2}} + \dots + \frac{1}{\Theta_{pn}} \right)^{-1} \quad (6)$$

If contact members are in the form of perfect, hard balls, then they touch each other in one point (contact point “A” in Fig. 4a). In fact, this point turns into a

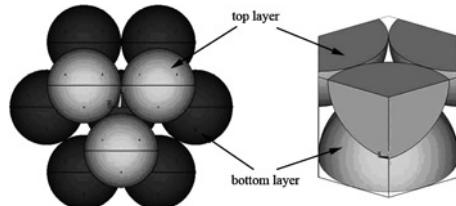
Fig. 4 Two-dimensional model of ideal balls (a) and ball-shape particle after deformation (b) filled thermally conductive polymer



small area of radius r of the filler particles contact area since contact materials are deformable (Fig. 4b).

The degree of this deformation influences the contact areas and consequently—the thermal contact resistance. To answer the question on how the contact state between particles affects the thermal conductivity of a polymer, the numerical modeling of silver-filled formulation was used [14, 15]. The numerical model was simplified by some assumptions: the identical, spherical-shaped silver particles were used as a filler and the unit cell has maximal possible packing efficiency ($V_{\text{filler}}/V_{\text{total}} = 74.05\%$) in three-dimensional for spherical particles, i.e., hexagonal close-packed (HCP) or face-centered cubic (FCC) crystal equivalent three-dimensional structure was analyzed (see Fig. 5). Additionally it was assumed that particles contact each other in the ideal pure metallic contact area (the impurity, roughness, etc. were neglected).

Fig. 5 Close-packed structure (left) and analyzed unit cell (right)



The unit cell shown in Fig. 5 was analyzed in ANSYS 8.0 commercial finite element method software. The two different contact area radii were taken into account represented by the relation to particle radius r/R where r is contact area radius and R is particle radius, shown in Fig. 4b. The simulation results were gathered in Table 1.

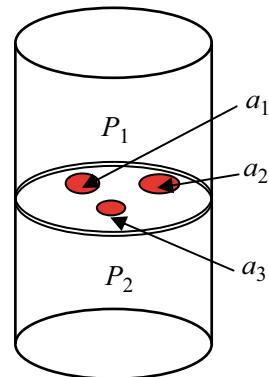
Table 1 Numerical simulation results

r/R	λ (W/m×K)
0.05	49.12
0.02	26.46

As shown in Table 1, the contact area between filler particles has a strong influence on thermal conductivity of the adhesive. This area depends on stresses between particles, which occur due to shrinkage reaction of resin during curing. The role of polymer matrix for this seems to be crucial. The stress forces may decrease thermal contact resistance of adjoined particles of the filler, but the calculation of these stresses is quite complex, because of its relaxation in time due to viscoelastic behavior of the resin [16–18].

Instead of identical, close-packed, spherical-shaped filler particles, the heat flow can also be analyzed when particles touch each other by flat surfaces. If the contact areas are nominally flat and hard, then they touch each other maximum in three points (Fig. 6). In fact, these points become the small areas (*a*-spots) since contact member materials are deformable [19]. If the contact surfaces are perfectly clean, then the total metallic contact area A_c is the sum of all *a*-spot areas.

Fig. 6 Two cylindrical filler particles P_1 and P_2 forming the apparent contact area
 $A_c = a_1 + a_2 + a_3$



The steady-state heat flow from the top of the two-particle unit of higher temperature (T_1) to the bottom of lower temperature (T_2) is constricted through small conducting *a*-spots. Therefore, in the case of adhesives, when polymer matrix fills the contact areas (except A_c surface), the total thermal resistance of such units is the sum of the bulk thermal resistances $\Theta_{\lambda 1}$ and $\Theta_{\lambda 2}$ of both contact members, the constriction resistances Θ_{c1} and Θ_{c2} of contact members, and real thermal resistance of *a*-spot bump Θ_b (if *a*-spot is treated as three-dimensional structure). This effect of heat flow constriction through the *a*-spot is identified with thermal constriction resistance.

The thermal constriction resistances of two cylindrical copper particles (Fig. 6) were calculated using the finite element method [20]. Due to symmetry, only quarter of the system was analyzed. The computer calculations were performed assuming that the thermal conductivity of polymer matrix can be neglected, the contact surface is perfectly clean, without oxides and alien films, the *a*-spot is positioned on the axis of the contact system and the radius a of this metallic contact area is equal to $0.0005R \dots 0.05R$, where $R = 10$ mm is the radius of the contact members as well as of apparent contact area. The length of each contact member is 40 mm. The gap

between both contact members ranges from 0.1 up to 10 μm . These values were chosen in such a way that the thermal constriction resistance could be regarded as “long resistance.”

The result of calculation shows (Figs. 7 and 8) that both the gap between filler particles and the contact area influence thermal constriction resistance significantly. In fact, this resistance can be treated as the thermal contact resistance Θ_{TC} . For analyzed two-particle unit in temperature of 393 K (for copper $\lambda = 364.1$), the bulk resistance of both filler particles is about 0.7 K/W, while the thermal contact resistance $\Theta_{TC} = 221 \text{ K/W}$ for both the radius of a -spot and the gap equal 10 μm

Fig. 7 The thermal constriction resistance of two-particle unit vs. temperature of heat source, with different gap distance between filler particles; the radius of a -spot is 0.01 mm [20]

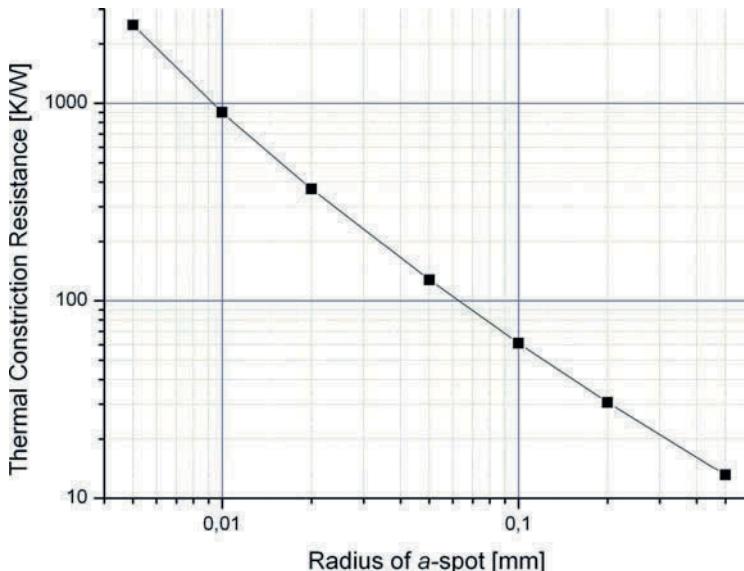
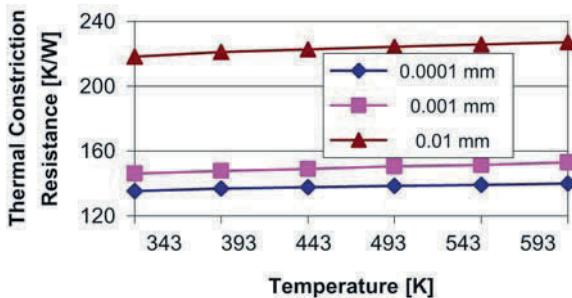


Fig. 8 The thermal constriction resistance of two-particle unit vs. contact area with different heat source temperature; the gap distance between filler particles is 10 μm [20]

(from Fig. 7). It means that the total resistance of thermally conductive composite depends mainly on the values of thermal contact resistance and numbers of contacts between filler particles making a chain in the direction of temperature gradient (Θ_{TC} and n in Equation 5). The bulk resistance plays a minimal role.

Thermal conductive components usually are applied between two surfaces with different temperatures, e.g., between heat-generating component and the eventual heat sink. Some of those materials consist of permanent bonds, such as adhesive, but often these are only bolted between other elements. In such cases, the calculation of heat transfer is very difficult, as the real contact area is uncontrolled, a small fraction of the nominal or apparent area. This is mainly due to nonflat shape of contacting surfaces, surface roughness, deformation possibility of contact materials under pressure, additional materials between contacting surfaces, etc. In such a case of heat transport, the thermal resistance Θ_T can be expressed by the formula

$$\Theta_T = \Theta_{IR'} + \Theta_{TA} + \Theta_{IR''} \quad (7)$$

where $\Theta_{IR'}$ and $\Theta_{IR''}$ are interfacial thermal resistance values between a composite and surface elements. When thermally conductive adhesives are used for bonding of two elements in the way of heat transport, the thermal resistance Θ_T of a joint can be defined as

$$\Theta_T = \Theta_{BR'} + \Theta_{TA} + \Theta_{BR''} \quad (8)$$

where $\Theta_{BR'}$ and $\Theta_{BR''}$ stand for the bond thermal resistance of contacts between adhesive layer and joined elements, while Θ_{TA} represents the total thermal resistance of the adhesive (Equation 6). In such a case, an analysis of bond thermal resistance influence on Θ_T can be done [21–23].

3 Thermal Conductivity Measurements

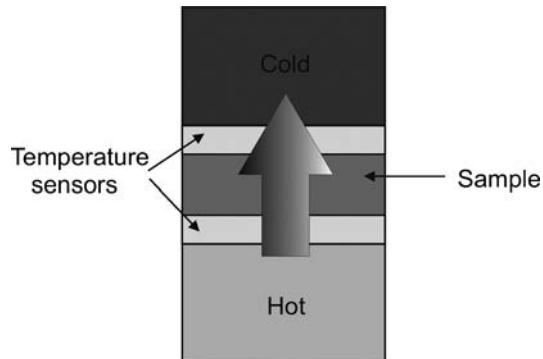
Thermally conductive composites reported in the literature are characterized by thermal conductivity not higher than a few $\text{W}/\text{m}\times\text{K}$ for formulations with both, commonly used filler materials (like silver [7, 24]) and special materials (diamond powder [19] or carbon fiber [25]). But there are also reports about polymer-matrix and silver-filled adhesives with a few times higher thermal conductivity [26–28]. Probably various methods of thermal conductivity measurement and their errors as well as low accuracy cause significant differences in measured thermal conductivity of similar adhesive formulations. To analyze thermal data of adhesives, the knowledge about used measurement method is necessary.

The history of heat-transfer measurement began in the second decade of the last century [29]. Currently there are two main categories of techniques to measure thermal conductivity, steady-state techniques, and transient techniques. The heat flow method and the guarded hot-plate method are the best examples of steady-state techniques. Hot wire and laser flash can be listed as the most popular measurement methods of the transient technique, but also the 3ω , photoacoustic method, the

pulsed photothermal displacement technique, and the thermal-wave technique are in use [30].

The steady-state methods are based on establishing a steady temperature gradient over a known thickness of a sample and on controlling the heat flow from one side to the other. The determination of the thermal conductivity of the insulation follows from basic law of heat flow (Equation 2). The principle of heat flow method is shown in Fig. 9.

Fig. 9 The principle of heat flow method



Among the steady-state methods, the guarded hot plate is the most common technique for measuring the thermal conductivity of insulation or poor thermal conductor materials. Its operation is based on establishing a steady temperature gradient over a known thickness of a sample and on controlling the heat flow from one side to the other. The determination of the thermal conductivity of the insulation follows from the applicable approximation heat flow law in one dimension. The specimen is placed between a heat source and a “cold plate” or heat sink, both of which contain temperature sensors very close to their surfaces. The sensors measure the temperature drop across the sample after a steady-state heat flow has been established. The heat source is surrounded by a guard heater, maintained at the heat source temperature that minimizes heat losses out of the edges. In this method, the thermal resistance at the sample interfaces must be minimized and well controlled for consistent results. The most commonly employed instruments create a steady-state temperature gradient across a specimen of the material by sandwiching it between isothermal hot and cold plates (Fig. 10).

There are a few modifications of the basic guarded hot plate method. In contrast to a conventional system that used uniformly distributed heaters, the line heat source was applied, which improved measurement accuracy by determining the errors resulting from heat gains or losses at the edges of the specimens [31]. The method's error was diminished by more stable operation at liquid nitrogen temperatures [32] and measurement under vacuum conditions [33]. But generally, the heat flow method are based on American ASTM C 518-98 and European ISO 8301 Standards, while guarded hot plate method—on American ASTM C 177-97 and European ISO 8302 Standards [34].

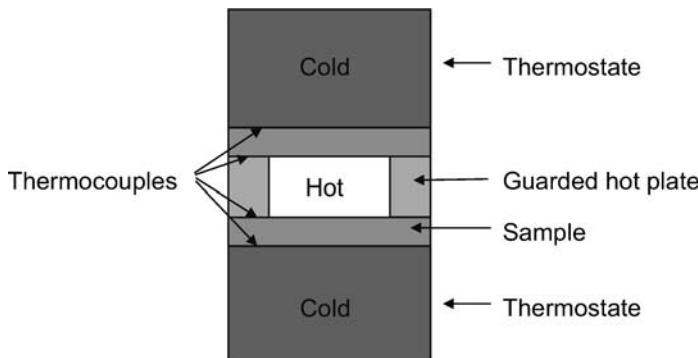


Fig. 10 The principle of the guarded hot plate method measurement

As the main disadvantage of the steady-state methods, the long time required to obtain a single measurement is mentioned. But it is worth to know that there are many sources of errors, which can influence dramatically the results of thermal conductivity measurement, especially when materials with low λ (like thermally conductive polymer-base materials) are tested. The high accuracy of measurement ought to meet at least the following requirements:

1. Heat is transferred only by conduction in a measurement setup
2. The heat flow is simply measured (e.g., by temperature drop)
3. The thermal resistances are taken into consideration
4. Data of measured as well as reference materials are independent from temperature, or thermal conductivity changes vs. temperature (see Fig. 2) are taken into consideration.

An experimental setup which fulfills most of the above listed requirements is shown in Fig. 11. The system is based on ASTM E 1530 Standard (Standard Test Method for Evaluating the Resistance to Thermal Transmission of Materials by the Guarded Heat Flow Meter Technique) consists of heater, heat sink, a pair of contact members (cylinders with a known thermal conductivity λ), as well as temperature sensors [35]. The whole setup is placed in high vacuum environment in order to minimize the effect of heat dissipation through convection. The radiation is out of the temperature work range and is therefore neglected and the heat conduction is a dominating factor in transport heat energy. When the experimental setup reaches the steady state then by extrapolating the readings of the temperature sensors it is possible to find out the temperature drop ΔT at the contact interface [36], which is shown in Fig. 12.

The experimental setup [7] consists of a pair of iron contact members ($\lambda_{Fe} = 80.4 \text{ W/m}\times\text{K}$) with a diameter of 60 mm with six thermal sensors (T_1-T_6), a bulb halogen heater immersed in a bottom cylinder and a water cooling system on the top. In the vacuum conditions, the readings of the temperature sensors were taken

Fig. 11 An experimental setup (from [7])

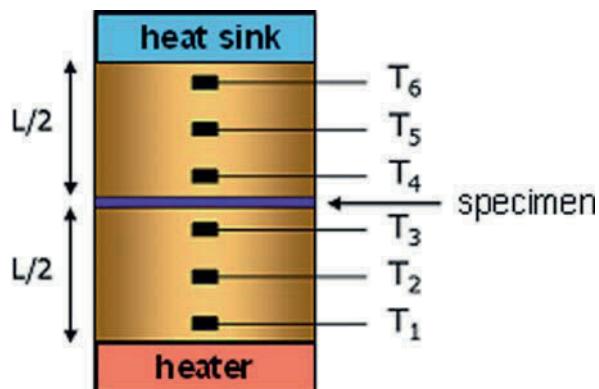


Fig. 12 Thermal distribution in steady state; T_1-T_6 —measured temperature, $\Delta T_{TC1,2}$ —temperature gradients in contact areas (unknown); ΔT_b —temperature gradient on measured sample (unknown), ΔT —the calculated temperature drop between a pair of iron contact members

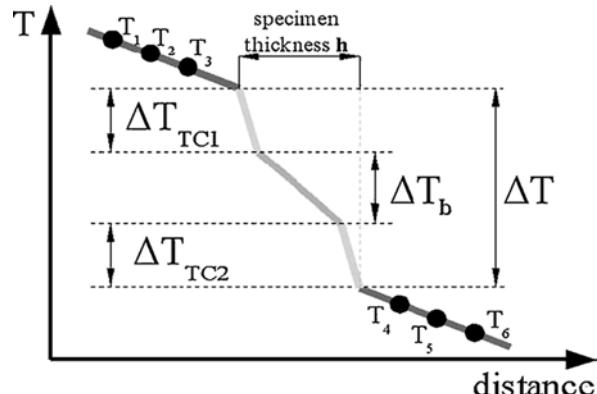
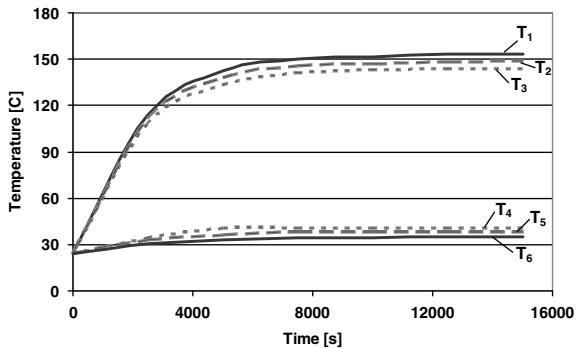


Fig. 13 The temperature distribution inside the contact members



into calculation when the experimental setup reached the steady state (Fig. 13) and then by extrapolating, it is possible to find out the temperature drop ΔT (Fig. 12) at the tested specimen.

To eliminate the influence of interfacial thermal resistances (Θ_{IR}' and Θ_{IR}'' in Equation 7), it is assumed that both those values are equal and it is necessary to perform at least two measurements with different thickness of specimens. Changing thickness of a sample from h to $h + \Delta h$ causes the increase of thermal resistance of a material but the thermal contact resistances remain unchanged. On the basis of Equations 2 and 6, it is possible to extract the thermal conductivity of the material tested [23]:

$$\lambda_b = \frac{\Delta h}{A} \cdot \left(\frac{\Delta T_2}{q_2} - \frac{\Delta T_1}{q_1} \right)^{-1} \quad (9)$$

where q_1 and q_2 are heat flows evaluated based on Equation 2 taking into account temperature gradients and thermal conductivity λ of the contact members for both samples and A is the contact area between the contact member and a measured sample. The measuring set was used successfully for thermal conductivity measurement of pure materials (e.g., polytetrafluoroethylene (PTFE)—0.246 W/m×K) as well as polymer composites filled with silver flakes (below 3 W/m×K).

As mentioned, it is very important to know uncertainty of results, especially when thermally conductive polymer-base materials are measured. The limiting error of the above described measurement method is given by [37]:

$$\Delta_{le} \lambda_b = \sum C_i \Delta_{le} X_i \quad (10)$$

where $\Delta_{le} \lambda_b$ is the limiting error of thermal conductivity of tested material, c_i the partial differential and $\Delta_{le} X_i$ is the limiting error of each variable. Uncertainty U of λ_b measurement with 95% probability is equal to:

$$U(\lambda_b) = 1.15 \sqrt{\sum (c_i \cdot \Delta_{le} X_i)^2} \quad (11)$$

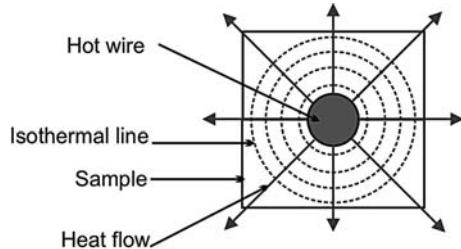
In the case of polymer-based material filled with carbon nanotubes its thermal conductivity λ_b equals 1.43 ± 1.1 W/m×K with 95% probability [37] was measured.

The transient methods measure a response as a signal is sent out to create heat in the sample. During the time of measurement the change in temperature is monitored.

For measurement of thermal conductivity of composites (also nanocomposites filled with carbon nanotubes [38]) the transient hot-wire technique can be used. This method is a transient dynamic technique based on a linear heat source (Fig. 14) of infinite length and infinitesimal diameter and on the measurement of

the temperature rise at a defined distance from the linear heat source embedded in the test material. As an electric current of fixed intensity flows through the wire, the thermal conductivity can be derived from the resulting temperature change over a known time interval.

Fig. 14 The radial heat flow model of the hot-wire measurement method



If the temperature of a hot wire on the boundary surface between a reference medium and a medium of unknown thermal conductivity, λ , is measured by a thermocouple at times, t_1 and t_2 , respectively, the thermal conductivity of the unknown medium can be obtained from the following equation [34]:

$$\lambda = K_1 \frac{P}{[T(t_2) - T(t_1)]} \ln \left(\frac{t_2}{t_1} \right) - H \quad (12)$$

where K_1 is a constant depending on the electric power supplied to the heating wire (P), H is a constant term that depends on the insulating material of the probe, and T the temperature.

In this technique, small diameter wires are immersed in the fluid and used simultaneously as electrical resistance heaters and as resistance thermometers to measure the resulting temperature rise due to the resistance heating. Two hot wires of differing length are operated in a differential mode to eliminate axial conduction effects due to the large diameter leads attached to the ends of each hot wire. Based on the transient line-source model, the thermal conductivity can be found from the slope of the measured linear temperature rise as a function of elapsed time, while the thermal diffusivity can be found from the intercept of this same linear temperature rise curve [39].

The transient hot-wire technique has been modified and nowadays four variations of the method are in use [40]: hot-wire standard technique, hot-wire resistance technique, two thermocouple technique, and the hot-wire parallel technique. The theoretical model is the same, and the basic difference among these variations lies in the temperature measurement procedure. It makes possible to determine the thermal conductivity of a wide range of materials, including cured thermally conductive composites.

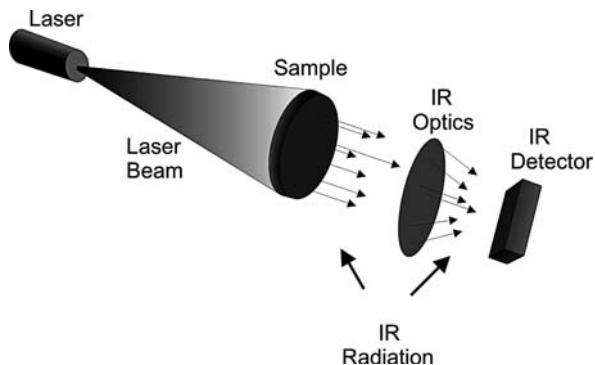
There are also a number of presently existing transient methods of measuring indirectly thermal conductivity—by measurement of material's thermal diffusivity. The relation between the thermal conductivity λ (W/m×K) and the thermal diffusivity α (m²/s) is given by

$$\lambda = \alpha \cdot \rho \cdot c_p \quad (13)$$

where ρ (kg/m^3) is the density and c_p ($\text{J}/\text{kg} \times \text{K}$) is the specific heat of the tested composite.

Among the transient methods, the so-called flash method is the most popular in determining the thermal diffusivity. The method was described in 1960 [41] and was developed during the next years [42–45]. In this method, a pulse of energy is absorbed instantaneously on the front face of a small disc-shaped specimen and the subsequent temperature change of the rear face is recorded. The front surface of the sample must be uniformly irradiated at a short time compared to the rise time of the back surface temperature. Figure 15 illustrates the principle of flash method with a laser beam as an energy source.

Fig. 15 The principle of thermal diffusivity measurement by using a laser flash method

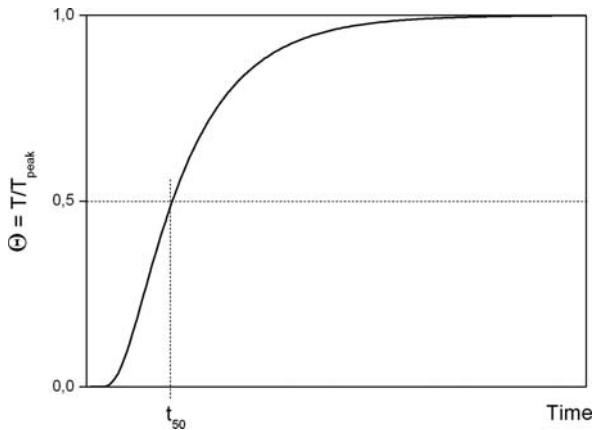


In determination of a thermal diffusivity of tested material, the magnitude of the temperature rise recorded on the back side of the sample and the amount of energy are not required. Only the shape of the time–temperature curve is used in further analysis [46]. The thermal diffusivity can be directly calculated from the curve and the sample thickness. If assumed that the sample is perfectly insulated from the environment during the test (there is no heat exchange) and the whole energy is absorbed instantaneously (zero pulse width) in a very thin layer of the sample material then the thermal diffusivity can be evaluated by using the simple relation [41]:

$$\alpha = 0.1388 \frac{d^2}{t_{50}} \quad (14)$$

where d (m) is a sample thickness and t_{50} (s) is the time required for the back face of the sample to reach half of the peak value of temperature (Fig. 16). This method of calculation of the thermal diffusivity describes the ideal case, but because of the heat loss through radiation (especially when the measurements are done at high temperature) and nonzero pulse width, the other types of analyses are also used [42, 47].

Fig. 16 Dimensionless plot of rear surface temperature history



The flash methods are distinguished in view of the source of different heat pulse, but today laser beam is the most popular as a heat source. The measured sample may be in the form of a disc with a diameter of a dozen or so millimeters and a few millimeters in thickness. The laser beam energy absorption depends on the sample surface. To minimize the mirror effect, the side on which the laser beam would hit may be coated with carbon. Under the carbon layer, the sample can be coated with gold. This layer is used to enhance the thermal contact and to prevent direct transmission of the laser beam through the specimen made with transparent polymer [48]. In spite of such procedure, the accuracy of the laser flash method cannot be too high. There are a lot of errors in the method caused by uncontrolled thermal conditions of the heated sample side, heat loss because of convection (immersing sample in air), contact with thermal sensor on the sample rear side and measuring the specific heat of tested composite (necessary for thermal conductivity calculation—Equation 13).

To avoid uncontrolled thermal conditions caused by laser beam reflection from a polymer composite surface, the electron beam as a more flexible heat source can be used. The absorptivity of an electron beam is extremely high in comparison to a laser beam and almost the whole energy may be absorbed on sample surface. Additionally this heat source enables exact temperature control of heated sample's surface. For this purpose, the specially designed buffer of high conductivity material (e.g., pure copper) ought to be used to provide the heat pulse on a sample surface and protect it against being destroyed by high-power electron beam. It is also necessary as the main problem with using an electron beam in the flash method is the impossibility of direct operation on the surface of no electrically conducting polymer-based materials. The measuring process is conducted in vacuum conditions and heat is transferred only by conduction in the measuring setup. The possible setup for testing thermally conductive adhesives is presented in Fig. 17 [49]. The top surface of the experimental setup was irradiated during test by pulse of defocused and deflected

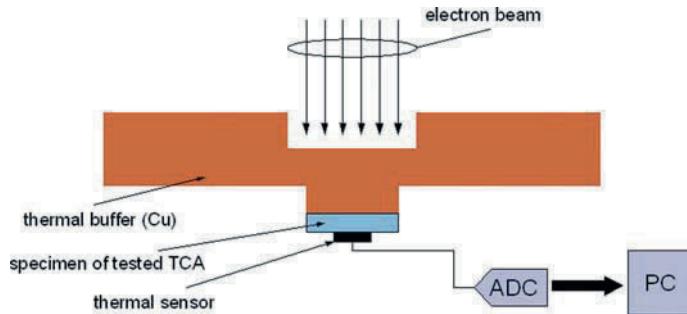
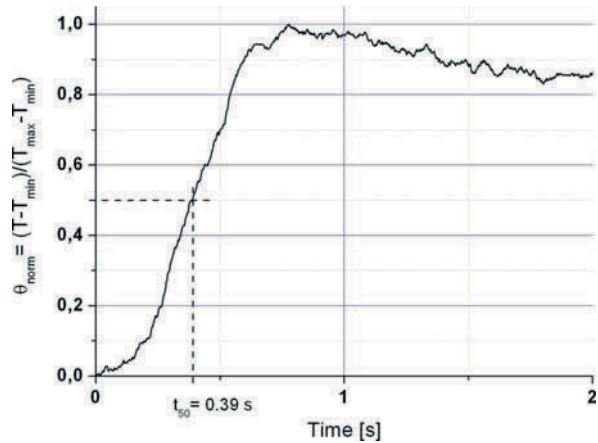


Fig. 17 Experimental setup designed for thermal diffusivity measurements of thermally conductive adhesives (TCA) using electron beam

electron beam to obtain uniform energy distribution on the heated surface. The electron beam pulse shape and length was controlled by using a timer. The temperature was monitored on the back surface of a tested material specimen, which was located at the bottom side of the buffer.

Figure 18 presents the experimental results (normalized temperature on the rear face of sample vs. time) [49]. Based on Equation (14) and data from Fig. 18, the thermal diffusivity for tested sample was calculated as $2.00 \times 10^{-7} \text{ m}^2/\text{s}$.

Fig. 18 The experimental result (normalized temperature measured on the rear face of sample vs. time) obtained for 0.75 mm thick adhesive sample containing 70 wt% of Al_2O_3 as a filler



The photoacoustic technique is another transient method for thermal conductivity of thin layer measurement. In this technique, a heating source (normally a laser beam) is periodically irradiated on a sample surface. The acoustic response of the gas above the sample is measured and related to the thermal properties of the sample. The method for single layer on a substrate was explained by Rosencwaig and Gersho [50] and also multilayered materials were measured [51].

A schematic diagram of the measuring setup is shown in Fig. 19 [52]. The laser beam is sinusoidally modulated by an acoustic-optical modulator AOM driven by

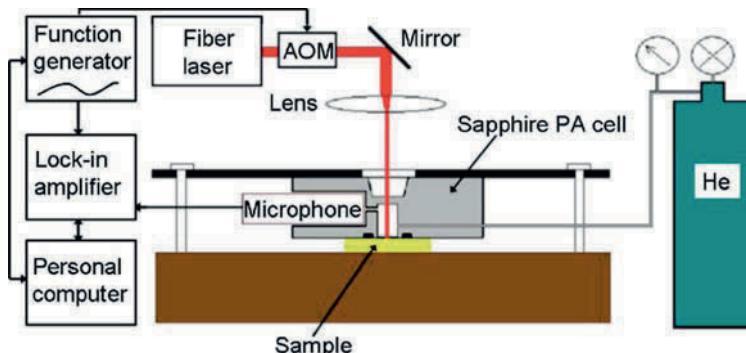


Fig. 19 The schematic diagram of the system for photoacoustic method [52] Courtesy of American Institute of Physics

a function generator. After being reflected and focused, the laser beam is directed onto the sample mounted at the bottom of the photoacoustic (PA) cell. This cell is pressurized by flowing compressed gas (He). A microphone, which is built into the PA cell, senses the acoustic signal and transfers it to a lock-in amplifier, where the amplitude and phase of the acoustic signal are measured (instead of a microphone, a piezoelectric transducer can also be used [53]). A personal computer is used for data acquisition and control of the experiment.

A tested sample can consist of any arbitrary number of layers and a backing material. Absorption of the laser beam is allowed in any layer and in more than one layer. The transient temperature field in the multilayer sample and gas can be derived by solving a set of one-dimensional heat conduction equations, and the transient temperature in the gas is related to the pressure, which is measured experimentally. Because the transient temperature in the gas is related to the thermal properties of the sample, measuring the pressure allows determination of the thermal quantities, usually the interfacial thermal resistance Θ_{IR} and thermal diffusivity.

Figure 19 presents equipment that was used for measurement of thermal conductance across multiwalled carbon nanotube array interfaces, but also liquid samples [54, 55], transparent materials [56], and other materials [57] including thin films [58] can be measured.

For thermal properties, estimation of dielectric composites and the three omega (3ω) thermal conductivity measurement method may be used. The method employs a metallic strip [59] or wire [60] in intimate contact with the specimen surface. An AC electrical current modulated at regular frequency ω is induced to flow in the strip causing heat generation in the strip. The heating has both a DC component which changes the average temperature of the specimen and an AC component at 2ω , which generates thermal waves in the specimen. Because the electrical resistance of the strip depends on the temperature, the resistance will be modulated at 2ω as well. Therefore, there will be an AC voltage drop across the ends of the strip at 3ω , $V_{3\omega}$, which is proportional to the AC temperature variation of the strip at 2ω , $T_{2\omega}$. $T_{2\omega}$ depends on thermal conductivity λ of measured material. Thus, it is possible to extract λ from a measurement of $V_{3\omega}$ versus ω [61].

The 3ω method can also be used for measuring thermal conductivity of quite different materials such as electrically conducting liquids [62] or the phase-change material [63].

The above-described transient method of thermal properties measurement do not exhaust all possibilities of measurement. But other methods either are modifications of the presented ones (by source of heat or signal sensors) or cannot be used for measurement of thermal conductivity of nanocomposites or more general—thermal interface materials.

For commercially available composites and quick comparison of the products, uniform methods for determining the thermal conductivity are proposed. National Fenestration Rating Council [64] recommends ASTM standards. In accordance with these standards, at least three samples shall be measured and the mean value reported. The test shall be done in accordance with ASTM C177, C518, C1114, E1225, E1530, or E1952, as applicable for the specific material. The test equipment shall be calibrated at least once each year as per the recommendations in the appropriate ASTM documents. The choice of standard depends on the possibility of the heat transfer by measured sample (Table 2).

Table 2 Summary of thermal conductivity standard series ASTM and their applicability [64]

Standard	Conductivity range	Conductance range	Min ΔT	Remarks
C177 ^a		< 16 W/m ² K	10 K	Thermally homogeneous
C518 ^b		< 10 W/m ² K	10 K	Thermally homogeneous
C1114 ^c			None	Thermally homogeneous, for specimens with moisture
E1225 ^d	0.2–200 W/m K		90–1300 K	Thermally homogeneous, Opaque
E1530 ^e		25–500 W/m ² K	150–600 K	Thermally homogeneous
E1952 ^f	0.1–1.0 W/m K		None	Thermally homogeneous

^aASTM C177—Standard Test Method for Steady-State Heat Flux Measurements and Thermal Transmission Properties by Means of the Guarded-Hot-Plate Apparatus

^bASTM C518—Standard Test Method for Steady-State Thermal Transmission Properties by Means of the Heat Flow Meter Apparatus

^cASTM C1114—Standard Test Method for Steady-State Thermal Transmission Properties by Means of the Thin-Heater Apparatus

^dASTM E1225—Standard Test Method for Thermal Conductivity of Solids by Means of the Guarded-Comparative-Longitudinal Heat Flow Technique

^eASTM E 1530—Standard Test Method for Evaluating the Resistance to Thermal Transmission of Materials by the Guarded Heat Flow Meter Technique

^fASTM E1952—Standard Test Method for Thermal Conductivity and Thermal Diffusivity by Modulated Temperature Differential Scanning Calorimetry

4 Composites with Metallic Fillers

Since the 1990s, isotropically conductive adhesives have been widely used in electronic packaging. Such composites consist of a polymer resin and electrically conductive fillers, mostly silver, although gold, nickel, and copper are also used. Silver is unique among all cost-effective metals by nature of its oxide, which unlike other metal oxides, is a good conductor. Silver with thermal conductivity $420 \text{ W/m}\times\text{K}$ has been considered to be the most suitable material also as the thermally conductive filler. In fact, the typically formulated electrically conductive adhesives with silver filler in the form of micronized flakes can also be used as thermally conductive adhesive.

Equation 5 points out that the less the contact between particles, the greater the thermal conductivity of the composite. This requires bigger sized fillers. It was found that at the alumina filler with a fixed loading level of 50 wt%, the thermal conductivity increases about 1.7 times when fillers of $<10 \mu\text{m}$ average particle size were replaced by much bigger ones—not higher than $149 \mu\text{m}$ (-100 mesh) [65]. This way of composite thermal conductivity improvement is not suitable as the miniaturized packaging scale and technology of composite dispensing need dimension of filler particles much less than a few micrometers. According to Equation 6 it is natural that the more the conducting path, the lower the thermal resistance of the composite. This can be provided by higher filler loadings. In Fig. 20, the thermal conductivity of the epoxy adhesives as a function of the volume loading of AlN filler is presented. The nonlinear conductivity increasing with an increase of the filler content suggests that additional paths for heat transport are being formed not only by new chains, but also by multiplication of thermal contacts between filler particles in the whole network of particles. This way of conductivity improvement is also limited because of a need for the proper viscosity of formulation and mechanical strength of adhesion in the case of adhesives.

As was stated earlier, bigger particles of the filler cause higher thermal conductivity of the composite. It means that with the use of nanosized silver particles, the thermal contact resistance (Θ_{TC}) between filler particles increases significantly and the thermal resistance (Θ_p) of the singular contact path between two surfaces also increases.

Nanosilver is used in today's miniaturized electronics mostly for conductive microstructures and contacts with dimensions in the range of tens of micrometers. Such lines, or patterns can be made by inkjet technology using a liquid containing nanosized particles of silver [67]. There are many methods of producing nanosilver particles, even in the size below 8 nm [68–73]. During the production process, all particles are uniformly dispersed without possibility of agglomeration. For this each particle is protected by a special kind of layer, usually an organic component, and single particles remain separated (Fig. 21a). The protection material occurs in the range of several percent of total metal mass (in case of Ag particles with average diameter of 6 nm—not higher than 4% [71]). As a result, this protective layer on single nanoparticles makes both *DC* electrical conduction impossible (at relatively low voltage) and heat transfer difficult. To obtain a good conductivity, each type of initial product needs additional energy—mostly thermal by heating.

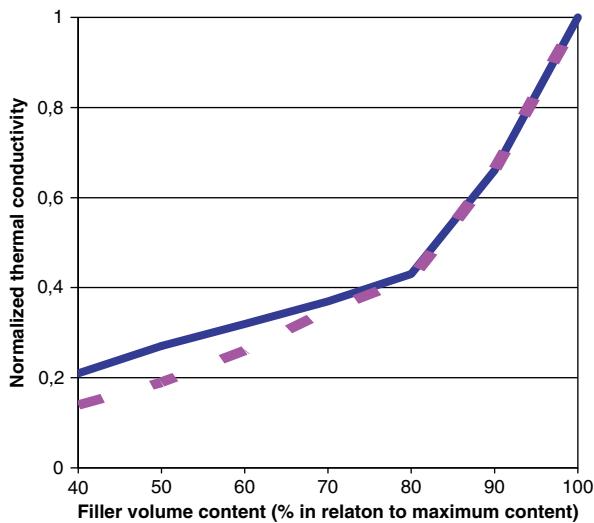


Fig. 20 Thermal conductivity of adhesive with electrically isolated filler vs. filler volume content; continuous line: mix of AlN whiskers and AlN particles with maximum content of 60 vol% (according to [48]), dotted line: AlN particles with maximum content of 62 vol% (according to [66])

Figure 21b presents the microstructure of a nanosized Ag sintered at 280°C [74]. It can be seen that the structure changed into a dense network of Ag with micropores. Such structure surely conducts heat much better than before thermal process. Energy dispersion X-ray (EDX) analysis shows that almost all the organic components have burned off after the thermal process [75].

In case of thermally conductive composites nanosized silver is added to a microcomposite rather than used as only a filler. It was stated that adding of silver nanoparticles to formulation with micro-sized silver particles may improve the composite thermal conductivity, even 2.55 times [26]. The conventional formulation

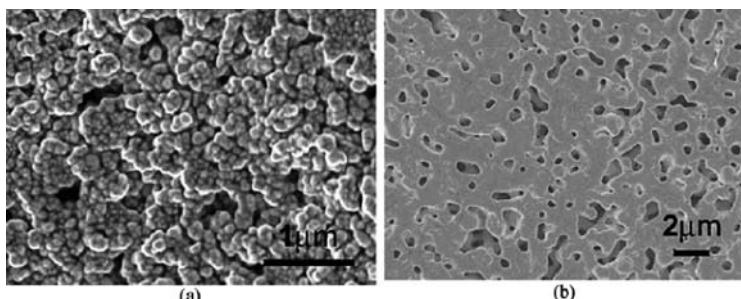
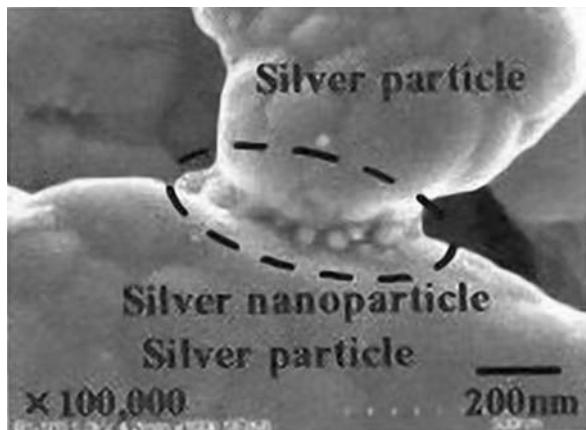


Fig. 21 SEM images of the nanosized Ag (a) before and (b) after sintering on silicon substrate at 280°C [74]

with silver flake was tested as reference. It contains silver particles of 1–10 μm size (88 wt%), epoxy resin (9 wt%), and solvent (3 wt%). The hybrid silver formulation contains micro- and nanosized silver- together of 90 wt%, and 10 wt% of phenol resin and a solvent. The size of silver microparticles is 3–5 μm , while silver nanoparticles with a coating material of 3–7 nm in diameter. Adding the silver nanoparticles makes the contact resistance between basic particles lower owing to the fusion process. The mechanism of the curing process of hybrid silver formulation can be explained by the following steps. In the initial stage, the formulation contains resin, solvent, silver particles (micro-sized), and silver nanoparticles with a coating material. During the first step of the curing process, the solvent starts to vaporize by heating. Next, the coating material begins to decompose and vaporize. Without the coating material, nanoparticles at high temperature (200°C) fuse together simultaneously. The nanoparticles have grown from their initial sizes of 3–7 nm to 20–50 nm. However, what is crucial for heat flow, the shapes of new particles depend on the surface of their large-sized neighbors. The SEM image (Fig. 22) shows that new contacts between particles have extremely large areas resulting in low thermal contact resistance Θ_λ .

Fig. 22 SEM image of adhesive containing silver nanopaste [26]



There are also experiments with doping silver nanowires to enhance the electrical conductance of composites containing micro-sized silver particles [76, 77]. In the system composed of silver particles and nanowires, these slender wires are excellent materials to contact the particles together. From Fig. 23a, it is apparent that the nanowires are spanning over the gap between two flake silver particles and forming a conductive pathway. Unfortunately, for the heat transport, the role of such a single pathway is not important. In the experiment [76], wires of 10–50 μm long with average diameter of about 100 nm were used. It is easy to calculate from Equation 4, that the thermal bulk resistance Θ_λ of such structure is relatively high (for silver nanowires with diameters about 27 nm and lengths more than 70 μm [78] the

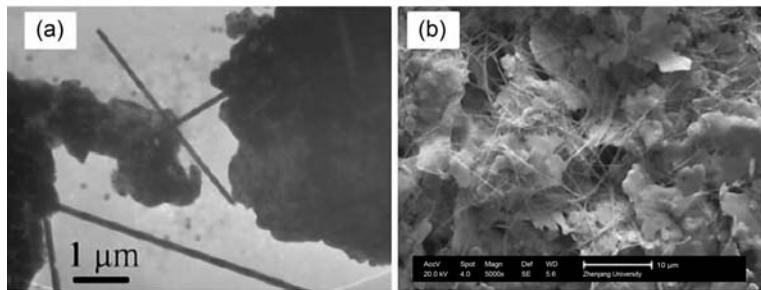


Fig. 23 Transmission electron microscopic (a) and scanning electron microscopic (b) images of local conductive pathways of flake silver particles doping nanowires [76]

resistance would be at least 20 times higher). Additionally, the contact area of the wire with the surface of the particle probably is in the range not higher than dozens of square nanometers, which make the value of thermal contact resistance (Θ_{TC} in Equation 5) also high. But the conductive pathways could be increased in number (Fig. 23b) which, according to Equation 6, causes lower value of total thermal resistance of composite Θ_{TA} and better heat transfer.

There are also investigations of composite in which nanosilver particles saturate polymer nanofibers [79, 80]. The polymer nanofibers were made in the electrospinning process and nanoparticles of silver were added during this process as a thermal conductivity promoter. It was found that presence of nanosilver slightly improves the thermal conductivity of the composite, especially in the range of 90–100°C. The mechanism of better heat transport is unknown, as the SEM picture (Fig. 24) reveals only separate particles of silver attached to the fiber structure.

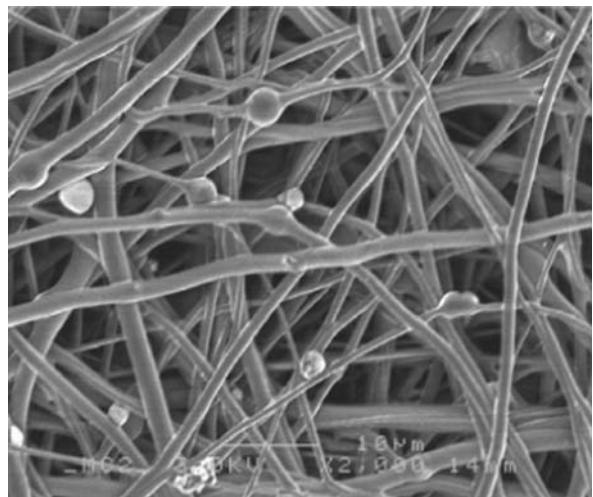


Fig. 24 Scanning electron microscopy image of polymer fibers embedded with nanosilver particles [80]
Courtesy of IEEE

5 Composites with Carbon Allotropies Fillers

Thermally conductive composites filled with the carbon allotropes, including diamond, graphite, the buckminster fullerene, carbon fibers, and carbon nanotube seem to be very promising because of high bulk thermal conductivity of such materials.

Unlike most electrical insulators, genuine diamonds are a very good thermal conductor because of the strong covalent bonding within the crystal. Specially purified synthetic diamonds could have the highest thermal conductivity of 2000–2500 W/m×K, which are five times more than any known solid element at room temperature. Unfortunately, the thermal measurement for micron-sized synthetic diamond powder as filler showed relatively a low thermal conductivity in comparison with micron-sized silver. This is believed to be due to the presence of nitrogen and other impurities in synthetic diamond [24]. On the other hand, thermal conductivity of diamond is strongly dependent on the microstructure of the material [81] and additionally is correlated to the grain size even for the bulk material [82]. Probably there are effects of structural defects such as stacking faults, twins, or dislocations from synthetic processes. Typically, where the grain size is in excess of 30–50 μm, the effects of grain boundaries on thermal conductivity become insignificant, whereas submicron grain sizes can reduce the thermal conductivity by more than two orders of magnitude. Because of this, the using of diamond with very low dimension crystals (nanosized) seems to be unfounded. Such conclusions were confirmed by experiments [83] in which thermal conductivity of composites was measured in a steady heat flow in the temperature range of 50–200°C in a vacuum. The results are presented in Fig. 25 as the function of different diamond grains content in the matrix of the natural single-crystal diamond powder with the fractal size of 10–14 μm.

From the histogram, the thermal conductivity of composites produced from 5 to 7 μm crystallites (region 4) is lower than that of composites synthesized from 10 to 14 μm crystallites (region 6) with the maximum thermal conductivity of about 500 W/m K. Adding nanodiamonds decreases component conductivity up to less than 10 W/m KK (region 1). This is likely to be due to the presence of a large number of grain boundaries, at which phonons are scattered.

Similar results were observed by testing of composites consisting of ultrafine (≈ 6 nm) diamond crystallites embedded into graphitic carbon matrix [84]. In such composites, thermal conductivity may increase by more than 5 times from a structure containing a small amount of carbon matrices (Fig. 26) up to components with 50% of graphite.

Without a carbon matrix, the thermal contact between diamond nanoparticles is assumed to be very poor, so the very low thermal conductivity obtained for this specimen seems logical (Equation 5). With increase of the graphite matrix volume more carbon fills the pores between diamond grains, so the composite conductivity increases with decreasing size of the remaining pores. It shows that graphite with the value of thermal conductivity smaller by at least one order of magnitude plays a much more important role in heat transport than nanosized diamond.

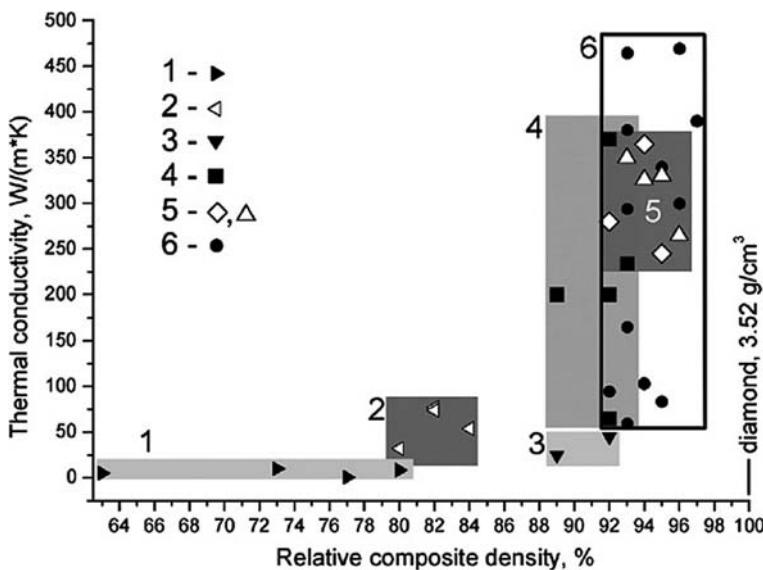


Fig. 25 The dependence of the thermal conductivity of composites sintered from micro- and nanodiamond powders on the density normalized to that of single-crystal diamond (3.52 g/cm^3). Regions showing the type of the sample composition: (1)—100% nanodiamonds, (2)—50% $10\text{--}14 \mu\text{m}$ diamonds and 50% nanodiamonds, (3)—85% $10\text{--}14 \mu\text{m}$ diamonds and 15% nanodiamonds, (4)—100% $5\text{--}7 \mu\text{m}$ diamonds, (5)—99.5% $10\text{--}14 \mu\text{m}$ diamonds and 0.5% fullerenes combined with data on 99% $10\text{--}14 \mu\text{m}$ diamonds and 1% nanodiamonds, (6)—100% $10\text{--}14 \mu\text{m}$ diamonds [83]

Fig. 26 Schematic presentation of composite structure with nanodiamond crystals

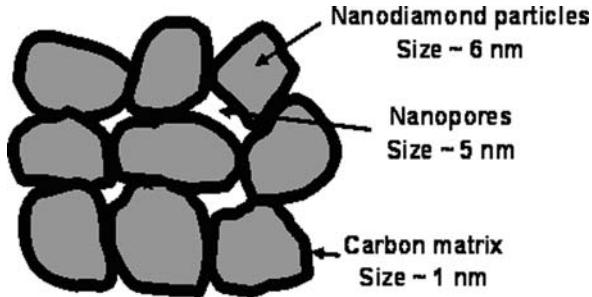


Figure 25 presents thermal conductivity of composites with adding of fullerene (region 5). Fullerenes and nanodiamonds added in small amounts (<1% mass) to powders of $10\text{--}14 \mu\text{m}$ diamond crystallites do not change markedly the thermal conductivity values, even though that fullerene can be characterized by so high thermal conductivity as single-wall carbon nanotubes [85]. It is also reported [86] that the conductivity of hard carbon prepared from fullerene C_{60} does not exceed a few $\text{W}/\text{m} \times \text{K}$ at room temperature.

The use of carbon nanotubes (CNT) as high-performance nanofillers to implement a new generation of high thermally conductive composites seems promising. Carbon nanotubes were discovered in 1991 by Sumio Iijima [3], first as multiwall structures with outer diameters of 4–30 nm and a length of up to 1 μm , which consisted of two or more seamless graphene cylinders concentrically arranged. Single-wall carbon nanotubes, which are seamless cylinders each made of a single graphene sheet, were first reported 2 years later. Their diameters range from 0.4 to 2 to 3 nm, and their length is usually of the micrometer order, i.e., with a very high aspect ratio.

Numerous strategies for the synthesis of carbon nanotubes and more control over their characteristics (number of walls, diameter, length), while improving samples and reproducibility were undertaken [87–89]. Basically, there are “high temperature” methods, such as the electric-arc discharge, laser ablation, and “low temperature” ones, such as the chemical vapor deposition, with many technical variants for each one.

The high value of thermal conductivity of CNTs is the most desirable feature, which can improve the heat transport in composites. The thermal conductivity of individual multiwall carbon nanotube is estimated to fall in the range of 2000–over 3000 $\text{W/m}\times\text{K}$ [90–92]. As the measurements of thermal properties of a single nanotube are extremely difficult and conventional methods cannot be used for this, the values of λ presented in literature are different. Additionally, experimental results show [92] that the thermal conductivity of a carbon nanotube depends on both temperature and tube diameter. From Fig. 27 it can be seen that the λ value increases with decreasing tube diameter and increases with an increase in temperature and appears to have an asymptote near 320 K. The thermal conductivity values of one order of magnitude smaller were measured for bulk samples where the nanotubes have been aligned by filtration in a magnetic field. The measurement was taken in the direction parallel to the tubes [93].

The first composites with carbon nanotubes were proposed in the last decade of the twentieth century. Polymers such as epoxy, thermoplastics, gels as well as

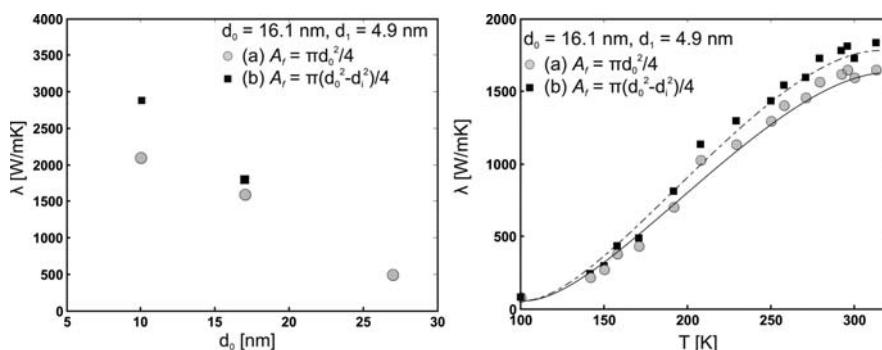


Fig. 27 Thermal conductivity of a single carbon nanotube at different diameters at room temperature (left) and different temperature (right) [92]

poly (methyl methacrylate) have been used as the matrix [94]. The dispersion process requires a crucial technology to develop a nanocomposite because the specific surface area of carbon nanotubes is too large to well disperse in the matrix resin. According to theoretical calculations [95], the specific surface area of single-wall nanotube may reach a value of $1315\text{ m}^2/\text{g}$. For multiwalled tube the value is lower, but even for higher tube the diameter amounts to a few hundreds of m^2/g . If the filler is not properly dispersed, the thermal property of composite cannot be significantly increased. Several processing methods are used for this reason, many of them based on improving nanotubes/matrix interactions [96–98], such as an ultrasonic bath, melt-mixing, in-situ polymerization or solution processing. To prepare carbon nanotubes for the proper dispersion on a macroscopic scale, some preprocessing is required. It can be purified to eliminate non-nanotubes material followed by deagglomeration for dispersing individual nanotubes, or chemical functionalization. However, the intensive chemical treatment may introduce more defects and reduce the intrinsic thermal conductivity of carbon nanotubes. The chemical treatment usually consists of many steps. For example, Fig. 28 presents the overall procedure modification of carbon nanotubes by an oxidation agent for improving the disparity between matrix and the filler.

In fact, even using complicated processes for improving nanotubes–matrix interactions it is impossible to incorporate the filler content higher than a few weight percent. It is reported that the dispersion of 5 wt% CNTs in an epoxy resin was obtained by an ultrasonic treatment. Although the carbon nanotubes were well separated, they remained poorly distributed [99]. In such CNT content, the viscosity of the composite is more than ten times higher in comparison to pure matrix [94]. Generally, sonication method is only suitable for very low viscous matrix materials and small volumes because ultrasonic devices have a high impact of energy. A proper way to apply the sonication technique to produce nanocomposites with carbon nanotubes is to disperse them in an appropriate solvent (i.e., acetone, ethanol etc.), which first allows the agglomerates to be separated due to vibrational energy. The suspension can later be mixed with the polymer and the solvent can simply be evaporated by heating.

For electrical current transport, the content level of the filler in which percolation threshold is achieved plays a crucial role. A review and analysis of electrical percolation in carbon nanotube polymer composites from 2009 [100] brings more information concerning the percolation threshold, but in most cases, the current is measured with the carbon nanotubes content in the level of 0.01–0.4 vol% [101, 102], or 0.01–1.5 wt% [103, 104].

There are no observed correlations between “electrical” percolation threshold and heat transport. The experimental results of heat transport show that a random dispersion of multiwall nanotubes in an organic fluid increases its thermal conductivity, by more than 2.5 times at approximately 1% volume fraction of nanotubes [105]. For the same value of nanotube loading a 125% increase is reported in the conductivity of industrial use epoxy due to a random dispersion of single-walled

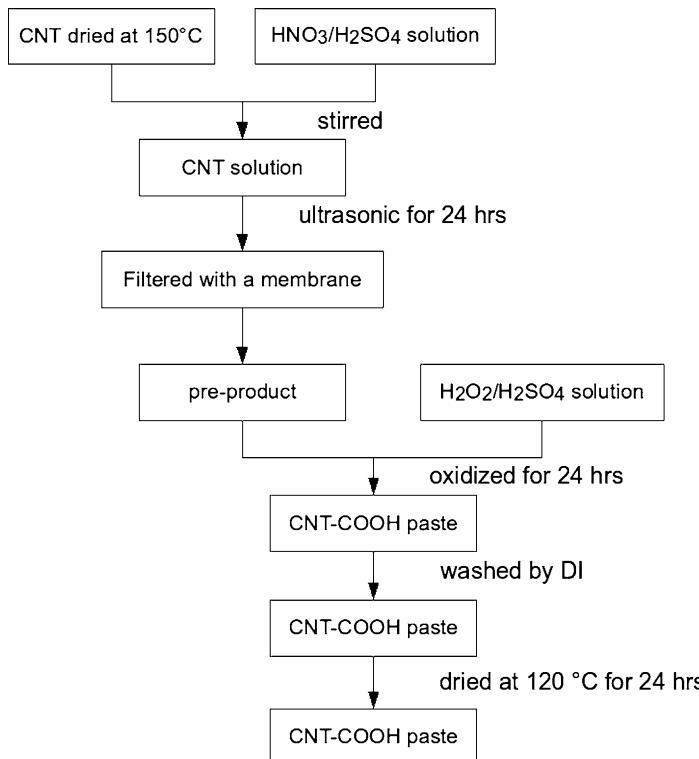


Fig. 28 Experimental procedure for acid-modified carbon nanotubes (CNT) [96]

nanotubes [106]. Nevertheless, the thermal conductivity of the composites filled with single- or multiwalled carbon nanotube is lower than for composites with microsized fillers. When proper methods of thermal conductivity measurement were used, the λ value reached only 1 W/m K or less [107–110]. This is lower than calculated value based on the model of the heat transport with neglecting all interactions between the nanotubes [111]. For contents of “average in shape” multiwall nanotubes in a single weight percent, the model predicts the thermal conductivity of a composite on the level of 1.8 W/m × K. It was calculated that effective conductivity increases linearly with increasing the nanotube loading, the conductivity is not very sensitive to the nanotube length (increases marginally with the increase in length) and the conductivity can change drastically with a change in the diameter and the smaller diameter nanotubes can significantly increase the overall conductivity.

6 Conclusive Remarks and Prospects

Temperature control is critical for both the operating performance and long-term reliability of single chip packages. The high junction-to-ambient thermal resistance resulting from an air-cooled heat sink provides inadequate heat removal capability at the necessary junction temperatures. The International Technology Roadmap for Semiconductors in 2007 Edition [112] projects power density and junction-to-ambient thermal resistance for high-performance chips at the nanometer generation on the level of $>1 \text{ W/mm}^2$ and $<0.2^\circ\text{C/W}$, respectively. The main bottlenecks in reducing the junction-to-ambient thermal resistance are the thermal resistances of the thermal interface material and the heat sink. There is a need for TIMs that provide the highest possible thermal conductivity, are mechanically stable during chip operation, have good adhesion, and conform to fill the gaps between two rough surfaces. The report [112] predicts that to address these needs, new TIMs containing carbon nanotubes will be elaborated.

In fact, carbon nanotubes due to their outstanding thermal properties have been the best candidate for a heat-transport application. Unfortunately, the thermally conductive composites with CNTs are characterized by thermal conductivity not higher than a few $\text{W/m}\times\text{K}$. Theoretically, according to Equations 5 and 6, the improvement of thermal conductivity of nanocomposites saturated by filler with very low thermal resistance Θ_λ can be done by:

- by decreasing the thermal contact resistance between particles forming a chain for heat transport in the direction of temperature gradient (Θ_{TC});
- by lessening the numbers contact resistance between particles responsible for heat transport
- by increasing the number parallel contact path in the direction of temperature gradient.

All these requirements would be fulfilled with high content of carbon nanotube in the polymer matrix. Unfortunately, the results presented in literature show that it is impossible to incorporate the CNT in composite higher than a few weight percent. In such cases there is not too many contact points between singular carbon nanotube and even applying a polymer matrix with high cure shrinkage would not bring about a considerable improvement of a thermal conductivity of the TIM.

The diminishing of contacts can be achieved by a stable, conductive network formed by the CNT fillers. It seems to be necessary to develop a systematic method of producing various multiple junctions of nanotubes to establish a three-dimension network instead of the set of singular structures. There is information about experiments aimed at establishing a junction between crossed nanotubes, the first attempt at doing such web-like CNTs have been achieved with electron beam, ion beam, laser, or other techniques of nanowelding [113].

As an example, Fig. 29 presents the single-walled carbon nanotubes joined by electron beam welding to form molecular junctions. In experiments, single-wall nanotubes were dispersed ultrasonically in ethanol and deposited onto

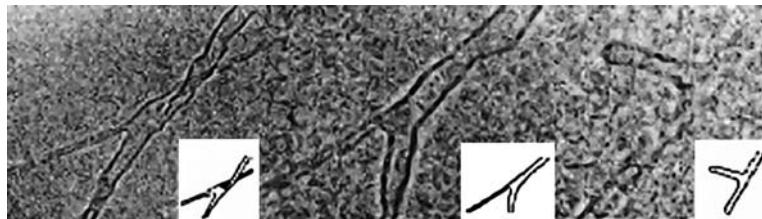


Fig. 29 Electron microscopy images of single-wall nanotubes welding under intense electron beam [85]

porous carbon grids for transmission electron microscopy (TEM) observations. TEM observations were performed under high voltage of 1.25 MeV at specimen temperatures of 800°C. From the random crisscrossing distribution of individual nanotubes and nanotube bundles on the specimen grid, several contact points could be identified where tubes were crossing and touching each other. After a few minutes of irradiating and annealing, different shapes of junctions were established, even the tube diameters within the junction were not the same [114].

Using empirical potential molecular dynamics, it was demonstrated that also ion irradiation should result in welding of crossed nanotubes. A simulation of the irradiation of crossed single-wall carbon nanotubes was carried out, both suspended and deposited on substrates, with 0.4–1 keV argon ions. The optimum Ar ion energies should be in the range of 0.4–0.6 keV, whereas the optimum irradiation dose should be about 10^{15} cm^{-2} [115]. In practice, the junctions were obtained as the result of carbon ion irradiation. In the experiment [116], the multiwall carbon nanotubes were transferred onto the porous carbon micro-grid for transmission electron microscopy examination with an accelerating voltage of 160 kV. Second, this porous carbon micro-grid filled with unirradiated nanotubes was managed to be irradiated perpendicularly by carbon ion beam with an energy of 50 keV in a vacuum chamber. The current and dose of carbon ion beam were about $10 \mu\text{A}$ and $1 \times 10^{17} \text{ cm}^{-2}$, respectively. The results were presented in Fig. 30. Figure 30a shows the image of a typical web-like structure which contains four junctions (indicated by A, B, C, and D). High-resolution transmission electronic microscope images taken at the regions of A ("T" shape), B, and C ("Y" shape) junctions were shown in Fig. 30b, c. The merging region between two carbon nanowires was a uniform amorphous solid structure without any evident sign of the interface presented in an atomic scale.

The increasing of number parallel contact path in direction of temperature gradient can be achieved by different types of microfillers or nanofillers. Figure 31 presents an example of mixed fillers (micro-sized silver flakes and CNT). The low content of nanotubes (about 0.5 wt%) weakly influenced the total thermal conductivity of the composite yet, and efforts of research labs are aimed at achieving the significant increase of the CNT content [117].

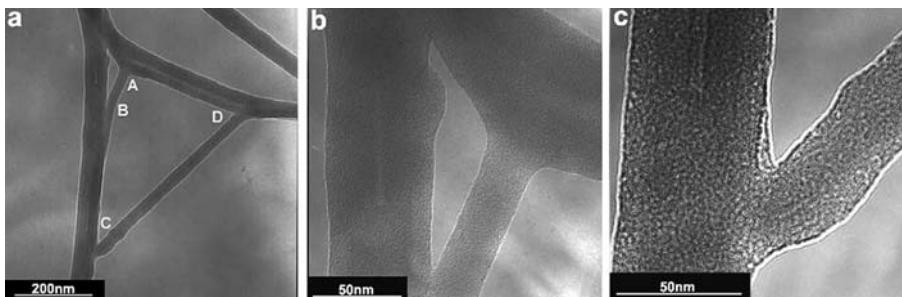
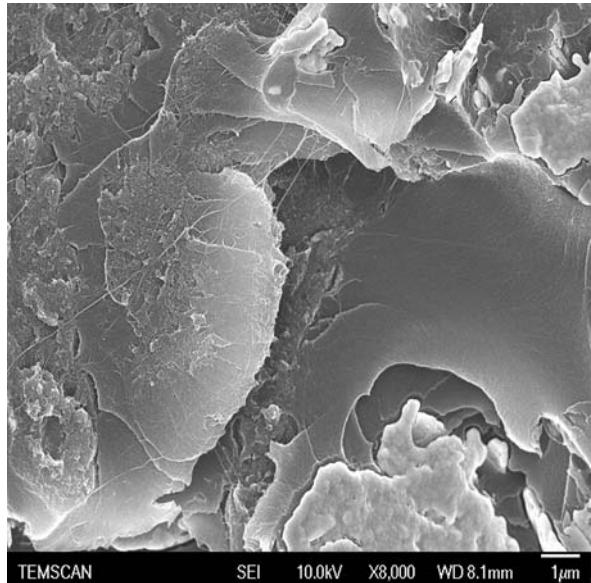


Fig. 30 (a) Web-like structure with junctions of amorphous carbon nanowires (indicated by A, B, C, and D). (b), (c) High-resolution transmission electron microscopy images for junctions at A and B showed in (a) [116]

Fig. 31 Scanning electron microscopy image of composite containing silver flakes and carbon nanotubes as the filler Courtesy of CANOPY Project partners[117]



Also adding of nanofillers can improve the thermal conductivity. It was observed that the presence of nanosized carbon black may improve the thermal conductivity of epoxy-based composite with 2 wt% carbon nanotubes [97] by ca. 20%. The carbon black with particle diameters of 22–60 nm was added to the mixture after it was sonicated in acetone. The prepared mixture was put into a vacuum chamber at 100°C for 4–5 h to remove the acetone and the trapped air.

References

1. J. Liu, T. Wang, B. Carlberg, M. Inoue. Recent Progress of Thermal Interface Materials. 2nd Electronics System integration Technology Conference, Greenwich 2008:351.
2. F. Sarvar, D. C. Whalley, P. P. Conway. Thermal Interface Materials – A review of the State of Art. 1st Electronics Systemintegration Technology Conference, Dresden 2006:1292.
3. S. Iijima. Carbon nanotubes: Past, present, and future. *Physica B* 2002; 323:1.
4. D. R. Paul, L. M. Robeson. Polymer nanotechnology: Nanocomposites. *Polymer* 2008; 49:3187.
5. Efunda – engineering fundamentals; <http://www.efunda.com>.
6. A. Damasceni, L. Dei, F. Guasti. Thermal behaviour of silver-filled epoxy adhesives; technological implications in microelectronics. *Journal of Thermal Analysis and Calorimetry* 2001; 66:223.
7. T. Falat, J. Felba, A. Wymyslowski. Improved Method for Thermal Conductivity Measurement of Polymer based Materials for Electronic Packaging. 28th International Conference of International Microelectronics and Packaging Society – Poland Chapter, Wroclaw 2004:219.
8. R. Prasher. Thermal interface materials: Historical perspective, status, and future directions. *Proceedings of the IEEE* 2006; 94(8):1571.
9. C. -W. Nana, R. Birringer, D. R. Clarke, H. Gleiter. Effective thermal conductivity of particulate composites with interfacial thermal resistance. *Journal of Applied Physics* 1997; 81(10):6692.
10. M. Y. Koledintseva, R. E. DuBroff, R. W. Schwartz A Maxwell-Garnett model for dielectric mixtures containing conducting particles at optical frequencies. *Progress In Electromagnetics Research* 2006; 63:223–242.
11. P. Mallet, C. A. Guérin, A. Sentenac. Maxwell-Garnett mixing rule in the presence of multiple scattering: Derivation and accuracy. *Physical Review B* 2005; 72:14205.
12. O. Levy, D. Stroud. Maxwell Garnett theory for mixtures of anisotropic inclusions: Application to conducting polymers. *Physical Review B* 1997; 56(13):8035.
13. M. Rong, M. Zhang, H. Liu, H. Zeng. Synthesis of silver nanoparticles and their self-organization behavior in epoxy resin. *Polymer* 1999; 40(22):6169.
14. T. Falat. Heat Transfer Analysis in Composite Materials filled with Micro and Nano-Sized Particles, Doctoral Thesis (in Polish), Faculty of Microsystem Electronics and Photonics, Wroclaw University of Technology, Wroclaw 2006.
15. J. Felba, T. Falat. Thermally Conductive Adhesives for Microelectronics – Barriers of Heat Transport. 6th International IEEE Conference on Polymers and Adhesives in Microelectronics and Photonics, Polytronic 2007:228.
16. T. Falat, J. Felba. Numerical Prediction of Influence Matrix and Filler Properties on Thermal Conductivity of Copper Filled TCA. 6th International IEEE Conference on Polymers and Adhesives in Microelectronics and Photonics, Polytronic 2007:114.
17. J. Felba, T. Falat, A. Wymyslowski. Influence of thermo-mechanical properties of polymer matrices on the thermal conductivity of adhesives for microelectronic packaging. *Materials Science-Poland* 2007; 25(1):45.
18. T. Falat, J. Felba, A. Wymyslowski, K. M. B. Jansen, J. S. Nakka. Viscoelastic Characterization of Polymer Matrix of Thermally Conductive Adhesives. 1st Electronics Systemintegration Technology Conference, Dresden 2006:773.
19. R. Holm. *Electric Contacts – Theory and Application*. Springer, Berlin, 1967.
20. A. Wymyslowski, K. Friedel, J. Felba, T. Falat. An Experimental-Numerical Approach to Thermal Contact Resistance. 9th International Workshop on Thermal Investigations of ICs and Systems, Aix-en-Provence 2003:161.
21. C. V. Madhusudana Thermal Contact Conductance. Springer, New York, 1996.
22. D. P. H. Hasselman, K. Y. Donaldson, F. D. Barlow, A. A. Elshabini, G. H. Schiroky, J. P. Yaskoff, R. L. Dietz. Interfacial thermal resistance and temperature dependence of

- three adhesive for electronic packaging. *IEEE Transactions on Components and Packaging Technologies* 2000; 23(4):633.
- 23. A. Wymyslowski, T. Falat, K. Friedel, J. Felba. Numerical Simulation and Experiment Verification of the Thermal Contact Properties of the Polymer Bonds. 5th International Conference on Thermal and Mechanical Simulation and Experiments in Microelectronics and Microsystems, Brussels 2004:177.
 - 24. J. C. Bolger. Prediction and Measurement of Thermal Conductivity of Diamond Filled Adhesives. 42nd Electronic Components and Technology Conference, 1992:219.
 - 25. H. Li, K. I. Jacob, C. P. Wong. An improvement of thermal conductivity of underfill materials for flip-chip packages. *IEEE Transactions on Advanced Packaging* 2003; 26(1):25.
 - 26. Y. Ukita, K. Tateyama, M. Segawa, Y. Tojo, H. Gotoh, K. Oosako. Lead Free Mount Adhesive using Silver Nanoparticles Applied to Power Discrete Package. International Symposium on Microelectronics IMAPS, Long Beach 2004, session WA7.
 - 27. S. Macek, D. Rocak, P. Sebo, P. Stefanik. The use of Polymeric Adhesives in Bonding Power Hybrid Circuits to Heat Sinks. 23rd International Spring Seminar on Electronics Technology ISSE, Balatonfured 2000:185.
 - 28. R. Dietz, P. Robinson, M. Bartholomew, M. Firmstone. High Power Application with Advanced High k Thermoplastic Adhesives. 11th European Microelectronics Conference, Venice 1997:486.
 - 29. R. R. Zarr. A history of testing heat insulators at the national institute standards and technology. *ASHRAE Transactions* 2001; 107:Pt 2.
 - 30. Y. He. Rapid thermal conductivity measurement with a hot disk sensor. Part 1. Theoretical considerations. *Thermochimica Acta* 2005; 436:122–129.
 - 31. R. P. Tye. Thermal conductivity. *Nature* 1964; 205:636.
 - 32. D. R. Smith, J. G. Hust, L. J. Van Poolen. A Guarded-Hot-Plate Apparatus for Measuring Effective Thermal Conductivity of Insulations Between 80 K and 360 K. US National Bureau of Standards 1982, NBSIR 81-1657.
 - 33. C. Stacey. NPL Vacuum Guarded Hot-Plate for Measuring Thermal Conductivity and Total Hemispherical Emittance of Insulation Materials. *Insulation Materials: Testing and Applications* 2002; 4, ASTM STP 1426.
 - 34. A. Franco. An apparatus for the routine measurement of thermal conductivity of materials for building application based on a transient hot-wire method. *Applied Thermal Engineering* 2007; 27:2495–2504.
 - 35. J. J. Salgon, F. Robbe-Valloire, J. Blouet, J. Bransier. A mechanical and geometrical approach to thermal contact resistance. *International Journal of Heat Mass Transfer* 1997; 40(5):1121.
 - 36. C. L. Yeh, C. Y. Wen, Y. F. Chen, S. H. Yeh, C. H. Wu. An experimental investigation of thermal contact conductance across bolted joints. *Experimental Thermal and Fluid Science* 2001; 25:349.
 - 37. T. Falat, B. Platek, S. Tesarski, J. Felba. An Approach to Measurement and Evaluation of the Thermal Conductivity of the Thermal Adhesives in Electronic Packaging. 32nd International Spring Seminar on Electronics Technology, Brno 2009.
 - 38. M. J. Assael, K. D. Antoniadis, D. Tzetzis. The use of the transient hot-wire technique for measurement of the thermal conductivity of an epoxy-resin reinforced with glass fibres and/or carbon multi-walled nanotubes. *Composites Science and Technology* 2008; 68:3178–3183.
 - 39. H. M. Roder, R. A. Perkins, A. Laescke. Absolute steady-state thermal conductivity measurements by use of a transient hot-wire system. *Journal Research National Institute of Standards and Technology* 2000; 105(2):221.
 - 40. W. N. dos Santos Advances on the hot wire technique. *Journal of the European Ceramic Society* 2008; 28:15.
 - 41. W. J. Parker, R. J. Jenkins, C. P. Butler, G. L. Abbott. Flash method of determining thermal diffusivity, heat capacity, and thermal conductivity. *Journal of Applied Physics* 1961; 32:1679.

42. R. D. Cowan. Pulse method of measuring thermal diffusivity at high temperatures. *Journal of Applied Physics* 1963; 34:926.
43. R. E. Taylor, J. A. Cape. Finite pulse-time effects in the flash diffusivity technique. *Applied Physics Letters* 1964; 5:212.
44. D. A. Watt. Theory of thermal diffusivity by pulse technique. *British Journal of Applied Physics* 1966; 17:231.
45. A. J. Walter, R. M. Dell, P. C. Burgess. The measurement of thermal diffusivities using a pulsed electron beam. *Rev. Int. Hautes Tempér. Réfract.* 1970; 7:271.
46. Y. S. Touloukian, R. W. Powell, C. Y. Ho, M. C. Nicolaou. Thermal Diffusivity. *Thermophysical Properties of Matter*. 1973, 10, IFI/PLENUM, New York – Washington.
47. I. I. I. L. M. Clark, R. E. Taylor. Radiation loss in the flash method for thermal diffusivity. *Journal of Applied Physics* 1975; 46:714.
48. Y. Xu, D. D. L. Chung, C. Mroz. Thermally conducting aluminium nitride polymer-matrix composites. *Composites A* 2001; 32:1749.
49. T. Falat, J. Felba. Electron beam as a heat source in thermal diffusivity measurement of thermally conductive adhesives. *Electrotechnica & Electronica* 2006; 5–6:189.
50. A. Rosencwaig, A. Gersho. Theory of the photoacoustic effect with solids. *Journal of Applied Physics* 1976; 47:64.
51. H. Hu, X. Wang, X. Xu. Generalized theory of the photoacoustic effect in a multilayer material. *Journal of Applied Physics* 1999; 86:3953.
52. B. A. Cola, J. Xu, Ch. Cheng, X. Xu, T. S. Fisher, H. Hu. Photoacoustic characterization of carbon nanotube array thermal interfaces. *Journal of Applied Physics* 2007; 101:054313.
53. L. Sun, S. -Y. Zhang, Y. -Z. Zhao, Z. -Q. Li, L. -P. Cheng. Thermal diffusivity of composites determined by photoacoustic piezoelectric technique. *Review of Scientific Instruments* 2003; 74:834.
54. J. A. Balderas-López. Thermal effusivity measurements for liquids: A self-consistent photoacoustic methodology. *Review of Scientific Instruments* 2007; 78:064901.
55. N. F. Leite, L. C. M. Miranda. Thermal property measurements of liquid samples using photoacoustic detection. *Review of Scientific Instruments* 1992; 63:4398.
56. J. A. Balderas-López. Self-normalized photoacoustic technique for thermal diffusivity measurements of transparent materials. *Review of Scientific Instruments* 2008; 79:024901.
57. J. Philip. A photoacoustic scanning method to determine thermal effusivity of solid samples. *Review of Scientific Instruments* 1996; 67:3621.
58. S. Govorkov, W. Ruderman, M. W. Horn, R. B. Goodman, M. Rothschild. A new method for measuring thermal conductivity of thin films. *Review of Scientific Instruments* 1997; 68:3828.
59. J. H. Kim, A. Feldman, D. Novotny. Application of the three omega thermal conductivity measurement method to a film on a substrate of finite thickness. *Journal of Applied Physics* 1999; 86(7):3959.
60. C. Dames, G. Chen. 1ω , 2ω , and 3ω methods for measurements of thermal properties. *Review of Scientific Instruments* 2005; 76:124902.
61. D. G. Cahill, R. O. Pohl. Thermal conductivity of amorphous solids above the plateau. *Physical Review B* 1987; 35(8):4067–4073.
62. S. R. Choi, J. Kim, D. Kim. 3ω method to measure thermal properties of electrically conducting small-volume liquid. *Review of Scientific Instruments* 2007; 78:084902.
63. W. P. Risk, C. T. Rettner, S. Raoux. In situ 3ω techniques for measuring thermal conductivity of phase-change materials. *Review of Scientific Instruments* 2008; 79:026108.
64. National Fenestration Rating Council, Procedure for Determining Thermo-Physical Properties of Materials, www.nfrc.org
65. L. Fan, B. Su, J. Qu, C. P. Wong. Effects of Nano-Sized Particles on Electrical and Thermal Conductivities of Polymer Composites. 9th International Symposium on Advanced Packaging Materials; Processes, Properties and Interfaces 2004:193.

66. P. Bujard, J. P. Ansermet. Thermally Conductive Aluminium Nitride-Filled Epoxy Resin. 5th Semiconductor Thermal and Temperature Measurement Symposium 1989:126.
67. J. Felba, H. Schaefer. Materials and Technology for Conductive Microstructures (in Nanopackaging: Nanotechnologies and Electronics Packaging, ed. J. E. Morris), Springer, 2008.
68. R. M. Tilaki, A. Iraji Zad, S. M. Mahdavi. Stability, size and optical properties of silver nanoparticles prepared by laser ablation in different carrier media. *Applied Physics A* 2006; 84:215.
69. H. Saito, Y. Matsuba. Liquid Wiring Technology by Ink-jet Printing Using NanoPaste. 35th International Symposium on Microelectronics IMAPS, San Diego, 2006.
70. <http://www.harima.co.jp/products/electronics>
71. A. Moscicki, J. Felba, T. Sobierajski, J. Kudzia, A. Arp, W. Meyer. Electrically Conductive Formulations Filled Nano Size Silver Filler for Ink-Jet Technology. 5th International IEEE Conference on Polymers and Adhesives in Microelectronic and Photonics, Wroclaw 2005:40.
72. A. Moscicki, J. Felba, W. Dudzinski. Conductive Microstructures and Connections for Microelectronics Made by Ink-Jet Technology. 1st Electronics Systemintegration Technology Conference, Dresden 2006:511.
73. M. Nakamoto, M. Yamamoto, Y. Kashiwagi, H. Kakiuchi, T. Tsujimoto, Y. A. Yoshida. Variety of Silver Nanoparticle Pastes for Fine Electronic Circuit Pattern Formation. 6th International IEEE Conference on Polymers and Adhesives in Microelectronic and Photonics, Tokyo 2007:105.
74. J. G. Bai, Z. Z. Zhang, J. N. Calata, G. -Q. Lu. Low-temperature sintered nanoscale silver as a novel semiconductor device-metallized substrate interconnect material. *IEEE Transactions on Components and Packaging Technologies* 2006; 29(3):589.
75. J. G. Bai, J. N. Calata, T. G. Lei, G. Q. Lu, K. D. Creehan. Lead-free Die-attachment with High-temperature Capability by Low-temperature Nanosilver Paste Sintering. 35th International Symposium on Microelectronics IMAPS, San Diego 2006.
76. C. Chen, L. Wang, R. Li, G. Jiang, H. Yu, T. Chen. Effect of silver nanowires on electrical conductance of system composed of silver particles. *Journal of Materials Science* 2007; 42(9):3172.
77. H. P. Wu, J. F. Liu, X. J. Wu, M. Y. Ge, Y. W. Wang, G. Q. Zhang, J. Z. Jiang. High conductivity of isotropic conductive adhesives filled with silver nanowires. *International Journal of Adhesion & Adhesives* 2006; 26:617.
78. A. Graff, D. Wagner, H. Dittlbacher, U. Kreibig. Silver nanowires. *European Physical Journal D* 2005; 34:263.
79. J. Liu, M. Olugbenga Olorunyomi, X. Lu, W. X. Wang, T. Aronsson, D. Shangguan. New Nano-Thermal Interface Material for Heat Removal in Electronic Packaging. 1st Electronics System integration Technology Conference, Dresden 2006:1.
80. J. Liu, M. Olugbenga Olorunyomi, X. Li, D. Shangguan. Manufacturing and Characterization of Nano Silver Particles Based Thermal Interface Material. 57th Electronic Components and Technology Conference 2007:475.
81. J. E. Graebner, S. Jin. Chemical vapor deposited diamond for thermal management. *JOM* 1998; 50(6):52.
82. C. J. Brierley. Thermal Management with Diamond. *IEE Colloquium on Diamond in Electronics and Optics* 1993.
83. S. V. Kidalov, F. M. Shakhev, A. Ya. Vul'. Thermal conductivity of nanocomposites based on diamonds and nanodiamonds. *Diamond & Related Materials* 2007; 16(12):2063.
84. A. Vlasov, V. Ralchenko, S. Gordeev, D. Zakharov, I. Vlasov, A. Karabutov, P. Belobrov. Thermal properties of diamond/carbon composites. *Diamond and Related Materials* 2000; 9(3–6):1104.

85. T. Kawamura, Y. Kangawa, K. Kakimoto. Investigation of the thermal conductivity of a fullerene peapod by molecular dynamics simulation. *Journal of Crystal Growth* 2008; 310:2301.
86. A. Smontara, M. Saint Paul, J. C. Lasjaunias, A. Bilusic, N. Kitamura. Thermal and acoustic transport properties of hard carbon formed from C₆₀ fullerene. *Physica B* 2002; 316–317:250.
87. M. Terrones. Carbon nanotubes: Synthesis and properties, electronic devices and other emerging applications. *International Materials Reviews* 2004; 49(6):325.
88. V. N. Popov. Carbon nanotubes: Properties and application. *Materials Science and Engineering R* 2004; 43:61.
89. Y. Yadav, V. Kunduru, S. Prasad. Carbon Nanotubes: Synthesis and Characterization (in Nanopackaging: Nanotechnologies and Electronics Packaging, ed. J. E. Morris), Springer, 2008.
90. P. Kim, L. Shi, A. Majumdar, P. L. McEuen. Mesoscopic thermal transport and energy dissipation in carbon nanotubes. *Physica B* 2002; 323:67.
91. M. Pecht, R. Agarwal, P. McCluskey, T. Dishongh, S. Javadpour, R. Mahajan. *Electronic Packaging Materials and their Properties*. CRC Press LLC, Washington, 1999.
92. M. Fujii, X. Zhang, H. Xie, H. Ago, K. Takahashi, T. Ikuta, H. Abe, T. Shimizu. Measuring the thermal conductivity of a single carbon nanotube. *Physical Review Letters* 2005; 95:065502.
93. J. Hone, M. C. Liaguno, M. J. Biercuk, A. T. Johnson, B. Batlogg, Z. Benes, J. E. Fischer. Thermal properties of carbon nanotubes and nanotube-based materials. *Applied Physics A* 2002; 74:339.
94. S. Bal, S. S. Samal. Carbon nanotube reinforced polymer composites—A state of the art. *Bulletin of Materials Science* 2007; 30(4):379–386.
95. A. Peigney, Ch. Laurent, E. Flahaut, R. R. Bacsa, A. Rousset. Specific surface area of carbon nanotubes and bundles of carbon nanotubes. *Carbon* 2001; 39:507.
96. T. -M. Lee, K. -C. Chiou, F. -P. Tseng, C. -C. Huang. High Thermal Efficiency Carbon Nanotube-Resin Matrix for Thermal Interface Materials. 55th Electronic Component & Technology Conference, Lake Buena Vista 2005:55.
97. K. Zhang, G. -W. Xiao, C. K. Y. Wong, H. -W. Gu, M. M. F. Yuen, P. C. H. Chan, B. Xu. Study on Thermal Interface Material with Carbon Nanotubes and Carbon Black in High-Brightness LED Packaging with Flip-Chip Technology. 55th Electronic Component & Technology Conference, Lake Buena Vista 2005:60.
98. E. Titus, N. Ali, G. Cabral, J. Gracio, P. Ramesh Babu, M. J. Jackson. Chemically functionalized carbon nanotubes and their characterization using thermogravimetric analysis, Fourier transform infrared, and Raman spectroscopy. *Journal of Materials Engineering and Performance* 2006; 15(2):182.
99. L. S. Schadler, S. C. Giannaris, P. M. Ajayan. Load transfer in carbon nanotube epoxy composites. *Applied Physics Letters* 1998; 73:3842.
100. W. Bauhofer, J. Z. Kovacs. A review and analysis of electrical percolation in carbon nanotube polymer composites. *Composites Science and Technology* 2009; 69:1486–1498.
101. L. Chang, K. Friedrich, L. Ye, P. Toro. Evaluation and visualization of the percolating networks in multi-wall carbon nanotube/epoxy composites. *Journal of Materials Science* 2009; 44:4003–4012.
102. J. Wescott, P. Kung, A. Mati. Conductivity of carbon nanotube polymer composites. *Applied Physics Letters* 2007; 90:033116.
103. H. -S. Kim, B. H. Park, J. -S. Yoon, H. -J. Jin. Thermal and electrical properties of poly(L-lactide)-graft-multiwalled carbon nanotube composites. *European Polymer Journal* 2007; 43:1729–1735.
104. J. Z. Kovacs, B. S. Velagala, K. Schulte, W. Bauhofer. Two percolation thresholds in carbon nanotube epoxy composites. *Composites Science and Technology* 2007; 67:922–928.

105. S. U. S. Choi, Z. G. Zhang, W. Yu, F. E. Lockwood, E. A. Grulke. Anomalous thermal conductivity enhancement in nanotube suspensions. *Applied Physics Letters* 2001; 79(14):2252.
106. M. J. Biercuk, M. C. Llaguno, M. Radosavljevic, J. K. Hyun, A. T. Johnson, J. E. Fischer. Carbon nanotubes for thermal management. *Applied Physics Letters* 2002; 80:2767.
107. W. -T. Hong, N. -H. Tai. Investigations on the thermal conductivity of composites reinforced with carbon nanotubes. *Diamond & Related Materials* 2008; 17:1577.
108. E. T. Thostenson, T. -W. Chou. Processing-structure-multi-functional property relationship in carbon nanotube/epoxy composites. *Carbon* 2006; 44:3022.
109. F. H. Gojny, M. H. G. , Wichmann, B. Fiedler, I. A. Kinloch, W. Bauhofer, A. H. Windle, K. Schulte. Evaluation and identification of electrical and thermal conduction mechanisms in carbon nanotube/epoxy composites. *Polymer* 2006; 47:2036.
110. Y. S. Song, J. R. Youn. Evaluation of effective thermal conductivity for carbon nanotube/polymer composites using control volume finite element method. *Carbon* 2006; 44:710–717.
111. A. Bagchi, S. Nomura. On the effective thermal conductivity of carbon nanotube reinforced polymer composites. *Composites Science and Technology* 2006; 66:1703.
112. International Technology Roadmap for Semiconductor; <http://www.itrs.net/Links/2007ITRS/Home2007.htm>.
113. S. Sahin, M. Yavuz, Y. N. Zhou. Introduction to Nanojoining (in Microjoining and Nanojoining, ed. Y. N. Zhou), Woodhead Publishing Limited, 2008.
114. J.-C. Charlier, M. Terrones, F. Banhart, N. Grobert, H. Terrones, P. M. Ajayan. Experimental observation and quantum modeling of electron irradiation on single-wall carbon nanotubes. *IEEE Transactions on Nanotechnology* 2003; 2(4):349.
115. A. V. Krasheninnikov, K. Nordlund, J. Keinonen, F. Banhart. Making junctions between carbon nanotubes using an ion beam. *Nuclear Instruments and Methods in Physics Research B* 2003; 202:224.
116. Z. Wang, L. Yu, W. Zhang, Y. Ding, Y. Li, J. Han, Z. Zhu, H. Xu, G. He, Y. Chen, G. Hu. Amorphous molecular junctions produced by ion irradiation on carbon nanotubes. *Physics Letters A* 2004; 324:321.
117. EUREKA/EURIPIDES project “Carbon Nanotubes/epoxy composites” (2007–2010), acronym CANOPY.

Physical Properties and Mechanical Behavior of Carbon Nano-tubes (CNTs) and Carbon Nano-fibers (CNFs) as Thermal Interface Materials (TIMs) for High-Power Integrated Circuit (IC) Packages: Review and Extension

Yi Zhang, Ephraim Suhir, and Claire Gu

Abstract Carbon nano-tubes (CNTs) and carbon nano-fibers (CNFs) are characterized by a number of unique and extraordinarily attractive physical/mechanical properties. Examples are high Young's modulus, high tensile strength, low coefficient of thermal expansion (CTE), and high thermal conductivity. These unique properties of CNTs and CNFs lead to many potentially attractive applications of these quasi-one-dimensional nano-structural elements in many fields of applied science and engineering. One of such applications is the use of CNT/CNF arrays as thermal interface materials (TIM) in electronic, opto-electronic, and photonic systems. In order to be employed in such a capacity, CNTs/CNFs must possess not only high thermal conductivity but also a number of other crucial physical/mechanical properties: be able to make good contact with attaching (usually hot) surface, possess good "anchoring" strength at their root cross-sections (i.e., good "adhesion" to their substrate), withstand high pressure (up to 100 psi), exhibit acceptable consistency in their heights and diameters, demonstrate adequate mechanical performance in the post-buckling conditions, exhibit good wettability to the embedding low-modulus material, if any, etc. Good manufacturability and low cost of CNT/CNF-based heat-spreaders and other heat removing devices are also necessary for practical applications in high-power electronic devices.

In this chapter we address several major experimental techniques that have been developed during the last decade and are being used to characterize the physical/mechanical properties of the individual CNTs/CNTs and CNT/CNF arrays, and to understand their physical (mechanical and thermal) behaviors and performance. We also discuss some results that have been obtained using these techniques. In the extension part of this chapter, we show how analytical stress modeling can be used to describe the mechanical behavior of CNTs/CNFs embedded into low-modulus elastic media. This is done to improve both the thermal and the mechanical performance of the CNT/CNF arrays. The developed models can be employed for the analysis and rational physical design of the structures in question.

Y. Zhang (✉)

Department of Electrical Engineering, University of California, Santa Cruz, CA, USA
e-mail: lyrixzhang@hotmail.com

1 Introduction

Since their discovery in 1991 [1], carbon nano-tubes (CNTs) have developed into a distinct branch of nano-technology, nano-material, and nano-mechanical engineering with numerous unique and useful applications. Carbon nano-tubes (CNTs) are unique tubular structures of nanometer diameter (say, <1–200 nm) and large length/diameter ratio (say, 500–1000 or so).

The following two major types of CNTs are distinguished:

- (1) *Single-wall* CNTs (SWCNTs) that consist of just one layer of tubular graphene and
- (2) *Multiwall* CNTs (MWCNTs) that consist of two or more concentric shells of carbon and a hollow inner capillary. The separation between the adjacent shells in MWCNTs is 0.34 nm.

There is also a third type of MWCNTs, in which there is no hollow inner capillary in the middle. Instead, the inner few layers of shells are linked together and form a “bamboo-like” structure. This type of CNTs is usually referred to as *carbon nanofibers* (CNFs).

Whether it is a SWCNT, a MWCNT, or a CNF, the unique quasi-one-dimensional structure and the graphene/graphite arrangement of carbon atoms in the carbon shells in these structural elements result in some extraordinary properties of the CNT/CNF material: exceptionally high Young’s modulus, significant high tensile strength, and extraordinarily high thermal conductivity. These unique properties lead to many attractive applications of CNTs and CNFs.

Thermal interface material (TIM) is one of the most attractive CNT/CNF applications [2–5]. Theoretically thermal conductivity of ideally structured SWCNTs can be as high as 6000 W/mK. This is even higher than that of the diamond [6]. Vertically aligned CNT/CNF arrays produced by using one of the *chemical vapor deposition* (CVD) techniques are viewed today as the most promising nano-structures and the most attractive fabrication technology, primarily because this technique enables one to produce consistent arrays with stable geometry [2].

Many theoretical and experimental studies have been conducted to evaluate the *thermal performance* of CNTs and CNFs. It has been established [7–9] that a single CNT exhibits thermal conductivities ranging from 1600 to 6000 W/mK. It is the common perception today that the predicted high thermal conductivity of single CNTs and CNFs might be retained in CNT/CNF arrays as well. This circumstance will lead to many breakthroughs in thermal management of IC and photonic devices. Particularly, it is expected that CNT/CNF-based systems might replace in the near future the existing conventional TIMs.

It should be pointed out, however, that in order to use CNTs or CNFs as effective and robust TIMs, these structural elements must possess, in addition to high thermal conductivity, also adequate mechanical properties, good “anchoring” strength (i.e., good “adhesion” to their substrates), ease in manufacturing, low enough cost, good wettability by the low-modulus embedding material, if any, etc.

There are several significant *challenges* one needs to overcome to make CNT/CNF systems suitable for TIM applications in electronic and photonic engineering. This is, first of all, the development and implementation of the well-controlled synthetic process of nano-materials and nano-structures. Second of all, it is the development of simple and reliable characterization techniques for both individual CNTs/CNFs and their “collective” properties, when arrays of CNT/CNFs are fabricated and used. Finally, one has to integrate these novel materials into the existing products, designs, and manufacturing technologies, or, if the advantages of the nano-technology-based TIMs are proven to be overwhelming, one has to be able to replace the existing designs, both in terms of better performance and lower cost. It is not easy to overcome the conservatism of the industry.

The ability to measure the physical (mechanical, thermal) properties of CNT/CNF systems (“*metroability*” of nano-systems) is of an obvious practical importance as well. It is also an obvious challenge. The major reason for such a challenge is due, of course, to the ultra small size of these structural elements. Specialized techniques and tools are needed even to pick up and install individual CNTs/CNFs into a metrological tool. In practical applications billions of CNTs/CNFs have to be grown and CNT-based films (arrays) have to be fabricated to provide transfer of an appreciable amount of heat. There is a crucial need therefore for the appropriate characterization methods and adequate equipment that would enable one to carry out, with satisfactory accuracy and a timely fashion, thermal and mechanical measurements of the “collective properties” of CNTs films and arrays.

In the review that follows we address some existing *measurement techniques*, instrumentation, and equipment that were developed during the recent years to characterize the mechanical properties of the individual CNTs/CNFs as well as the physical (mechanical and thermal) properties of CNT/CNF arrays. We report also on some practically useful results obtained using these techniques. In the extension part of this chapter we show how analytical *stress modeling* can be used to describe the mechanical behavior of CNTs/CNFs embedded into low-modulus elastic media. The developed models can be employed for the analysis and rational physical design of the structures in question.

2 Young's Modulus of Individual CNTs/CNFs

Young's modulus (YM) is an important material characteristic of any material, and therefore substantial effort has been dedicated to its evaluation [10–24] of the YM of CNTs/CNFs. It can be evaluated particularly by generating thermal vibrations of CNTs/CNFs and using *high resolution tunneling electron microscope* (HRTEM) [10–14] to record the vibration characteristics. The YM can be then evaluated from the measured vibration frequency and the CNT/CNF geometry. The mechanical behavior of individual CNTs/CNFs was investigated also using the nano-probe manipulation method [21–23]. It has been found that the YM of a CNT/CNF might be diameter dependent. Based on some other literature, primarily theoretical, data, the YM of individual CNTs/CNFs was found to be in a rather wide range from 2

TPa to 200 GPa, when the CNT diameter changes from 5 to 40 nm. This information agrees satisfactorily with the experimental data obtained using HRTEM and assisted by an external oscillated electrical field [14]. It has been found also that the YM of the CNTs decreases exponentially with an increase in the CNT diameter [14].

CNFs produced by *plasma enhanced chemical vapor deposition* (PECVD) are characterized by much larger diameters (about 100 nm) than CNTs. As is known, PECVD is a process mainly to grow CNTs or CNFs from a gas state (vapor) to a solid state on some substrates. In the case of CNT/CNF growth three metals are used as catalysts: cobalt, nickel, or iron. Because of the large diameter and also because of the far-from-ideal “bamboo-like” morphology of the CNFs, their measured effective YM is considerably lower, by orders of magnitude, than the CNTs.

3 Tunneling Electron Microscopy (TEM)

Tunneling electron microscopy (TEM) is perhaps the most powerful tool currently available to observe the structure of nano-scale materials. This tool is able to provide real space resolution better than 0.1 nm, as well as atomic-scale structural information. Compared with other techniques, TEM is viewed as one of the best candidates to characterize the mechanical behavior of individual CNTs/CNFs.

Different approaches have been developed today to integrate the nano-structured materials (nano-tubes, nano-fibers, nano-wires) into the TEM high vacuum chamber with other additional specialized fixtures to obtain the *in situ* measurements of the CNTs/CNFs properties.

4 Thermal Vibration Method

The first measurements conducted to evaluate the YM of individual CNTs were carried out by *thermal vibration method* in the TEM chamber by Treacy et al. back in 1996 [11]. The CNT used was synthesized by an arc-charged method. This method generally provides the CNT with the best structural quality. The CNT was mounted onto a specially designed holder, and put into the vacuum chamber of the TEM. The high resolution of TEM enabled one to clearly see the structure and the dimensions of the CNTs. At room temperature (300 K), the image was blurred at the free end of the CNT; however, the supporting end had good focus. After the temperature was increased to 600 K, one could clearly observe that the tip became more blurry. It was concluded that the CNT was vibrating due to thermally induced vibrations (*Brownian motion*).

It has been found that the thermal vibration amplitude depends on the temperature: it has been 30 and 40 nm at 300 and 600K, respectively. Detailed measurements of the change in the vibration amplitude with temperature are shown in Fig. 1. These measurements confirm the thermal origin of vibrations because of the Brownian motion.

When computing the YM of an individual CNT based on the thermal vibration phenomenon, the CNT was assumed to be a homogeneous cylindrical cantilever of

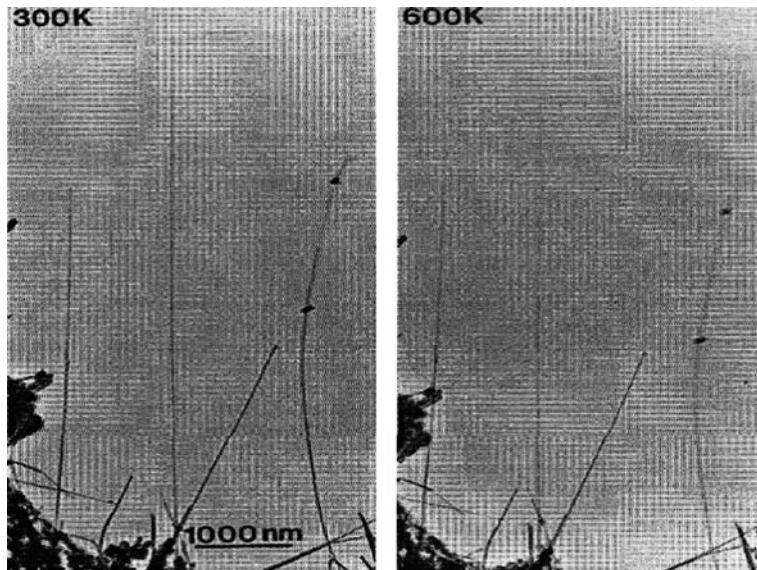


Fig. 1 TEM images show that the freestanding CNT vibrates at different temperatures (300 and 600 K) due to thermal effects [11]

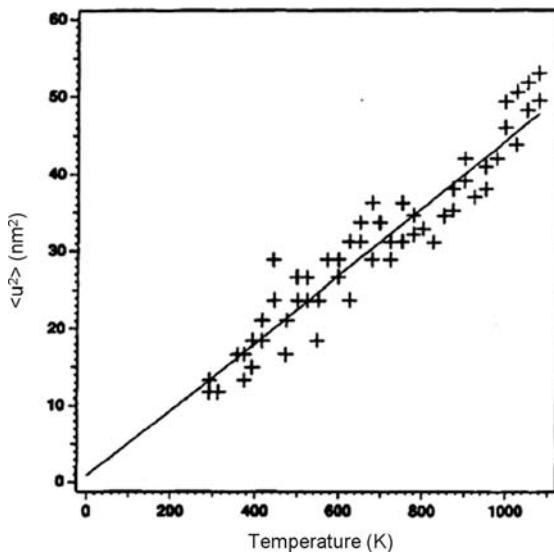
certain length and diameter. It could be shown that the vibration amplitude, u , at the CNT tip is related to the CNT Young's modulus (YM) and the vibration energy W as $W = (1/2)cu^2$ (see [11]). Here c is the effective spring constant, which is proportional to the YM and is affected by the CNT structure and geometry.

The vibration amplitude could be obtained from Fig. 2 for different temperatures. As to the additional vibration energy, it could be obtained from the analysis of the Brownian motion as an average $= kT$. Here k is Boltzmann's constant and T is the absolute temperature. Then one could easily calculate the YM of an individual CNT at different vibration modes. CNTs of different lengths and different inner and outer diameters were measured, and the Young's moduli were calculated. The calculated results indicated that the YM of CNT had a variation between 1.0 and 3.7 TPa, i.e., much higher than in the case of a typical CNF. The variations in the measured YM could be attributed to the uncertainties in the measurements of the CNT dimensions. Nonetheless, these data agree satisfactorily with the theoretical predictions.

5 Method Based on External Electric Field-Induced Vibrations

Poncharal et al. [15] developed a methodology for determining the YM of individual CNTs based on the observed, using TEM, electromechanical resonances. The CNTs were synthesized by an arc-charged method. The CNT was placed on a specially designed fixture that enabled one to apply the voltage across the CNT. Because of the small diameter of the CNT, the charges were accumulated on the CNT tip.

Fig. 2 Relationships between the heating temperature and the CNT thermal vibration amplitudes. The CNT was heated up to 800°C with 25°C as step. A linear relationship can be observed. This gives a YM value around 3.7 TPa for these particular CNTs [11]

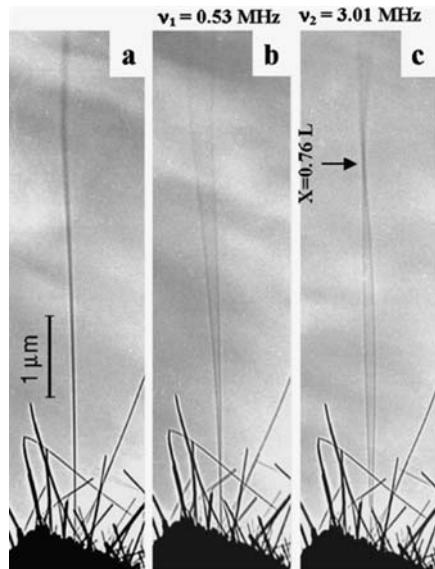


The CNTs experienced bending due to the electrostatic forces. By oscillating the voltage, one does so with the charges as well. The force direction also varies. When the applied frequency becomes close to the intrinsic (free) resonance frequency of the CNT, the CNT becomes subjected to elevated amplitudes of mechanical resonant vibrations.

For the first harmonic resonance, one can observe vibrations with the largest amplitude, as shown in Fig. 3. In order to calculate the YM of the CNT, an appropriate theoretical model has to be developed and applied. Based on the size of the CNT, one needs to decide whether the continuous elasticity theory is applicable to the case in question or a molecular dynamic approach should be considered. The authors calculated the node length of the second harmonic resonance, the frequency ratio, and the shape (deflected curve) of the oscillating CNT. They did that based on the elastic theory and compared the calculated data with the actually observed data. They found excellent agreement between the computed and the measured data. It has been concluded therefore that the model based on the continuous elastic approximation is still valid for the CNT data analysis. In this connection it should be pointed out that many well-established methods and approaches of engineering structural analysis and theory of elasticity that have been developed for continuous elastic media seem still be applicable to nano-structures and nano-materials despite their next-to-molecular size. It would be advisable, of course, to compare, whenever possible, the results based on engineering and applied mechanics with the results based on applied physics and quantum mechanics for the given nano-structure and nano-material.

Assuming that a CNT can be idealized as a continuous hollow cantilever with one end fixed on the substrate, the resonance frequency can be found as [15]

Fig. 3 A selected carbon nano-tube at (a) stationary, (b) the first harmonic resonance ($n_1 = 1.21$ MHz), and (c) the second harmonic resonance ($n_2 = 5.06$ MHz). The vibration of the CNT is caused by the alternating external electrical field [15]



$$v = \frac{\beta^2}{8\pi} \frac{1}{L^2} \sqrt{\frac{(D^2 + D_i^2) E}{\rho}}$$

Here D is the diameter of the CNT, L is the CNT length, ρ is the material's density, and E is the bending (Young's) modulus, which depends on the shape of the CNT. To apply the frequency of the excitations that matches the intrinsic resonance frequency of the CNT is not easy: one needs to omit the nonlinear term in the theoretical model as introduced in [15]. This is due to the fact that nonlinear frequencies are affected by the excitation force and, for this reason, are not characteristics of the structure and the material only. On the other hand, one needs to observe the largest oscillation amplitude in order to determine the intrinsic resonance frequency. As shown in Fig. 4, the oscillation amplitude is symmetric with respect to the central intrinsic frequency, with the integral 0.12 MHz.

In the CNT case, the bending of a CNT is determined not only by the YM of the material, but also by the wall thickness and tube diameter. The in situ TEM experiments measured the bending modulus. Figure 4 shows a group of experimentally measured bending moduli of MWCNTs. It is apparent that the smaller-size nano-tubes have a bending modulus (i.e., YM computed from the bending tests) approaching the YM obtained for tensile deformations, while the difference in these moduli for the larger-size CNTs turned out to be much smaller. This size-dependent property could be attributed to the possible inelastic deformations of the CNTs.

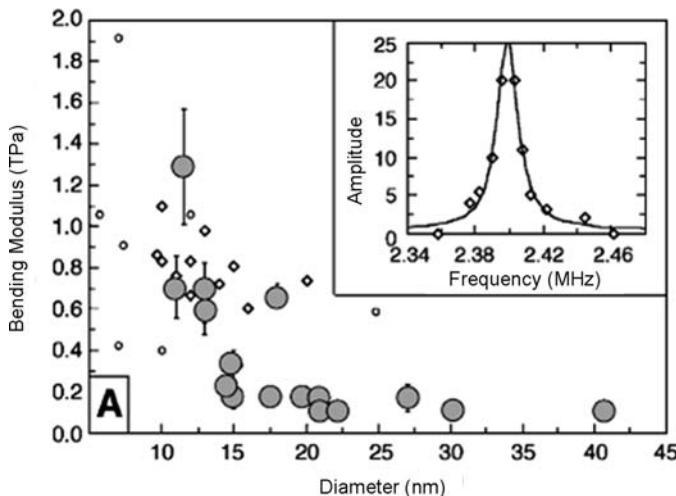


Fig. 4 Bending modulus of the multiwalled carbon nano-tube produced by arc discharge as a function of the outer diameter of the nano-tube. The larger CNT typically possesses a smaller YM compared with the thinner CNT. The inset shows the CNT vibration amplitude around intrinsic vibration frequency [15]

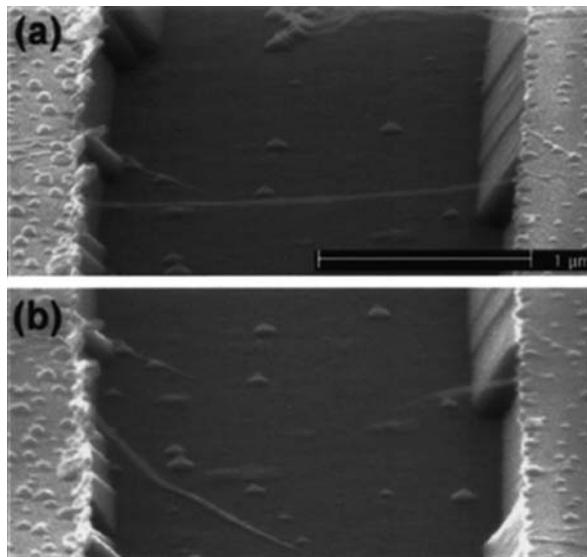
6 Measurements Using Scanning Probe Microscope (SPM)

The YM of an individual CNT can also be evaluated by using *scanning probe microscope* (SPM), i.e., the SPM method [22–25]. The mechanical behavior of individual CNTs was investigated using the *nano-probe manipulation* method. The observed CNT structural failure was due to buckling of graphene layers. In the conducted experiments, the SWNT bundle rope was suspended on a specialized designed fixture, and the *atomic force microscope* (AFM) tip was launched onto it, thereby applying sideward pulling forces.

SWCNT “ropes” suspended on the designed Si fixture are shown in Fig. 5a, b. It was reported that the CNT was able to sustain reversibly many cycles of elastic elongation up to 6%. By assuming that the elongation is transmitted directly to the individual CNTs, the corresponding tensile strength was calculated to be above 45 GPa. This number is in satisfactory agreement with that for multiwalled tubes, although the details of the strain distribution could not be revealed in this experiment [26].

A direct tensile, rather than sideways, pull of a multiwall tube or a rope (i.e., bending deformations) has a clear advantage due to simpler load distribution and a straightforward way of measuring the YM. An important step in this direction has been recently reported by M. Yu et al. [27]. In this work, the CNT was attached between two atomic force microscope (AFM) tips and the tensile pulling force was applied directly onto the CNT with the displacement and force relationship recorded. The direct tensile experiments for MWNTs report tensile strengths in the

Fig. 5 Ropes of SWCNT suspended on the designed Si fixture, the CNT before and after sideways pulling were shown in (a) and (b) separately [26]



range of 11–63 GPa. The measured data did not show clear dependence of this strength on the outer diameter of the CNT shell. The nano-tube broke in the outermost layer, and the analysis of the stress–strain curves indicated that the YM for this layer was between 270 and 950 GPa. The measured normal strain at failure can be as high as 12%.

7 Method Using CNT Buckling

In addition to the deformations due to vibrations, deflections, and bending discussed above, the deformations associated with *buckling* (loss of elastic stability under the action of compressive loading) is another effective method to determine the YM of CNT/CNF materials and structures. This method was used to characterize the CNT by integrating it with the SPM tip [25].

The concept is more or less straightforward: one needs to measure the buckling forces of the CNT and calculate the YM based on the Euler formula. The preparation of test specimens is not trivial, however, to say the least. The experiment was conducted in the SEM chamber, so one could record the displacement of the CNT by comparing the images. The forces were measured by the cantilever force sensor (Figs. 6 and 7).

The measured force vs. displacement relationship is shown in Fig. 8. The measured forces are proportional to the displacements. At the positions indicated by the thick arrows, the force abruptly decreases. This decrease should be interpreted as CNT buckling. The strain energy of the nano-tubes is released at this point. The Young's moduli of the 20 shells of MWCNT and the processed nano-tubes, from

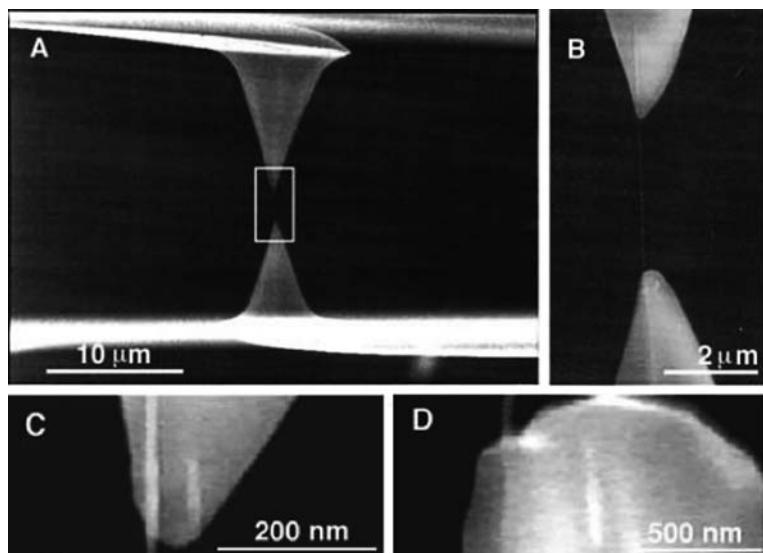


Fig. 6 SEM image of two oppositely aligned AFM tips holding a MWCNT attached at both ends on the AFM silicon tip surface by electron beam deposition of carbonaceous material. The lower AFM tip in the image is on a soft cantilever whose deflection is used to determine the applied force on the MWCNT. **B–D:** Large magnification SEM image of the indicated [27]

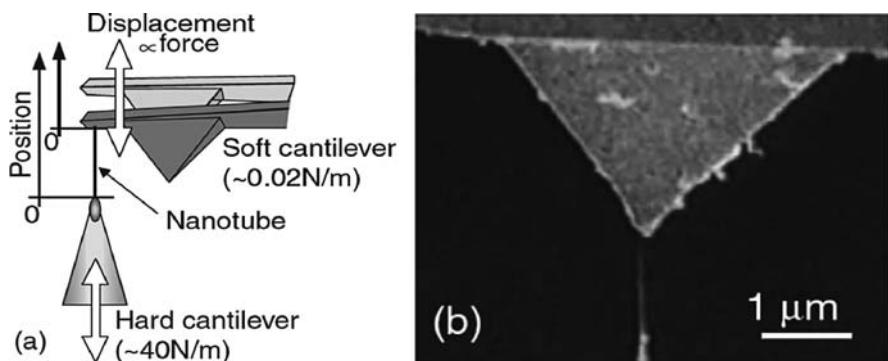
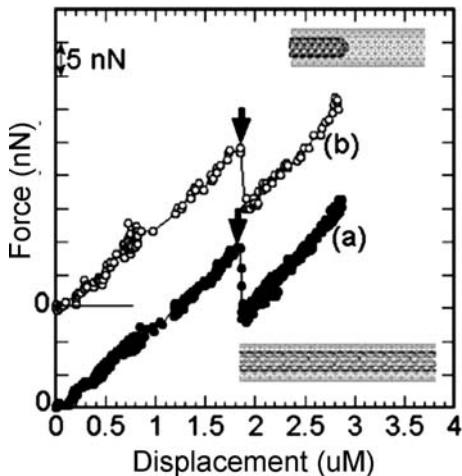


Fig. 7 Experimental setup schematic for using AFM to measure the direct tensile strength of individual CNT, which is attached onto the AFM tip, the actual SEM images for the testing sample is shown in (b) [25]

which 14 inner shells were extracted, are estimated to be 0.77 and 0.80 TPa, respectively. These data follow from the Euler's buckling model. The results imply that this model is applicable to the analysis of the axial buckling behaviors of the CNTs. It could also be applicable also for the characterization of mechanical properties of other types of nano-wires.

Fig. 8 The measured relationship between pulling force and the atomic force microscope (AFM) tip displacement, which is the same as the CNT extension length under tensile stress [25]



8 Buckling of CNT Shells

Shell buckling strength (critical loading) is another important mechanical property of the CNT/CNF material and structure. Ultra high resolution AFM can be employed to characterize CNT shell buckling. M. Yu et al. [27] developed a method for measuring the shell buckling strength by applying the nano-indenter instead of the more conventional nano-manipulation technique.

With a specially designed CNT array, which was synthesized on an aluminum substrate, one can obtain well-oriented and vertically aligned CNT array with an aluminum matrix as filling materials to provide structure rigidity [27]. The synthesized CNTs had a relative large diameter ($20 \sim 200$ nm) with well-secured space in-between (100 nm). With CNTs exposed out from the Al surface for a very short distance, one could make sure that the measured results are from the shell buckling (“local buckling”) rather than of CNT “global” buckling or bending. The size of the nano-indenter tip is much bigger than the AFM tip and the CNT, so that when it was launched onto the CNT matrix surface; tens of CNTs were making direct contact. So, the measured results show the group appearance for the CNT, and one can extract the actual shell buckling strength (critical load) for each one with appropriate model. The compressed CNTs are shown in Fig. 9.

The load–displacement curves obtained in three sets of experiments are shown in Fig. 10. Each plot represents a loading–unloading cycle. The loading portion consists of three stages: (1) an initial linear increase, followed by a (2) sudden drop in the slope and the curve becoming flat, and a (3) third stage comprising an increasing load. The sudden decrease in the slope is the signature of the shell buckling, which indicates the collapse process (Fig. 9). The critical buckling load has been consistently measured to be between 2 and $2.5 \mu\text{N}$ from multiple experiments. After buckling, the adjacent nano-tubes get into contact with the indenter tip. This results

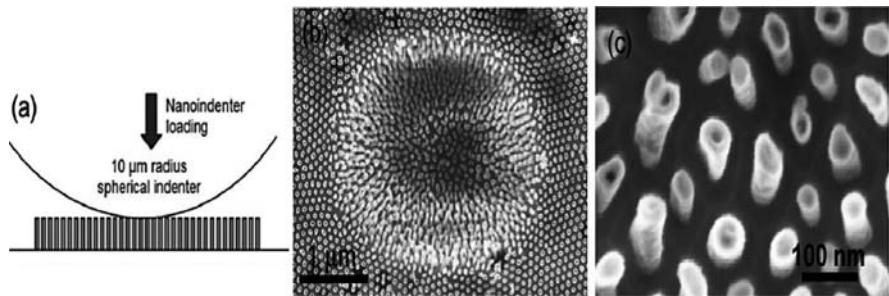


Fig. 9 Nano-indentation of individual carbon nano-tubes. (a) Schematic illustration of the experiment, in which a Berkovich indenter with 100 nm tip radius vertically compresses a multiwalled carbon nano-tube. (b) Schematic illustration of shell buckling in nano-tubes. (c) An in situ scanning image of the sample surface obtained immediately after compressing a nano-tube [27]

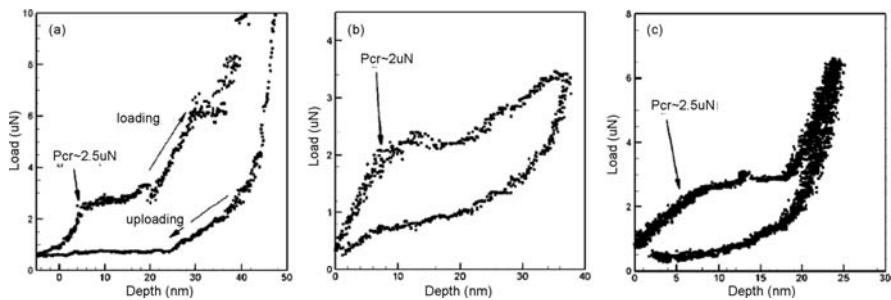


Fig. 10 Representative load–displacement data from a loading–unloading cycle. (a) Displacement-controlled indent on a 100 nm long nano-tube, (b) load-controlled indent on a 100 nm long nano-tube, and (c) load-controlled indent on a 50 nm long nano-tube. In all cases, buckling, evident by a distinct drop slope, occurs between 2 and 2.5 μN . The critical buckling load was repeatable to within $\pm 0.25 \mu\text{N}$ between experiments [27]

in an increase in the load, as shown in Fig. 10 (third stage). In each plot, the position of zero displacement corresponds to a nonzero load because a small preload was used to detect the location of the sample surface.

9 Effective Young's Modulus of CNT/CNF Arrays

9.1 Stress–Strain Relationship of PECVD-Synthesized CNT Film

In addition to the methods that have been developed to characterize the YM of individual CNT, there has been significant interest of utilizing CNT as TIMs or compressive foams [28, 29]. Within these applications, one is interested in studying mechanical properties of the CNT systems as quasi-bulk materials. The CNT/CNF

array (CNFA) “brush” is treated [30], when such an approach is used, as a continuous elastic layer. As is known, the elastic response of a continuous, homogeneous, and isotropic elastic medium depends on two constants – YM and Poisson’s ratio (or on two Lame constants, which are derivative of the above two). While the Poisson’s ratio changes in relatively narrow limits, the YM might be quite different for CNT/CNF materials obtained by using different synthesis methods. On the other hand, it is the YM that plays the most important role in the mechanical behavior and performance of materials. Although many phenomena of the CNT/CNF materials behavior can be adequately described and explained only on the basis of quantum mechanics, it is often assumed [31, 32] that methods of the theory-of-elasticity and methods of engineering mechanics can be successfully employed to evaluate the YM of the CNT/CNF materials.

The main objective of the study, whose major findings are described below, was to develop a practical methodology for the evaluation of the effective Young’s modulus (EYM) of vertically aligned PECVD-synthesized CNFA. This could be done on the basis of the measured compressive force vs. axial displacement in the post-buckling mode.

There are several reasons why the axial displacements should be large enough to ensure that the CNFA behaves in the post-buckling mode conditions. First, the force-displacement relationships in the post-buckling mode can be obtained in a very wide range of measured forces and displacements. Second, exact solutions exist (Euler’s “elastica”) for the evaluation of the highly geometrically nonlinear bending deformations of flexible rods (wires). Treating CNFs as nonlinear wires, one can use the “elastica” solutions to predict the EYM from the experimentally obtained force-displacement relationship. If the material exhibits nonlinear stress–strain relationship, i.e., if the EYM is stress dependent, then this dependency could be obtained as a suitable correction to the YM value obtained on the basis of the “elastica” solution. Finally, it is important that the interfacial pressure does not change significantly with the change in the axial displacements of the CNFA. This indeed takes place, when the compressed CNFA is operated in the post-buckling mode not very much remote from the buckling condition. Since in many practical applications CNFs are intended to perform in the post-buckling mode, the evaluation of their behavior in such a mode enables one to obtain valuable information for these structural elements in the use conditions.

The samples used in our experiments were synthesized using plasma enhanced chemical vapor deposition (PECVD). The fabrication procedures of CNFs in an embedded vertical array were described in detail in [33, 34]. Briefly, a 30-nm thick layer of Ti was deposited on a Si substrate, followed by a 35-nm layer of Ni catalyst, both deposited by ion beam sputtering. The CNF array was grown using a PECVD process with C_2H_2 as the hydrocarbon source and NH_3 as a diluent/etchant gas [33]. The table version of the Instron tester (Model 5542) was used in our analysis to measure the compressive force. The CNFA was placed on the lower anvil of the Instron machine, while the displacements of the upper anvil were automatically and thoroughly controlled, particularly, when approaching the CNFA surface with a speed of about 5 $\mu m/min$. The maximum capacity of the Instron and the maximum

force applied during the experimentation was about 0.5 N (~ 500 gf), and the resolution of the measured force was about 0.01 N. This provided sufficient accuracy to distinguish the changes in the compressive strain.

The axial displacements of the CNFA were on the micrometric level. The regular Instron equipment loses its accuracy, when measuring such small axial displacements. This is because the measured displacements become comparable with the displacements of the beam in the Instron's load cell. For this reason, the *Capacitec equipment* was employed in our experimental setup to measure the relative displacements of the Instron anvils and, hence, the axial displacements of the CNFA. The Capacitec equipment uses a non-contact way of determining the displacement by measuring, with a rather high accuracy, the change in the electrical capacity of the air gap between the Instron anvils. The anvils play in this situation the role of two electrodes. The strain signal is imported into the Instron device for calibration. By initializing the beginning strain point to correspond to zero displacement, one is able to measure the strain signal accurately and to transfer this signal into the displacement data, which has a higher resolution at ~ 0.5 μm . By using a combination of Instron and Capacitec devices, we were able to measure the force vs. displacement with a rather high resolution and, owing to that, to calculate the EYM with a high accuracy.

The quality and the structure of the CNFs obtained, as has been indicated, by using PECVD, were observed under the scanning electron microscope (SEM). An example of the SEM image is shown in Fig. 11. The length (thickness) of the CNFA was about 15 μm with about 2 μm variation. The diameter of the CNTs was about 80~100 nm. The samples had a high CNF coverage density ($\sim 10^8$ CNF/cm 2), and the space between the two adjacent CNFs was about 1 μm . CNFA sample with lower density was used to obtain such a relatively larger space. The properties of

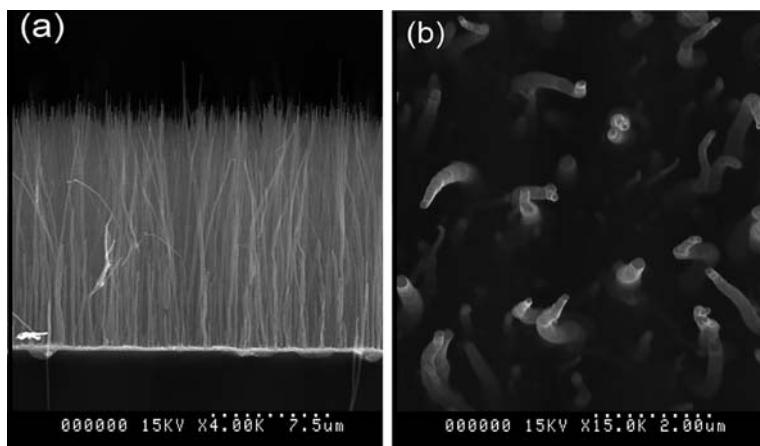


Fig. 11 Microscopic picture for a typical CNT/CNF samples used in the experiment. (a) SEM cross-section view of the CNTs sample. (b) SEM top view of the CNT sample shows the density of the CNT generated. The scale bar of these images is 7.5 and 2 μm

the CNFA could be controlled by varying the synthesis conditions. The size of the CNFA in our test samples was 2×2 (length \times width) cm².

The measured signals were the compressive force and the axial displacement. Based on the developed predictive model developed for a flexible rod (Euler's "elastica"), one is able to obtain the relationship between the T/T_e and δ/l ratios, where T is the compressive force, T_e is the buckling force [32, 35] (that could be determined assuming that an individual CNF could be treated as a cantilever rod clamped at the CNF substrate), δ is the axial displacement of an individual CNF, and l is its length. As a result of measuring the force-displacement relationship using the Instron-Capacitec device, one could obtain the relationship between the applied force, T , and the relative axial displacement, δ/l . Based on this relationship, if one assumes the YM (E_0) at low strain levels of the CNF array to be constant, then it could be directly calculated from [32, 35].

The actual measurements indicated that the experimentally obtained T/T_e vs. δ/l relationships deviate from the predictions based on the "elastica" solution, i.e., on the solution obtained for strain-independent YM. This deviation should be attributed to the nonlinear stress-strain relationship for the CNFA material. To account for such nonlinearity, we use the following approximation for the "nonlinear" EYM of the CNFA:

$$E \cong E_0 + \gamma\sigma$$

Here, σ is the compressive stress and γ is the factor of nonlinearity. Both the stress and the factor of nonlinearity could be obtained from the compressive stress vs. displacement relationship. In order to determine the best fit γ value, a series of experimental stress vs. displacement curves was generated (Fig. 2). These curves enable one to approximate ("fit") the experimental results by theoretical curves, and to obtain the appropriate γ values that result in the best agreement between the experimental and the theoretical data. The γ values obtained in such a fashion could be then used to predict the EYM of the CNFA.

In Fig. 12, some simulation results are shown for different values of the parameter γ , along with the experimentally obtained curves for a particular specimen. As evident from the processed data, the theoretical curve with the parameter of nonlinearity of $\gamma = 20$ satisfactorily matches the experimentally obtained curve. Hence, it is this γ value that could be used in the case in question to evaluate the EYM for CNFA of a particular batch ("run").

Five samples from the same wafer ("batch") were prepared. Because of the same synthesis conditions, these samples were assumed to be identical. SEM images were taken for all the samples. Figure 1 contains, as an example, the SEM image for one of the samples. During the experiment, all the test parameters were kept the same for all the samples. The relationship between the compressive forces and the displacement was recorded for each sample and introduced into the calculation method (Fig. 3a). As a result, one could obtain the relationship between the calculated YM and the δ/l ratio for the tested samples. The EYM of the CNFA at 30% deflection is about 90 GPa and exhibits clearly pronounced nonlinear behavior.

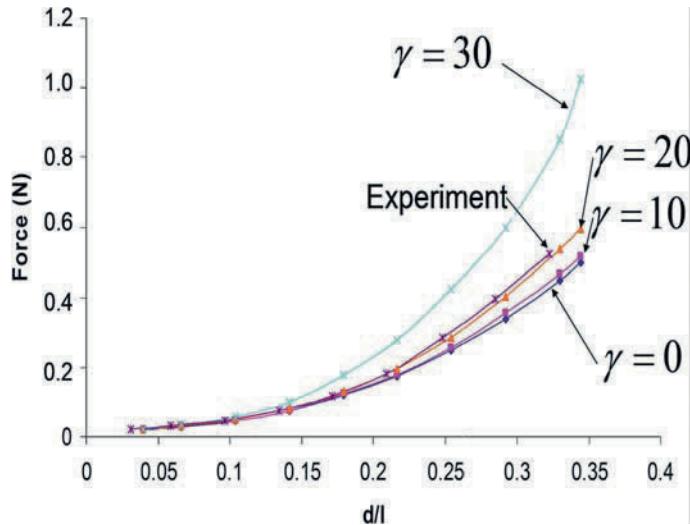


Fig. 12 The simulated relationship between the applied force, T , and the displacement-to-length ratio δ/l corresponding to different nonlinear coefficient. The non-linear coefficient $\gamma = 20$ fits best (for the given specimen) the experimental curve

Some scattering of the experimental data should be attributed not only to some differences in the test samples. The error bar in Fig. 13 characterizes the system measurement error. For some samples, the measured data still fall outside of the error bar. Thus the samples of the same run (“batch”) possess some minor sample-to-sample variations. But one could nevertheless assume that these samples possess similar mechanical properties.

CNFs in our tests were synthesized at five different deposition conditions. The catalyst density was kept the same in order to obtain the same CNFA density for all the different synthesis conditions. The SEM images of all the samples were taken. The diameter and density of the CNFAs remained the same for all the samples. There was, however, about $2\sim 3 \mu\text{m}$ variation in the length of the CNFs in the CNFA. The growth time was controlled to be the same, in order to make sure that this parameter does not affect the CNF characteristics. Figure 4 shows the EYM for the five tested samples. The computed EYM varies from 5 to 35 GPa for different samples. The CNF synthesized at optimized growth condition typically possess better atomic structure than those synthesized at other conditions. This would make to CNFs synthesized at optimized condition “stiffer,” i.e., their YM would be higher, when subjected to a compressive force.

From Fig. 14, one could observe that under small strain, like 10% in compression, the measured YM of the CNFs appears to be very similar to each other. It is only when larger compression is achieved, the nonlinearity of the measured YM of samples begins to get manifested clearly. For the samples of the same batch, we used, as a suitable characteristic of the material’s nonlinearity, the parameter γ averaged over the CNF length. This parameter is certainly different for samples taken from

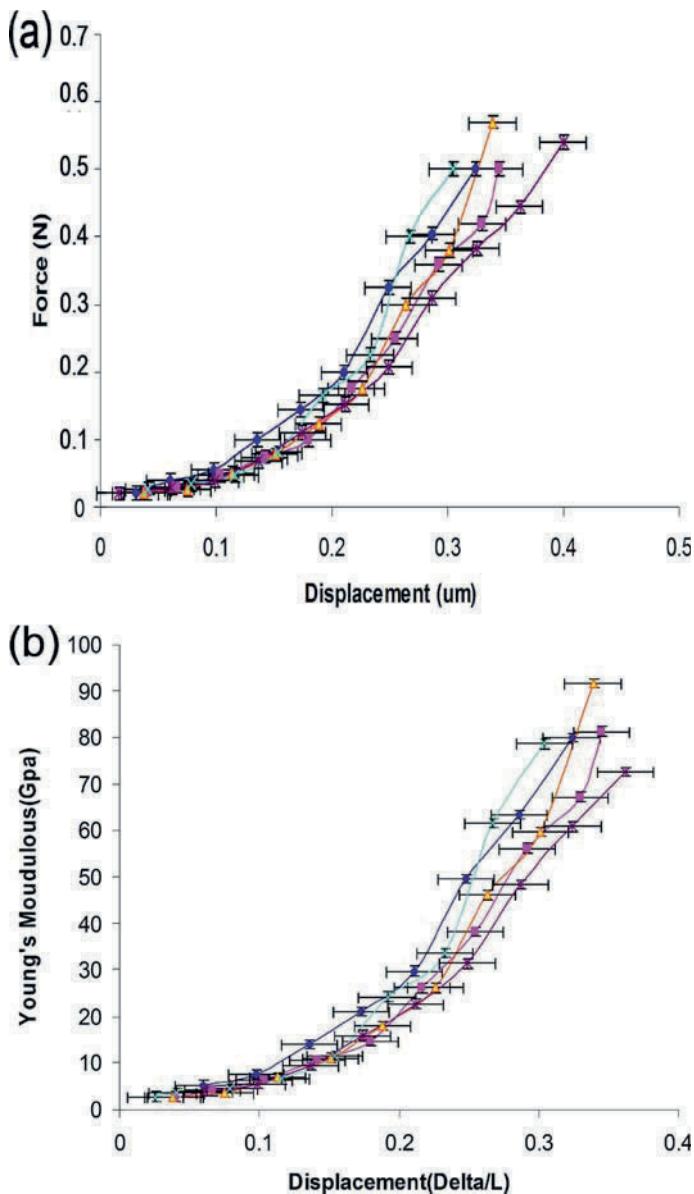


Fig. 13 (a) Measured compressive force as a function of the δ/l ratio for five samples taken from the same wafer (run). (b) Calculated YM of these samples based on the measured compressive force in (a), it reaches 90 GPa corresponding to a 30% lateral displacement of the CNT array

different batches and is determined by both the quality of the CNFs and the density distribution.

It is noteworthy that the developed experimental methodology along with the predictive stress models provides a rather high resolution for distinguishing between

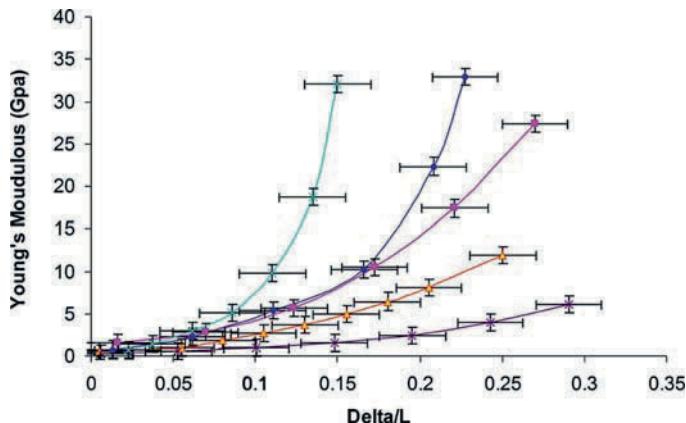


Fig. 14 Calculated Young's modulus vs. δ/l ratios for samples taken from different “runs” (different growth conditions). This data indicate that different growth conditions result, in effect, in considerably different materials

the mechanical properties of CNFA synthesized under different conditions. Despite of some unavoidable “noises” in our experiments, the developed methodology for the EYM evaluation could be used, in particular, to establish which method for monitoring the CNF growth has the highest potential to obtain the CNFs of the highest quality possible.

9.2 Stress–Strain Relationship of TECVD-Synthesized CNT Film

Compared with the well-aligned CNT synthesized by PECVD, the TECVD synthesized CNT is more entangled and appears as matrix. So the mechanical properties of these two types of film should behave differently. Previously, millimeter long carbon nano-tubes synthesized by TECVD have been used as super compressive foam [19]. The experiment was demonstrated on the same Instron equipment used by us. The CNTs array was placed between the anvils. With the loading compressive pressure, the CNT film was compressed. It was discovered that the CNT film could perform like foam, which means the CNTs array would bound back to its original length or maybe reasonably shorter. In addition, the CNTs array was also compressed repetitively by a few thousand times under the same condition and they buckled like the regular spring. Figure 15 shows the SEM images of the CNTs array under different cycles of compressive load. The CNT perform like spring. Also the force vs. displacement curve was presented as well. Based on the measured stress–strain relationship, one is allowed to calculate the effective YM of CNT film similar to the method introduced before.

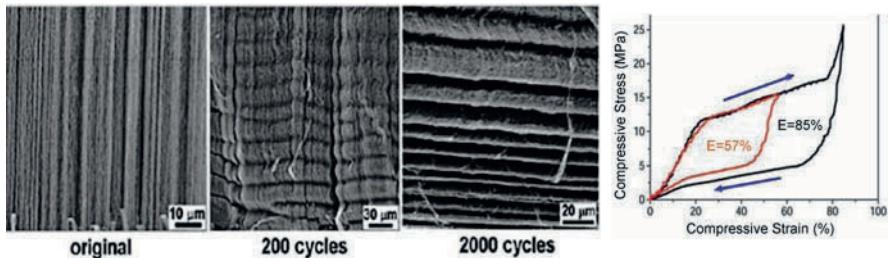


Fig. 15 SEM images for the CNTs array under different cycles of compressive load and the measured stress–strain curve [19]

9.3 Effective Young's Modulus has a Certain Limit

We have been able to build a model to calculate the effective Young's modulus (YM) of the CNT array from the force vs. displacement provided by Instron and Capacitec. Here with the same methodology, we calculated the effective YM of the CNT array from the relationship between force and displacement of CNT. The load cell was launched onto the CNT array as described before and the maximum force applied on the CNT array is about 50 gf. With this amount of force, the CNT array will not be damaged. It is confirmed by the scanning electronic microscope images, with the same sample placed in the same location. Instead of applying the force on the sample one time as before, same force is applied to the sample a few times and the YM is calculated accordingly. It is interesting to observe that the effective YM of the CNT reaches a limit when the sample is pressed by the load cell for a few times (Fig. 16). In other words, the effective YM of the CNT array

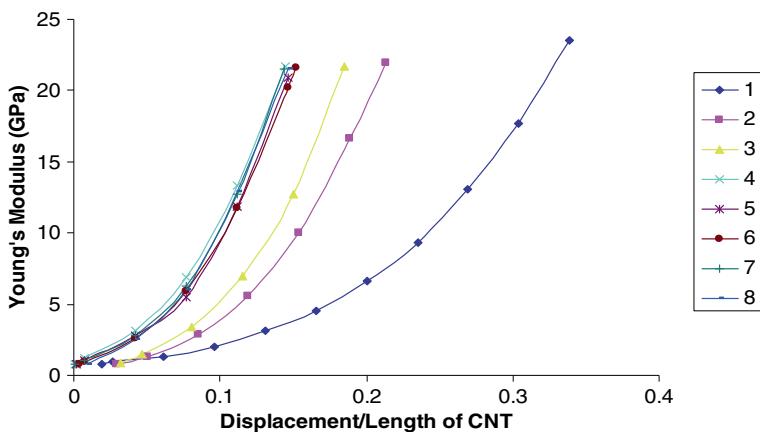


Fig. 16 The Young's modulus converges to a certain limit when the sample is compressed with 50 gf for a few times

was increasing and finally saturated with certain amount of compressions. Although the amount of forces applied here is relatively small, however, it is already large to cause the CNT film to deflect. The measurement technique is accurate, since one can clearly observe the difference in mechanical properties at each time of compression.

More interestingly, it was discovered that the effective YM of the PECVD-synthesized CNT tends to increase to certain limit under the repetitive load. By setting different strain point (displacement/length), the final YM of the CNT tends to increase to a certain limit and gets saturated. However, compared with the previous publication in [36], it was discovered that the YM of the CNTs array tends to decrease with the repetitive load. The exact reason that caused that difference is still unknown. However, it could possibly attribute to the mechanical properties difference and different buckling mechanisms (Fig. 17).

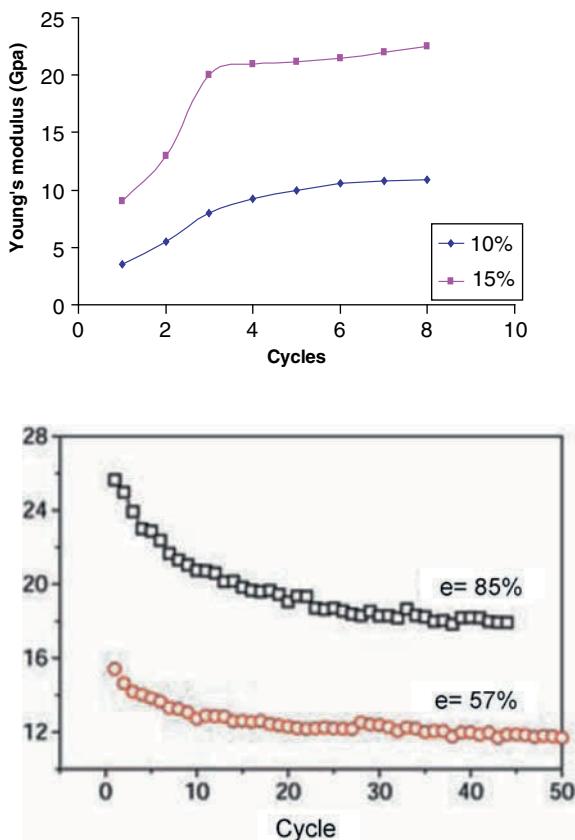


Fig. 17 The At different displacement position, both plots show that the Young's modulus of CNF array and of the individual CNFs begin to increase for PECVD-synthesized CNT; however, it shows the opposite for TECVD synthesized CNT [19]

10 CNT/CNF-Based TIMS: Requirements for Physical (Mechanical) Properties

10.1 CNT/CNF Compliance

We have indicated above that the CNTs are good candidates for the *TIM applications* mainly because of their good thermal properties. It is equally important that the employed CNT/CNF arrays possess adequate physical/mechanical properties, especially those that directly affect their performance as TIMs.

The reason of bringing in a TIM is to fill the gaps between two rough solid surfaces (bond lines) and thereby reducing the thermal resistance of the heat removing device. CNT/CNF arrays can be grown on copper(Cu) substrates or on Cu films. They can be grown, of course, on other materials as well. Cu has excellent thermal conductivity, and is generally used as heat sink, however, due to the rigidity of Cu, the surface roughness cannot be filled. That is why one needs a material with good compliance (softness) to compensate for the “hardness” of Cu.

Thermal resistance of CNFs as TIMs includes the bulk resistance of CNFs and contact resistances between the CNFs tips and the hot contact surface. Due to the good thermal properties of CNFs, the physical nature of the contact resistance tends to take precedence. CNFs also need to possess good compliance properties as conventional *phase changing materials* (PCMs) to change the orientation of the CNFs array according to the contacting surface roughness and configuration. The nano-scale dimension nature of CNFs allows them to fit well with the surface roughness and compliance nature of CNFs and make good contact with the hot surface under certain pressures. The CNT array synthesized by thermal enhanced chemical vapor deposition (ECVD) has also been demonstrated to have super foam-like nature and could be compressed by up to 85% of its original length, when it is experiencing a 13 MPa compressive pressure [36].

In the section, we elaborate on the compliance properties of CNT/CNF arrays and their effect on the thermal properties of the CNF arrays [28, 29], as well as on CNF arrays filled with Cu as TIMs. The CNF-Cu composites are fabricated by *electrochemical deposition* (ECD), by varying the time of the sample going through the ECD, amount of filling Cu could be controlled and left different numbers of CNFs, and CNFs with different length to be exposed. Briefly, a thin layer of titanium (30 nm) was deposited on a Cu substrate, followed by a 35-nm layer of nickel catalyst, both deposited by ion beam sputtering. The CNF array was grown using a PECVD process with acetylene (C_2H_2) as the hydrocarbon source and NH_3 as a diluent/etchant gas. The quality and the structure of the CNF could be observed under the *scanning electron microscope* (SEM) and *high resolution transmission electron microscope* (HRTEM). The length of the CNF arrays was about 20 μm with a 3~5 μm variation (see Fig. 18). The assessed diameter of the CNFs was about 100 nm (see Fig. 18b). The sample had a high CNFs coverage density ($\sim 10^9$ CNF/ cm^2), as one could see from Fig. 18b.

Sibling samples synthesized with the same conditions as in Fig. 18a, b were used to fabricate the CNF–Cu composites. They were prepared in this work by Cu

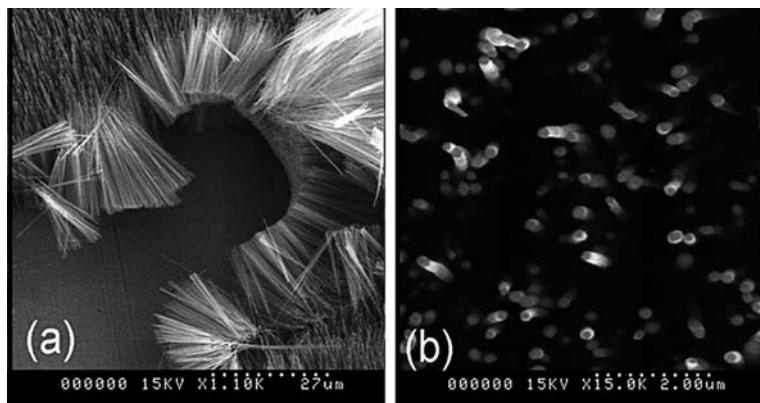


Fig. 18 *Tilt view and top view of the CNTs samples synthesized on Cu substrate*

ECD using a three-electrode configuration in a Cu-containing electrolyte bath [37]. The titanium–copper (Ti–Cu) substrate, on which the CNF array is grown, acts as the working electrode for Cu deposition. Titanium is chosen due to its favorable electrochemical properties that lead to good filling characteristics compared to other metals, such as chromium. The potential and current control of the electrolyte bath with respect to the working electrode allows us to control deposition rate and amount of Cu deposited.

By controlling the deposition time, the thickness of ECD Cu layer filled into the gap between the CNFs could be varied and left different amount of CNFs and CNFs with different length exposed. For 20-min soaking time of CNFs array in the electrolyte bath, the 30° tilt view indicates the CNF protruding from the Cu matrix is about 4~6 μm (Fig. 19a), and the density of the CNF array decreased to approximately 5 μm⁻² (Fig. 19b). It is interesting to observe that after the ECD process and the soak of CNF in Cu-containing solution, the CNF array still possess the vertical aligned nature. Furthermore, another sibling sample was soaked in the solution for 25 min. In this case, a relative thicker ECD Cu layer is obtained, the SEM images indicate that the length of the CNF protruding out is about 2~3 μm, there are a few CNFs relatively longer, but they are not the majority part of the CNFs. One could clearly observe the good “rock” like crystallized Cu. The CNFs are nicely surrounded by the ECD Cu. Visually there is no significant amount of voids introduced by the ECD process. If one continues to do the ECD deposition, eventually, all the CNT will be buried and it ends up with a relative smooth Cu surface (Fig. 19e, f).

Usually, there would be a following electro-polish step, which helps one obtain the smooth CNF–Cu surface and leave exposed nano-fibers protruding from the surrounding Cu matrix. However, we observed that with our CNFs array synthesized on Cu substrate, after ECD process for 30 min, all the CNFs are covered by the ECD Cu, and we surprisingly observe an excellent flat Cu surface (Fig. 19f), the surface roughness is measured to be ~100 nm. The SEM image of the side view of the

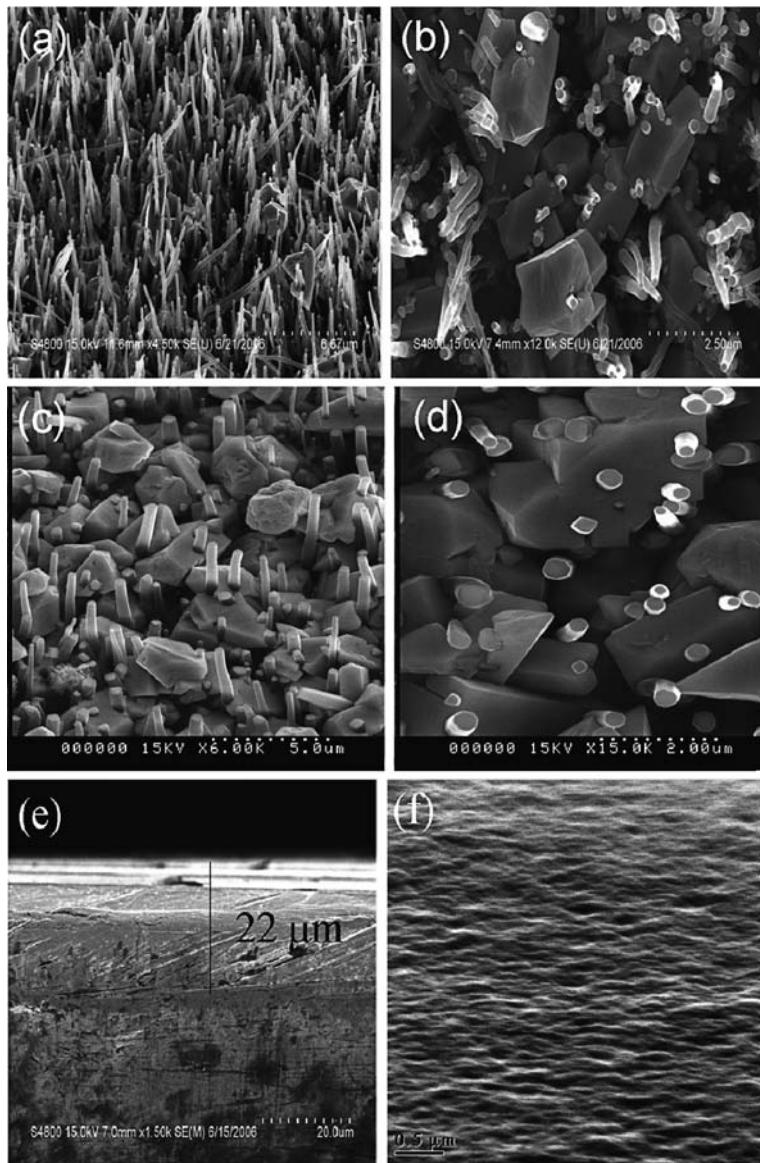


Fig. 19 SEM images for the ECD-prepared CNT samples. The sibling CNTs possess the sample length, however, after the ECD bath for different period of time, the gap between the CNTs was filled with Cu for different thickness. So the amount of exposed CNT appears differently and also rigidity of the CNT sample has an affect on the CNTs thermal performance

sample indicates that the ECD Cu layer is about 22 μm thick, which matches the length of the bare CNF array in Fig. 18a).

After the CNFs array and CNF–Cu composites are prepared, in order to study the compliance property of CNFs array and its effect on the thermal properties of CNFs as TIMs, the total thermal resistance for each group of samples are measured based on the recently developed photo-thermal method. The experimental setup that used to measure the effective thermal resistance of CNT array has been described in detail earlier. The phase shift signals measured for CNF array, CNF–Cu composites with different ECD copper thickness are shown in Fig. 20. The calibrated phase shifts of the samples are obtained by subtracting the system reference phase shift from the phase shift of the thermal radiation from the Ni film. The calibrated phase shift signal characterizes the thermal properties of the CNFs and CNF composite samples. In the experiment; the phase shift is measured over a wide frequency range. Different values of the unknown properties are used to calculate the theoretical phase shift and compare with the experimental result. The properties giving the best fit of the experimental data are taken as the properties of the CNF composites sample.

Based on the calibrated phase shift fitting and the employed physical model, the effective total thermal resistance of CNF, CNF–Cu with 20, 25, and 30 min deposition are measured to be 0.13 ± 0.01 , 0.18 ± 0.01 , 0.25 ± 0.01 , and $0.32 \pm 0.01 \text{ cm}^2 \text{ K/W}$, respectively. The total thermal resistance of the sample were calculated by considering the actual heating radius of the laser spot ($A = 1 \text{ mm}$ [2]). From the measured thermal resistance, one can observe that the bare CNFs array sample has

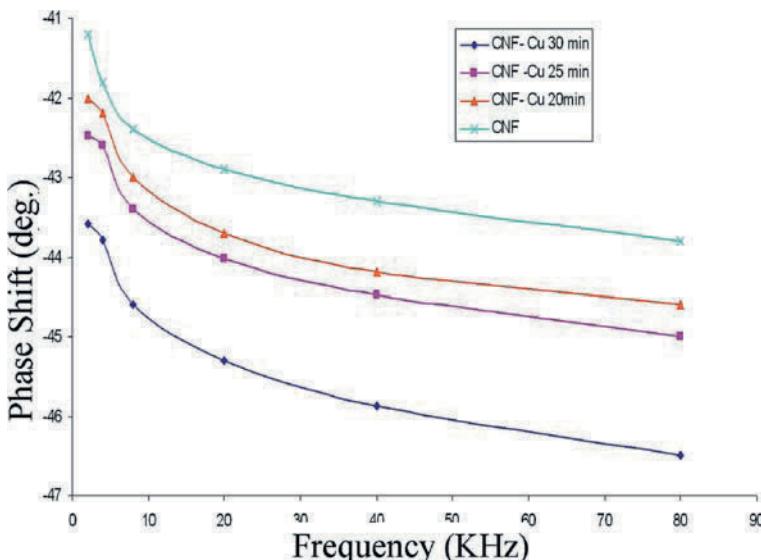


Fig. 20 Phase shift measurement results for CNT–Cu sample under different time period of Cu bath

the smallest thermal resistance, which means the CNFs are making good contact with the Ni hot surface by filling the air gaps and transfer the heat efficiently away. By depositing Cu into the gap between CNFs, the number of CNF protruding out of the Cu matrix is reduced, the length of the CNFs array are getting shorter, in which case, the compliance nature of CNF is strongly restricted and not be able to make good contact with the contacting surface. Furthermore, from Fig. 6.13, one could observe the rough ECD Cu surface, which is also counted for the worse thermal properties of the composite.

However, although there is a relative flat ECD Cu surface after 30-min deposition, the measured total thermal resistance for the CNF–Cu with 30 min deposition possessed the largest thermal resistance ($0.32 \pm 0.01 \text{ cm}^2 \text{ K/W}$). Clearly, there is no phase changing properties for Cu, although the thermal conductivity of Cu is excellent, it does not make good contact with the hot surface compared with CNF. The phase shift would still allow us to calculate the total thermal resistance of the CNF–Cu composite. However, major part is dominated by the interface resistance. In this case, it would not be very helpful to distinguish the interface resistance with the bulk resistance of the CNF–Cu composite.

The relationship between the applied forces and CNF–Cu thermal performance as TIMs is also investigated. In our experiment, a load cell is placed against a Cu substrate to measure the forces applied on the CNF composite samples. The applied pressure is increased gradually and the total effective thermal resistance is calculated corresponding to phase delay of the CNFs composites sample (Fig. 20). For each point in this figure, the phase delay signal is acquired for the entire frequency region, and the corresponding thermal resistance is calculated. One can see that the total effective thermal resistance including the interface resistance is reduced and reached its steady-state value at a pressure of 80 psi. With high pressure applied on the CNF and CNF composites, it ensures that better contact between CNF arrays; CNF filling materials and the Ni film make the bulk resistance to be dominating the total thermal resistance. As a result, the heat transfer from the Ni film to the Cu substrate is mostly affected by the CNF array composite, and makes it possible to measure the thermo-physical properties of CNF array composite. If no pressure or low pressure is applied on the ZnSe, the CNF composite will have poor contact with the Ni film due to the surface roughness, i.e., less number of CNF and filling materials would be in contact with Ni film, leading to a large contact resistance between CNF arrays and the Ni film. Thus the compliance properties of CNFs array are very important for utilizing it as TIMs.

10.2 Bonding Strength of CNTs to Its Substrate

It goes without saying that satisfactory *adhesion* of the CNF array (CNFA) to its substrate is critical for making the CNF-based technology practical. Accordingly, the analysis that follows is aimed at emphasizing importance of CNF adhesion to its substrate as TIM and also the easy-to-use and effective experimental method for the evaluation of such an adhesion. Note that the adhesion strength of a single CNF to

its substrate was addressed qualitatively, apparently for the first time, by Cui et al. [38] and by Chen et al. [39]. This was done in connection with the use of PECVD-synthesized CNFs as probe tips in the atomic force microscopy imaging equipment. In the experiments described by these investigators, individual CNFs were directly grown on tipless cantilevers. In the reported observation, the CNF probe was operated on a continuous scan mode for 8 h, and no degradation in image resolution has been observed. The CNFA in our tests, however, consisted of billions of CNFs. The obtained information characterizes therefore the performance of an ensemble of a plurality of CNFs. In order to translate the obtained experimental data into the corresponding shearing stresses, we developed a simple analytical stress model that enables one to calculate the magnitude and the distribution of the interfacial shearing stress from the measured (given) external force. Testing has been conducted using specially designed shear-off test specimens.

In the analysis we determine the maximum effective *shear stress-at-failure* for a CNFA fabricated on a thick Cu substrate. We use the term “effective” to emphasize that we address a plurality of CNFs and treat the CNFA “brush” as a sort of a continuous bonding layer. This approach enabled us to use an analytical stress model that we developed for the evaluation of the interfacial shearing stress in an assembly with a continuous bonding layer. The model used in our analysis is a modification of the models that were suggested earlier for the evaluation of the interfacial thermally induced stresses in thermally mismatched assemblies [40–43]. The samples in experiments were also synthesized using plasma enhanced chemical vapor deposition (PECVD).

A $20 \times 20 \times 1$ mm copper plate was used in the experimental specimens as a central block. A 20×10 mm copper foil was placed on, and then pressed onto the central block, so that the edges of the foil had the same boundary as the central copper block. Two identical CNTA samples grown on $20 \times 20 \times 1$ mm copper substrates were prepared and pressed against the copper foil. Only a portion of the surface of the CNFA was actually in contact with the copper foil and was able to provide adhesion. The test specimen was placed on the Instron’s lower anvil. Then the upper anvil with the load cell was lowered to approach the copper central block as shown in Fig. 21. Once the load cell got into contact with the central block, the Instron tester, with the Capacitec equipment mounted on it, started to record the forces and the corresponding displacements.

The copper foil used in our experiments was strong enough to hold in place the top surface of the CNFA. The thickness of the CNFA “brush” (the lengths of the CNFs) was about $30 \mu\text{m}$, i.e., considerably thicker/longer than the thickness of the copper foil, so that the CNFs were able to deform in a beam-like fashion, while the CNFA “brush” as a whole experienced shear deformations.

The experimental data obtained for five CNTF test specimens are shown in Fig. 19. The force applied to the sample gradually increases until the interfacial failure occurs, and the CNF is sheared off from the substrate. If the test specimen were ideally well assembled, both the CNFA samples would have been simultaneously separated from the central block. In reality, however, only one of CNFA interfaces was typically broken.

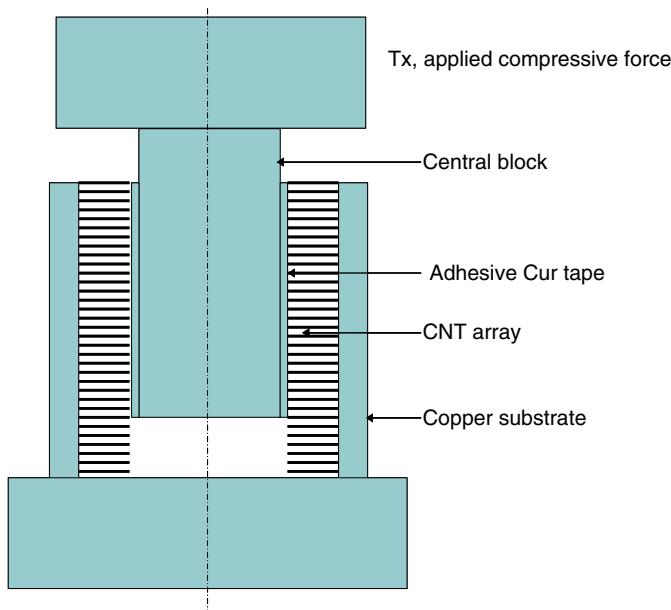


Fig. 21 Schematics of the test specimen assembly for the shear stress test

The force-at-failure varied in our tests appreciably from specimen to specimen. The most important source of uncertainty in our tests is the actual percentage of the contact area between the CNFs in the copper foil. If the CNFA is not attached well to the copper foil, the CNFs in the CNFA will not be deformed and would not be sheared off from the substrate (Fig. 23). Clearly, with more CNFs in contact with the copper foil, the force holding the CNFA on its substrate would be larger and, as a result, the induced shear-off stress would be higher. As one could see from Fig. 22, the maximum force-at-failure varied in our experiments from 2 to 5 kg, depending on the CNFA bonding strength and on how many CNFs were in contact with the copper foil. The failure site was observed under a SEM, after the test samples were taken out from the Instron tester. One could clearly see that about 80% of CNFs were in contact with the copper foil. This case corresponded to the 5 kg force value. Unfortunately, there was no good way to quantitatively correlate the percentage of the CNTs in contact with the copper foil and the corresponding shear-off stress. This stress was calculated from the measured shearing force on the basis of the formula (see Appendix):

$$\tau(x) = \tau_{\max} e^{-kx}$$

where

$$\tau_{\max} = kT$$

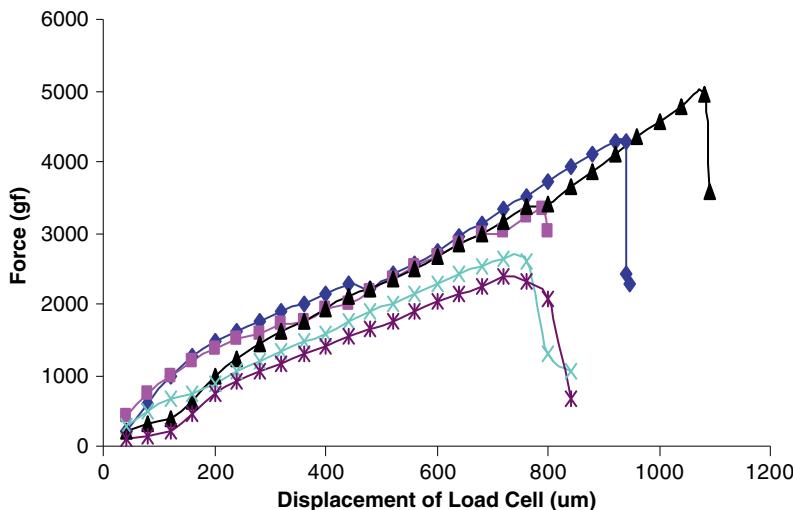


Fig. 22 The relationship between the forces applied on the surface of the vertical aligned CNF array. With different amount of CNFs in contact with the copper tape, the maximum breaking forces for the samples are different. The displacement is actually the displacement of the load cell pushed up by the CNFs

is the maximum shearing stress, T is the measured force, and k is the factor of the interfacial shearing stress [40–43]. The origin of the coordinate, x , is at the assembly end, at which the external force was applied. The calculated shear-off stress was 300 psi for the case of a 5 kg force-at-failure. Note that only about half of the sample was in contact with the copper foil. SEM images were subsequently taken for the tested samples. Figure 23a shows the top view for the CNF samples prior to the shear-off testing. The density of the CNFA was about $10^9/\text{cm}^2$, the average diameter of the CNFs was about 80~100 nm, the lengths of the particular CNFs in the CNTA had a variation of about 3 μm . This would cause a small percentage of CNFs that would not be in contact with the copper foil. After the shear-off testing, part of the CNFs was sheared off completely from the substrate. With a close look at the substrate, one could observe a ring-shaped bottom part of the remaining CNFs. Only a few very short CNFs, which did not have contact with the copper foil, still remained on the substrate. The tilt view at the CNF roots of the CNTs remained on the substrate (Fig. 23c) which shows that the roots of the CNFs have graphitic walls with hollow cores at the center.

Although the PECVD-synthesized CNFs have a typical bamboo structure, they still have a partially hollow core structure, as CNFs synthesized by other methods do. In Fig. 23d, one can see sheared-off CNFs on the copper foil. The close end of these CNTs has a bamboo-like structure, which resembles the roots showed in Fig. 23c. The most likely breaking point takes place at the connection part of the bamboo-like CNFs.

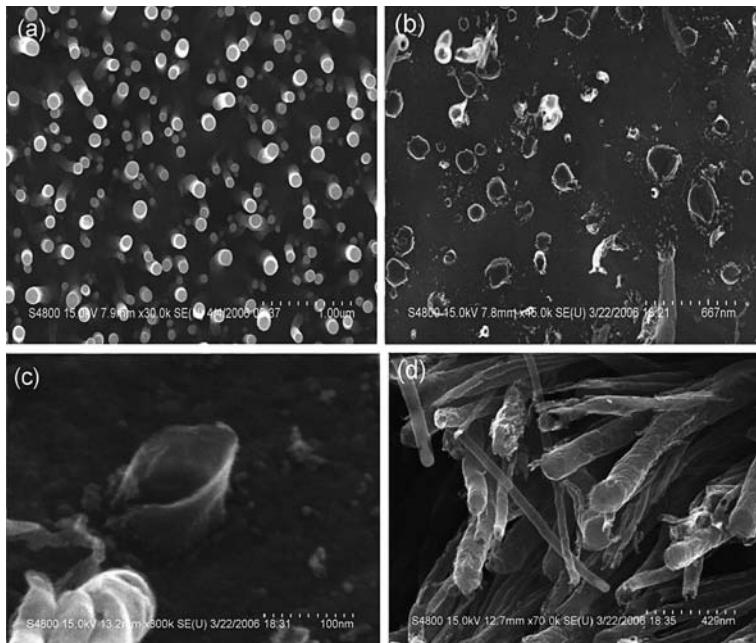


Fig. 23 SEM images reveal the *top* view of the CNFs array before (**a**) and after (**b**) the shear-off test. Most of the CNFs are sheared off from the substrate. (**c**) is the *tilt* view of the CNF left on the original substrate, it reveals that our CNFs have the real hollow center and graphitic outside walls. (**d**) shows the CNFs sheared off from the substrate and stay on the copper tape

11 Conclusions

We have addressed some of the existing methods that have been developed and applied for the characterization of the physical/mechanical properties of CNTs and CNFs. We have particularly addressed those methods and properties that have to do with the thermal performance of these structural elements as parts of the new generation of TIMs.

Appendix

Calculated Interfacial Shearing Stress from the Measured Shearing Force in a Bi-material Assembly

Let a bi-material assembly be subjected to a shearing force, \hat{T} . The objective of the analysis that follows is to evaluate the magnitude and the distribution of the interfacial shearing stress, $\tau(x)$, from the measured force, \hat{T} . We proceed from the following approximated formulas for the interfacial longitudinal displacements for the assembly components #1 and #2:

$$\begin{aligned} u_1(x) &= -\lambda_1 \int_0^x T(\xi) d\xi + \kappa_1 \tau(x) \\ u_2(x) &= \lambda_2 \int_0^x T(\xi) d\xi - \kappa_2 \tau(x) \end{aligned} \quad (1)$$

where

$$T(x) = \int_{-l}^x \tau(\xi) d\xi \quad (2)$$

is the force acting in the cross-section, x , of the assembly, $\tau(x)$ is the shearing stress acting in the same cross-section,

$$\lambda_1 = \frac{1 - \nu_1}{E_1 h_1}, \lambda_2 = \frac{1 - \nu_2}{E_2 h_2} \quad (3)$$

are the axial compliances of the assembly components, h_1 and h_2 are their thicknesses, E_1 and E_2 are Young's moduli of the component materials, ν_1 and ν_2 are Poisson's ratios,

$$\kappa_1 = \frac{h_1}{3G_1}, \kappa_2 = \frac{h_2}{3G_2} \quad (4)$$

are the interfacial compliances of the assembly components,

$$G_1 = \frac{E_1}{2(1 + \nu_1)}, G_2 = \frac{E_2}{2(1 + \nu_2)} \quad (5)$$

are the shear moduli of the component materials, and l is the assembly length. The origin of the coordinate x is at the cross-section where the force \hat{T} is applied. The first terms in the right parts of the formulas (1) are due to the forces $T(x)$ and are evaluated based on the Hooke's law. The second terms account for the fact that the interfacial shearing displacements are somewhat larger than the displacements of the inner points of the given cross-section.

The compatibility condition for the displacements (1) can be written as

$$u_1(x) = u_2(x) - \kappa_0 \tau(x) \quad (6)$$

where

$$\kappa_0 = \frac{h_0}{\varepsilon_0} \quad (7)$$

is the interfacial compliances of the bonding layer,

$$G_0 = \frac{E_0}{2(1 + \nu_0)} \quad (8)$$

is the shear modulus of the bonding material, and E_0 and ν_0 are the elastic constants of this material.

Introducing the formulas (1) into the displacement compatibility conditions (6), we have

$$-\lambda \int_0^x T(\xi) d\xi + \kappa \tau(x) = 0 \quad (9)$$

where the following notation is used:

$$\lambda = \lambda_1 + \lambda_2, \kappa = \kappa_0 + \kappa_1 + \kappa_2 \quad (10)$$

Here λ is the total axial compliance of the assembly, and k is its total interfacial compliance. From (9) we find, by differentiation:

$$\tau'(x) - k^2 T(x) = 0 \quad (11)$$

where

$$k = \sqrt{\frac{\lambda}{\kappa}} \quad (12)$$

is the parameter of the interfacial shearing stress.

The obvious boundary conditions for the force $T(x)$ are

$$T(0) = \hat{T}, T(l) = 0 \quad (13)$$

Then the equation (11) results in the following boundary conditions for the interfacial shearing stress:

$$\tau'(0) = k^2 \hat{T}, \tau'(l) = 0 \quad (14)$$

The next differentiation of the equation (11), considering the formula (2), yields

$$\tau''(x) - k^2 \tau(x) = 0 \quad (15)$$

This equation has the following solution:

$$\tau(x) = C_1 \sinh kx + C_2 \cosh kx \quad (16)$$

Here C_1 and C_2 are constants of integration. Introducing the sought solution (16) into the boundary conditions (14), we conclude that these constants are expressed as

$$C_1 = k\hat{T}, C_2 = -k\hat{T} \coth kl \quad (17)$$

so that the solution (16) leads to the following simple formula for the distributed shearing stress:

$$\tau(x) = -k\hat{T} \frac{\cosh [k(l-x)]}{\sinh kl} \quad (18)$$

After introducing this formula into the formula (2), we obtain the following expression for the distributed force, $T(x)$:

$$T(x) = \hat{T} \frac{\tanh [k(l-x)]}{\sinh kl} \quad (19)$$

One could easily check that the conditions (13) are fulfilled for this force. The solution (18) yields

$$\tau(0) = -k\hat{T} \coth kl, \tau(l) = -\frac{k\hat{T}}{\sinh kl} \quad (20)$$

These formulas indicate that for sufficiently long (large l value) and/or stiff (large k value) assembles, when the product kl is, say $kl = 2.5$, the stress $\tau(l)$ is next-to-zero, and the stress $\tau(0)$ at the virgin is

$$\tau(0) = \tau_{\max} = -k\hat{T} \quad (21)$$

For large enough kl values, the formula (18) can be simplified:

$$\tau(x) = -k\hat{T}e^{-kx} \quad (22)$$

This result indicates that the interfacial shearing stress is the largest at the edge, where the force is applied, and decreases exponentially with an increase in the distance x from this edge. The sign “-” in front of the formulas (18) and (22) indicates that the direction of the interfacial shearing stress is opposite to the direction of the external shearing force \hat{T} .

References

1. S. Iijima, Nature 354, 56, 1991.
2. Q. Ngo, BA Cruden, AM Cassell, G Sims, M Meyyappan, J Li, and CY Yang, Nano Lett. 4, 2403, 2004.
3. J. Xu and TS Fisher, Int. J. Heat Mass Transfer 49, 1658–1666, 2006.
4. X.W. Wang, Z Zhong, and J Xu, J. Appl. Phys. 97, 064302, 2005.
5. Y. Xu, Y Zhang, and XW Wang, J. Appl. Phys. 100, 074302, 2006.
6. R.S. Ruoff and DC Loretz, Carbon 33, 925, 1995.
7. S. Berber, Y Kwon, and D Tománek, Phys. Rev. Lett. 84, 4613, 2000.
8. S. Maruyama, Microscale Thermophys. Eng. 7, 41, 2003.

9. M.A. Osman and D Srivastava, *Nanotechnology* 12, 21, 2001.
10. O. Lourie and HD Wagner, *JMR* 13(9), 2418–2422, 1998.
11. M.M.J. Treacy, TW Ebbesen, and JM Gibson, *Nature* 381, 678, 1996.
12. N. Yao and V Lordi, “Young’s Modulus of Single Walled Carbon Nanotubes”, *J. Appl. Phys.* 84, 1939, 1998.
13. E. Hernandez, C Gose, P Bernier, and A Rubio, “Elastic Properties of c and bxcyzn Composite Nanotubes”, *Phys. Rev. Lett.* 80, 4502, 1998.
14. A. Krishnan, E Dujardin, TW Ebbesen, PN Yianilos, and MMJ Treacy, “Young’s Modulus of Single Walled Nanotubes”, *Phys. Rev. B* 58, 14013, 1998.
15. P. Poncharal, ZL Wang, D Ugarte, and WA de Heer, *Science* 283, 1513, 1999.
16. J. Gaillard, M Skove, and AM Rao, “Mechanical Properties of Chemical Vapor Deposition-Grown Multiwalled Carbon Nanotubes”, *Appl. Phys. Lett.* 86, 233109, 2005.
17. C. Wei and D Srivastava, “Nanomechanics of Carbon Nanofibers: Structural And Elastic Properties”, *Appl. Phys. Lett.* 85, 2208, 2004.
18. J.P. Salvetat et al. “Elastic Modulus of Ordered and Disordered Multi-Walled Carbon Nanotubes”, *Adv. Mater.* 11, 161, 1999.
19. A. Cao, P Dickrell, WG Sawyer, M Ghasemi-Nejhad, and P Ajayan, “Super Compressible Foam like Carbon Nanotube Film”, *Science* 310, 1307, 2005.
20. X. Zhou, JJ Zhou, and ZC Ou-Yang, “Strain Energy and Young’s Modulus of Single-Wall Carbon Nanotubes Calculated From Electronic Energy-Band Theory”, *Phys. Rev. B* 62, 13692, 2000.
21. C.Q. Ru, “Effective Bending Stiffness of Carbon Nanotubes”, *Phys. Rev. B* 62, 9973, 2000.
22. M.F. Yu, O Lourie, MJ Dyer, K Moloni, TF Kelly, and RS Ruoff, *Science* 287, 637, 2000.
23. T. Kuzumaki and Y Mistuda, *Appl. Phys. Lett.* 85, 1250, 2004.
24. M. Hishio, S Akita, and Y Nakayama, *Jpn. J. Appl. Phys.* 44, L1097, 2005.
25. H.D. Wagner, O Lourie, Y Feldman, and R Tenne, *Appl. Phys. Lett.* 72, 188, 1998.
26. D.A. Walters, LM Ericson, MJ Casavant, J Liu, DT Colbert, KA Smith, and RE Smalley, *Appl. Phys. Lett.* 74, 3803, 1999.
27. M. Yu, O Lourie, M Dyer, K Moloni, T Kelly, and RS Ruoff, *Science* 287, 637, 2000.
28. Y. Zhang, E Suhir, Y Xu, and C Gu *JMR* 21(11), Nov. 2006.
29. Y. Zhang, Y Xu, C Gu, and E Suhir, “Predicted Shear-off Stress in Bonded Assemblies: Review and Extension”, *IEEE COMT ASTR Workshop*, 2006, San Francisco, CA
30. J. Li, C Papadopoulos, JM Xu, and M Moskovits, *Appl. Phys. Lett.* 75, 367, 1999.
31. V.M. Harik, *Solid State Comm.* 120, 331, 2001.
32. S.P. Timoshenko and J Gere, “Theory of Elastic Stability”, McGraw-Hill, New York, 1988.
33. B.A. Cruden, AM Cassell, Q Ye, and M Meyyappan, *J. Appl. Phys.* 94, 4070, 2003.
34. J. Li, R Steven, L Delzeit, HT Ng, A Cassell, J Han, and M Meyyappan, *Appl. Phys. Lett.* 81, 910, 2002.
35. E. Suhir, “Structural Analysis in Microelectronic and Fiber Optic Systems: Basic Principles of Engineering Elasticity and Fundamentals of Structural Analysis”, Van Nostrand Reinhold, New York, 1991.
36. Y. Zhang, Y Xu, and E Suhir, *J. Phys. D: Appl. Phys.* 39, 4878, 2006.
37. Y. Zhang, Y Xu, and E Suhir *JMR* 21(11), Nov. 2006.
38. H. Cui, SV Kalinin, X Yang, and DH Lowndes, *Nano Lett.* 4, 2157, 2004.
39. I.C. Chen, LH Chen, XR Ye, C Daraio, S Jin, CA Orme, A Quist, and R Lal., *Appl. Phys. Lett.* 88, 153102, 2006.
40. E. Suhir, “Interfacial Stresses in Bi-Metal Thermostats”, *ASME J. Appl. Mech.* 56, 595, 1989.
41. E. Suhir, “Thermomechanical Stress Modeling in Microelectronics and Photonics”, *Electronic Cooling* 7(4) 2001.
42. E. Suhir, “Thermal Stress in an Adhesively Bonded Joint with a Low Modulus Adhesive Layer at the Ends”, *J. Appl. Phys.* 93, 3657, 2003.
43. E. Suhir, “Analysis of Interfacial Thermal Stresses in a Tri-Material Assembly”, *J. App. Phys.* 89, 3685, 2001.

On-Chip Thermal Management and Hot-Spot Remediation

Avram Bar-Cohen and Peng Wang

Abstract The rapid emergence of nanoelectronics, with the consequent rise in transistor density and switching speed, has led to a steep increase in die heat flux and growing concern over the emergence of on-chip “hot spots.” The application of on-chip high heat flux cooling techniques provides a viable direction for the thermal management of such nanoelectronic components. Following a review of the relevant passive and active thermal management techniques, the physical phenomena underpinning the most promising on-chip thermal management approaches are described. Attention is devoted to thin-film and miniaturized thermoelectric coolers, orthotropic TIMs/heat spreaders, and phase-change microgap coolers for hot-spot remediation and thermal management of these nanoelectronic chips.

Keywords Nanoelectronics · Hot spots · Electronic cooling · Thermal management · Thermoelectric coolers · Microcoolers · Thermal interface materials (TIM)

1 Introduction

The Moore’s law progression in semiconductor technology, leading to shrinking feature size, increasing transistor density, faster circuit speeds, and higher chip performance, continues unabated. As shown in Fig. 1, the 2005 International Technology Roadmap for Semiconductors (ITRS) [1], predicts a continuous decrease in transistor size, which can be expected to lead to functioning 20 nm devices at more than a billion transistors per square centimeter ($10^9/\text{cm}^2$) by 2010 and to continue down to 10 nm, along with a rise in transistor density toward 10 billion transistors/ cm^2 ($10^{10}/\text{cm}^2$), by 2018. These changes in semiconductor technology can be expected

A. Bar-Cohen (✉)

Department of Mechanical Engineering, University of Maryland, College Park, MD, USA
e-mail: abc@umd.edu

to lead to ever faster and more computationally complex chips. Moreover, despite the extensive efforts that have been made to reduce both capacitance and operating voltage on the chip, the chip power dissipation – at the very cutting edge of the technology – continues to rise. In the absence of more aggressive thermal control techniques, the elevated temperatures resulting from higher power dissipation can be expected to decrease CMOS transistor switching speeds and accelerate some key failure mechanisms, thus compromising the performance and reliability of such advanced semiconductor devices.

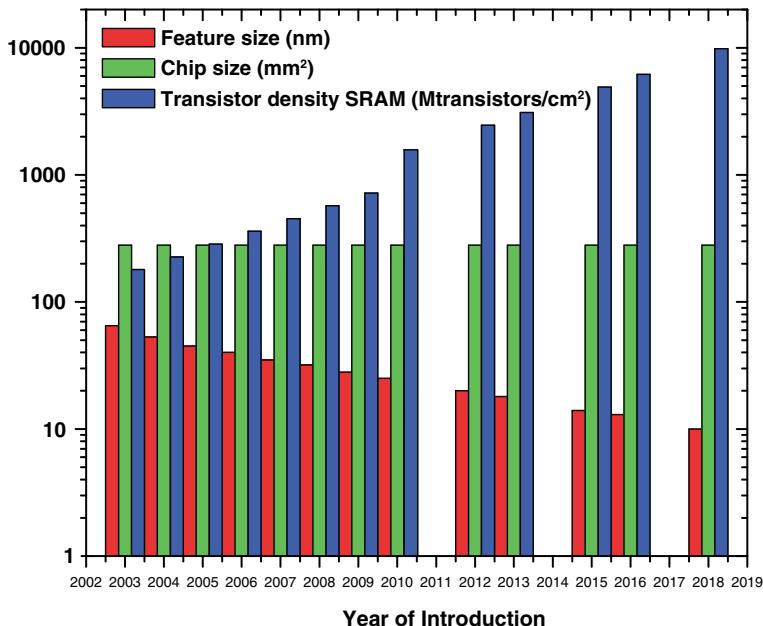


Fig. 1 The 2005 ITRS predictions of feature size, chip size, and transistor density for high-performance microprocessor chips [1]

The Moore's law-driven transition to the widespread use of nanoscaled electronics poses three specific thermal management challenges, which have motivated significant recent activity in microprocessor cooling [2–7]. Foremost among these is the drive to improve component speed, which has motivated circuit designers to compress the “core” of the microprocessor to an ever smaller size. Along with the reduced “time-of-flight” between transistors, this spatial compression leads disproportionately to high heat flux in the “core” areas of the silicon chip. Second, higher interconnect current densities, due to higher chip power, and the extreme interconnect aspect ratios, resulting from the high transistor densities, result in rapidly increasing interconnect temperatures in today's IC technology. The problem is aggravated by the trend to replace the SiO₂ interconnect passivation layer with

lower dielectric constant materials, such as novel organic and porous dielectrics, which also possess lower thermal conductivity and greatly impede the conduction of heat away from the interconnect and transistor. Third, the use of novel nanoscaled electronics technologies, e.g., narrow channels, further aggravates the local temperature rise around individual transistors. Figure 2 shows the temperature contour around a single nanoscaled transistor and indicates there is nanometer scale hot spot in the transistor drain region [8].

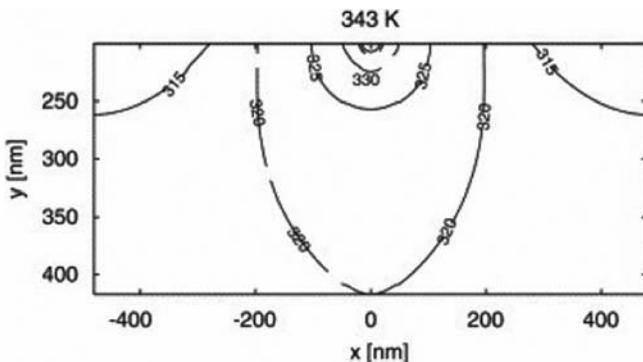


Fig. 2 The temperature field from a device-like hot spot in bulk 90 nm silicon transistor [8]

Under the combined influence of these trends, chip power dissipation and heat flux are expected to increase further over the next decade. According to the International Electronics Manufacturing Initiative Technology Roadmap (iNEMI 2006) [9], the maximum chip power dissipation is projected to be 510 W and the maximum chip heat flux to be 300 W/cm^2 for the high-power automotive devices in 2015, as indicated in Fig. 3.

Moreover, as seen in Fig. 4, these advanced nanoelectronic chips are characterized by substantial nonuniformities in power dissipation, resulting in localized, high heat flux “hot spots” with a large on-chip temperature gradient [3]. Because chip thermal management must ensure that all junction temperatures in the microprocessor do not exceed an application-driven maximum temperature, typically in the range of 90–110°C, it is often these hot spots, not the entire chip power dissipation, that drive the thermal design. This leads to two undesirable consequences: (1) nonuniform heat generation limits the total heat dissipation that can be managed by a conventional thermal solution and, thus, a much more aggressive thermal solution than would be required for uniform heating, is necessary and (2) the focus on controlling the temperature of the hot spot can lead to over-design of the microprocessor cooling solution. As shown in Fig. 5, the 2006 iNEMI Roadmap predicted that desktop PC second-level thermal solution requirements for a uniformly powered chip will become more demanding, pushing the limits of current air-cooled systems. Once the nonuniform distribution of power across the chip is considered,

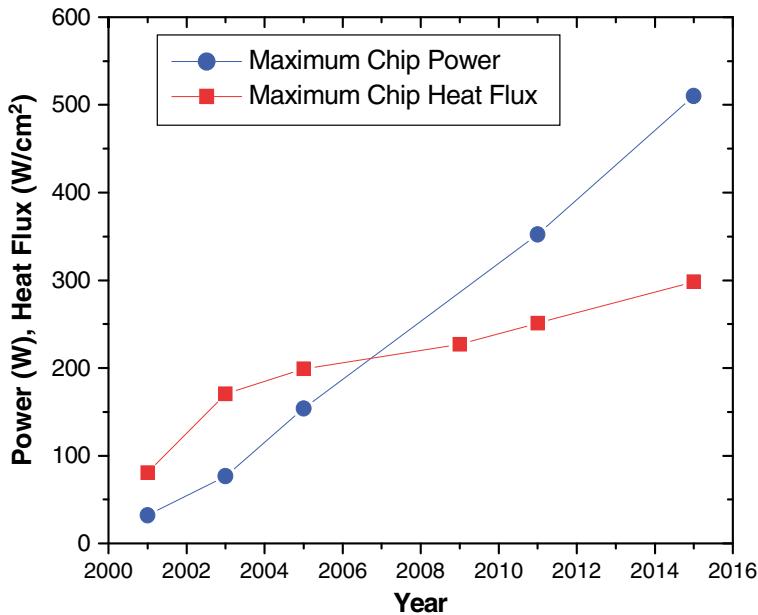


Fig. 3 The 2006 iNEMI prediction of chip power dissipation and heat flux for the highest power automotive devices [9]

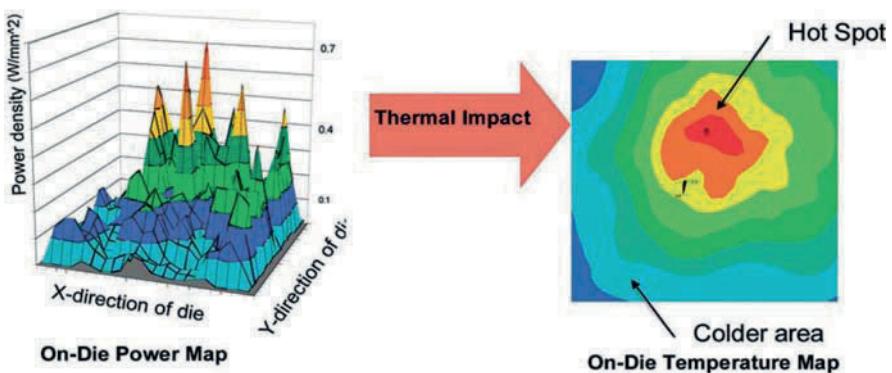


Fig. 4 Schematic illustrating typical die power map

air-cooling system will have to be replaced with more aggressive cooling solutions. Due to its complexity, on-chip hot-spot cooling has become one of the most active and challenging research areas in thermal management of electronic devices and packages.

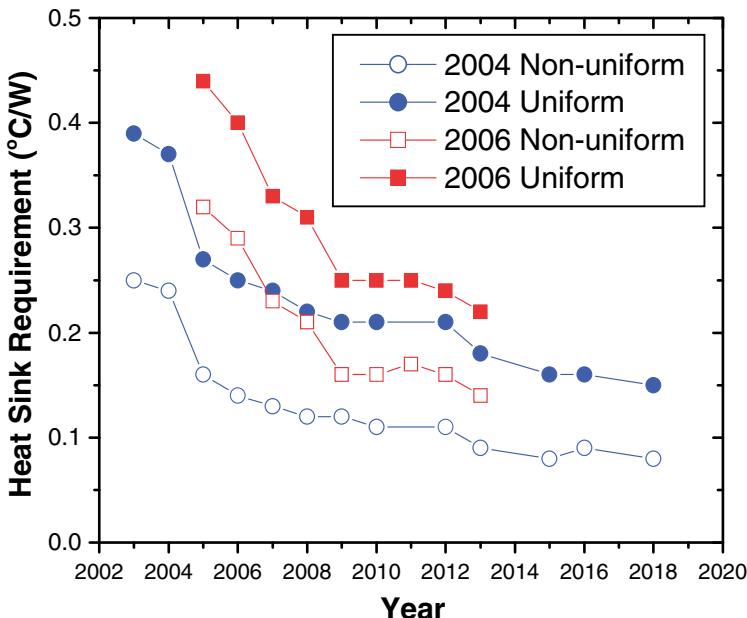


Fig. 5 The 2006 iNEMI predictions of desktop PC second-level thermal solution requirement

1.1 Potential Hot-Spot Cooling Solutions

Thermal management for high-flux electronic silicon chips can be classified into two strategies: passive cooling and active cooling, both of which have continued to be extensively studied during the past few years.

1.1.1 Passive Cooling Solutions

Passive cooling solutions are those that do not have moving parts and generally require no external electric power to activate or assist thermal transport. These techniques mainly rely on heat spreading in high-conductivity materials, such as bonded spreaders or diamond coatings on the silicon chip, and/or on vapor transport along with evaporation and condensation in tubes and channels, to transport the heat from the high-flux regions to the areas of lower heat flux. Due to the high thermal conductivity of silicon ($\sim 150 \text{ W/m K}$), only modest spreading improvements can be obtained from the use of traditional spreading materials, such as copper (390 W/m K), beryllia (250 W/m K), aluminum nitride (220 W/m K), or other composites [10]. Alternatively, diamond is a very attractive material for passive cooling of high-flux regions on a chip because it has the highest thermal conductivity of any known materials and also has a very high electric resistance ($\sim 10^8 \Omega\text{m}$). The

thermal conductivity is in the range of 1500–2100 W/m K for single-crystal diamond fabricated by the high-pressure synthesis method and 500–1300 W/m K for polycrystalline diamond fabricated by CVD low-pressure synthesis [11]. Deposition of diamond on substrates, such as silicon, is a reasonably mature technology and there are now multiple techniques that provide high-quality single-crystal or polycrystalline diamond films. The fabrication of a diamond heat-spreading layer on silicon's active region includes direct growth of diamond on the silicon substrate or bonding of a polished diamond film onto the silicon substrate. Figure 6 shows an example of diamond deposited directly on the aluminum metallization layers of a silicon chip. However, the contamination with impurities that may occur in the silicon wafer during diamond deposition, associated with the carbon, nitrogen, oxygen, hydrogen, and other elements diffusing into the device wafer from reactive gases, have – thus far – kept diamond deposition from becoming the technique of choice. To avoid these difficulties, diamond is often bonded to silicon using a gold–tin eutectic alloy solder film. However, the thermal contact resistance at the silicon substrate/diamond interface and metallization layers/diamond interface and poor adhesion of diamond to the metallization layer, aggravated by cyclic thermally induced stress variations, has restricted the application of diamond as a reliable heat spreader.

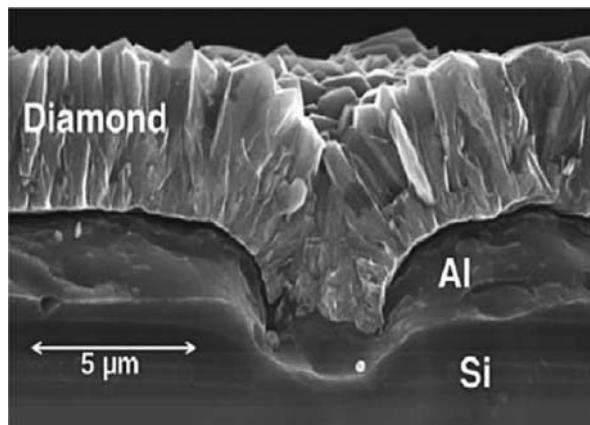


Fig. 6 Diamond deposited on aluminum metallization in a silicon substrate using a microwave plasma technique [12]

Since the thermal conductivity of silicon is relatively high, while the heat-transfer resistance through the thickness is relatively small, a heat pipe micromachined in the silicon substrate or a flat-plate heat pipe attached to the silicon chip is advantageous in providing in-plane spreading of the heat released by localized regions of high heat flux [13, 14]. A heat pipe is a passive heat-transfer device with an extremely high effective thermal conductivity, resulting from the evaporation and

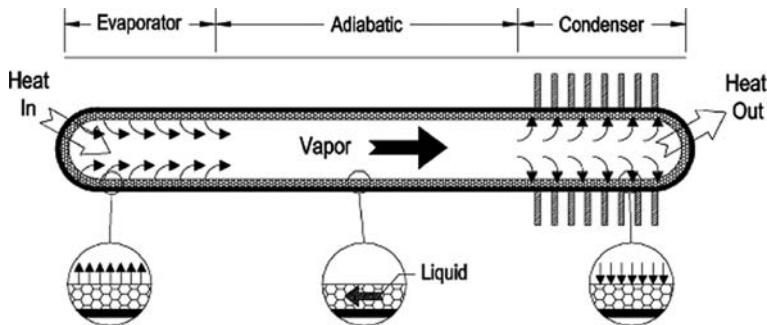


Fig. 7 Structure of typical wicked heat pipe

condensation of a suitable working fluid within an evacuated, hermetically sealed enclosure. Figure 7 is a side view of a heat pipe showing the wick and the vapor/liquid flow characteristics.

In the past 10 years, extensive research has demonstrated that heat-spreading performance comparable to that of diamond substrates can be obtained [15–17]. In addition, the factors that affect the performance of micro or miniature heat pipes such as the frictional vapor flow, the vapor space, the shape and dimensions of the microchannels, the operating temperature, the disjoining pressure, and evaporating heat resistance have been systematically investigated. These investigations have provided insight into the design of highly efficient heat pipes, which can meet the new requirements of high heat flux cooling. Adkins at Sandia National Laboratories, Albuquerque, NM, investigated a heat-pipe heat spreader embedded in a silicon substrate as an alternative to the conductive cooling of integrated circuits using diamond films. In their design miniaturized heat pipes, created by $35\text{ }\mu\text{m}$ grooves in the silicon wafer, function as highly efficient heat spreaders, collecting heat from the localized hot spots and dissipating the heat over the entire chip surface, with an effective thermal conductivity of at least 800 W/m K [18, 19].

Recently research has been focused on achieving further increases in performance by using thermally driven pulsating two-phase flows [20, 21], new capillary structures [22], and MEMS-based heat pipes [23, 24]. Plesch et al. reported that flat miniature heat pipes with axial grooves, using water as the working fluid, were capable of sustaining heat fluxes on the order of 40 W/cm^2 [25]. Ma et al. fabricated the heat pipes with microscaled sintered powder wicks, which can remove heat fluxes of up to 80 W/cm^2 without any sign of evaporator dry out [26]. Gillot fabricated flat miniature heat pipes with microcapillary grooves inside a silicon substrate and the heat removal capability was reported to be 110 W/cm^2 [27]. Using miniaturized heat pipes Lin et al. reported that a cooling heat flux of 140 W/cm^2 was achieved [28]. While heat pipes possess certain inherent advantages and current research may raise the observed thermal performance limits, heat removal rates of the currently available heat pipes are generally still not suitable for controlling the temperature of the most severe, high-flux hot spots on the chip.

1.1.2 Active Cooling Solutions

Active cooling solutions usually involve moving parts and require the input of electric energy for their operation. The most common active cooling solution is air-cooled, forced convection heat sinks, which have been long used for a wide range of electronic equipment, including office and desktop computers. However, this conventional active cooling method has very limited capability for dealing with high-flux zones on microelectronic chips, due to its inherently low heat-transfer coefficients. Compared to air cooling, the use of liquid coolants has many advantages such as high thermal conductivity, high specific heat, low viscosity, and high latent heat of evaporation for two-phase application. As a result, different active liquid-cooling technologies, such as microchannel heat sinks and direct jet impingement, have been developed due to the higher heat-transfer coefficient and high cooling flux achieved as compared to air-cooling heat sink.

With microfabrication techniques developed by the electronics industry, it is possible to manufacture microscaled three-dimensional structures, typically consisting of closely spaced parallel channels with rectangular, trapezoidal, or triangular cross sections and hydraulic diameters ranging from 100 to 1000 μm (see Fig. 8). Such microchannel “compact heat exchangers” may be fabricated in the chip itself or in the heat sink to which a chip or array of chips is attached. Microchannel heat sinks

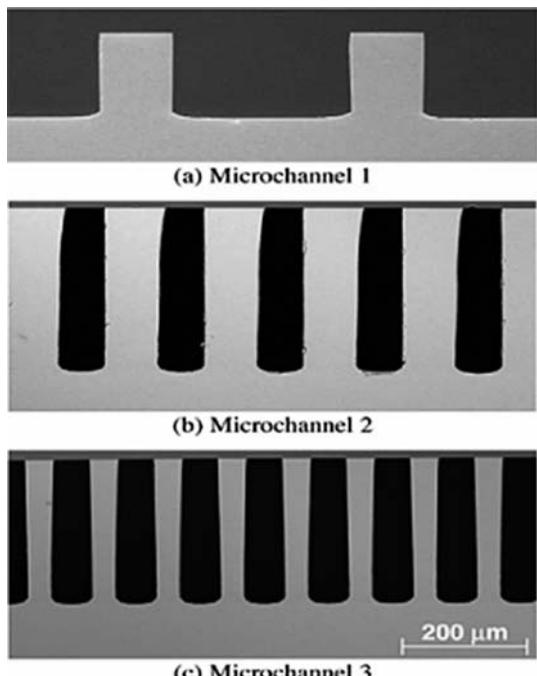


Fig. 8 SEM photos of cross sections of three different microchannels [30]

can be used either with single-phase flow, where heat is transferred from the electronic chip via sensible heat gain to the coolant, or with two-phase flow, which also utilizes the latent heat of the coolant during liquid/vapor phase change. In both thermal transport modes, microchannel coolers are capable of significantly increasing the wetted surface area and of thinning the thermal boundary layers, thus achieving very high heat-transfer rates. Single-phase liquid-cooling systems utilize a pump to actively circulate the liquid to the microchannels and have been studied for many years. In their 1981 pioneering work Tuckerman and Pease used single-phase convection with water flowing through a microchanneled silicon chip to demonstrate heat removal of 790 W/cm^2 with a flow rate of $1 \times 10^{-5} \text{ m}^3/\text{s}$ and the pressure drop below $3.45 \times 10^5 \text{ Pa}$ [29]. Although the high heat flux capability is promising, and considerable pressure reductions and improvements in flow distribution have been achieved in recent years, the integration of such a microchannel cooler into a closed-loop plumbing/pumping system and justifying the required pumping power has proven challenging. Moreover, the industry's resistance to flowing liquids directly through active microchanneled chips has necessitated the use of thermal interface materials for the attachment of the microchannel coolers and considerably raised the thermal resistance of this thermal management approach.

Two-phase microchannel cooling, exploiting the latent heat of the coolant to reduce the flow rates and pumping power requirements and the high heat-transfer coefficients associated with boiling, has received growing attention, focusing on reductions in pumping power and dealing with the performance limitations posed by flow instabilities and maldistribution. Bowers et al. showed that two-phase microchannels, operating with water flow rates of less than $1.08 \times 10^{-6} \text{ m}^3/\text{s}$ and pressure drops of $3.45 \times 10^5 \text{ Pa}$ could remove more than 200 W/cm^2 [31]. Mudawar also demonstrated a heat removal rate of 361 W/cm^2 with two-phase forced convective cooling on an enhanced surface with FC-72 as the dielectric coolant [32]. However, in a recent study of water-cooled two-phase microchannel coolers, Prasher found the biggest challenge to the use of two-phase microchannel coolers in hot-spot cooling is the nonuniformity in flow distribution and the resulting large temperature nonuniformities [30]. The testing results show that the worst-case temperature fluctuations are on the order of 30°C for $0.6 \text{ W } 400 \times 400 \mu\text{m}$ hot-spot heating condition, on the order of 20°C for $0.4 \text{ W } 400 \times 400 \mu\text{m}$ hot-spot heating condition, and on the order of 15°C for the uniform heating condition, showing that temperature fluctuations depend on the power being dissipated from localized hot spots. In particular, poor flow distribution in two-phase microchannels might lead to less flow in the high-flux regions, leading to localized dry out on the hot spot, which will result in large and rapid increase in the hot-spot temperature.

Jet impingement cooling, with high-velocity liquid streams directly impinging onto the hot surface, is an alternative to microchannel coolers. This method offers several potential advantages, such as high heat-transfer coefficients associated with the evaporation of thin liquid films and the ability to form patterns during the liquid distribution through an array of jets. The jets can be fabricated by circular or slot-shaped orifices or nozzles of various cross sections and can flow into a gaseous medium (vapor or air) forming the so-called free-jets or into a liquid medium,

leading to a “submerged jet.” As a final distinction, jet impingement cooling of electronic components may involve forced convection alone or localized flow boiling, with or without net vapor generation [33–35]. Although electronic cooling applications will require the use of dielectric liquids, much of the available data is for jet impingement of water, including the work by Zhang demonstrating that two-phase jet impingement is capable of removing a heat flux of more than 100 W/cm^2 at water flow rates below $2.5 \times 10^{-7} \text{ m}^3/\text{s}$ [36], use of boiling macrojets for removal of a heat flux over 400 W/cm^2 [37], and the results by Kiper, using microscaled direct water impingement from an orifice plate to cool VLSI circuits dissipating more than 500 W/cm^2 [38]. The need to achieve this level of performance with dielectric coolants and concerns over the reliability, complexity, volume, weight, and cost of such jet impingement systems have posed significant barriers to the successful commercial implementation of these approaches.

1.1.3 Solid-State Cooling Solutions

In recent years there has been increased interest in the application of solid-state thermoelectric coolers for high-flux thermal management because of their compact structure, fast response, high-flux spot-cooling capability, and high reliability, and their absence of moving parts [39–46]. Another major advantage of a thermoelectric cooler is that it can be miniaturized and integrated into the chip package. A thermoelectric cooler is based on the Peltier effect and consists of N-type and P-type thermoelectric elements, as shown in Fig. 9, where the N-type and P-type thermoelectric elements are joined by metallic connectors at the top and bottom. When a

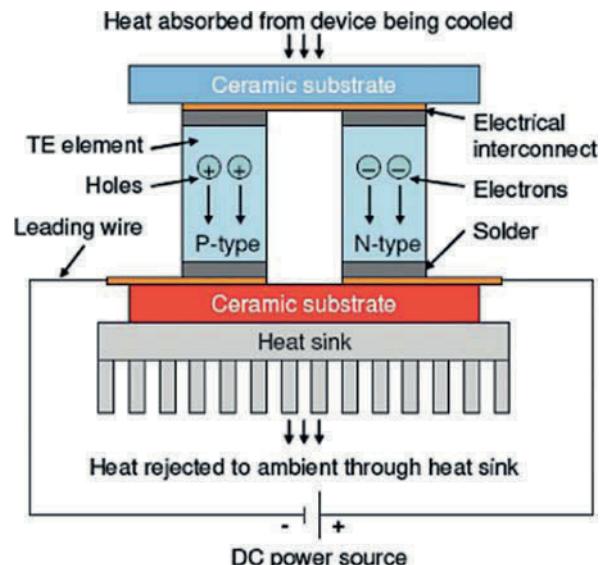


Fig. 9 Schematic diagram of a thermoelectric cooler

DC current goes through these thermoelectric element/metal contacts, heat is either released or absorbed in the contact region depending on the direction of the current. Attaching the TEC cold junction to a working device makes it possible to lower its temperature below its surroundings and possibly even below the ambient temperature.

In the conventional thermoelectric cooler (TEC) design, the maximum achievable temperature reduction across the thermoelectric cooler can be determined by:

$$\Delta T_{\max} = \frac{S^2 T_c^2}{2\rho k} \quad (1)$$

and the maximum achievable cooling flux on the cold side of thermoelectric cooler by Equation (2)

$$q''_{\max} = \frac{S^2 T_c^2}{2\rho d} \quad (2)$$

where S is the Seebeck coefficient, k is the thermal conductivity, ρ is the electrical conductivity, T_c is the absolute temperature at the cold side and d is the thickness of thermoelectric elements [47–49]. Thus, the largest temperature reduction is attained for thermoelectric materials when the Seebeck coefficient is large and the electrical resistivity and the thermal conductivity are as small as possible. Equation (2) indicates that the maximum cooling flux is inversely proportional to thermoelectric element thickness and thus the main advantage of going to thin-film thermoelectric coolers (TFTECs) is the dramatic increase in cooling heat flux. As shown in Fig. 10, Fleurial estimated that the heat flux of several hundred Watt per square centimeter could be removed with thin film thermoelectric coolers when thermoelectric element thickness is on the order of 20 μm [50].

Recent attempts to improve the cooling performance of TECs have focused on low-dimensional nanostructured superlattices which are capable of suppressing the thermal conductivity through phonon trapping and improving the TEC's figure of merit Z ($Z=S^2/k\rho$). Venkatasubramanian reported thin-film superlattice $\text{Bi}_2\text{Te}_3/\text{Sb}_2\text{Te}_3$ coolers capable of providing up to 32°C net cooling measured on the cold side of the coolers and a maximum estimated cooling heat flux of 700 W/cm^2 at room temperature [51]. Harman demonstrated a thin-film cooler based on quantum dot n-type $\text{PbSeTe}/\text{PbTe}$ superlattice structure, which provided 43.7°C net cooling at room temperature [52]. Fan and Shakouri demonstrated net cooling of up to 2.5°C at room temperature and 7°C at 100°C ambient temperature and a maximum cooling heat flux as high as 680 W/cm^2 on the surface of a microcooler for p-type thin-film superlattice SiGeC/Si microcoolers TEC device miniaturization [53], to extract greater performance from existing bulk thermoelectric materials, is the other approach to TEC improvement currently receiving attention. Using ceramic thinning technology, Semenyuk developed and commercialized a 130 μm -thick miniaturized Bi_2Te_3 thermoelectric cooler which can provide more than 100 W/cm^2 cooling heat

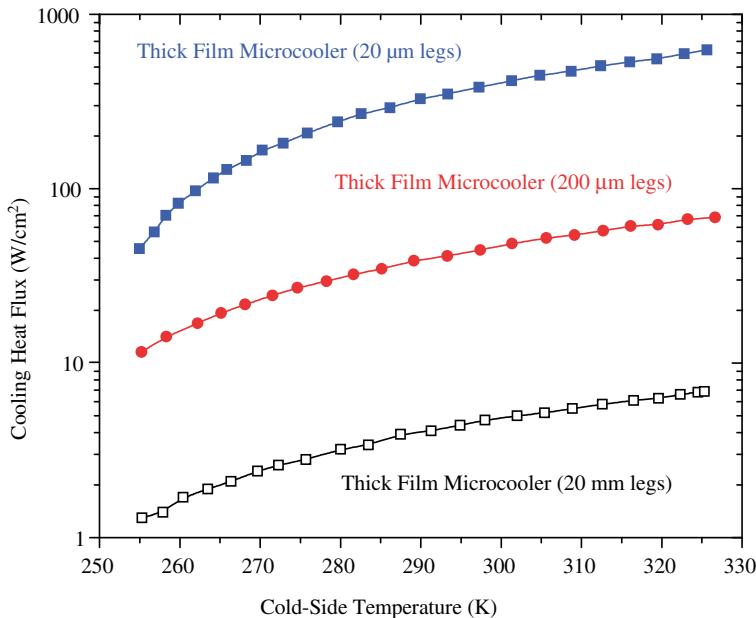


Fig. 10 Cooling heat flux as a function of thermoelectric leg thickness when diamond is used as the TEC substrate and bulk Bi₂Te₃ as thermoelectric material ($ZT = 0.9$)

flux at the cold side junction [54–57]. These reported results show solid-state thermoelectric coolers that provide a powerful alternative to traditional high-flux coolers and offer great promise for reducing the severity of on-chip hot spots.

Succeeding sections of this chapter will address the possible application of these proposed approaches to the thermal management of nanoelectronic hot spots. Attention will be focused on silicon thermoelectric coolers, mini-contact, miniaturized TECs that overcome the present heat flux limitations of conventional Bi₂Te₃ devices, two-phase micro-gap coolers that can directly cool high heat flux chips without the deleterious effects of contact resistance, and the use of anisotropic thermal interface materials (TIM) to enhance the effectiveness of both passive and active cooling, in suppressing the temperature rise in high-flux regions of the chip.

2 On-Chip Hot-Spot Cooling Using Thermoelectric Microcoolers

Since the discovery of the Peltier effect in 1834, extensive research has been done to develop solid-state refrigeration devices and solid-state energy generators based on thermoelectric effects. However, only limited applications of thermoelectric cooling existed until the middle 1950s, when it was discovered that doped semiconductors could achieve better thermoelectric properties than metallic materials. After

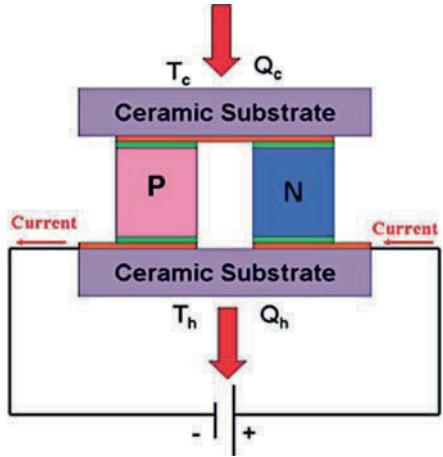
several decades of research, the efficiency of thermoelectric cooling devices still reached only about 10% of Carnot efficiency in comparison to the 30% efficiency typical of vapor compression refrigerators. The progress in thermoelectrics research declined until the early 1990s, when theoretical predication indicated that low-dimensional structures, such as two-dimensional superlattices, could be used to produce high-performance thermoelectric materials. Since then extensive studies, theoretical as well as experimental, have been conducted to explore new materials and new designs for fabrication of high-performance thermoelectric coolers. Since 2000, the demand for hot-spot cooling in microprocessors, as well as in power electronic, RF, and laser components, has generated renewed interest in exploring novel thermoelectric microcoolers, which can provide localized, high-flux cooling capability.

2.1 Principle of Conventional Thermoelectric Cooler (TEC)

The principle of conventional thermoelectric coolers was developed 50 years ago by Ioffe and coworkers [58]. A typical TEC consists of an array of n-type and p-type thermoelectric elements, two ceramic substrates that provide mechanical support for the TEC, electric conductors that provide a serial electric connection for the thermoelectric elements and electric contacts to lead wires, solders that join the thermoelectric elements, and lead wires that are connected to the ending conductors and deliver power from a DC electrical source. When semiconductor TEC devices are assembled, the array of heavily doped p-type and n-type semiconductor elements is soldered to ceramic substrates so that they are connected electrically in series and thermally in parallel. As is known, electrons can move freely in electrical conductors but less in semiconductors. When the electrons leave an electric conductor and enter the p-type semiconductor, they drop down to a lower energy level and release heat at the interface. However, as the electrons move across the bonded interface from the p-type semiconductor into the electric conductor, they transition to a higher energy level and absorb heat. When these electrons then move into the n-type semiconductor, they rise further in energy level so that additional heat is absorbed. When the electrons leave the n-type semiconductor and enter the conductor they drop down to a lower energy level and release heat during the process. Thus, in this thermoelectric circuit, heat is always absorbed when electrons enter n-type semiconductor or leave p-type semiconductor, and heat is always released when electrons enter p-type semiconductor or leave n-type semiconductor. The precise heat pumping capacity of a TEC is proportional to the current and is dependent on the element geometry, number of couples, and thermoelectric properties of the materials.

Figure 11 shows the basic configuration of a thermoelectric cooler with one p-type element and one n-type element. For a single-stage TEC, as shown in Fig. 11, the amount of heat that can be pumped at the cold side of a TEC is the net of three contributions. If we assume to have perfect thermal interfaces at the cold side and

Fig. 11 Sketch of a thermoelectric cooler



the hot side and neglect the Thomson effect, the net cooling power on the cold side of the TEC can be expressed by:

$$q_c = (S_p - S_n) IT_c - K (T_h - T_c) - \frac{1}{2} I^2 R_e \quad (3)$$

where R is the overall electrical resistance of the TEC:

$$R_e = \rho_p \frac{L_p}{A_p} + \rho_n \frac{L_n}{A_n} \quad (4)$$

K is the overall thermal conductance of the TEC

$$K = k_p \frac{L_p}{A_p} + k_n \frac{L_n}{A_n} \quad (5)$$

and S , k , ρ , A , and L represent the Seebeck coefficient, thermal conductivity, electrical resistivity, the cross-section area, and the thickness of thermoelectric element, respectively. The p and n denote p-type and n-type thermoelectric materials, while T_h and T_c are the temperatures at the cold side and the hot side of the TEC, respectively. The overall cooling rate is driven by the Peltier cooling (the first term) and reduced by the heat flowing back from the hot side to the cold side of the TEC (the second term) and Joule heating in the element (the third term).

When a current is applied to the TEC, a voltage drop is generated, which includes the resistive voltages and the Seebeck voltages across the thermoelectric elements and is given by:

$$V = (S_p - S_n) (T_h - T_c) + IR \quad (6)$$

Therefore, the electric power consumption of the TEC system is equal to

$$W = (S_p - S_n)(T_h - T_c)I + I^2R \quad (7)$$

The coefficient of performance (COP) is used to describe the cooling efficiency of the TEC, and is the ratio of the net cooling power at the cold side and the power consumption of the system and is given by:

$$\text{COP} = \frac{q_c}{W} \frac{(S_p - S_n)IT_c - K(T_h - T_c) - \frac{1}{2}I^2R_e}{(S_p - S_n)(T_h - T_c)I + I^2R} \quad (8)$$

The maximum cooling rate that can be achieved by this TEC device can be determined by differentiating the cooling rate given in Equation (3) to find the optimal current I_{opt} :

$$\left(\frac{dq_c}{dI} \right)_{\text{opt}} = 0 \Rightarrow I_{\text{opt}} = \frac{(S_p - S_n)T_c}{R_e} \quad (9)$$

When $I = I_{\text{opt}}$, the heat removal rate attains its maximum, value given by:

$$q_{\max} = \frac{1}{2} \frac{(S_p - S_n)^2 T_c^2}{R_e} - K(T_h - T_c) \quad (10)$$

To determine the largest temperature reduction, i.e., the deepest cooling, that the TEC can achieve, the heat removed from the cold side is set equal to zero, $q_{\max} = 0$, yielding the maximum temperature difference across the TEC, i.e., from the cold side to the hot side of the thermoelectric cooler:

$$\Delta T_{\max} = \frac{(S_p - S_n)^2 T_c^2}{2KR_e} = \frac{ZT_c^2}{2} \quad (11)$$

where Z is known as the TEC figure of merit, given by

$$Z = \left[\frac{(S_p - S_n)}{(k_p \rho_p)^{0.5} + (k_n \rho_n)^{0.5}} \right]^2 \quad (12)$$

For simplification, it is frequently assumed that the semiconductor Seebeck coefficients of the two materials are equal but opposite in sign, i.e., $S_p = -S_n = S$, and that the thermal conductivity, electrical resistivity, and dimensions are equal for the two elements, i.e., $\rho_p = \rho_n = \rho$, $k_p = k_n = k$, $A_p = A_n = A$, $L_p = L_n = L$, so Equation (12) can be simplified as

$$Z = \frac{S^2}{k\rho} \quad (13)$$

Then the maximum achievable temperature difference at the optimized current, $I = I_{\text{opt}}$, when the heat flow is zero, can be calculated as

$$\Delta T_{\max} = \frac{S^2 T_c^2}{2k\rho} = \frac{1}{2} Z T_c^2 \quad (14)$$

Similarly, the maximum cooling rate at the optimized current ($I = I_{\text{opt}}$), when the hot side and cold side are equal in temperature can be calculated as:

$$q_{\max} = \frac{S^2 T_c^2}{2R_e} = \frac{S^2 T_c^2 A}{2d\rho} \quad (15)$$

The corresponding cooling heat flux at the cold side of the TEC can be calculated as:

$$q''_{\max} = \frac{q_{\max}}{2A} = \frac{S^2 T_c^2}{2d\rho} \quad (16)$$

It is interesting to note that the maximum achievable cooling ΔT_{\max} only depends on the figure of merit Z and the temperature of the cold side of the TEC, but it does not change with the cross-sectional area or thickness of the TEC elements. To attain the lowest temperature, it is desirable to have a thermoelectric material with a high value of Z . Alternatively, when thermal conduction is not an important parasitic effect, the power factor P , defined by Equation (17), is often used to characterize the thermoelectric properties of a given material:

$$P = \frac{S^2}{\rho}$$

The maximum COP can be found by differentiating the COP given in Equation (8) to find the COP-optimal current $I_{\text{COP,opt}}$:

$$\left(\frac{d\text{COP}}{dI} \right)_{\text{opt}} = 0 \Rightarrow I_{\text{COP,opt}} = \frac{(S_p - S_n)(T_h - T_c)}{R_e(1 + ZT_{\text{ave}})^{0.5} - 1} \quad (18)$$

When $I = I_{\text{COP,opt}}$, the maximum value of COP can be calculated:

$$\text{COP}_{\max} = \frac{T_c}{T_h - T_c} \frac{(1 + ZT_{\text{ave}})^{0.5} - T_h/T_c}{(1 + ZT_{\text{ave}})^{0.5} + 1} \quad (19)$$

where $T_{\text{ave}} = (T_h + T_c)/2$ is the mean temperature of the TEC.

2.2 Thermoelectric Cooling Materials and Devices

As described in the last section, for conventional thermoelectric applications, the best thermoelectric materials are those providing the highest Z values, usually associated with high Seebeck coefficients, high electrical conductivities and low thermal conductivities. So far, the best thermoelectric material properties are found in heavily doped semiconductors. In semiconductors, the thermal conductivity is established by the flow of both electrons and phonons, but with much of the thermal transport ascribed to phonons. The phonon thermal conductivity can be reduced without causing significant reduction in the electrical conductivity. A common approach to reduce the phonon thermal conductivity is through alloying or doping because the mass difference scattering in alloys or doped semiconductors reduces the lattice thermal conductivity significantly without much degradation in the electrical conductivity.

Bismuth telluride-based compounds, using alloys of Bi_2Te_3 with Sb_2Te_3 (p-type) and Bi_2Te_3 with Bi_2Se_3 (n-type), are the best commercial state-of-the-art materials for thermoelectric cooling, with the highest values of the figure of merit ZT . In bulk materials, a ZT of 0.75 for p-type $(\text{BiSb})_2\text{Te}_3$ at room temperature was reported about 40 years ago. Since the 1960s, much effort has been made to raise the ZT of bulk materials based on bismuth telluride by doping or alloying other elements in various fabrication processes. Recently, a ZT of 1.14 at 300 K has been reported for the p-type $(\text{Bi}_{0.25}\text{Sb}_{0.75})_2(\text{Te}_{0.97}\text{Se}_{0.03})_3$ alloy [59]. By annealing the ingots prepared by the Bridgman method, Yamashita et al. have most recently achieved a significant increase in the ZT value to 1.19 at 298 K for the n-type $\text{Bi}_2(\text{Te}_{0.94}\text{Se}_{0.06})_3$ and 1.41 at 308 K for the p-type $(\text{Bi}_{0.25}\text{Sb}_{0.75})_2\text{Te}_3$ alloy, so that both ZT values exceed 1 [60–62]. In addition to alloying, several other approaches have been proposed to enhance the ZT value, through either improving electrical conductivity or reducing thermal conductivity. In this respect, low-dimensional materials, such as quantum wells, superlattices, quantum wires, and quantum dots offer new ways to manipulate the electron and phonon properties of a given material. Some experiments have demonstrated that superlattice thermoelectric materials can achieve ZT values greater than 2.0. These inspiring results show the feasibility of applying thermoelectric materials to high-flux cooling applications.

Thermoelectric coolers (TECs) have traditionally been fabricated using bulk bismuth telluride materials and traditional processing techniques such as hot pressing sintering and extrusion. Commercial bulk thermoelectric coolers are made from such thermoelectric elements, typically several millimeters in thickness and combined into arrays that span 1.8×3.4 mm (and 2.4 mm thick) to 62×62 mm (and 5.8 mm thick). For such commercial TE modules, the maximum cooling at room temperature is about 70°C , with relatively low cooling heat flux of 5–10 W/cm², which makes it impossible to use such TECs for high-flux hot-spot cooling application [63]. Because the maximum cooling heat flux of a TEC is inversely proportional to the thickness of its elements, there have been extensive studies focusing on microscaled thin-film TECs, miniaturized TECs, and nanosstructured superlattice TECs, which could realize high-flux cooling requirement

for on-chip hot-spot reduction. In recent years, significant progresses have been reported in making microscaled thermoelectric coolers, as described in the three sections that follow.

2.2.1 Thin-Film Thermoelectric Coolers (TFTECs)

It is widely accepted that thin-film thermoelectric coolers (TFTECs) have great potential for high-flux cooling because of the dramatic theoretical increase in cooling heat flux made possible by decreasing thermoelectric element thickness as given by Equation (16). Among the available deposition methods, electrochemical deposition is very attractive from an application perspective, due to its ability to deposit thin films at high deposition rates of tens of microns per hour and at the much lower batch processing cost when compared to other state-of-the-art thin-film fabrication processes [64]. Snyder et al. used electrochemical-MEMS technique to fabricate thin-film thermoelectric microcoolers, which contains 63 n-type Bi_2Te_3 elements and 63 p-type $\text{Bi}_{2-x}\text{Sb}_x\text{Te}_3$ elements, each element being 20 μm in thickness and 60 μm in diameter with bridging metal interconnects, as shown in Fig. 12 [65]. Unfortunately, the defect structure in this MEMS TEC produced a high concentration of low-mobility carriers, yielding an estimated Z value of 3.2×10^{-5} (1/K) and producing a maximum cooling of 2°C and a maximum cooling heat flux of 7 W/cm².

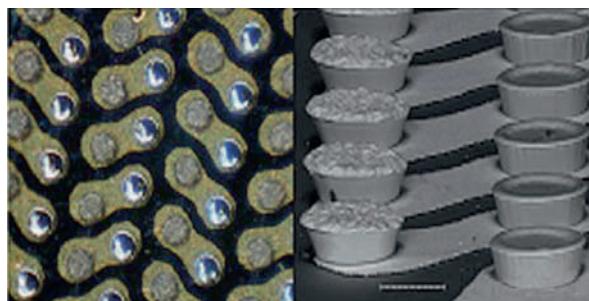


Fig. 12 Thin-film thermoelectric microcooler fabricated by MEMS [65]

Using co-evaporation as the deposition method, da Silva et al. fabricated thin-film thermoelectric microcoolers, which provided 60 n-type and p-type thermoelectric element pairs, with the thickness and width of the elements approximately 4.5 and 40 μm , respectively, as shown in Fig. 13. In da Silva's most recent work, the Seebeck coefficient, electrical resistivity, and power factor values of 149 $\mu\text{V/K}$, 1.25×10^{-5} $\Omega\text{ m}$, and 1.78 mW/K²m, respectively, were achieved for p-type Sb_2Te_3 thin films at the optimized deposition temperature [66]. However, the overall thermoelectric cooling performance achieved with such improved thin films has not been reported.

Böttner et al. developed a two-wafer process to fabricate thin film thermoelectric coolers. Figure 14 depicts a schematic drawing of the two wafer process on the left

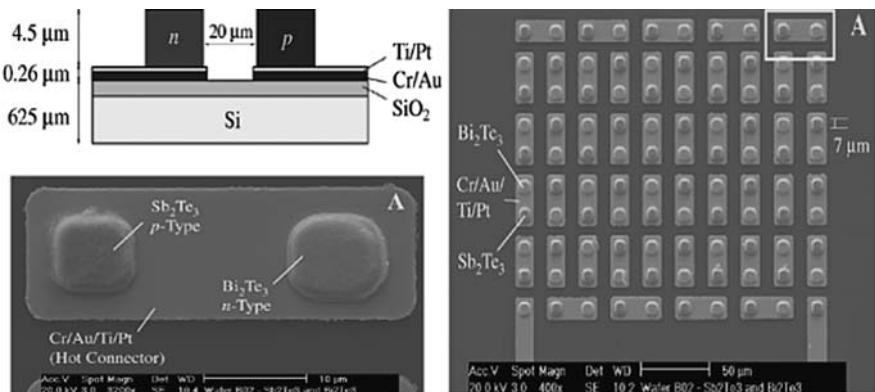


Fig. 13 Bi_2Te_3 and Sb_2Te_3 films deposited on $\text{Cr}/\text{Au}/\text{Ti}/\text{Pt}$ bottom connectors [66]

and a schematic drawing of the resulting device on the right. The polycrystalline n-type $\text{Bi}_2(\text{Se}, \text{Te})_3$ and p-type $(\text{Bi}, \text{Sb})_2\text{Te}_3$ materials were deposited by co-sputtering from 99.995% element targets (Bi , Sb , Te). However, these alloys were not grown very well in thin-film form due to delivery problems of the Se -target suppliers. For these coolers, the thickness of n-type Bi_2Te_3 elements and p-type $(\text{Bi}, \text{Sb})_2\text{Te}_3$ elements is about 20 μm as shown in Fig. 14 [67]. While he initially reported a net cooling of 11°C at 60°C ambient, more recently, Böttner et al. reported achieving a maximum temperature differences of nearly 48°C at an applied current of 2.1 A, and a maximum cooling flux of $\sim 100 \text{ W/cm}^2$ for the complete device [68].

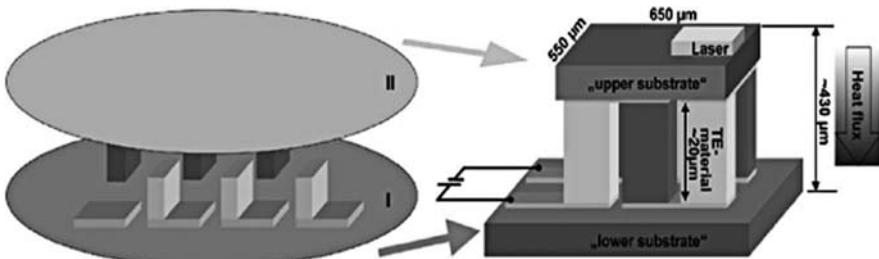


Fig. 14 Micro-Peltier cooler. *Left:* Schematic drawing of the developed two wafer (I,II) concept. *Right:* Schematic drawing of the thermoelectric cooler used for telecommunication device [67]

Zou et al. found that direct vapor deposition of bismuth telluride compounds is made difficult by the large difference in the vapor pressure between antimony, bismuth, and tellurium, which could result in noncongruence and in a lack of stoichiometry. In his work, Sb_2Te_3 films were deposited by co-evaporation of antimony and tellurium and Bi_2Te_3 thin films by co-evaporation of bismuth and tellurium onto heated, clean glass substrates. The figure of merit Z for the p-type Sb_2Te_3 film and n-type Bi_2Te_3 film was calculated and found to be approximately 1.04×10^{-3} at

room temperature, corresponding to ZT of 0.32. The maximum temperature difference measured between the hot and cold ends was 15.5°C at a current of 55 mA, showing a promising procedure for fabricating thermoelectric microcooler [69].

In using the state-of-the-art deposition techniques to develop thermoelectric thin films, a primary difficulty is to maintain the stoichiometry of the bismuth telluride compounds. For example, the problem of resputtering during the film growth is present in sputter deposition, while differences in volatility of the component elements pose difficulty in vacuum evaporation. A large deviation from stoichiometry arises in vapor deposition because the constituent elements in the target exhibit dissimilar sticking coefficients on the substrate. In addition, there is a tendency for re-evaporation of certain elements from the deposited thin films because of their higher vapor pressure. Therefore, the thermoelectric properties of these thin films reported in the above publications vary widely and the figure of merit (ZT) is always much small than ~ 1.0 for bulk bismuth telluride material. As shown in Table 1, to date thin film thermoelectric coolers are still not well developed and complete characterization of the material properties of the various TEC thin films is lacking.

Table 1 Summary of cooling performance of Bi_2Te_3 -based thin-film TEC

Author (year)	Growth method	ΔT_{\max} ($^{\circ}\text{C}$)	q''_{\max} (W/cm^2)	TE properties
Snyder (2002)	Electrochemical deposition	2@ 80°C	7	$S = 60\text{--}100 \mu\text{V/K}$ $ZT=0.01$ (estimated)
Zuo (2002)	Co-sputtering	15.5@ 25°C	N/A	p-type: $S = 160 \mu\text{V/K}$ $\rho = 3.12 \times 10^{-5} \Omega \text{ m}$ n-type: $S = -200 \mu\text{V/K}$ $\rho = 1.29 \times 10^{-5} \Omega \text{ m}$
da Silva (2005)	Co-evaporation	1.0@ 25°C	N/A	p-type: $S = 228 \mu\text{V/K}$ $\rho = 2.83 \times 10^{-5} \Omega \text{ m}$ n-type: $S = -149 \mu\text{V/K}$, $\rho = 1.25 \times 10^{-5} \Omega \text{ m}$
Böttner (2005)	Co-sputtering	48@ 25°C	100@ 25°C	p-type: $S = 180 \mu\text{V/K}$ $\rho = 1.30 \times 10^{-5} \Omega \text{ m}$ n-type: $S = -175 \mu\text{V/K}$ $\rho = 1.95 \times 10^{-5} \Omega \text{ m}$

2.2.2 Bulk Miniaturized Thermoelectric Coolers

Although thin-film deposition technology has an advantage for mass production, currently it appears not to provide thermoelectric cooling performance comparable to that available in bulk TECs, due to the difficulty in controlling thin-film growth

conditions to obtain the desired stoichiometry and a defect-free microstructure. Alternatively, attention can be turned to miniaturized thermoelectric coolers, based on thinning of bulk materials, which can reduce thermoelectric element thickness down to tens of microns and, at the same time, maintain the excellent thermoelectric properties of the bulk materials. Table 2 shows the progress in developing bulk miniaturized TECs since the 1960s [70].

Table 2 Thermoelectric cooling of bulk Bi_2Te_3 -based miniaturized TEC

Year	TEC configuration	TE element		ΔT_{\max} (°C)	q''_{\max} (W/cm ²)
		Thickness (μm)	TE properties		
1967	Single TE couple	130	$ZT = 0.54$	38 @30°C	95
1994	20 couple TECs	100	$ZT = 0.78$	50 @30°C	100
1997	120 couple TECs	200	$ZT = 0.78$	67 @30°C	65
2002	18 couple TECs	200	$ZT = 0.90$	70.6 @30°C 91.8 @85°C	80 98
2006	18 couple TECs	130	$ZT = 0.90$	64.2 @30°C 83.5 @85°C	110 132

In 1967 Semenyuk began developing bulk bismuth telluride material, which could be used to form very short thermoelectric elements [71]. Following years of development, in 2006 Semenyuk thinned the thermoelectric element down to 130 μm, as shown in Fig. 15. The extruded p-type and n-type bismuth telluride thermoelectric materials were in the form of rods with Z values of $3.02 \times 10^{-3} \text{ K}^{-1}$, corresponding to ZT value of 0.9. The 200 μm thick p-type and n-type slices were initially cut from the rods using electroerosion process. Then the slices were lapped to the final thickness of 130 μm, etched electrochemically, and nickel plated.

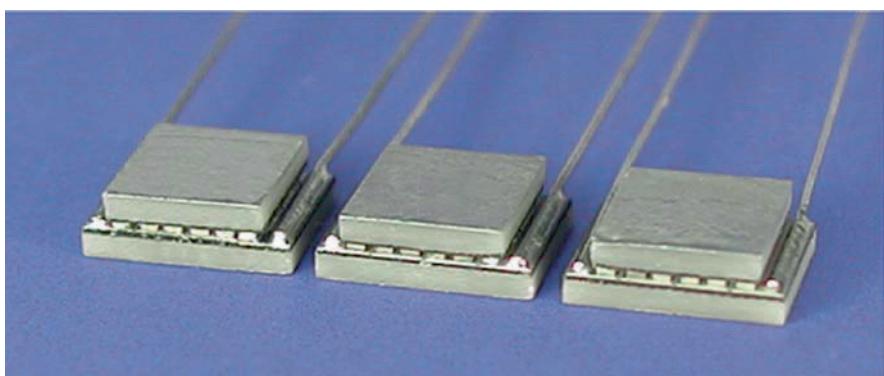


Fig. 15 Thermion TECs with 130 μm thick TE elements

AlN ceramic substrates were used with metal patterns obtained by standard microelectronics processing, including vacuum deposition of thin-film adhesive layers followed by electrochemical growth of thick copper films through photoresist processing, nickel plating, and finally Tin solder electrodeposition. The modules were found to provide a maximum cooling of 64.2 and 83.5°C and a maximum cooling flux of 110 and 132 W/cm² when operated at 30°C and 85°C, respectively, offering somewhat lower temperature reductions than achieved by the miniaturized TEC with 200 μm thick elements but improving the cooling heat flux by 30% under both conditions.

2.2.3 Nanostructured Thermoelectric Cooler

Nanostructured low-dimensional materials, such as superlattices, quantum wells, quantum wires, and quantum dots, offer opportunities to manipulate the electron and phonon properties of a given material leading to new ways to increase *ZT*. In a low-dimensional n-type material, the Fermi level is lower and the Seebeck coefficient is higher than that for corresponding bulk semiconductors with the same electron concentration, enhancing the value of the product of S^2/ρ , the power factor. Dresselhaus and coworkers theoretically predicted that the use of quantum well nanostructures could increase the power factor via quantum size effects, which improve the electron performance by taking advantage of sharp features in the electron density of states and result in *ZT* values in the range of 2–3 [72]. In addition, thermal conductivity can be reduced due to significantly modified phonon dispersion and enhanced phonon scattering mechanisms using the short period superlattice to impede phonon transport without excessively restricting the carrier flow [73, 74]. Experimental studies have demonstrated significant thermal conductivity reduction in a wide variety of nanostructured superlattices [75], leading to significant enhancements of the thermoelectric figure of merit for Bi₂Te₃/Sb₂Se₃ and PbTe/PbTeSe superlattice nanomaterials [76, 77]. Table 3 compares the reported power factor, *ZT*, and thermal conductivity of these nanostructured materials with that of their corresponding bulk materials at room temperature. It is clear that thermal conductivity reduction plays a significant role in the reported *ZT* enhancement and consequently there is only a small improvement in the power factor. It is interesting to note that these nanostructured materials can also be used to achieve multilayer thermionic (TI) cooling,

Table 3 Thermoelectric properties of nanostructured materials with high *ZT* [80]

Thermoelectric properties at 25°C	PbTe–PbSeTe quantum dot superlattices	PbTe–PbSe bulk alloy	Bi ₂ Te ₃ –Sb ₂ Te ₃ superlattices	Bi ₂ Te ₃ –Sb ₂ Te ₃ bulk alloy
Figure of merit <i>ZT</i>	1.6	0.35	2.4	1.0
Thermal conductivity (W/m K)	0.6	2.5	0.5	1.45
Power factor (mW/K ² m)	3.2	2.8	4.0	5.0

[78, 79] which allows for reduced parasitic Joule heating, since transport through the thin barriers is largely ballistic.

Venkatasubramanian used metal–organic chemical vapor deposition (MOCVD) to epitaxially grow a 5 μm thick $\text{Bi}_2\text{Te}_3/\text{Sb}_2\text{Te}_3$ nanostructured superlattice on GaAs substrates in 2001 [81]. These are phonon-blocking/electron-transmitting superlattices, which are produced by alternately depositing thin (1–4 nm) films of Bi_2Te_3 and Sb_2Te_3 . ZT was reported to be 2.4 for p-type $\text{Bi}_2\text{Te}_3/\text{Sb}_2\text{Te}_3$ superlattices and 1.4 for n-type $\text{Bi}_2\text{Te}_3/\text{Bi}_2\text{Te}_{2.83}\text{Se}_{0.17}$ at room temperature. This high ZT was explained by a reduction of the lattice thermal conductivity due to scattering of the phonons at the superlattice interfaces. The maximum cooling of 32.2 and 40°C was measured using an infrared camera and the maximum cooling flux of 585 and 700 W/cm^2 was estimated for p-type $\text{Bi}_2\text{Te}_3/\text{Sb}_2\text{Te}_3$ superlattice at the temperatures of 25°C and 80°C, respectively, with a response time of only about 5 μs .

More recently, PbSeTe-based quantum dot superlattice grown by molecular beam epitaxy (MBE) was reported by Herman's group for thermoelectric cooling applications [39]. The superlattice thin film with a thickness of approximately 100 μm was grown on BaF_2 substrates. The developed superlattice thin-film n-type PbSeTe/PbTe has a ZT of 1.6–2.0 at room temperature and achieved a maximum cooling of 43.7°C at 700 mA, under vacuum conditions, as shown in Fig. 16. It should be noted that theoretically increasing the ZT value will increase the maximum temperature differential of the TEC through $\Delta T_{\max} = 0.5ZT_c^2$. However, for on-chip hot-spot cooling,

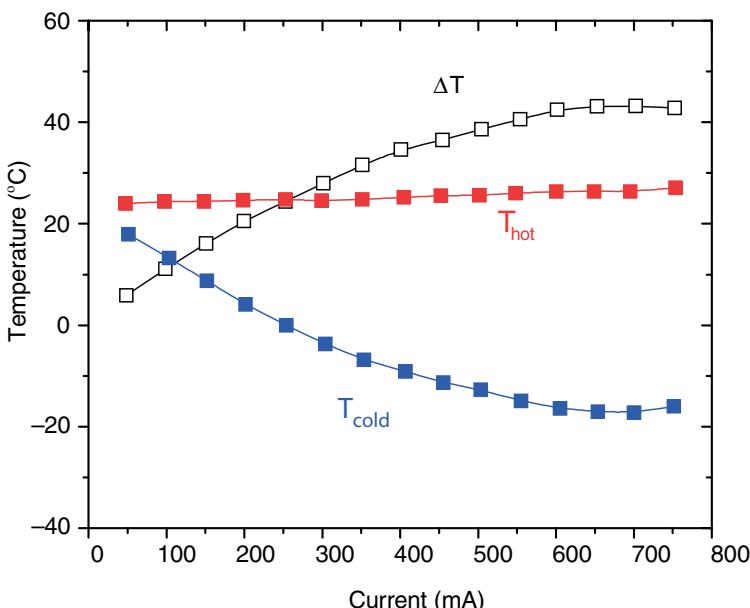


Fig. 16 Thermoelectric cooling characteristics of one-leg device made from n-type PbSeTe/PbTe superlattice thermoelectric cooler. ΔT represents measured data points of temperature differential between the hot junction temperature T_{hot} and T_{cold} cold junction temperature

ZT does not appear to be the relevant figure of merit. It was found that the materials with the same ZT will not necessarily provide equal degrees of hot-spot cooling and, among three material parameters which determine ZT value, increase of Seebeck coefficient is most effective to improve hot-spot cooling than decrease of electrical resistivity and thermal conductivity of the thermoelectric materials [82].

Shakouri and coworkers fabricated thin-film SiGe/Si, and SiGeC/Si thermoelectric microcoolers based on superlattice nanostructures using molecular beam epitaxy (MBE) [83–87]. As the SiGe/Si superlattice microcoolers can be monolithically integrated with microelectronic components to achieve localized cooling and temperature control, such devices could provide advantages for on-chip hot-spot cooling. The microcooler structure is based on cross-plane electrical transport and the main part of the microcooler is a 3 μm thick strain-compensated SiGe/Si superlattice, as shown in Fig. 17. It consists of 200 periods of 12 nm $\text{Si}_{0.9}\text{Ge}_{0.1}$ /3 nm Si, doped with boron to about $6 \times 10^{19} \text{ cm}^{-3}$. A maximum cooling of 4.5°C at 25°C, 7°C at 100°C and 14°C at 250°C was demonstrated. The maximum cooling heat flux increases with decreasing microcooler size, increasing from 120 to 680 W/cm² when the microcooler sizes reduces from 100 × 100 μm to 60 × 60 μm . Table 4 is the summary of cooling performance of thin-film nano-structured superlattice microcoolers developed since 2000.

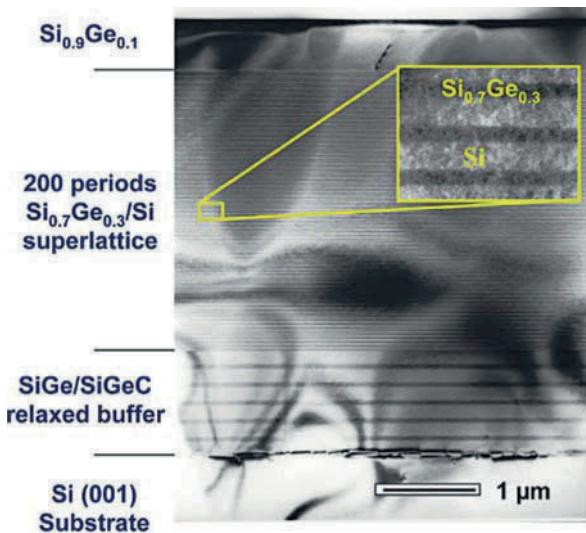


Fig. 17 Transmission electron micrograph of 3 μm thick $\text{Si}_{0.7}\text{Ge}_{0.3}/\text{Si}$ nano-structured superlattice grown on a buffer layer

2.2.4 Silicon Thermoelectric Materials and Microcoolers

While single-crystal silicon has been the key semiconductor material for much of the microelectronics era, silicon's thermoelectric potential has been largely ignored because of its high thermal conductivity and thus low value of the traditional TEC figure of merit, ZT (≈ 0.017) [88, 89]. However, silicon thermoelectric microcoolers, when formed on the back of the silicon chip for hot-spot cooling, provide

Table 4 Summary of cooling performance of nanostructured superlattice TEC

Year	Superlattice	ΔT_{\max} (K)	q_{\max} (W/cm ²)	Properties
LaBounty (2000)	InGaAs/InGaAsP	1.2@25°C 2.3@90°C	N/A	N/A
Fan (2002)	SiGe/Si	4.5@25°C	680@25°C	$S = 200 \mu\text{V/K}$
	SiGeC/Si	7.0@100°C 14.0@250°C		$k = 6.8\text{--}8.7 \text{ W/m K}$ $P = 2.2 \text{ mW/K}^2\text{m}$ $ZT = 0.085$
Venkatasubramanian (2001)	Bi ₂ Te ₃ /Sb ₂ Te ₃	32.2@25°C 40@80°C	585@25°C 700@80°C	$ZT = 2.4$ $P = 4.0 \text{ mW/K}^2\text{m}$
	PbSeTe/PbTe	43.7 @25°C	N/A	$ZT = 1.6$ $P = 3.2 \text{ mW/K}^2\text{m}$
Zhang (2003)	AlGaAs/GaAs	0.8@25°C 2.0@100°C	N/A	N/A

unique advantages over TFTECs. As may be seen in Table 5, which provides the thermal and electrical properties for three conventional thermoelectric materials – bulk Bi₂Te₃ alloy bulk SiGe alloy and single-crystal silicon – at room temperature, single-crystal silicon appears to offer the highest power factor of the materials shown, due to its high Seebeck coefficient and low electrical resistivity, and thus constitutes a very viable candidate for high-flux on-chip cooling.

Table 5 Typical values on the thermoelectric properties for Bi₂Te₃, SiGe and single-crystal silicon at room temperature [90]

Material	Seebeck coefficient, S ($\mu\text{V/K}$)	Electrical resistivity, ρ ($\mu\Omega \text{ m}$)	Thermal conductivity, k (W/m K)	Figure of merit	Power factor P (mW/K ² m)
Bi ₂ Te ₃ (n-type)	-240	10	2.02	$Z = 2.85 \times 10^{-3}$ $ZT = 0.86$	5.76
Bi ₂ Te ₃ (p-type)	162	5.5	2.06	$Z = 2.32 \times 10^{-3}$ $ZT = 0.70$	4.77
SiGe (n-type)	-136	10.1	4.45	$Z = 0.328 \times 10^{-3}$ $ZT = 0.1$	1.83
SiGe (p-type)	144	13.2	4.80	$Z = 0.413 \times 10^{-3}$ $ZT = 0.12$	1.57
Silicon (p-type)	450	35	150	$Z = 0.039 \times 10^{-3}$ $ZT = 0.012$	5.79

Zhang and Shakouri developed a silicon thermoelectric microcooler using bulk silicon which is p-type boron doped at a doping concentration of around 10¹⁹ cm⁻³ [91]. The device structure, which was fabricated with standard microfabrication techniques: dry etch, lithography, metal evaporation, etc., is illustrated in Fig. 18. They experimentally demonstrated the ability of such silicon thermoelectric

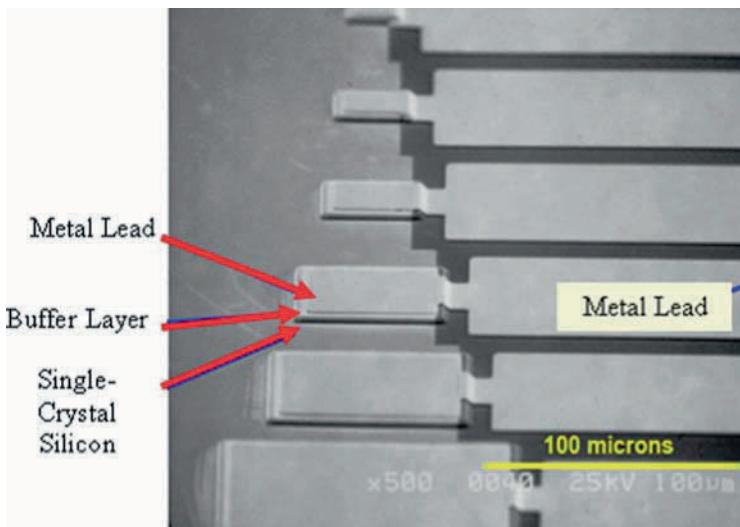


Fig. 18 A SEM photo of silicon microcooler [91]

microcoolers to achieve a surface temperature reduction of 1.2°C and a cooling flux of 580 W/cm^2 for a $40 \mu\text{m} \times 40 \mu\text{m}$ microcooler operating at room temperature at an optimized current of 0.1 A . In anticipation of the application of these microcoolers to the thermal management of microprocessor hot spots, and using an analytical model – with the embedded temperature dependence of electrical resistivity, thermal conductivity, and Seebeck coefficient based on values reported for single-crystal silicon [92, 93], Bar-Cohen et al. predicted the achievable temperature reduction, cooling heat flux, and parametric sensitivities of such thermoelectric microcoolers at 100°C [94]. The results displayed in Fig. 19 reveal that, in the absence of parasitic effects – Joule heating from electric contact resistance and heat conduction from metal lead, the silicon microcooler with the described configuration could achieve a maximum temperature reduction of 6.2°C on the microcooler at the optimum current of 0.9 A .

Figure 20 shows the maximum attainable temperature reduction on the microcooler, at an operating temperature of 100°C , for various doping concentrations with the microcooler size ranging from $20 \times 20 \mu\text{m}$ to $100 \times 100 \mu\text{m}$. The highest maximum temperature reduction of 6.2°C is achieved at a doping concentration of $2.5 \times 10^{19} \text{ cm}^{-3}$, and is independent of microcooler size. However, as shown in Fig. 20(b), smaller microcoolers do achieve the optimal performance at lower currents. In Fig. 20(a), the maximum average temperature reduction over the entire microcooler surface (average cooling) is also included for comparison. It is found that the maximum average cooling is approximately 30% lower than the maximum peak cooling.

One of the main advantages of silicon microcoolers is the very high cooling heat flux made possible by the high power factor for silicon. As with any thermoelectric cooler, the maximum cooling flux is achieved at a negligibly small temperature

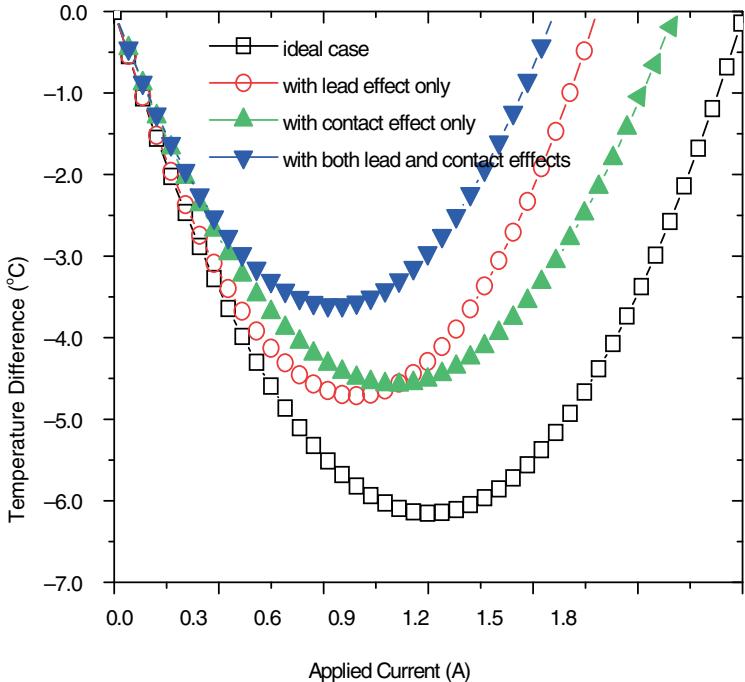


Fig. 19 Variation of silicon microcooler performance with applied current

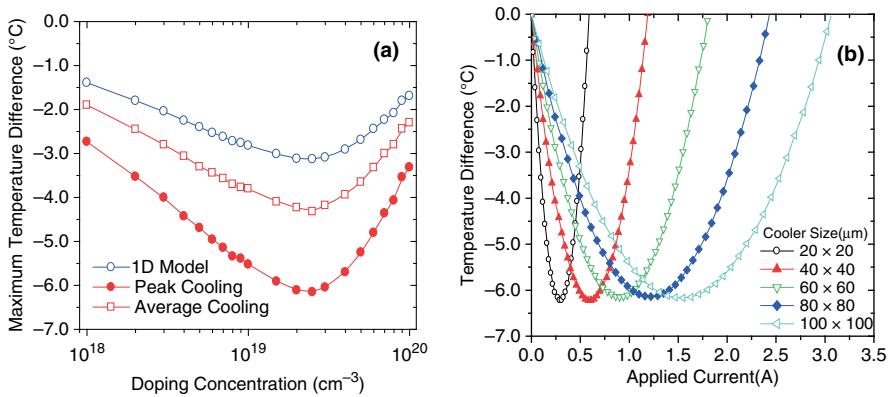


Fig. 20 (a) Variation of maximum temperature difference with the doping concentration for the ideal case and (b) Dependence of temperature difference on the applied current for different microcooler sizes at 100°C

reduction, while the greatest temperature reduction is achieved with negligibly small heat flux. For the present microcooler configuration, Figure 21 shows that the maximum cooling heat flux attains a predicted maximum value of 1 kW/cm^2 for $100 \times 100 \mu\text{m}$ microcooler and 6 kW/cm^2 for $20 \times 20 \mu\text{m}$ microcooler. These

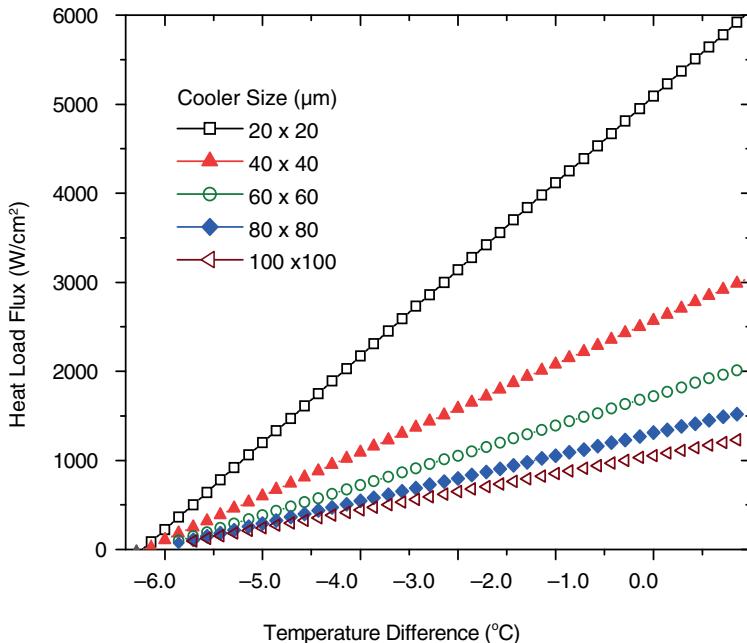


Fig. 21 Variation of heat load flux with temperature difference on the microcooler at 100°C

results support the expectation that silicon microcoolers provide a very promising approach to high heat-flux spot cooling in silicon microprocessors.

2.3 Hot-Spot Cooling Using Silicon Thermoelectric Microcooler

The concept of silicon thermoelectric microcooler for on-chip hotspot cooling, fabricated on the back of the silicon chip, is illustrated in Fig. 22, which displays a single microcooler, activated by an electric current entering the silicon chip through the metal contact and the silicon cap, flowing laterally through the chip, and exiting at the ground electrode located on the periphery of the chip. In a thermoelectric circuit, the flow of electrons across the interface between dissimilar materials, each with a distinct Seebeck coefficient, induces the Peltier effect, providing localized cooling when the direction of the current flow is from the low Seebeck coefficient to the high Seebeck coefficient material. The flow of electric current also serves to transport the absorbed heat away from that junction and to deposit that heat at a secondary interface where the electric current flows from the high Seebeck coefficient to the low Seebeck coefficient material. Joule heating, associated with the resistance to electric current in the thermoelectric circuit, and heat conduction from the hot junction to the cold junction of the thermoelectric circuit limit the thermoelectric cooling that can be achieved. The possible use of silicon thermoelectric

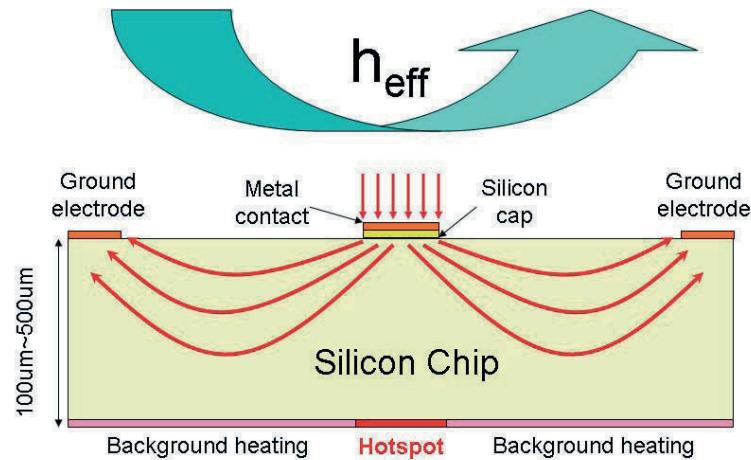


Fig. 22 Silicon thermoelectric microcooler for on-chip hot-spot cooling (The *arrows* indicate the direction for electric current)

microcoolers for the remediation of on-chip hot spots is facilitated by the use of well-established metal-on-silicon fabrication techniques, yielding a very low thermal contact resistance between the electrodes and the chip. In addition, incorporation of the silicon chip into the thermoelectric circuit makes it possible to transfer the absorbed energy via the electric current to the edge of the chip, far from the location of the hot spot, thus substantially reducing the detrimental effect of thermoelectric heating on the temperature of the active circuitry.

Referring to the structure of the on-chip silicon microcooler depicted in Fig. 22, it may be seen that Peltier cooling occurs at the junction between the metal contact and the silicon cap which is highly doped silicon with a doping concentration of more than $1 \times 10^{20} \text{ cm}^{-3}$ and again at the silicon cap/silicon chip interface, and that Peltier heating is encountered at the silicon chip/ground electrode interface, located on the periphery of the chip, where the electrons must shed some of their energy in entering the highly conductive metal. The overall Peltier heat-transfer (cooling) rate of the silicon microcooler can be expressed as:

$$q_{\text{TE} \times c} = -S_{\text{Si}} T_c I \quad (20)$$

where T_c is the absolute temperature at the microcooler, S_{Si} the Seebeck coefficient of the silicon chip, and I the applied current. Similarly, the Peltier heating rate at the silicon chip/ground electrode interface can be represented as:

$$q_{\text{TE}, h} = S_{\text{Si}} T_{\text{ed}} I \quad (21)$$

where T_{ed} is the absolute temperature at the ground electrode. In addition to volumetric Joule heating inside the silicon chip, the silicon cap, and the metal contact, these parasitic effects also arise at both the metal contact/silicon cap interface

and the silicon chip/ground electrode interface. The interfacial Joule heating at the metal contact/silicon cap interface is given by:

$$q_{\text{contact}} = I^2 R_{\text{cont}} = I^2 \rho_c / A_{\text{cont}} \quad (22)$$

where R_{cont} is the electric contact resistance, A_{cont} the cross-sectional area of metal contact, and ρ_c the specific electric contact resistance at this interface. Equation (22) applies as well at the peripheral ground electrode/silicon chip interface, with the appropriately adjusted contact area and the specific electric contact resistance.

In the work of Bar-Cohen [95] a test vehicle was used, consisting of a 12×12 mm silicon chip with 70 W/cm^2 background heat flux and $70 \times 70 \mu\text{m}$ hotspot, with a heat flux of 680 W/cm^2 , located at the center of the active side of the chip, to explore the hot-spot cooling potential of this approach. In their extensive modeling studies, the back of the chip was cooled with a high-performance 25°C air-cooled heat sink, capable of producing an effective heat-transfer coefficient of $8700 \text{ W/m}^2 \text{ K}$ representing the combined effect of the heat sink, heat spreader, and thermal interface materials used for electronic packages [96, 97]. The thermal conductivity of the silicon chip is assumed to be 110 W/m K , appropriate for 100°C operating temperature [98].

The prediction of hot-spot cooling achievable with on-chip silicon microcooler, as described in Fig. 23, requires the solution of the three-dimensional Poisson's equation for the temperature distribution in a volume subjected to nonuniform heat generation, associated with the Joule heating in the silicon chip, heating and cooling boundary conditions, associated with Peltier cooling and Peltier heating on the back surface, along with the microprocessor heat generation on the front surface (active circuitry), i.e.,

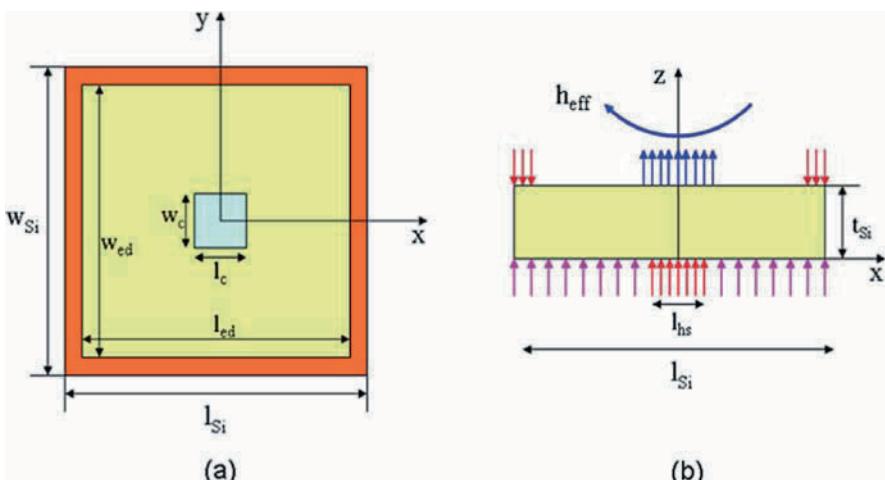


Fig. 23 (a) Coordinate system and (b) boundary conditions in the analytical model for the silicon chip integrated with silicon thermoelectric microcooler

$$\frac{\delta^2 T}{\delta x^2} + \frac{\delta^2 T}{\delta y^2} + \frac{\delta^2 T}{\delta z^2} + \frac{q_{Si}'''(x,y,z)}{k_{Si}} = 0 \quad (23)$$

where $q_{Si}'''(x,y,z)$ is the nonuniform volumetric heat generation due to the silicon Joule heating and k_{Si} is the thermal conductivity of the silicon chip. Unfortunately, solution of Equation (23) requires detailed knowledge of the internal heat generation function, $q_{Si}'''(x,y,z)$. Determination of this function requires a parallel solution of the Laplace's equation for the electric potential field, which will vary significantly with the geometries of the silicon chip and the silicon microcooler and the placement of the ground electrode. The resulting highly nonuniform heat generation function can be expected to render Equation (23) analytically unsolvable for all but the simplest approximations of $q_{Si}'''(x,y,z)$.

Alternatively, considering the common use of “allocation factors” in determining the performance of one-dimensional thermoelectric devices [99] and the successful application of this approach to silicon thermoelectric microcoolers in an earlier publication by the authors, it is possible to re-formulate Equation (23) in the Laplace's form by allocating an appropriate fraction of the Joule heating to the microcooler (α) and the hot spot (β), respectively. With this approach, the volumetric silicon Joule heating is replaced with modified boundary conditions at the microcooler and the hot spot, respectively, and the Poisson's equation can then be transformed into the Laplace's equation for this same domain, which can be solved analytically.

In subsequent sections, the hot-spot remediation capability of silicon microcoolers will be characterized by three distinct metrics, including:

- (1) ΔT – the temperature reduction anywhere in the studied domain that is achievable by activating the microcooler. This metric characterizes the intrinsic thermoelectric cooling capability of the silicon microcooler. It is generally applied to the hot spot or the microcooler in this study and given by:

$$\Delta T = T_{\text{cooler, on}} - T_{\text{cooler, off}} \quad (24)$$

- (2) $\Delta T_{\text{hot spot}}^*$ – the ratio of the temperature change at the hot spot due to activating the microcooler to the temperature rise engendered by the hot spot. This metric quantifies the hot-spot cooling effectiveness of the silicon microcooler and is defined as:

$$\Delta T_{\text{hot spot}}^* = \frac{T_{\text{hot spot on, cooler off}} - T_{\text{hot spot on, cooler on}}}{T_{\text{hot spot on, cooler off}} - T_{\text{hot spot off, cooler off}}} \quad (25)$$

For $\Delta T_{\text{hot spot}}^* = 1$, the temperature rise engendered by the hot spot can be completely removed by the microcooler. For $\Delta T_{\text{hot spot}}^* = 0$, the microcooler is totally ineffective and for $0 < \Delta T_{\text{hot spot}}^* < 1$, the microcooler can achieve partial success in reducing the hot-spot temperature. For $\Delta T_{\text{hot spot}}^* > 1$, the microcooler is capable of “overcooling” the hot spot relative to the base temperature of the silicon chip.

- (3) π – the thermal impact factor, which provides a measure of the power needed, P_{in} , to achieve a specified temperature reduction at the hot spot, $\Delta T_{\text{hot spot}}$. This dimensional metric ($\text{K}/\text{W}_{\text{elec}}$) can be expressed as:

$$\pi = \frac{-\Delta T_{\text{hot spot}}}{q_{\text{in}}} \quad (26)$$

Clearly, as π increases less electric power is required in order to achieve a specific temperature reduction at the hot spot.

Following the transformation above, Bar-Cohen and coworkers developed a three-dimensional analytical thermal model of on-chip hot-spot cooling to investigate the effectiveness of such silicon thermoelectric microcoolers for a wide range of hot-spot sizes and heat fluxes, microcooler sizes, silicon chip thicknesses, doping concentrations, and electric contact resistances. The analytical solution yields the temperature distribution in the silicon chip, under the influence of hot-spot heating and background heating from related circuitry on the active surface, Peltier cooling, Peltier heating and conductive/convective cooling on the opposite surface, volumetric Joule heating inside the silicon chip, and interfacial Joule heating at the electric contact created by the silicon microcooler. The analytical solution employs numerically derived allocation factors to redistribute the Joule heating inside the chip to the hot spot and the microcooler. Results obtained from a three-dimensional electro-thermal finite-element numerical simulation were used to validate and calibrate the analytical model. The parametric trends revealed by use of this analytical model are presented and discussed in subsequent sections.

2.3.1 Doping Concentration Effect

The thermoelectric properties of semiconductors are strongly dependent on doping concentration but modestly on the doping type [100–103]. Figure 24 shows that the electrical resistivity of silicon decreases with increasing doping concentration, while the Seebeck coefficient also displays an inverse relationship with doping concentration. Thus, increasing doping concentration results in lower electrical resistivity and, as a consequence, less Joule heating in the silicon chip, but, the associated decrease in the Seebeck coefficient leads to reduced thermoelectric cooling power.

The largest possible thermoelectric cooling power is attained by maximizing the thermoelectric power factor $P (= S^2 \rho)$, which for boron-doped single-crystal silicon at 100°C occurs at about $2.5 \times 10^{19} \text{ cm}^{-3}$. The variation of maximum hot-spot cooling with doping concentration for various microcooler sizes is presented in Fig. 24 for 100 μm thick chip and the specific electric contact resistance ranging from 1×10^{-7} to $1 \times 10^{-4} \Omega \text{ cm}^2$, revealing – as expected – that across the range of microcooler sizes studied, with increasing doping concentration, the hot-spot cooling increases until reaching a maximum value and then decreases with further increases in the doping concentration.

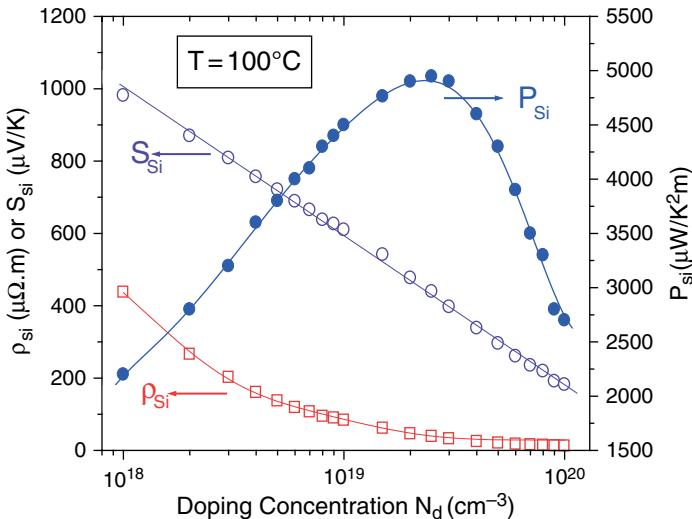


Fig. 24 Electrical resistivity (ρ), Seebeck coefficient (S), and power factor (P) as a function of boron doping concentration (N_d) in single-crystal silicon [104–106]

It is interesting to find that, despite the three-dimensional characteristic of heat spreading and electrical current spreading in the silicon chip surrounding the micro-cooler, for small electric contact resistance, e.g., $\rho_c < 1 \times 10^{-5} \Omega \text{ cm}^2$, the optimum

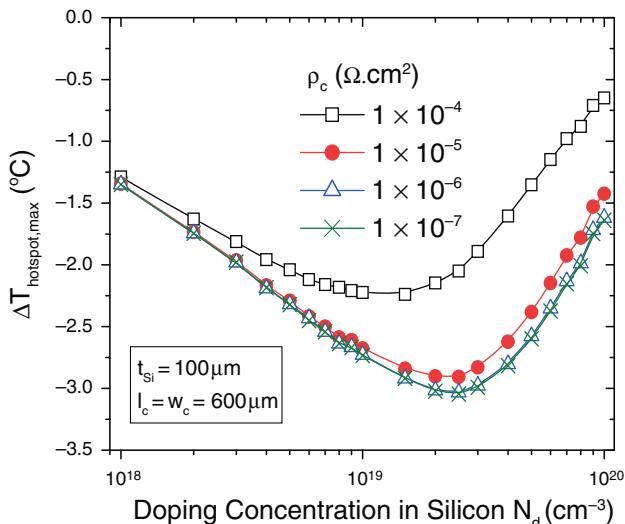


Fig. 25 Hot-spot cooling as a function of boron doping concentration for various specific electric contact resistances. The hot spot is $70 \times 70 \mu\text{m}$ with a heat flux of 680 W/cm^2 and the microcooler size is $600 \times 600 \mu\text{m}$. The microcooler size is $600 \times 600 \mu\text{m}$

doping concentration is nearly equal to $2.5 \times 10^{19} \text{ cm}^{-3}$, which yields the maximum power factor shown in Fig. 24. However, it is to be noted that the parasitic effect from larger electric contact resistance does have an influence on the optimum doping concentration, yielding a lower optimized doping concentration of $1.5 \times 10^{19} \text{ cm}^{-3}$ for a $600 \times 600 \mu\text{m}$ microcooler with the specific electric contact resistance of $1.0 \times 10^{-4} \Omega \text{ cm}^2$. It has been found that this trend becomes more pronounced as the microcooler size gets smaller.

It should be noted that the optimum doping level for a silicon thermoelectric microcooler is, thus, likely to be substantially higher than commonly used in semiconductor silicon chips. However, as is almost always the case for chip thermal management, the present analysis assumes that the back of the chip is used for cooling while the front is used for the active circuitry. Consequently, the doping concentration on the back side of the chip need not equal the more common doping concentration in the active semiconductor regions at the front of the chip, e.g., $1 \times 10^{16} \text{ cm}^{-3}$. Due to the far higher electrical resistivity in the active silicon layer, the electric current that is used to activate thermoelectric cooling is not expected to penetrate into this region.

2.3.2 Microcooler Size Effect

The effect of microcooler size on cooling performance involves the interplay of thermoelectric cooling by the microcooler and thermal diffusion from the hot spot to the microcooler. With decreasing microcooler size, the effective cooling flux and thus the temperature reduction at the microcooler increases, while the thermal resistance between the hot spot and the microcooler also increases. Consequently, this larger cooling flux at smaller microcoolers cannot effectively translate into larger temperature reduction at the hot spot. On the other hand, with smaller thermal resistances between the hotspot and the microcooler, the more modest cooling flux on larger microcoolers can be projected effectively onto the hot spot, narrowing the temperature difference between the hot spot and the microcooler. However, the modest cooling flux achievable on the larger microcoolers reduces the beneficial temperature reduction at both the hot spot and the microcooler. The competition between these two effects results in an optimum microcooler size.

Figure 26 displays this behavior and shows the temperature reductions at the hot spot and the microcooler for a wide range of microcooler sizes for $100 \mu\text{m}$ thick chip operating under the background and hot-spot heat fluxes of 70 W/cm^2 and 680 W/cm^2 , respectively. For each microcooler, the applied current is carefully optimized in order to achieve the maximum temperature reductions at the hot spot. It is seen that the temperature reduction at the hot spot first increases with microcooler size and, after reaching the maximum value of 3.0°C for $600 \times 600 \mu\text{m}$ microcooler, decreases with a further increase in microcooler size. Interestingly, the temperature reduction at the microcooler varies monotonically with microcooler size, yielding progressively larger temperature reductions, to as much as 3.9°C , as the microcooler dimension shrinks to $100 \times 100 \mu\text{m}$, which, however, only provides 1.6°C temperature reduction at the hot spot.

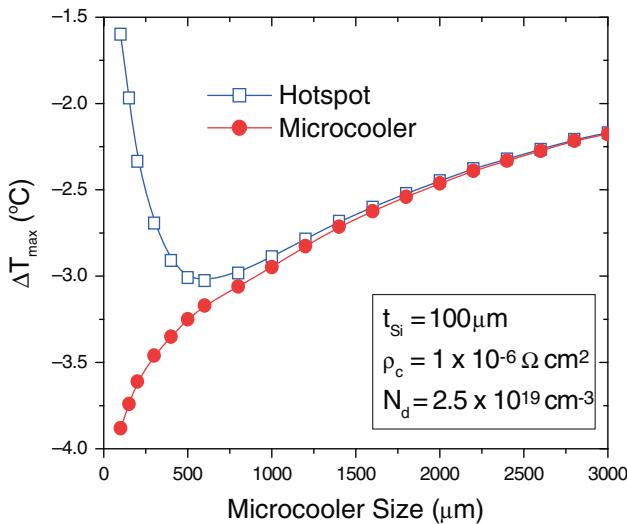


Fig. 26 Variation of temperature reductions at the hot spot and the microcooler with microcooler size. The hot spot is $70 \times 70 \mu\text{m}$ with a heat flux of 680 W/cm^2

2.3.3 Chip Thickness Effect

In the application of silicon microcoolers to hot-spot remediation, the silicon chip plays multiple roles, functioning as a thermoelectric material, to provide on-chip cooling and, at the same time, as an electrical conductor to transfer electrons from the ground electrode to the microcooler, and as a thermal conductor to provide a diffusion path for the heat generated in the chip to the ambient. Therefore, the chip thickness influences Joule heating distribution inside the chip, heat spreading from the hot spot, heat diffusion from the hot spot to the microcooler, and heat diffusion from the ground electrode, where Peltier heating occurs, to the hot spot. As the chip becomes thinner, the thermal resistance between the microcooler and the hot spot decreases, allowing the microcooler to achieve greater hot-spot temperature reductions, e.g., $2.05\text{--}3.03^{\circ}\text{C}$ as the chip thickness decreases from $500 \mu\text{m}$ to $100 \mu\text{m}$, for the conditions of Fig. 28. However, due to the smaller heat-spreading effect in thinner chips, the temperature rise engendered by the hot spot is also higher for thinner chips and increases with decreasing chip thickness from 2.2°C for a $500 \mu\text{m}$ thick chip to 2.9°C for $100 \mu\text{m}$ thick chip. These two trends compete with each other, yielding the maximum hot-spot cooling effectiveness at the chip thickness of $200 \mu\text{m}$, with $\Delta T_{\text{hot spot}}^* = 1.2$ as shown in Fig. 27. At this chip thickness, the silicon microcooler is, thus, capable of reducing the hot-spot temperature below the baseline temperature of the chip by approximately 0.5°C . Moreover, for the present $70 \times 70 \mu\text{m}$ hot spot with a heat flux of 680 W/cm^2 , the silicon microcooler is capable of completely suppressing or over-cooling the hot spot, with $\Delta T_{\text{hot spot}}^* \geq 1$, for the chip thicknesses between 100 and $475 \mu\text{m}$.

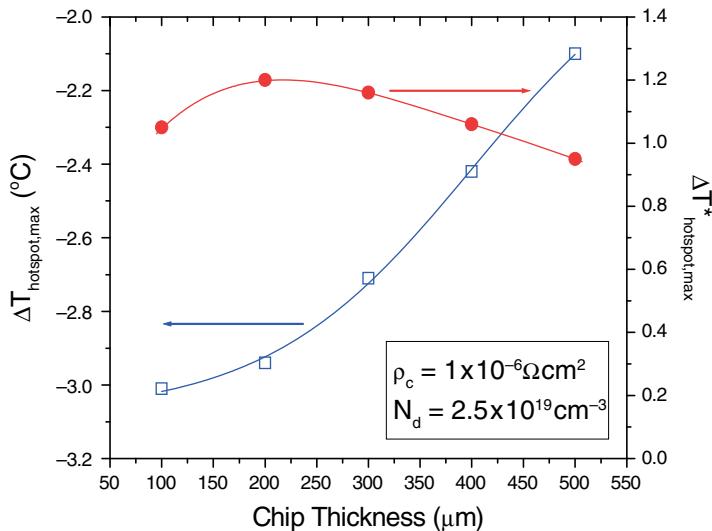


Fig. 27 Hot-spot cooling and hot-spot cooling effectiveness as a function of chip thickness. The hot spot is $70 \times 70 \mu\text{m}$ with a heat flux of 680 W/cm^2

It is interesting to find that the chip thickness also influences the optimum micro-cooler size. As shown in Figs. 28 and 29, the thicker the silicon chip the larger the optimized microcooler size. For $100 \mu\text{m}$ thick chip, the maximum hot-spot cooling of 3.0°C , for the conditions studied, is achieved with a $600 \times 600 \mu\text{m}$

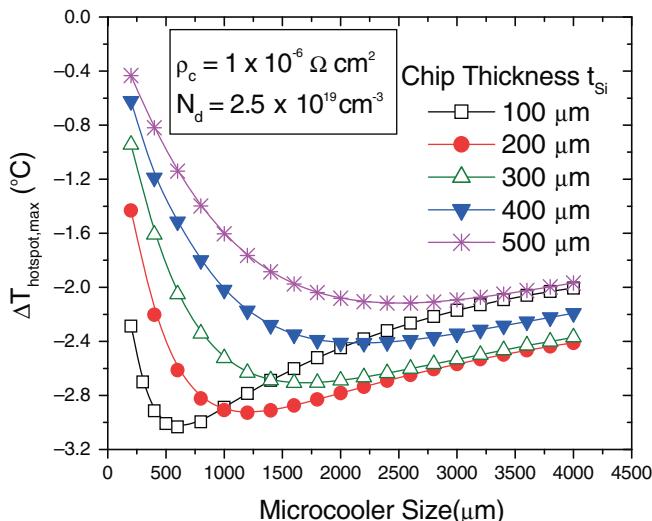


Fig. 28 Hot-spot cooling as a function of microcooler size for various chip thicknesses. The hot spot is $70 \times 70 \mu\text{m}$ with a heat flux of 680 W/cm^2

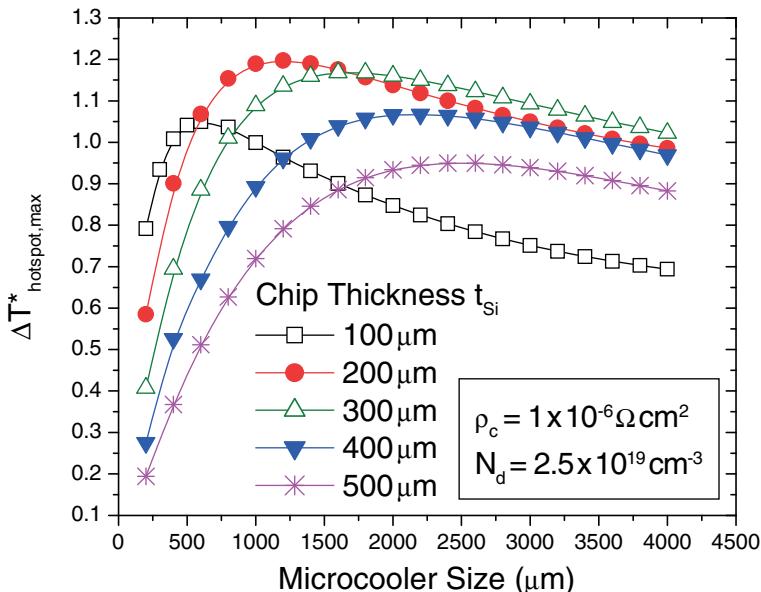


Fig. 29 Hot-spot cooling effectiveness as a function of microcooler size for various chip thicknesses. The hotspot is $70 \times 70 \mu\text{m}$ with a heat flux of 680 W/cm^2

microcooler, while for $500 \mu\text{m}$ thick chip, $2500 \times 2500 \mu\text{m}$ silicon microcooler is required in order to attain the maximum hot-spot cooling of 2.1°C . It should be noted that the optimum ratio of microcooler size to chip thickness is, thus, approximately 5.5, with a modest sensitivity to chip thickness, reaching 6.0 for $100 \mu\text{m}$, $200 \mu\text{m}$ and $300 \mu\text{m}$ thicknesses, dropping to 5.5 for the $400 \mu\text{m}$ thick chip, and to 5.0 for the $500 \mu\text{m}$ chip. This decreasing ratio can be related to the growing contributions of silicon Joule heating and Peltier heating to the hot-spot temperature as the optimized current – necessitated by the larger microcooler – increases.

Figure 30 shows the dependence of the thermal impact factor, π , of the silicon microcooler, on the microcooler size for different chip thicknesses, revealing that this factor and the relative benefit of input power decreases steeply with the microcooler size but more gently with the chip thickness. For example, $200 \times 200 \mu\text{m}$ microcooler can achieve a π of 17.0 in comparison with 1.2 for $4000 \times 4000 \mu\text{m}$ microcooler on $100 \mu\text{m}$ thick chip. With an increase of microcooler size, the effect of chip thickness on the thermal impact factor becomes less important. Consequently, this example for the specific parameters of the “test vehicle” used to explore the cooling potential of the silicon microcoolers, suggests that more generally the largest π values and the best returns on invested energy are attained when smaller microcoolers are used to remediate hot spots on thinner chips.

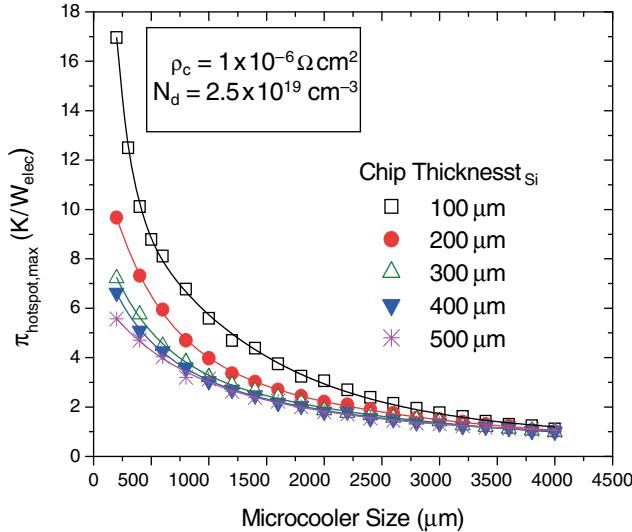


Fig. 30 Thermal impact factor as a function of microcooler size for various chip thicknesses. The hot spot is $70 \times 70 \mu\text{m}$ with a heat flux of 680 W/cm^2

2.3.4 Electric Contact Resistance Effect

The miniaturization of thermoelectric coolers tends to exacerbate the deleterious effects of the electric contact resistance, which is expected to occur at the interface between the metal contact and the silicon cap. The theoretical value of the specific electric contact resistance between highly doped silicon and a metal contact is in the range of $1 \times 10^{-9} \Omega \text{ cm}^2$ at room temperature or above [104]. However, due to process-related limitations, the typical specific electric contact resistance at such an interface usually ranges from 1×10^{-7} to $1 \times 10^{-5} \Omega \text{ cm}^2$, with significant batch to batch variations. Figure 31 shows the impact of the electric contact resistance on hot-spot cooling for different microcooler sizes on $100 \mu\text{m}$ thick chip. It should be noted that for a typical state-of-the-art thin-film process, which yields an average specific electric contact resistance of approximately $1 \times 10^{-6} \Omega \text{ cm}^2$ [105], the results displayed in Fig. 31 reveal that the electric contact resistance induced degradation in hot-spot cooling can be neglected.

More generally, as the specific electric contact resistance increases, hot-spot cooling performance is degraded, but the electric contact resistance has a larger impact for smaller microcooler sizes because the contact resistance is inversely proportional to the microcooler area. For an increase in the specific electric contact resistance from 1×10^{-9} to $1 \times 10^{-4} \Omega \text{ cm}^2$, hot-spot cooling will be degraded by a factor of 6.5 for $100 \times 100 \mu\text{m}$ microcooler but only by 5% for $3000 \times 3000 \mu\text{m}$ microcooler.

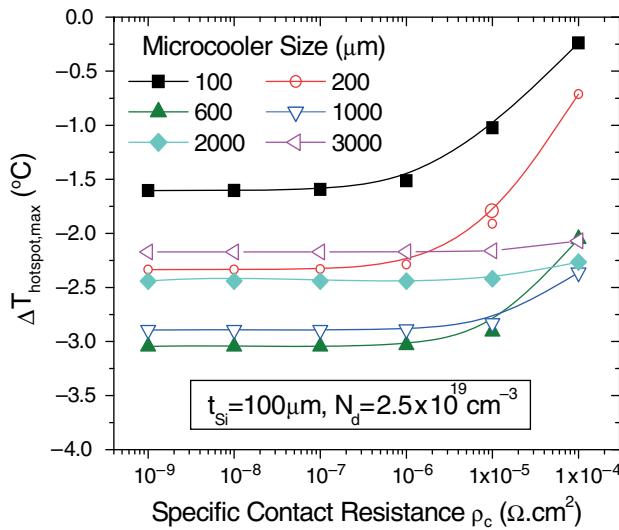


Fig. 31 Hot-spot cooling as a function of specific electric contact resistance for various microcooler sizes. The hot spot is $70 \times 70 \mu\text{m}$ with a heat flux of 680 W/cm^2

2.3.5 Hot-Spot Parameter Effect

Finally, attention is turned to the effects of the hot-spot parameters – namely, hot-spot size and hot-spot heat flux – on cooling performance, as evaluated by the three proposed metrics – ΔT , ΔT^* , and π . For each hot-spot size and hot-spot heat flux, the applied current, the microcooler size, and the doping concentration have been optimized in order to achieve the maximum hot-spot temperature reduction, while the specific electric contact resistance is fixed at $1 \times 10^{-6} \Omega \text{ cm}^2$. It was found that the optimized current and thus the optimized input power increase slightly with hot-spot size and hot-spot heat flux if the chip thickness and the doping concentration in the silicon chip remain constant. As may be seen in Figs. 32, 33, and 34 for $100 \mu\text{m}$ thick chip, the efficacy of the silicon microcooler varies with these hot-spot parameters in a complex manner. For example, the maximum temperature reduction at the hot spot, shown in Fig. 32, increases from 3.0°C for $70 \times 70 \mu\text{m}$ hot spot with 680 W/cm^2 heat flux to 3.90°C for $400 \times 400 \mu\text{m}$ hot spot with 1000 W/cm^2 heat flux, primarily because of the effect of the higher chip temperature (105 vs. 150°C) on the Peltier cooling rate. However, as seen in Fig. 33, the maximum cooling effectiveness decreases steeply with hot-spot size and hot-spot heat flux, from 1.05 for $70 \times 70 \mu\text{m}$ hot-spot with 680 W/cm^2 heat flux to 0.08 for $400 \times 400 \mu\text{m}$ hot-spot with 1000 W/cm^2 heat flux. Interestingly, since as the hot-spot size and the hot-spot heat flux increases, the maximum hotspot temperature reduction increases while the optimized input power remains nearly constant, π , the thermal impact factor, increases with the hot-spot size and the heat flux, as shown in Fig. 34. It should, thus, be understood that the silicon microcoolers can produce the largest cooling effect for constant input power when encountering large, high heat flux hot spots.

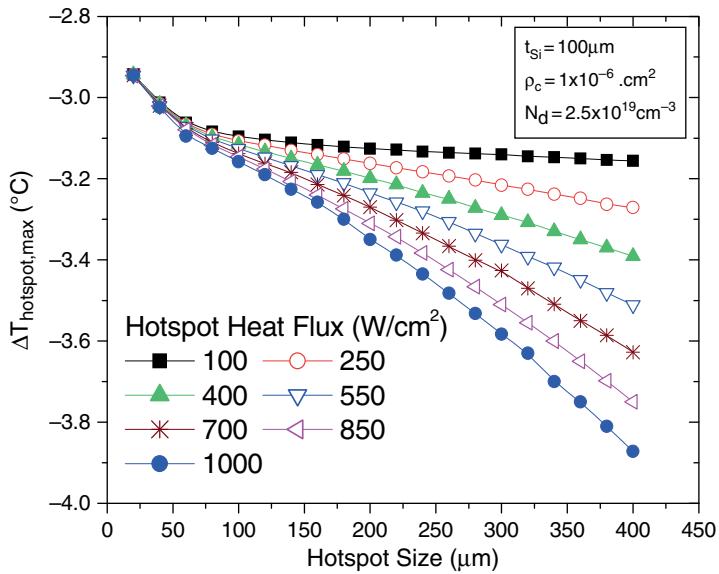


Fig. 32 Hot-spot temperature reduction as a function of hot-spot size and hot-spot heat flux

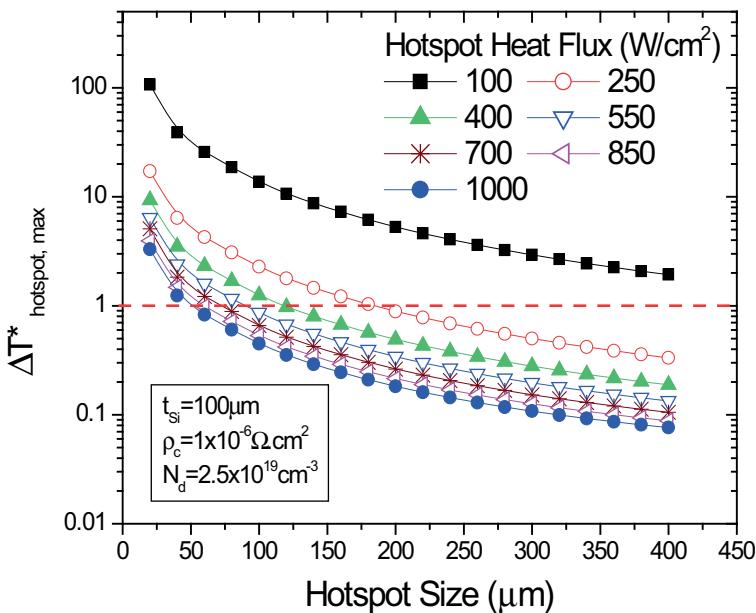


Fig. 33 Hot-spot cooling effectiveness as a function of hot-spot size and hot-spot heat flux

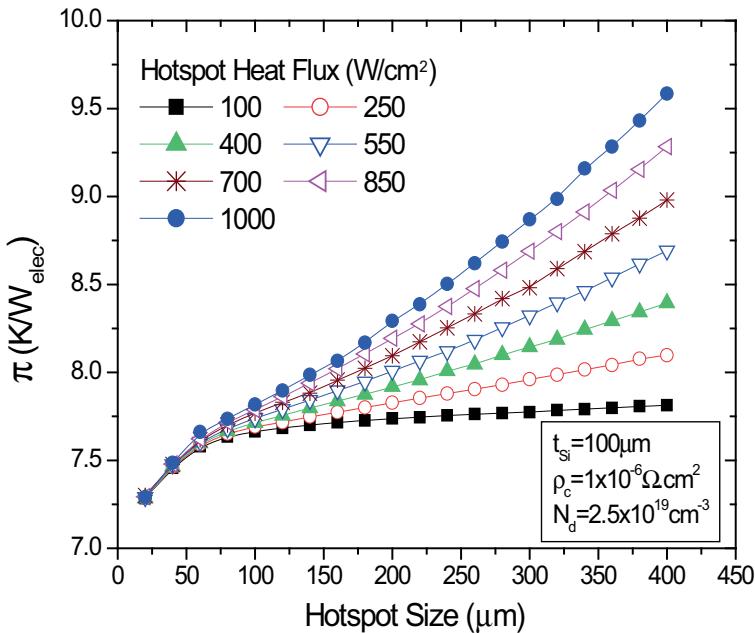


Fig. 34 Thermal impact factor as a function of hot-spot size and hot-spot heat flux. The microcooler size is $600 \times 600 \mu\text{m}$

2.4 Mini-Contact-Enhanced TEC for Hot-spot Cooling

Solid-state thermoelectric coolers (TECs), which are highly reliable, can be locally applied for spot cooling, and can be integrated with IC processing, have been proposed for hot-spot thermal management. However, the relatively low cooling heat flux, $5\text{--}10 \text{ W/cm}^2$, of conventional TEC modules severely limits the application of these devices to hot-spot remediation. Recently, a novel use of a mini-contact pad, which connects the TEC and the silicon chip, thus concentrating the thermoelectric cooling power on the top of the silicon chip to significantly improve hot-spot cooling performance, was proposed and investigated [106]. The physical phenomenon underpinning the use of mini-contact-enhanced thermoelectric coolers is displayed in Fig. 35, where the mini-contact is seen to concentrate the thermoelectric cooling power on the reduced cross-sectional area of the mini-contact tip. It can be expected that, to a first approximation, the smaller the cross-sectional area of the mini-contact tip, the larger the cooling flux on the top of the silicon chip. Moreover, the localization of the TEC cooling flux to the region most affected by the hot spot reduces the overall cooling requirements and the input power needed to effectively utilize the TEC.

To analyze and optimize the on-chip hot-spot cooling performance of the mini-contact-enhanced TEC, a three-dimensional numerical thermal model was

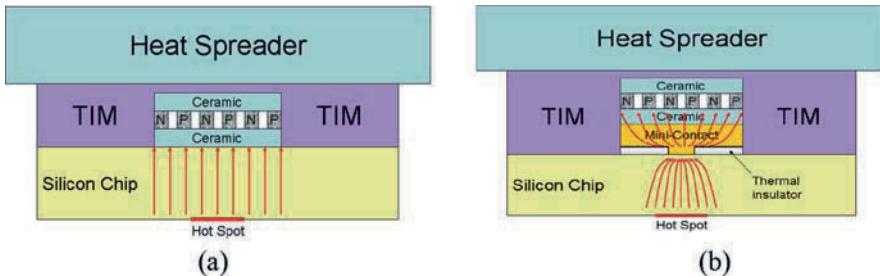


Fig. 35 Schematic of a TEC (consisting of numerous N-type and P-type TE elements sandwiched between two ceramic substrates) attached on silicon chip and embedded inside thermal interface material (TIM) of the chip package: (a) TEC without a mini-contact and (b) TEC with an integrated mini-contact. The arrows indicate the heat flow pattern in silicon chip and mini-contact. The detailed chip package is not shown in this figure

developed using the commercial finite element software ANSYSTM and applied to a typical chip package with a mini-contact-enhanced TEC, as shown in Fig. 36. This package consists of a silicon chip, thermal interface materials (TIM), a copper-integrated heat spreader (IHS), an air-cooled aluminum heat sink, and a miniaturized bismuth telluride-based TEC. The TEC is integrated with a mini-contact pad, attached on the top of the silicon chip, and then embedded inside the thermal interface materials, TIM1 and TIM2. These could be different thermal interface materials such as solder or thermal grease. The TEC consists of a 4×4 array of $400 \times 400 \times 20 \mu\text{m}$ thermoelectric elements that are sandwiched between two $50 \mu\text{m}$ thick ceramic substrates, and is $120 \mu\text{m}$ in overall height [107]. The copper

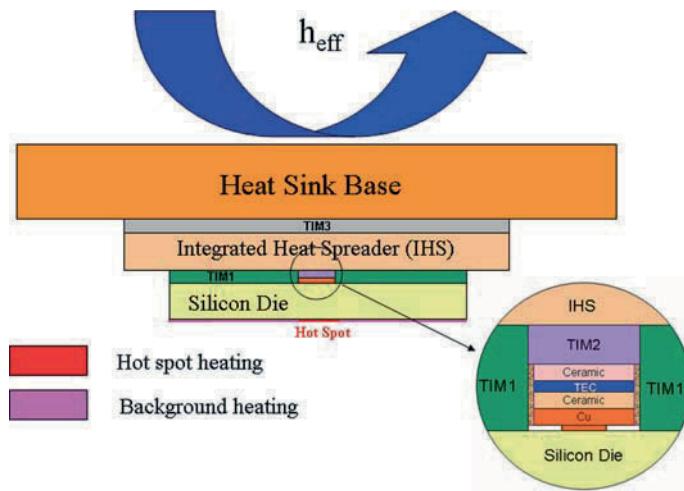


Fig. 36 Schematic of a typical chip package with a mini-contact-enhanced TEC

mini-contact pad features a 2.4 mm × 2.4 mm × 100 μm base to facilitate heat spreading, and a 2.4 mm × 2.4 mm × 50 μm tip to concentrate the thermoelectric cooling capability. For the purposes of this study, the thickness of TIM1 was held constant at 300 μm so that the mini-contact enhanced TEC and the TIM1 layer could be accommodated within the height of the TIM1 thermal interface layer.

When a mini-contact-enhanced TEC is integrated into the chip package, it introduces several thermal interfaces. In consideration of possible assembly procedures for such an enhanced TEC, it is assumed that the two most important thermal interfaces occur between the top ceramic substrate and the TIM2, and between the mini-contact tip and the silicon die, as indicated in Fig. 36. R_{c1} and R_{c2} are used to represent these two thermal contact resistances, with each varying from 1×10^{-7} to 1×10^{-4} $\text{m}^2 \text{K/W}$, the typical range reported for electronic package application [108, 109]. In the simulation, thermoelectric cooling rate, $Q_{\text{TE cooling}}$, is determined as the product of the Seebeck coefficient, temperature, and current, ST_cI , where S is the Seebeck coefficient of the thermoelectric material, T_c the cold-side junction temperature at the TEC, and I the applied electrical current on the TEC. Similarly, the Peltier heating rate at the heat rejection side of the device, $Q_{\text{TE heating}}$, is given by the product ST_hI , where T_h is the hot-side junction temperature at the TEC. Thermoelectric cooling and heating effects are represented as heat flux boundary conditions in the numerical simulation and directly added to the cold and hot sides of the TEC, respectively, while Joule heating is modeled as uniform volumetric heat generation inside the bismuth telluride elements. Joule heating from the electrical contact resistance is modeled as a surface boundary condition at the two TEC junctions [110]. Since the hot-spot cooling performance is strongly dependent on the input power supplied to the TEC, in the course of this simulation, various electric currents are applied to the TEC until the lowest hot-spot temperature is achieved and usually it takes about 10 min to complete a simulation using Pentium IV processor.

2.4.1 Effect of Input Power on TEC

Thermoelectric cooling performance is dependent on the applied current or input power to the TEC in a nonlinear manner, as Peltier cooling has a favorable linear dependence on electric current, while the parasitic Joule heating effect has a quadratic dependence on electric current. The competition of these two opposite contributions leads to an optimum current or input power at which the maximum hot-spot cooling, or lowest hot-spot temperature, can be achieved. Figure 37 demonstrates the variation of the hot-spot temperature with the input power to the TEC, for a TE element thickness of 20 μm and the mini-contact tip size of 1250 × 1250 μm . The thermal contact resistance is chosen to be 1×10^{-7} Km^2/W at both the mini-contact tip/silicon chip interface and the ceramic/TIM2 interfaces. For a 400 × 400 μm hot spot with a heat flux of 1250 W/cm^2 , if there is no TEC, the peak hot-spot temperature is found to reach 137.0°C. However, if the TEC is activated, the hot-spot temperature decreases steeply as the power increases, reaching a minimum of 120°C at approximately 10 W, which corresponds to a temperature reduction of

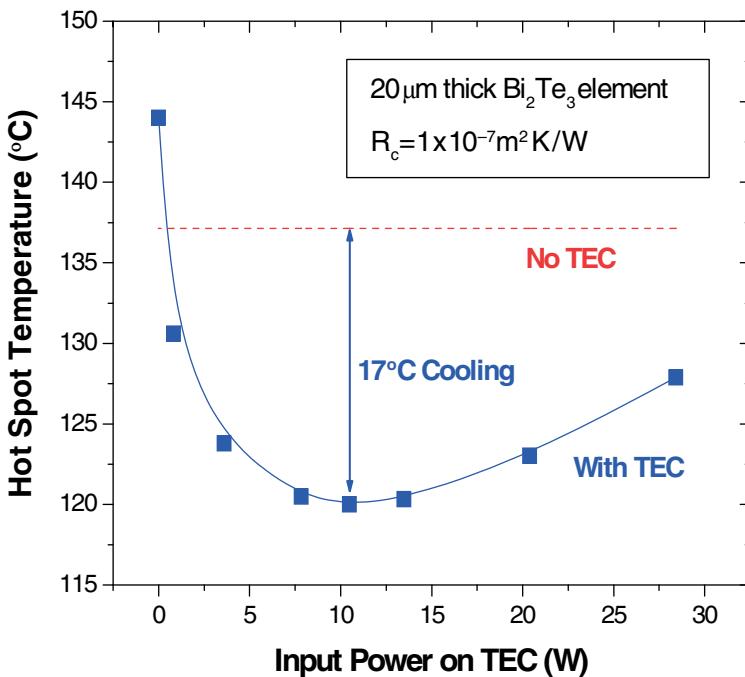


Fig. 37 Variation of typical hot-spot temperature with the input power on TEC (Thermoelectric element thickness = 20 μm , copper mini-contact tip size = $1250 \times 1250 \mu\text{m}$)

17.0°C at the hot spot compared to the temperatures encountered without the TEC, and then rises slowly as the power increases further. It is worth noting that if the TEC is present but not activated, the hot-spot temperature will increase by 7°C, due to the additional thermal resistance to heat flow created by the presence of the TEC.

2.4.2 Effect of Mini-Contact Size

The mini-contact pad, sandwiched between the TEC and the silicon chip, is used to concentrate the thermoelectric cooling rate on the top of the silicon chip. Its beneficial effect on hot-spot cooling could be limited by the heat-spreading resistance inside the mini-contact pad as well as inside the silicon chip. Consequently, care should be taken to optimize the geometric configuration to achieve the maximum hot-spot cooling performance. Figure 38 shows the temperature profiles achieved along a line bisecting the bottom of the silicon chip, with and without an embedded TEC, and revealing the characteristic “W”-shaped temperature profile created by the mini-contact-enhanced TEC cooler. It may be seen that if there is no hot spot on the chip and no TEC embedded inside the package, the peak chip temperature is about 109°C. However, if there is a $400 \times 400 \mu\text{m}$ hot spot with a heat flux of

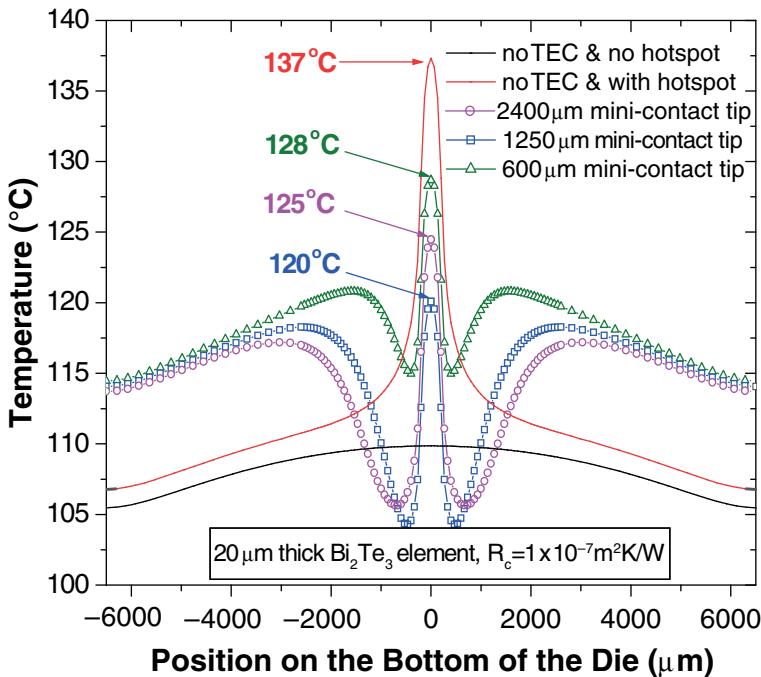


Fig. 38 Effect of mini-contact size on TEC-induced temperature profile (Thermoelectric element thickness = 20 μm , input power = 10 W)

1250 W/cm^2 , the peak chip temperature will increase to 137°C. Therefore, the concentrated heat flux leads to about a 28°C peak temperature rise on the chip. If the TEC with a thermoelectric element thickness of 20 μm is activated with a 10 W input and enhanced with a 600 \times 600 μm mini-contact pad, the temperature profile created around the high-flux region, shows a 9°C reduction in the hot-spot temperature (to 128°C), a low-temperature ring with temperatures below the uniform heat dissipation values, and a warm outer ring with slowly decaying temperatures in the radial direction. If the mini-contact tip grows to 1250 \times 1250 μm , the hot-spot temperature will reduce further, down to 120°C, resulting in 17°C maximum cooling at the hot spot. However, if we expand the mini-contact size further, to 2400 \times 2400 μm , the hot-spot cooling is limited to just 12°C. Obviously there exists an optimum mini-contact size for each configuration. The observed temperature increases in the outer ring or “side lobes” of the profile, as well as a very modest increase in the average chip temperature, is due to the additional power dissipation associated with the operation of the TEC device.

Figure 39 shows the heat flux distribution along the line bisecting the top surface of the silicon chip, just below the TEC enhanced with three different sizes of copper mini-contacts. On the top of the silicon chip and far away from the contact zone, the heat flux is approximately 70 W/cm^2 , the same as the background heat flux on

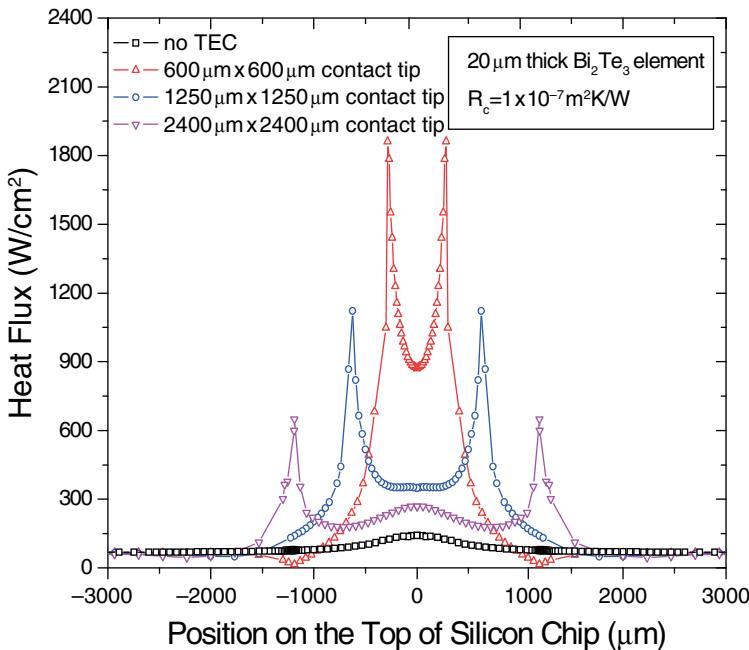


Fig. 39 Effect of mini-contact size on heat flux on the top of silicon chip (Thermoelectric element thickness = 20 μm , input power = 10 W)

the bottom of the silicon chip. However, at the interface between the mini-contact and the silicon chip, the heat flux increases significantly, indicating a strong cooling effect from the activated TEC. The heat flux averaged over the entire surface of the mini-contact/silicon chip interface increases from 250 to 640 W/cm^2 , and then to 1600 W/cm^2 when the copper mini-contact size decreases from $2400 \times 2400 \mu\text{m}$ to $1250 \times 1250 \mu\text{m}$, and then to $600 \times 600 \mu\text{m}$, suggesting that reduction of the mini-contact sizes can significantly increase the local cooling heat flux.

It is also interesting to note that the heat flux is highly nonuniform at the mini-contact/silicon chip interface, with the highest value occurring at the corner and the lowest value at the center. Although the cooling heat flux continues to increase with decreasing mini-contact size, the combined effect of the thermal resistance between the hot spot and the mini-contact area and this cooling flux results in the optimum mini-contact size previously shown in Fig. 38.

2.4.3 Effect of Thermoelectric Element Height

Thermoelectric element height is a key parameter for improving the hot-spot cooling performance as the maximum achievable cooling flux of the TEC is inversely proportional to the thermoelectric element thickness. Figure 40 illustrates the variation of hot-spot cooling with the mini-contact size for three different thermoelectric

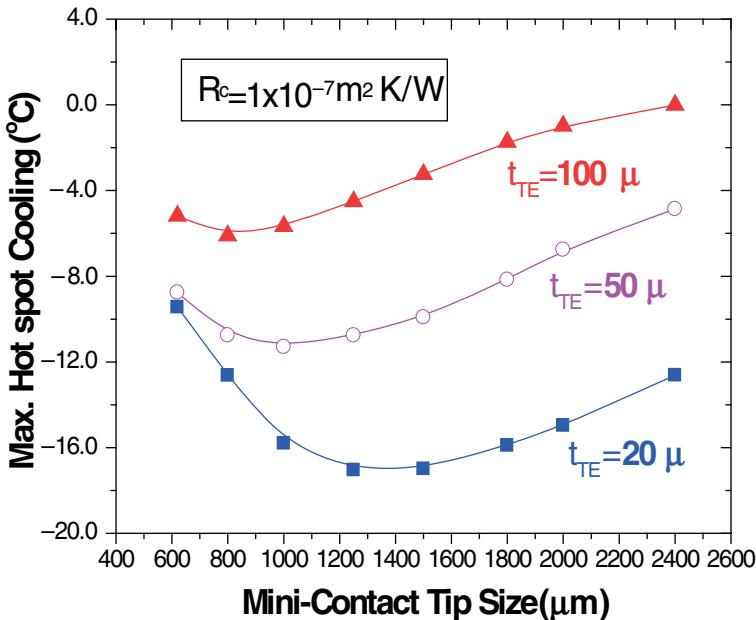


Fig. 40 Effect of thermoelectric element thickness t_{TE} on the maximum hot-spot cooling performance. The optimized input power of 10, 7.5, and 6.1 W is applied on the TEC with the TE element thickness of 20, 50, and 100 μm , respectively

element heights, under optimized input power on the TEC. As is expected, shorter thermoelectric elements allow the TEC to achieve better hot-spot temperature reductions, e.g., 6°C to 11.2°C and to 17.0°C as the thermoelectric element decreases from 100 to 50 μm and to 20 μm in thickness, using the optimum mini-contact tip size. Even when the mini-contact tip size is kept constant, shorter thermoelectric elements always yield better hot-spot cooling than longer elements due to the higher fluxes extracted by the cold side of the TEC. However, it is interesting to find that for fixed contact resistances the optimum mini-contact tip size increases with decreasing element height, from $800 \times 800 \mu\text{m}$ for a 100 μm thick element to $1000 \times 1000 \mu\text{m}$ for a 50 μm thick element, and to $1250 \times 1250 \mu\text{m}$ for a 20 μm high element.

As may be seen in Fig. 40, the TEC with 20 μm high elements has more dependence on the mini-contact size than the TECs with 50 μm or 100 μm high elements. It should be noted that the improvement provided by use of the mini-contact pad is larger with taller thermoelectric elements than with shorter elements. For the TEC with 20 μm thick elements, the addition of an optimally sized mini-contact pad improves the cooling by 4.3°C, from hot-spot cooling of 12.7°C with no mini-contact to 17°C with a $1250 \times 1250 \mu\text{m}$ mini-contact. However, for the TEC with 100 μm thick elements, the addition of an optimally sized mini-contact pad reduces the hot-spot temperature by an additional 6°C, from hot-spot cooling of 0.1°C with no mini-contact to 6.1°C with a $800 \times 800 \mu\text{m}$ mini-contact. Interestingly,

the deterioration in performance with suboptimum contact pads displays the reverse trend, with the hot-spot temperature rising steeply with reduced mini-contact area for the 20 μm thick TEC but more gradually for the 100 μm thick TEC.

2.4.4 Effect of Thermal Contact Resistance

Low thermal resistance interfaces are critical to mini-contact enhanced hot-spot cooling, since a high thermal resistance at the mini-contact/chip interface – where the cooling flux is highest – will significantly reduce the effectiveness of the mini-contact enhancement. Moreover, a bad thermal interface between the TEC and the TIM2 will impede the conduction of thermoelectric and Joule heat into the heat spreader and then into the heat sink and the ambient. Figure 41 displays the interplay between the thermal contact resistance and the achievable hot-spot cooling, with the assumption of equal thermal contact resistance at the two interfaces (e.g., $R_{c1} = R_{c2} = R_c$) and reveals that with increasing thermal contact resistance at both interfaces, the net cooling achievable on the hot spot diminishes. It may be seen that if the thermal contact resistance is 1×10^{-5} Km^2/W or higher, the hot-spot temperature will exceed 140°C and the embedded TEC will raise rather than lower the hot-spot temperature.

The thermal contact resistance also has an impact on optimized mini-contact size. As shown in Fig. 41, with increasing thermal contact resistance, the optimized mini-contact increases from $1250 \times 1250 \mu\text{m}$ for the thermal contact resistance of

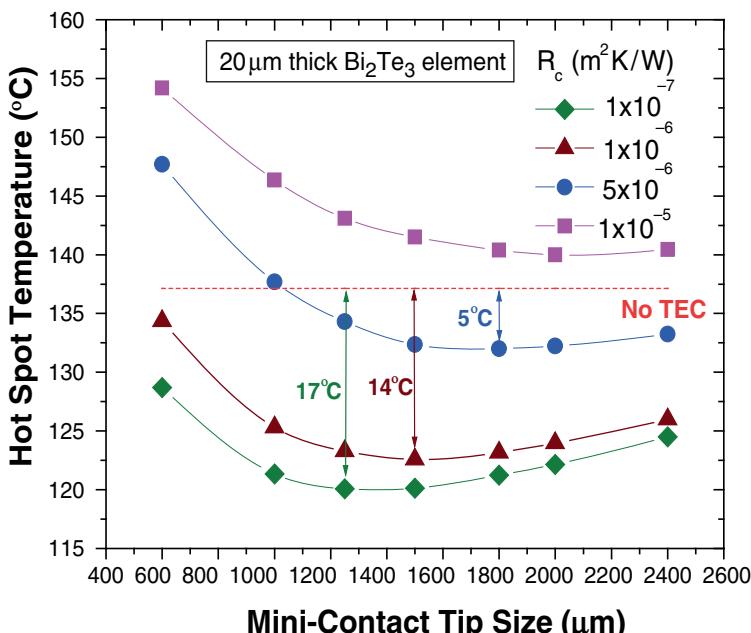


Fig. 41 Influence of thermal contact resistance on hot-spot cooling performance

1×10^{-7} Km²/W, representing a nearly perfect interface, to $2000 \times 2000 \mu\text{m}$ for a thermal contact resistance of 1×10^{-5} Km²/W, typical of thermal grease interfaces. It should be noted that if the thermal contact resistance is 5×10^{-5} Km²/W or lower, use of an optimized mini-contact tip always provides a lower hot-spot temperature than achieved without the mini-contact. However, the mini-contact is seen to provide diminishing returns as the contact resistances increase and to elevate the hot-spot temperatures for thermal contact resistances of 1×10^{-5} Km²/W or higher.

2.4.5 Experimental Demonstration

Thermal measurements were performed on the chip package test vehicle shown in Fig. 42 to quantify the spot cooling improvement provided by the mini-contact pad, and to determine its relationship to the TEC input power and power dissipation on the silicon chip. In this “proof of concept” experiment there are no micro-scaled hot spots and temperature sensors on the chip. Instead, a uniform heat flux is imposed on the bottom of the chip, by attaching four, thin-film heaters, and the mini-contact-enhanced TEC is used to locally cool that chip below the temperature of the surrounding silicon. The schematic of the experimental structure is shown in Fig. 42, displaying a $2.5 \text{ mm} \times 2.5 \text{ mm} \times 500 \mu\text{m}$ silicon chip attached to four, thin-film heaters used to simulate chip power dissipation. The copper mini-contact pad with various mini-contact tip sizes, varying from $0.8 \times 0.8 \text{ mm}$ to $3.6 \times 3.6 \text{ mm}$, was bonded onto the silicon chip using an indium-based solder. To facilitate the soldering process, a 200 nm thick Ni thin-film adhesion layer deposited on the silicon and then a 200 nm thick Au layer deposited onto the Ni layer by an e-beam process were used [111]. The copper mini-contact pad was then soldered onto the silicon by indium solder, which reacted with the Au thin film at around 160°C to form an AuIn₂ intermetallic compound.

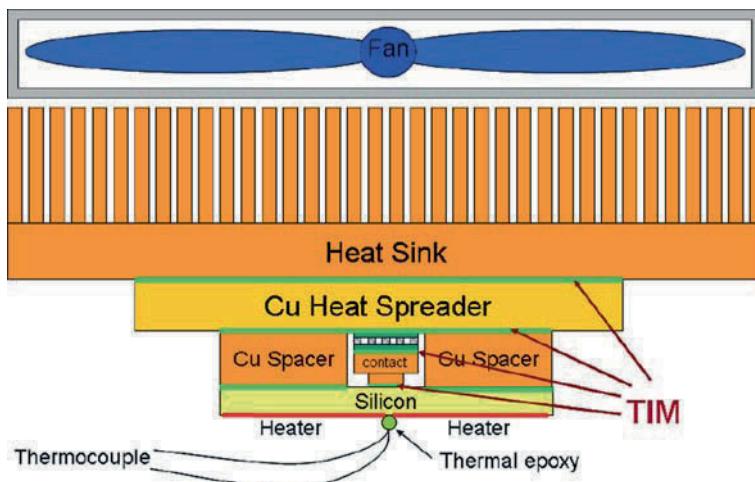


Fig. 42 Schematic of test vehicle for mini-contact-enhanced TEC for spot cooling

Miniaturized TECs from Thermion (model number: 1MC04-018-02-2200D) [112] with a dimension of $3.6 \times 3.6 \times 1.6$ mm and a total of 36 diced p-type and n-type 200 μm thick bismuth telluride thermoelectric elements, were used in these experiments. The thermal conductivity of bismuth telluride in the Thermion TEC is reported to be 1.3–1.4 W/m K, the Seebeck coefficient 200 $\mu\text{V/K}$, and the electrical resistivity 10 $\mu\Omega\text{ m}$, with a figure of merit value Z of $3 \times 10^{-3} \text{ K}^{-1}$ [113]. The two ceramic substrates are made of AlN, each with a thickness of 635 μm , and indium-tin solder was pre-tinned to the end faces to facilitate solder connections. In the present experiment, the TEC was attached to the mini-contact using thermal grease. The copper mini-contact pad and the TEC are accommodated inside the copper spacer. Above the copper spacer and the TEC, the copper heat spreader was attached using thermal grease. The heat spreader was then attached to the air-cooled copper heat sink.

Figure 43 shows the experimentally determined dependence of spot cooling on the TEC input power for a 500 μm thick silicon chip, with the mini-contact tip size kept at 1.8×1.8 mm and the power dissipation on the silicon chip varying from 0 to 67 W. It is found that the temperature reduction at the targeted spot (the center of the silicon chip) varies parabolically with the TEC input power, reflecting the competing mechanisms of rapidly improving Peltier cooling at lower input powers (or current) and progressively more damaging Joule heating, as well as reverse heat conduction, at the higher input powers. In this test vehicle, the silicon chip, thus, experiences its largest value of spot cooling at an input power of 3.5–5.5 W.

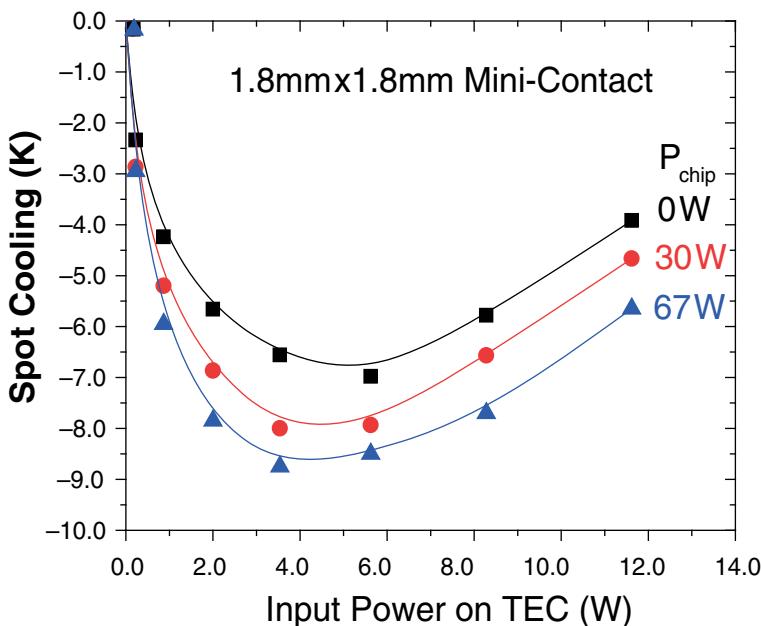


Fig. 43 Variation of measured spot cooling with TEC input power

It is interesting to note that the power dissipation on the silicon chip has some effect on spot-cooling performance and that higher chip power dissipation leads to greater achievable spot cooling on the silicon chip. For example, if there is no power dissipation on the silicon chip, a maximum cooling of about 7.0°C can be achieved. However, if the power dissipation on the chip is increased to 67 W, the maximum spot cooling increases to around 9°C . This improvement in cooling performance is related primarily to the increase in the cold-side junction temperature of the TEC, which raises the Peltier cooling rate. Interestingly, this Peltier cooling improvement also leads to lower values of the optimum TEC input power, since at the higher temperature more effective cooling can be achieved at lower current or power. This trend is illustrated in Fig. 8, where an increase of the power dissipation on the silicon wafer from 0 to 70 W, is seen to produce a decrease in the optimum TEC input power from 5.7 to 3.6 W.

The experimentally observed effect of the mini-contact tip size on the temperature reduction at the targeted spot is displayed in Fig. 44. For the three different power dissipations and a 500 μm thick chip, the maximum spot cooling is seen to display a parabolic dependence on the mini-contact tip size, showing very favorable improvements as the mini-contact tip size decreases in area from the “full coverage” limit, but ultimately reversing direction as the tip size shrinks below an optimum value and approaches point contact. The presence of an optimum tip size reflects

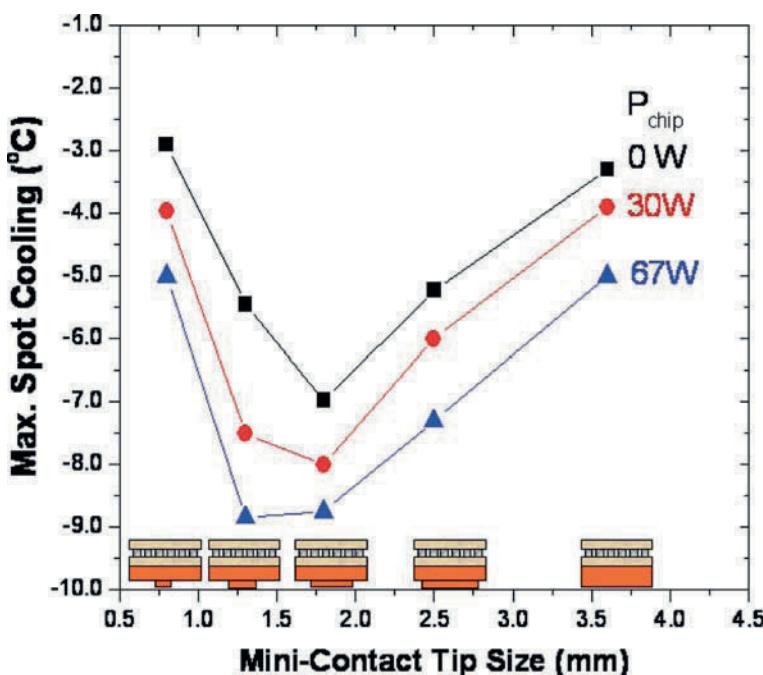


Fig. 44 Variation of measured maximum spot cooling on copper mini-contact size

the competing effects of the favorable cooling flux concentration and the parasitic spreading resistance in the mini-contact tip. As shown in Fig. 44, for the case of no power dissipation on the silicon chip, if the mini-contact is of the same size as the TEC base, the measured maximum spot cooling is about 3.3°C . However, if a 1.8×1.8 mm copper mini-contact is integrated onto the TEC, 7.1°C maximum spot cooling can be obtained, which results in 115% improvement on spot-cooling performance. Similarly, spot-cooling performance can be improved by 100 and 80% if the power dissipation of the silicon chip is 30 W and 67 W, respectively. It is interesting to note that the power dissipation on the silicon chip has an impact on the optimized mini-contact size and the larger the power dissipation on the silicon chip, the smaller the optimized mini-contact size. As clearly illustrated in Fig. 44, as the power dissipation on the silicon chip increases from 0 to 67 W, the optimum mini-contact size decreases from 1.8×1.8 mm to 1.3×1.3 mm.

2.5 Applications in Biomedical Systems

Temperature control is one of the most essential operations in many biomedical systems [114]. Active temperature control is capable of accurately and rapidly bringing the intended temperature above, below, or in between preset limits. In particular, in biomedicine thermal control systems are encountered frequently, for example, in electrophysiology [115], radiopharmaceutical synthesis [116], microbial studies [117], rapid thermal cycling of cells, [118] and DNA sequencing [119]. Wijngaards et al. developed a concept of an active micro-thermostat system using an integrated thin-film poly-SiGe thermoelectric cooler, capable of stabilizing the temperature of a suspended structure at ambient temperature and above. This opens up the way to a large number of applications that need to be thermally controllable in the range of $10\text{--}50^{\circ}\text{C}$, especially in the field of biomedicine.

Precise fluid temperature control in microfluidic channels is a requirement for many lab-on-a-chip and microreactor devices, especially in biotechnology where most processes are highly temperature sensitive. Microheaters integrated with microfluidic channels have been proposed but are of limited use for temperature control, as they can only be used for raising the fluid temperature. In addition, they are limited in creating large thermal gradients, in the order of 10 K/mm , due to thermal conduction in the substrate. One feasible method that scales well on the microscale is thermoelectric cooling. Maltezos demonstrated the concept of a Bi_2Te_3 -based thermoelectric microcooler integrated into a microfluidic channel in order to give rapid and localized fluid cooling [120]. They reported an on-chip thermoelectric refrigerator and heat exchanger for microfluidic devices. The microfluidic chamber was cycled between -3 and over 120°C , thus spanning water freezing and boiling, and the entire PCR temperature range. For smaller chambers, it was shown that it is possible to cool reagent from room temperature to freezing within $10\text{--}20\text{s}$, and to obtain relatively good temperature stability ($<\pm 0.2^{\circ}\text{C}$) over long periods of time. The ability to localize heating and cooling in microfluidic chambers and channels enable massive parallelization of chemical reactions

in which the temperature of each reaction vessel can be independently controlled. Thus, these thermal management systems enable the fabrication of complex chip-based chemical and biochemical reaction systems in which the temperature of many processes can be controlled independently.

3 On-Chip Hot-Spot Cooling Using Anisotropic Heat Spreader

Anisotropic heat spreaders, with a high in-plane thermal conductivity, provide a very promising passive approach for hot-spot remediation and constitute a very attractive component for active cooling of nanoelectronic chips. This anisotropy can be used to laterally spread the hot spot heat to cooler regions of the chip and thus can significantly reduce the hot-spot temperature. Recently, Bar-Cohen et al. used analytical and numerical models to investigate the feasibility of hot-spot remediation using the lateral heat-spreading capability of a thermal interface material (TIM) with orthotropic thermal conductivity [121]. It was found that when used together with existing cooling solutions, such materials, bonded directly to the silicon chip as shown in Fig. 45, can substantially reduce the temperature rise associated with a severe flux spot and more uniformly distribute the heat at the interface between the TIM and the next element in the heat-transfer path.

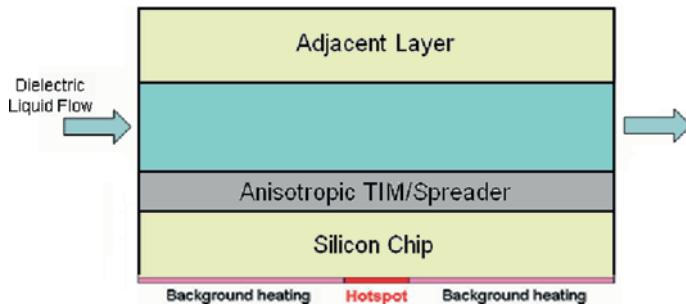


Fig. 45 Implementation of an anisotropic TIM/Spreader in a three-dimensional stack up

In their work, the bi-layer slab shown in Fig. 46 was used to investigate the hot-spot remediation provided by an orthotropic spreader attached directly to the back of a square chip with a single, centrally located, square flux spot. All external surfaces are assumed adiabatic, except for a heat flux boundary condition at the flux spot and a convective boundary condition on the back of the orthotropic spreader. Background heating on the active side of the chip is forgone because only the hot-spot temperature rise is sought. The boundary conditions are such that inclusion of background heating would simply elevate the entire temperature field. This effect becomes nontrivial if temperature-dependent material properties and heat-transfer coefficients are employed, but all properties and heat-transfer coefficients are taken to be constant in this analysis. The convective boundary condition represents the

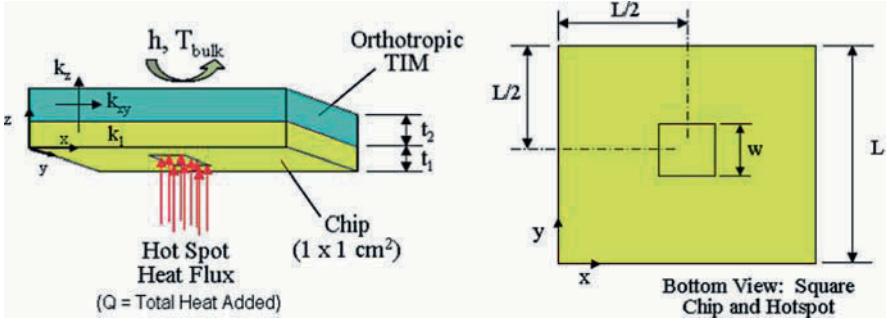


Fig. 46 Schematic of compound chip/spreader system

influence of the “global” cooling scheme and is modeled by a uniform heat-transfer coefficient on the back of the spreader.

Following Muzychka’ work [122], an analytical solution was developed using the separation of variables method and the excess temperature on the active side of the chip ($z = 0$) is given by the following equation:

$$\Delta T_{\text{bulk}}(x, y, z = 0) = A_0 + \sum_{m=1}^{\infty} A_m \cos(\lambda_m x) + \sum_{n=1}^{\infty} A_n \cos(\delta_n y) \dots \\ + \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{mn} \cos(\lambda_m x) \cos(\delta_n y) \quad (27)$$

The eigenvalues in Equation (27) are given by $\lambda_m = \frac{m\pi}{L}$, $\delta_n = \frac{n\pi}{L}$, and $\beta_{mn} = \sqrt{\lambda_m^2 + \delta_n^2}$ and the Fourier coefficients, A_0 , A_n , A_m , and A_{mn} , are found from the application of the boundary conditions in the z -direction. Imposing the boundary conditions at $z = 0$ results in expressions for the “A” coefficients as follows:

$$A_0 = \frac{Q}{L^2} \left(\frac{t_1}{k_1} + \frac{t_2}{k_2} + \frac{1}{h} \right) \quad (28)$$

$$A_m = \frac{2Q \left[\sin\left(\frac{L+w}{2}\lambda_m\right) - \sin\left(\frac{L-w}{2}\lambda_m\right) \right]}{L^2 w k_1 \lambda_m^2 \phi(\lambda_m)} \quad (29)$$

$$A_n = \frac{2Q \left[\sin\left(\frac{L+w}{2}\delta_n\right) - \sin\left(\frac{L-w}{2}\delta_n\right) \right]}{L^2 w k_1 \delta_n^2 \phi(\delta_n)} \quad (29)$$

$$A_{mn} = \frac{16Q \cos\left(\frac{\lambda_m L}{2}\right) \sin\left(\frac{\lambda_m w}{2}\right) \cos\left(\frac{\delta_n L}{2}\right) \sin\left(\frac{\delta_n w}{2}\right)}{L^2 w^2 k_1 \beta_{mn} \lambda_m \delta_n \phi(\beta_{mn})} \quad (29)$$

The parameter, ϕ , appearing throughout Equation (29), is a spreading parameter that is a function of a dummy variable ζ , with $\phi(\zeta)$ given by

$$\varphi(\zeta) = \frac{(\alpha e^{4\zeta t_1} - e^{2\zeta t_2}) + \psi [e^{2\zeta(2t_1+t_2)} - \alpha e^{2\zeta(t_1+t_2)}]}{(\alpha e^{4\zeta t_1} + e^{2\zeta t_2}) + \psi [e^{2\zeta(2t_1+t_2)} + \alpha e^{2\zeta(t_1+t_2)}]} \quad (30)$$

and ζ is replaced by λ_m , δ_n , or β_{mn} in Equation (30) as appropriate. The new parameters in Equation (30) are to be evaluated as $\psi = \frac{\zeta+h/k_2}{\zeta-h/k_2}$ and $\alpha = \frac{1-k_2/k_1}{1+k_2/k_1}$ where ζ is again replaced by λ_m , δ_n , or β_{mn} as appropriate. The conductivity of the second layer, k_2 , is assumed to be isotropic in the above expressions. However, as suggested in the literature, if either of the layers exhibits orthotropic conductivity, the solution for purely isotropic layers can be used when the following length scale and conductivity transformations are employed in the subject orthotropic layer:

$$\begin{aligned} k &\rightarrow k_{\text{eq}} = \sqrt{k_{xy}k_z} \\ t &\rightarrow t_{\text{eq}} = \frac{t}{\sqrt{k_z/k_{xy}}} \end{aligned} \quad (31)$$

Thus, k_2 and t_2 in Equations (28), (29) and (30) can be replaced by the transformations in Equation (31) to account for the anisotropicity in the spreader.

Equations (27), (28), (29), (30), and (31) provide a full solution for the excess temperature on the active side of the compound structure shown in Fig. 46. The overall resistance to heat transmission for the system shown in Fig. 46 is comprised of (1) the resistance to one-dimensional heat flow and (2) the spreading resistance. The first term in Equation (27) is the Fourier coefficient, A_0 , which is given by Equation (28) and is attributable to uniform one-dimensional conduction through the compound bi-layer slab. The three remaining terms are related to thermal spreading and thus vanish as the hot-spot size approaches the chip size. The total thermal resistance can be related to the average excess temperature at the hot spot through the following definition:

$$R_T = \frac{\overline{\Delta T_{\text{bulk}}}}{Q} \quad (32)$$

where R_T is the total thermal resistance, including thermal transport by both conduction and convection. The term $\overline{\Delta T_{\text{bulk}}}$ in Equation (32) is found by integrating Equation (27) over the flux-spot region and dividing by the flux-spot area, or, expressed mathematically:

$$\overline{\Delta T_{\text{bulk}}} = \frac{1}{A_s} \iint_{A_s} \Delta T_{\text{bulk}}(x, y, 0) dA_s \quad (33)$$

As noted, the total thermal resistance is also the sum of the one-dimensional resistance and the spreading resistance as follows:

$$R_T = R_{1D} + R_s \quad (34)$$

The one-dimensional resistance to heat conduction is easily found to be:

$$R_{1D} = \frac{t_1}{k_1 L^2} + \frac{t_2}{k_z L^2} + \frac{1}{h L^2} \quad (35)$$

Meanwhile, the spreading resistance, R_s , can be found by substituting Equations (32), (33), and (35) into Equation (34):

$$R_s = \frac{1}{A_s Q} \iint_{A_s} \Delta T_{\text{bulk}}(x, y, 0) dA_s - \left(\frac{t_1}{k_1 L^2} + \frac{t_2}{k_z L^2} + \frac{1}{h L^2} \right) \quad (36)$$

The integration in Equation (36) yields the following expression for the spreading resistance in a bi-layer structure with a centrally located flux spot:

$$R_s = \frac{1}{2 (w/2)^2 (L/2)^2 k_1} \left[\sum_{m=1}^{\infty} \frac{\sin^2(w\delta_m/2)}{\delta_m^3 \varphi(\delta_m)} + \sum_{n=1}^{\infty} \frac{\sin^2(w\lambda_n/2)}{\lambda_n^3 \varphi(\lambda_n)} \right] \dots \\ + \frac{1}{(w/2)^4 (L/2)^2 k_1} \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \frac{\sin^2(w\delta_m/2) \sin^2(w\lambda_n/2)}{\delta_m^2 \lambda_n^2 \beta_{mn} \varphi(\beta_{mn})} \quad (37)$$

where the eigenvalues are the same as those for Equation (27) and the parameter φ is given by Equation (30).

3.1 Effect of In-Plane Spreader Thermal Conductivity

Successful hot-spot remediation via an orthotropic spreader depends on the ability of the spreader to conduct heat laterally from local regions of high heat flux to other parts of the chip with lower thermal loads and is most directly influenced by its in-plane thermal conductivity, k_{xy} . This conclusion is supported by the thickness and conductivity transformations in Equation (31), which clearly show that any increase in k_{xy} for fixed values of k_z and t will result in larger values k_{eq} and t_{eq} , through the indicated square root dependence. Therefore, an increase in the conductivity ratio, k_{xy}/k_z , leads to an attendant decrease in the overall thermal resistance, and thus a decrease in the average hot-spot temperature.

A representative chip/spreader system with the parameter settings listed in Table 6 was used to determine the magnitude of the benefits of increasing k_{xy} . Please note that, unless otherwise specified, the parameters in Table 6 will be used throughout this section. The thru-plane conductivity of the spreader, k_z , was chosen to be 5 W/m K because this is a nominal thru-plane conductivity for some natural graphite materials as well as pyrolytic graphite [123]. A heat-transfer coefficient of 10,000 W/m² K was applied to represent the presence of an aggressive cooling approach (e.g., pool boiling or a microchannel cold plate) [124]. The in-plane conductivity was varied between 5 and 1800 W/m K in order to determine the effect that the degree of anisotropy had on hot-spot remediation.

Table 6 Parameter settings for k_{xy} variation

Parameter	Description	Value	Units
t_1	Chip thickness	250	μm
t_2	Spreader thickness	500	μm
k_1	Isotropic chip conductivity	163	W/m K
k_z	Thru-plane spreader conductivity	5	W/m K
k_{xy}	In-plane spreader conductivity	Varies	W/m K
q''	Hot-spot heat flux	1.4	kW/cm ²
h	Effective heat-transfer coefficient	10,000	W/m ² K
T_{bulk}	Ambient temperature for convective transfer	25	C
L	Chip size, square	1	cm
W	Hot-spot size, square	500	μm

The excess temperature profiles on the active side of the chip, subjected to a 1.4 kW/cm², 0.5 mm² flux spot, are shown for several different in-plane conductivities in Fig. 47. The Gaussian-like temperature profile for the isotropic spreader ($k_z = k_{xy} = 5$ W/m K) is seen to yield a hot-spot temperature that is nearly 47.5 K above ambient. Use of the orthotropic spreaders is seen to produce similar profiles but to considerably reduce the peak temperature at the center of the hot spot and over the adjacent silicon, while modestly elevating the temperature (~1°C), in the edge regions of the chip. As expected, increasing the in-plane conductivity, k_{xy} , decreases the excess temperature. Thus, for $k_z = 5$ and $k_{xy} = 350$, which is representative of natural graphite sheets, the hot-spot temperature is ~9.3°C below that obtained through use of the isotropic spreader. If the in-plane conductivity is further increased to 1800 W/m K, a hot-spot suppression of ~14.3°C is attained.

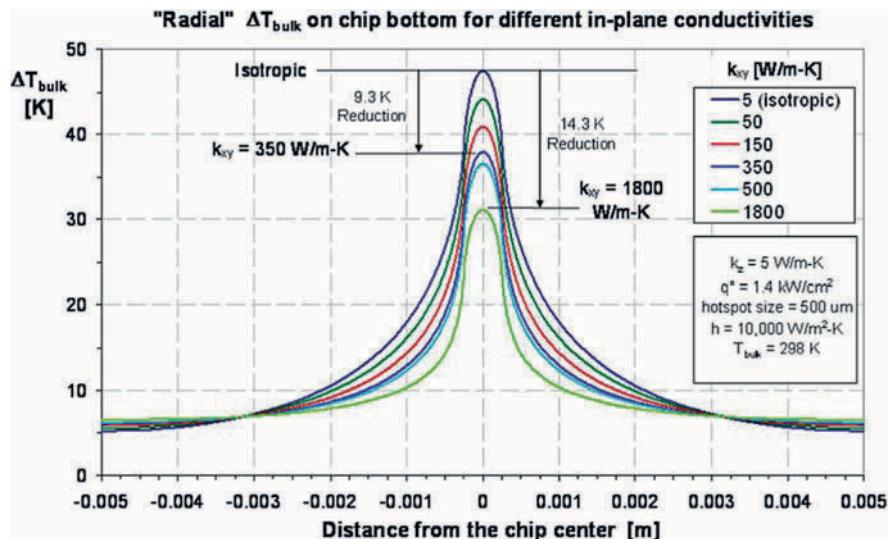


Fig. 47 Excess temperature profiles taken through the middle of the active chip surface for various k_{xy}

The variation of hot-spot remediation with the in-plane thermal conductivity of an orthotropic heat spreader, relative to a low-conductivity isotropic spreader, is shown in Fig. 48. The peak temperature reduction is seen to asymptotically approach 15°C as the in-plane conductivity increases towards 1800 W/m K. In order to put these results into context, the performance of several spreaders with varying conductivities are evaluated for the same geometric parameters and thermal boundary conditions, as listed in Table 6. With the isotropic spreader ($k = 5$ W/m K) as a baseline, the hot-spot cooling achieved by each alternative spreader is shown on the right side of Fig. 48. The first data point corresponds to an isotropic spreader with conductivity of 160 W/m K, representing the use of a thick silicon chip. It is found that increasing the thickness of the silicon provides approximately 18.3°C of cooling, which exceeds by 4°C the best hot-spot suppression achieved. However, it can be seen in Fig. 48 that an orthotropic spreader with $k_z = 10$ W/m K and $k_{xy} = 1700$ W/m K – which is characteristic of annealed pyrolytic graphite (APG) [125] can provide the same hot-spot suppression as the thicker silicon chip.

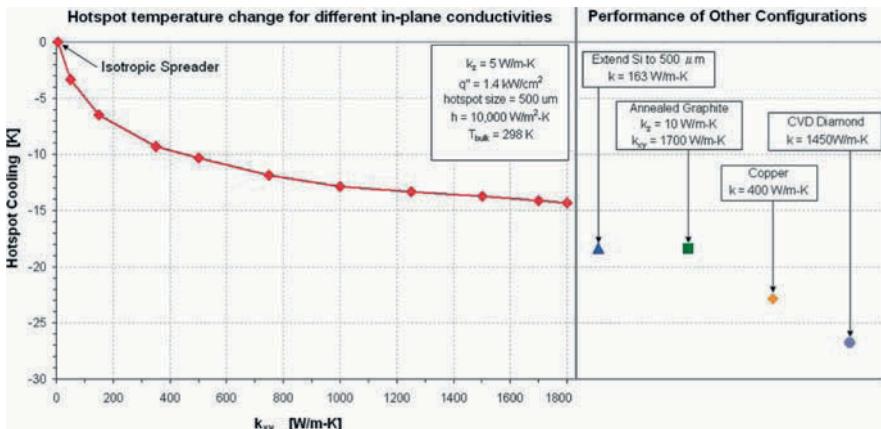


Fig. 48 Variation of hot-spot cooling with in-plane thermal conductivity in an orthotropic heat spreader

Implementation of a copper spreader with isotropic thermal conductivity of 400 W/m K provides further hot-spot suppression, reaching a temperature reduction of $\sim 23^\circ\text{C}$. Of the alternative spreaders considered, the best hot-spot suppression of $\sim 26.6^\circ\text{C}$ was provided by isotropic CVD diamond film with a thermal conductivity of 1450 W/m K in all three directions.

The above results show that increasing the in-plane thermal conductivity of an orthotropic spreader can provide substantial hotspot temperature reduction. For the parameters considered, the cooling performance of a highly orthotropic APG spreader matches the cooling provided by an equal thickness of pure silicon. This is an important result given the general reluctance of chip manufacturers to allocate valuable semiconductor-grade silicon for thermal management functions. Despite the good performance of the best highly orthotropic spreaders, an equally sized

copper spreader provides about 4°C better hot-spot remediation for the conditions examined. However, natural graphite orthotropic spreaders can be made extremely pliable and offer a weight advantage over copper spreaders of equal size since the density of natural graphite is approximately just 25% that of copper. This weight difference could be significant in weight-constrained mobile applications. Furthermore, highly orthotropic spreaders with low thru-plane conductivity have the ability to reduce hot-spot temperatures while simultaneously insulating adjacent layers of a chip stack.

Figure 49 shows the temperature rise of the silicon, copper, and APG spreaders as a function of location on the back of the spreader. The excess temperature displayed in Fig. 49 is the difference between the local temperature and the edge temperature; thus, the value of ΔT vanishes for all profiles as the edge of the spreader is approached. For the parameters considered it can be seen that the silicon and copper spreaders allow a maximum temperature variation of 8°C and 4.3°C on the back of the spreader, respectively, compared to a maximum variation of 0.05°C for the APG spreader. Consequently, in hybrid cooling systems – using two-phase cooling with an anisotropic spreader attached to the chip – an APG spreader would offer only a modest temperature rise on the rear of the spreader and yield a system that is less susceptible to local dryout or critical heat flux.

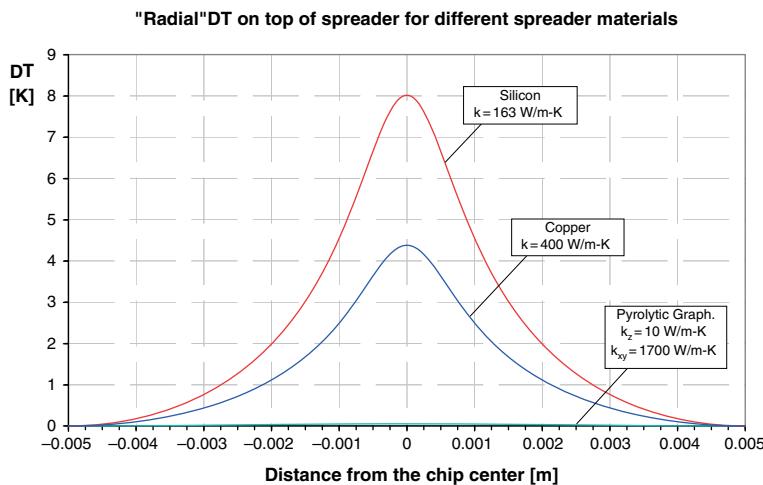


Fig. 49 Temperature rise on the back of silicon, copper, and APG spreaders

3.2 Variation of Spreader Thickness

The thickness of the orthotropic TIM/spreader will not only establish the magnitude of hot-spot cooling, but will also determine the viability of this thermal solution for volume-constrained three-dimensional stacks of hot-spot-laden chips. It is, therefore, important to explore the tradeoffs between spreader thickness and

cooling performance when assessing the merits of an anisotropic spreader. The impact of spreader thickness on hot-spot remediation is best understood in terms of the overall thermal resistance of the system, R_T , which is the sum of the spreading resistance, R_s , and the resistance to one-dimensional conduction and convection, R_{1D} . MATLAB codes were developed to aid in the evaluation of R_{1D} (Equation (35)) and R_s (Equation (37)) for various spreader thicknesses.

Using the model in Fig. 46 and the parameter settings in Table 6, the thickness of each of the spreaders represented in Fig. 48 was varied to determine the effect on R_{1D} , R_s , and hence R_T . Typical trends seen during this analysis are depicted in Fig. 50 where it is clear that the total thermal resistance generally experiences a minimum for some critical value of the spreader thickness. Near a spreader thickness of zero, the total thermal resistance of the system approaches that of a single layer of silicon. But, as the thickness increases, the one-dimensional resistance, R_{1D} , grows and the spreading resistance, R_s , falls leading to a nonmonotonic variation in the overall thermal resistance. When the negative slope of R_s equals the positive slope of R_{1D} the minimum thermal resistance, and hence minimum average hot-spot temperature, is attained. For any increase in spreader thicknesses beyond this optimum, the linear rise in R_{1D} is greater than the decrease in R_s and the total thermal resistance increases.

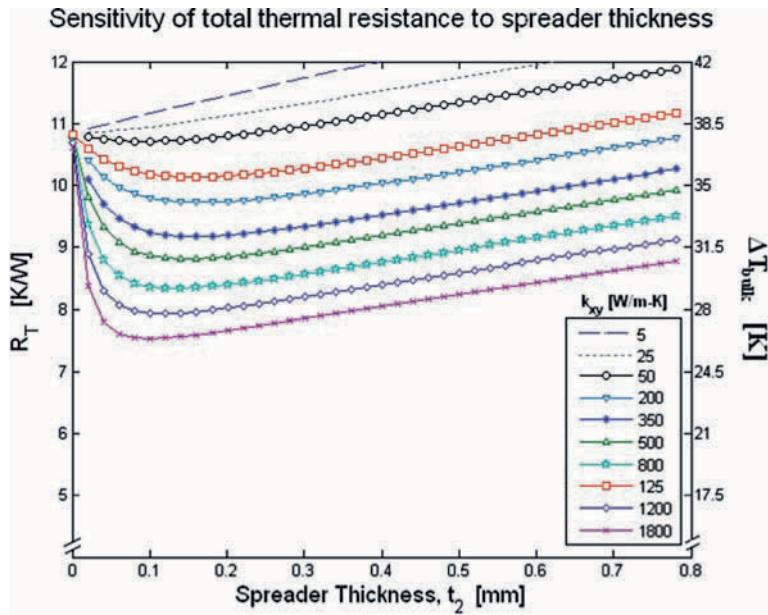


Fig. 50 Variation of R_T and average hot-spot excess temperature for increasing spreader thickness ($k_z = 5$ W/m K)

The resulting variation in overall thermal resistance for a spreader with a thru-thickness conductivity of 5 W/m K is shown in Fig. 50 for a range of in-plane

conductivities. As anticipated, the total thermal resistance for each of the k_{xy} curves approaches the resistance of a directly cooled chip, i.e., $R_T \rightarrow 10.84 \text{ K/W}$, as spreader thickness approaches zero. For increasing values of the in-plane conductivity (above 5 W/m K), the total thermal resistance and average hot-spot excess temperature generally decrease and display an optimum value. The specific optimum thickness varies with in-plane conductivity as shown in Fig. 51, exhibiting a peak optimum thickness of approximately 160 μm in the vicinity of $k_{xy} = 200 \text{ W/m K}$ and of approximately 100 μm at a thermal conductivity of 1800 W/m K. However, as seen for the dashed lines where $k_{xy} = 5$ and 25 W/m K, for low values of the in-plane conductivity, there is no optimum thickness and R_T continuously increases. For the parameters considered, it was found that the development of a monotonically increasing total thermal resistance occurs for $k_{xy} \leq 14.5 \text{ W/m K}$.

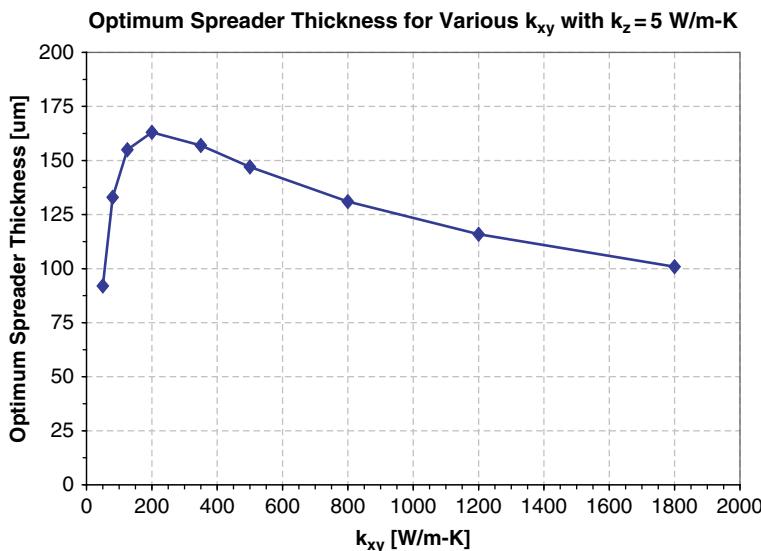


Fig. 51 Optimum spreader thickness varies with in-plane thermal conductivity

The variation of R_T with the TIM/spreader thickness, t_2 , for each of these alternatives can be seen in Fig. 52. It is found that the highly orthotropic APG exhibits similar behavior to that shown in Fig. 50, with a distinct minimum occurring at a spreader thickness of 157 μm , for the stated conditions. However, the silicon, copper, and diamond spreaders all exhibit a broad “plateau” for which the thermal resistance remains relatively constant with thickness (these spreaders do indeed have minimum values of R_T , but the minima occur beyond the 2 mm thickness at which plotting was stopped in Fig. 52).

Figures 51 and 52 reveal that highly orthotropic graphite TIMs/spreaders yield optimum hot-spot cooling performance at relatively low thicknesses – under 165 μm – for the conditions examined. Furthermore, it is interesting to note in Fig. 52 that the orthotropic APG spreader yields lower average hot-spot temperatures

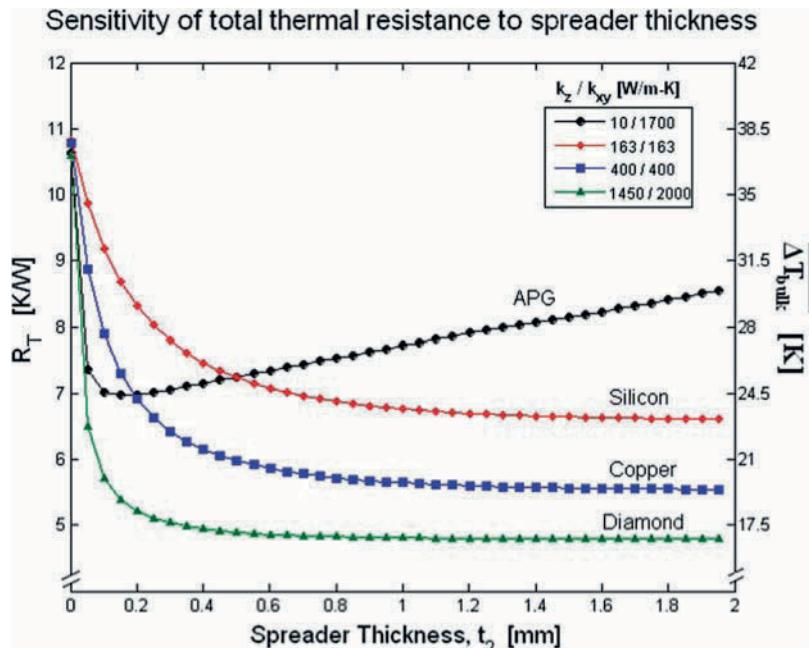


Fig. 52 Variation of R_T for increasing spreader thickness for alternative spreaders

than copper for thicknesses up to $\sim 200 \mu\text{m}$ and lower temperatures than silicon up to $\sim 500 \mu\text{m}$. Also, the minimum average hotspot excess temperature for APG is 24.4 K at 157 μm , which is only 5.0°C and 1.3°C hotter than that provided by nine times the thickness of copper and silicon (~ 1.4 mm), respectively. The exceptional performance of highly orthotropic TIMs/spreaders at low thickness may lead them to be favored over conventional heat-spreading materials in space constrained 3D chip stacks.

Up to this point, the thickness of the silicon chip in Fig. 46 has been fixed at 250 μm for all cases. The variation of R_T with spreader thickness for a spreader with $k_z = 5 \text{ W/m K}$ and $k_{xy} = 350 \text{ W/m K}$, where each of the plotted line represents a different chip thickness (all other parameters remain unchanged, see Table 6) is shown in Figure 53. The total thermal resistance is seen to decrease with increasing chip thickness. However, it is clear that the additional hot-spot cooling provided by an optimally thick spreader becomes less dramatic for greater chip thicknesses. This is more clearly shown in Fig. 54 where the R_T data for each plotted line in Fig. 53 has been normalized by the total thermal resistance that would exist if the spreader were removed and the bare chip were cooled directly (this resistance is called $R_{T,bare}$).

It is seen that an optimally thick spreader reduces $R_{T,bare}$ by $\sim 43\%$ when the chip is 75 μm thick but reduces $R_{T,bare}$ by only $\sim 7\%$ when the chip is 400 μm thick. The reduction in spreader effectiveness for increasing chip size is the result of more

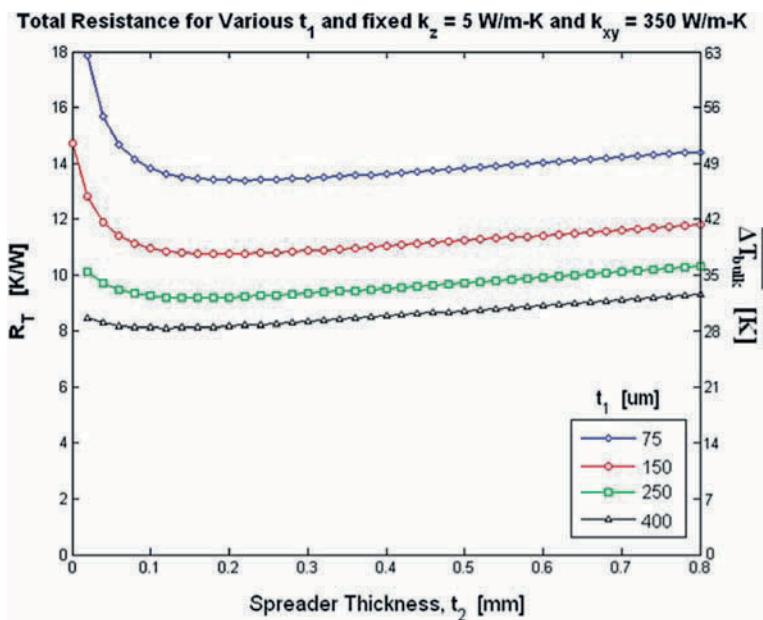


Fig. 53 Effect of chip thickness of spreader thickness variation for an orthotropic spreader with $k_z = 5 \text{ W/m K}$ and $k_{xy} = 350 \text{ W/m K}$

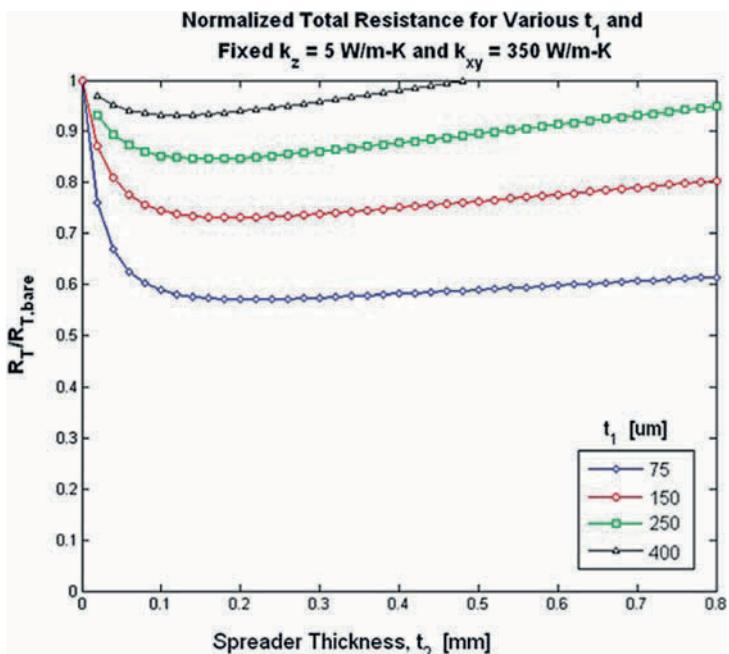


Fig. 54 Normalized total resistance for different chip thicknesses and an orthotropic spreader with $k_z = 5 \text{ W/m K}$ and $k_{xy} = 350 \text{ W/m K}$

effective heat spreading in the thicker silicon, which reduces the role played by the orthotropic spreaders. Alternatively, as chip thicknesses shrink – the more likely scenario as chip manufacturers strive to more efficiently utilize valuable silicon ingots and package electronics in thinner packages – the orthotropic TIM/spreaders can compensate for the loss of inherent spreading in these thinner silicon chips of the future.

Figure 54 also reveals that increasing chip thicknesses are accompanied by a steady decrease in the optimum spreader thickness for a given k_{xy} . In order to better understand this variation, the plot in Fig. 51 was reproduced for different values of chip thickness, t_1 . The results can be seen in Fig. 55 and it is clear that smaller chip thicknesses yield a larger optimum spreader thickness for a given k_{xy} . Also, thicker chips yield lower sensitivity of optimum thickness to in-plane conductivity, as evidenced by the suppression of the peak optimum thickness in Fig. 55 for larger values of t_1 .

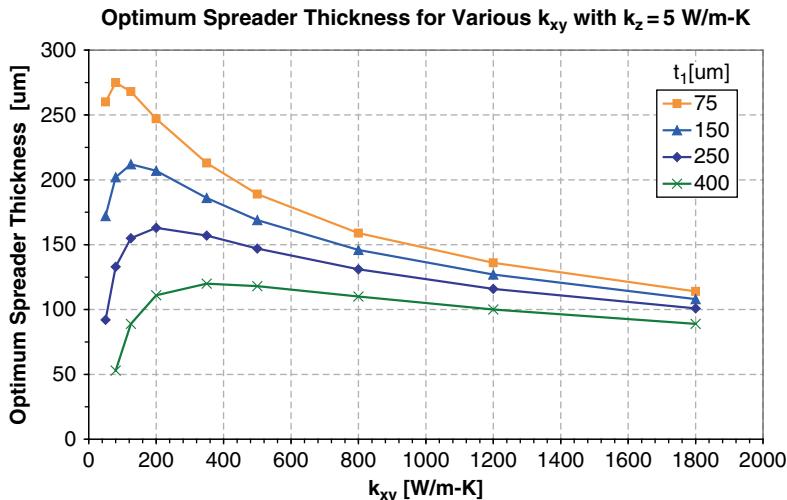


Fig. 55 Change in optimum spreader thickness for various k_{xy} and t_1

3.3 Numerical Simulations and Contact Resistance Variation

The parametric results above indicate that use of an orthotropic spreader is a promising approach to reducing hot-spot temperatures. However, it may be anticipated that the physical attachment of the orthotropic material to the back of the chip, as depicted in Fig. 46, may well result in the creation of a potentially significant and deleterious thermal contact resistance. Typical contact resistances for electronic packaging applications are reported to be in the range of

10^{-7} – 10^{-4} K m²/W. The thermal contact resistances of $\sim 10^{-7}$ K m²/W are representative of an excellent interface achieved by monolithic growth or eutectic interface attachment, while resistances of $\sim 10^{-4}$ K m²/W represent a relatively poor thermal interface achieved through the use of phase change materials and elastomeric pads [126]. While it is, thus, desirable to assess the impact of thermal contact resistance at the chip/spreader interface on hot-spot remediation, to the authors' knowledge, there is no analytical solution that explicitly accounts for this effect in a layered structure. Consequently, the parametric sensitivity of hot-spot temperature to contact resistance is explored through numerical simulations with ANSYS for contact resistances varying from 0 to 10^{-4} K m²/W.

Figure 56 shows the resulting heat flux vectors near the flux spot in the silicon chip for each contact resistance, with the results for perfect thermal contact on the top and poor thermal contact on the bottom, for the previously described conditions. Comparing the two heat flux plots reveals the subtle influence of contact resistance on the flow of heat in the system. In the case of perfect thermal contact, the heat flux vectors display a large thru-plane component, reflecting the relative ease with which heat can flow across the interface and into the spreader. Alternatively, the heat flux vectors for poor thermal contact exhibit a larger in-plane component, reflecting the additional resistance to heat flow across the interface. The contact resistance thus acts to more evenly distribute the heat flux imposed on the spreader, thereby reducing the spreader's effectiveness. Ultimately, the presence of the contact resistance results in larger peak and average temperature rises at the flux spot (these peak and average are 61.9 and 57.7°C for the perfect interface, respectively, with the poor thermal interface resulting in 96.2 and 91.6°C peak and average temperatures, respectively).

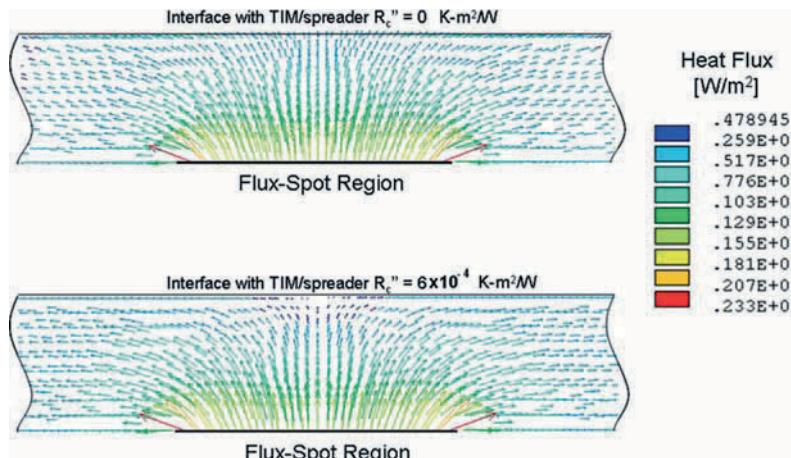


Fig. 56 Heat flux plots in the silicon chip for perfect and poor thermal contact

With the expectation that hot-spot temperatures should increase for escalating contact resistance, a total of 78 ANSYS simulations were run for contact resistances ranging from 0 to 10^{-4} K m²/W, with model parameters defined in Table 6. Figure 57 depicts the increase in hot-spot temperature over the stated contact resistance range for differing degrees of spreader anisotropy. The reader is reminded that 10^{-7} K m²/W is a very low contact resistance that may be achieved by monolithic growth on the back of the chip or through the use of a soldered interface. Alternatively, a poor interface, such as a lightly loaded interface with a phase change material or elastomeric pad, is represented by a contact resistance near 10^{-4} K m²/W.

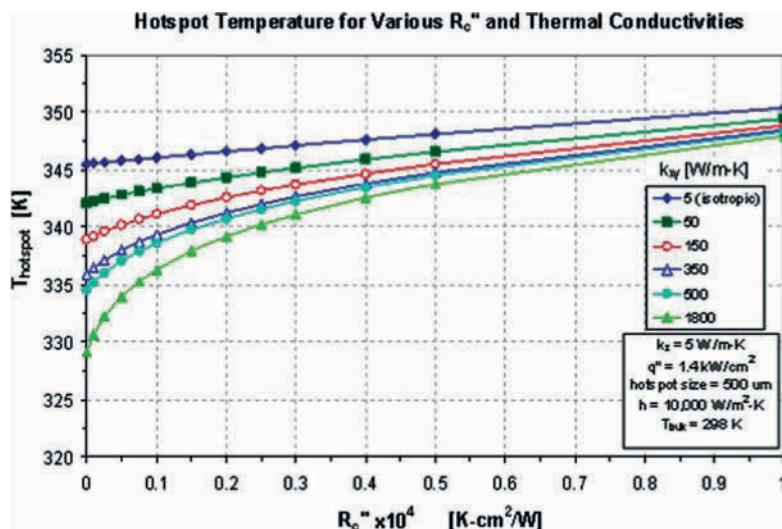


Fig. 57 Hot spot vs. contact resistance for various k_{xy}

For the conditions studied, it is found that the contact resistance has a significant effect on hot-spot temperature, particularly when extreme anisotropy is present in the spreader. In order for a spreader with $k_z = 5$ W/m K and $k_{xy} = 1800$ W/m K to provide at least 10°C better hot-spot cooling than the isotropic spreader, the contact resistance must be maintained below 0.1×10^{-4} K m²/W. However, even for a contact resistance of 0.5×10^{-4} K m²/W, a nearly 5°C temperature reduction can be achieved by the best orthotropic material and 3 or 4°C for more commonly available graphite TIM/spreaders.

3.4 Experimental Demonstration

Graphite has an anisotropic crystal structure that results in different properties in different directions. Its in-plane thermal conductivity can range from 140 to 1650 W/m K, while its thru-thickness thermal conductivity is much lower and ranges

between 2 and 10 W/m K. Based on the analytical and numerical results described in the previous section, it would appear that such anisotropy can be used advantageously to both reduce the severity of a hot spot and to spread the heat in the TIM, thus shielding a surface adjacent to the heat source from a high heat flux. Recently, Xiong et al. experimentally demonstrated that a thin graphite heat spreader can reduce the hot spot significantly [127]. In his experiment, two graphite heat spreaders with different in-plane thermal conductivities and thicknesses were tested and compared. The first graphite material had an in-plane thermal conductivity of 425 W/m K, a thru-thickness thermal conductivity of 3.2 W/m K and a thickness of 110 μm . The second material was an experimental grade of graphite with an in-plane thermal conductivity estimated to be greater than 1000 W/m K, a thru-thickness thermal conductivity estimated between 5 and 6 W/m K, and a thickness of 30 μm . A $50 \times 70 \times 1$ mm acrylonitrile butadiene styrene (ABS) plate was used as the substrate. A thin graphite heat spreader was adhesively bonded to the ABS substrate and a constant heat flux was applied over a limited area on one side, while the other side of the ABS substrate was exposed to the ambient for natural convection cooling. Three cases were investigated in their work:

- Case 1: without a heat spreader and with a power of 0.75 W on the hot spot.
- Case 2: with a 425 W/m K, 110 μm -thick graphite heat spreader and with a power of 3.0 W on the hot spot.
- Case 3: with a 1000 W/m K, 30 μm -thick graphite heat spreader and with a power of 3.0 W on the hot spot.

Figure 58 shows the IR thermal image of the ABS substrate without an attached heat spreader, at a power dissipation of only 0.75 W on the hot spot. Because of the low thermal conductivity of the ABS (~ 0.25 W/m K), the heat from the hot spot did not spread well and thus there is a localized hot spot immediately above the heat source and the temperature drops dramatically a short distance away from

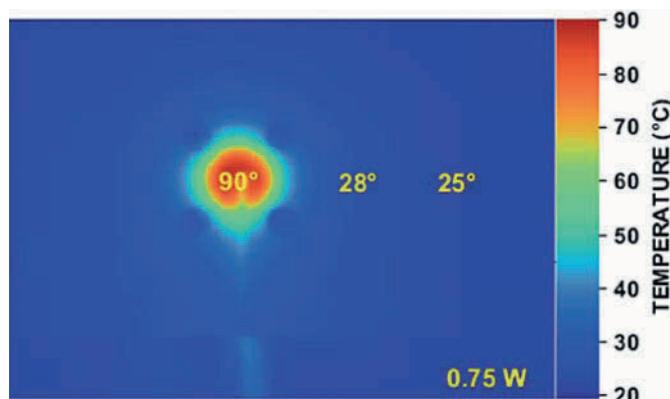


Fig. 58 Surface temperature distribution on the ABS substrate without heat spreader (0.75 W was applied on the hot spot)

the heat source. The hot-spot temperature is 90°C and most of the spreader is at a temperature less than 30°C.

Figure 59 is a thermal image of the ABS with an attached 425 W/m K, 110 μm -thick graphite heat spreader, while Fig. 66 is a thermal image of the ABS plate with a 1000 W/m K, 30 μm -thick graphite heat spreader. Note that in both cases, the excellent spreading effect of graphite spreader allowed the hot spot power to be increased to 3.0 W. These results demonstrate the significant heat-spreading effect of graphite heat spreaders. As shown in Figs. 59 and 60, with much higher power applied on the hot spot, the hot-spot temperatures are only 64 and 69°C respectively, much lower than that of the ABS without the spreader as shown in Fig. 58. The

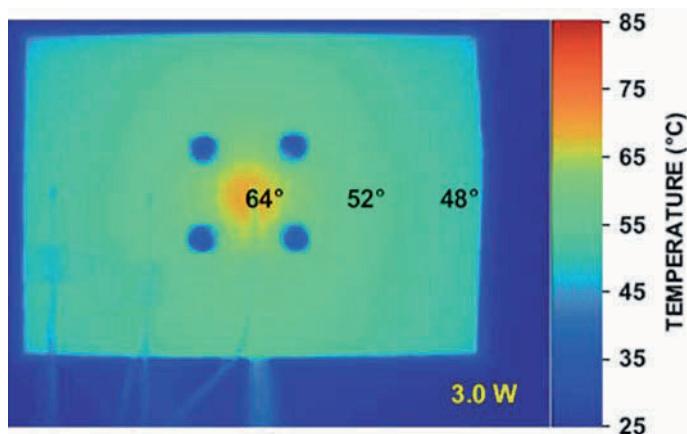


Fig. 59 Surface temperature distribution on an ABS substrate with an attached 425 W/m K graphite heat spreader (3.0 W was applied on the hot spot)

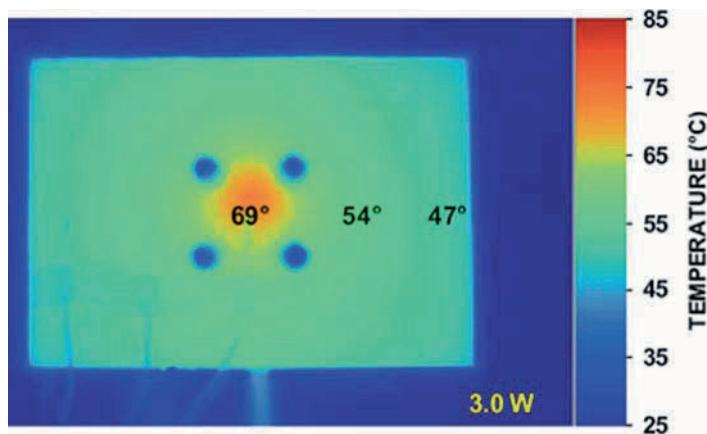


Fig. 60 Surface temperature distribution on an ABS substrate with an attached 1000 W/m K graphite heat spreader (3.0 W was applied on the hot spot)

temperature gradients across the surfaces are also much lower with the entire ABS plate now above 47°C. These results demonstrate that a very thin spreader made of graphite, with an in-plane thermal conductivity on the order of 500–1000 W/m K, or other highly orthotropic materials, is a practical thermal solutions for hot-spot remediation in both a passive cooling mode and an active cooling mode.

4 On-Chip Hot-Spot Cooling Using Micro-Gap Cooler

As is made abundantly clear in the previous sections of this chapter, the thermal interface resistance between the “cooling solution” and the chip poses a considerable barrier to effective remediation of on-chip hot spots. Direct liquid cooling techniques, which allow for direct contact between an inert dielectric liquid and the surface of the chip and eliminate the TIM (thermal interface material), hold great promise for hot-spot-driven thermal management of ICs. Moreover, use of phase-change processes, including pool boiling, gas-assisted evaporative cooling, jet impingement, and spray cooling, exploit the latent heat of these liquids to reduce the required mass flow rates and can provide the added advantage of inherently high heat-transfer coefficients.

However, direct cooling of microelectronic components imposes stringent chemical, electrical, and thermal requirements on the liquids to be used in this thermal control mode. Direct liquid cooling of microelectronic components requires compatibility between the liquid coolant and a system-specific combination of the chip, chip package, substrate, and printed circuit board materials, e.g., silicon, silicon dioxide, silicon nitride, alumina, o-rings, plastic encapsulants, solder, gold, and epoxy glass. In addition, a liquid coolant must possess the dielectric strength needed to provide electrical isolation between adjacent power/ground conductors and signal lines. Fortunately, 3 M’s family of perfluorinated liquids possess the required attributes and has been used extensively in the electronic industry [128, 129].

Recently, Bar-Cohen et al. proposed the use of micro-gap coolers to achieve a volume-efficient application of direct liquid cooling, while providing high heat-transfer coefficients – in the range needed to control the temperature of on-chip hot spots – on the back of the chip [130]. As shown in Fig. 61, in a micro-gap

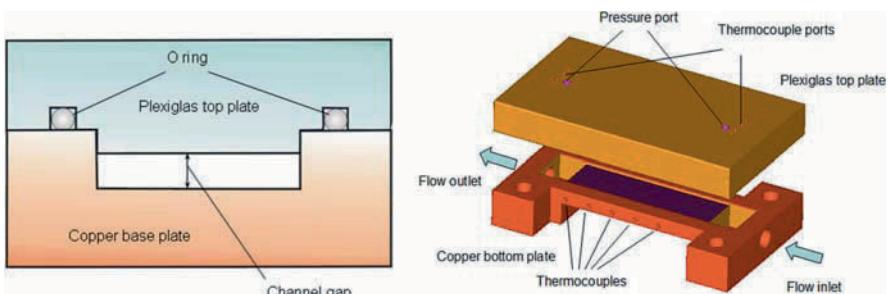


Fig. 61 Schematic diagram of micro-gap cooler

cooler a narrow sub-millimeter channel is created above the chip or substrate and liquid is pumped through the channel, thus removing the heat dissipated from the chip. Compared with a more conventional microchannel cooler, which needs to be attached with a TIM to the chip/substrate, micro-gap coolers require no attachment and no micromachining and could be a very attractive cooling approach for on-chip hot-spot remediation. Subsequent subsections deal with the empirical validation of this novel thermal management approach.

4.1 Single-Phase Experiments

Following single-phase heat-transfer experiments, with water and FC-72 as the working fluid, and successful comparison to the predicted heat-transfer coefficients obtained with established correlations and CFD simulation using IcePak [131], the researchers turned their attention to two-phase thermal transport in the microgap channel. The test results for the $210\text{ }\mu\text{m}$ gap two-phase heat-transfer experiment are shown in Fig. 62. In these experiments, the mass flux was varied from 130 to $660\text{ kg/m}^2\text{ s}$ in steps of $130\text{ kg/m}^2\text{ s}$ and the corresponding inlet fluid velocity was $0.0794, 0.159, 0.238, 0.317, 0.397\text{ m/s}$. The average heat-transfer coefficient was found to generally display a parabolic variation with the heat flux, increasing toward a peak value as the channel condition changed from subcooled to saturated flow boiling and then decreasing with further increases of heat flux.

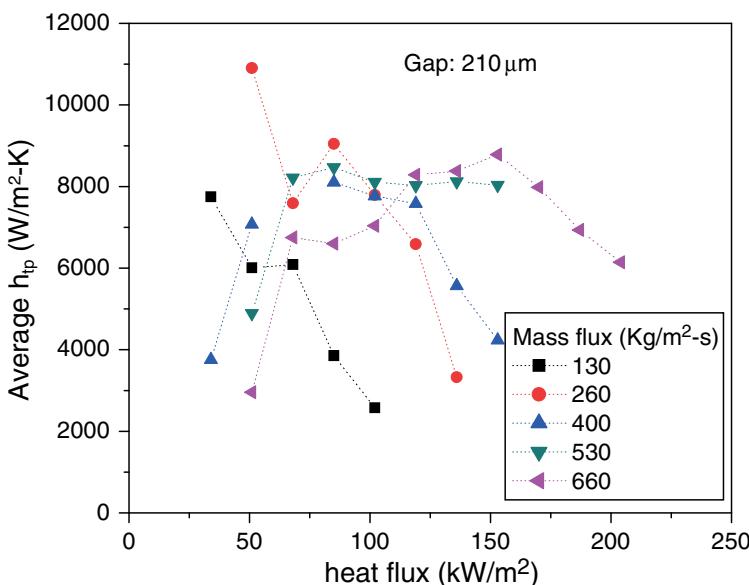


Fig. 62 Average heat-transfer coefficient vs. heat flux of $210\text{ }\mu\text{m}$ -gap cooler

The 110 μm gap channel shows some similarities to the heat-transfer coefficient variation observed in the 210 μm gap but – as shown in Fig. 63 – it appears to possess only the down-sloping branch of the previously observed parabolic trend. The peak heat-transfer coefficients attained in the 110 μm gap channel are higher than the 210 μm gap or the 500 μm gap channels, and the downward trend with heat flux moderates substantially as the mass flux increases.

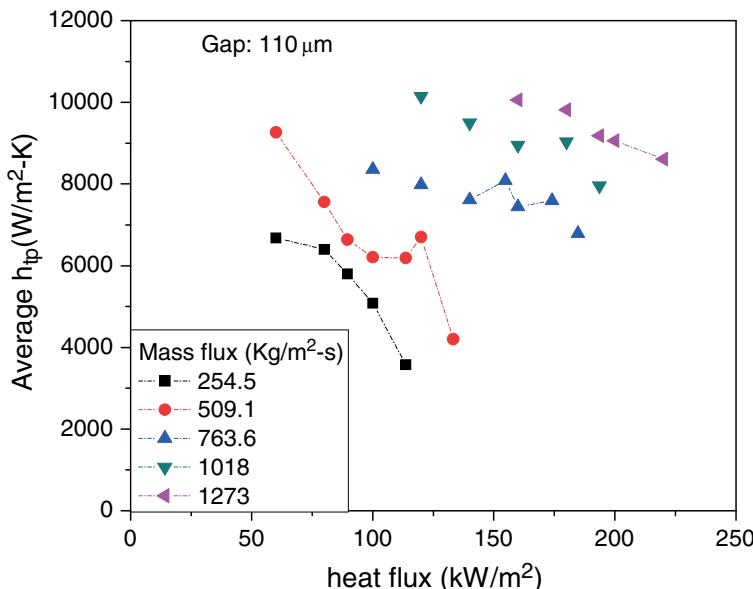


Fig. 63 Average heat-transfer coefficient vs. heat flux for 110 μm -gap cooler

The above and other reported results show that the area-averaged heat-transfer coefficients for FC-72, of between 7.5 and 15.5 $\text{kW}/\text{m}^2 \text{ K}$, can be attained in micro-gap channels of 110–500 μm . These values are significantly higher than the single phase FC-72 values, showing that two-phase micro-gap cooler can provide the high cooling capability needed for hot spot remediation.

4.2 Application to Hot Spot Remediation

To assess the efficacy of micro-gap cooler for thermal management of hot spots, it is instructive to simulate the thermal performance of a notional advanced semiconductor chip cooled by the micro-gap coolers. A 10×10 mm silicon chip, 500 μm thick, dissipating a uniform heat flux of $100 \text{ W}/\text{cm}^2$ across nearly all the active chip area serves as the test vehicle for this simulation. In keeping with the theme of this Chapter, the chip is assumed to possess a central, circular hot spot, varying from 100 to 400 μm in diameter and dissipating between 1 and 2 kW/cm^2 . It

is further assumed that the thermal conductivity of the silicon chip is invariant at 125 W/m K, that it is cooled from the back surface (opposite to that of the active circuitry) with heat-transfer coefficients that can vary from 5 to 20 kW/m² K, reflective of the values that can be potentially achieved with micro-gap coolers and that the liquid temperature is 22°C.

Figure 64 presents the three-dimensional temperature profile while Fig. 65 depicts the temperature along a diagonal, on the active face of the silicon chip for a baseline microgap-cooled chip configuration with a 400 μm hot spot, generating a 2 kW/cm² heat flux. As observed in Figs. 64 and 65, when this notional

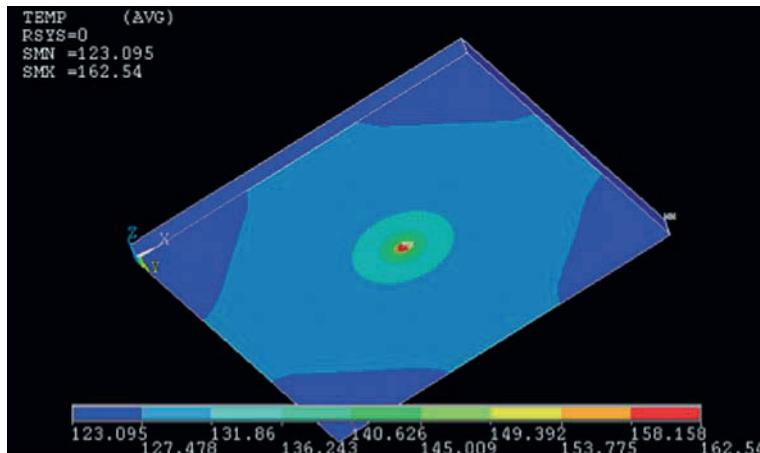


Fig. 64 Three-dimensional temperature profile for direct liquid cooling of advanced semiconductor chip [10 × 10 × 0.5 mm, $q'' = 100 \text{ W/cm}^2$, $q''_{hs} = 2 \text{ kW/cm}^2$, $d_{hs} = 400 \mu\text{m}$, $h = 10 \text{ kW/m}^2 \text{ K}$]

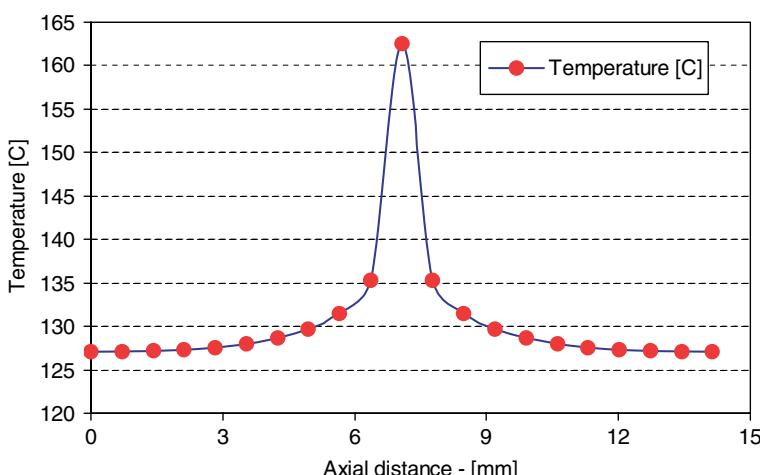


Fig. 65 Temperature distribution across the diagonal of the chip [$q_{chip} = 100 \text{ W/cm}^2$, $q_{spot} = 2000 \text{ W/cm}^2$, $h = 10000 \text{ W/m}^2 \text{ K}$, $T_c = 22^\circ\text{C}$]

baseline chip, with a very severe hot spot, is cooled by a micro-gap cooler with an h equal to $10 \text{ kW/m}^2 \text{ K}$, it experiences an elevated average temperature of approximately 130°C and a significant hot spot with a maximum temperature of 163°C , or some 33°C above the average chip temperature. The average and peak temperatures for various other combinations of the specified parameters are shown in Tables 7 and 8.

Tables 7 and 8 present the results for a hot-spot diameter of $100 \mu\text{m}$ and $400 \mu\text{m}$, respectively, with various hot-spot heat fluxes and a range of heat-transfer coefficients associated with the micro-gap coolers. In these tables, the first and second columns present the hot-spot flux and convective coefficients, respectively, while the next three columns provide: the average temperature on the active side of the chip, the average temperature on the cooled (back) side of the chip, and the average hot-spot temperatures. The last column in Tables 7 and 8 presents the maximum temperature on the active side of the silicon chip. Not surprisingly, the average chip temperature (on both the active and wetted surfaces) is seen to vary directly with the heat-transfer coefficient, while the on-chip hot-spot temperature rise is conduction limited and – for the fixed chip geometry and thermal conductivity – is driven by the heat flux and size of the hot spot. Thus, as seen in Table 8, while raising the heat-transfer coefficient to $20 \text{ kW/m}^2 \text{ K}$ lowers the average chip temperature to 77°C for a chip heat flux of 100 W/cm^2 , the on-chip temperature rise for a 1 kW/cm^2 , $400 \mu\text{m}$ hot spot, remains at approximately 15°C across the range of heat-transfer coefficients from 5 kW to $20 \text{ kW/cm}^2 \text{ K}$. However, since the peak chip temperature

Table 7 Temperatures for $100 \mu\text{m}$ hot-spot diameter for various heat flux and cooling conditions

$q''_{\text{spot}} [\text{W/cm}^2]$	$h [\text{W/m}^2 \text{ K}]$	$T_{\text{chip}} [^\circ\text{C}]$	$T_{\text{conv}} [^\circ\text{C}]$	$T_{\text{spot}} [^\circ\text{C}]$	$T_{\text{spot_max}} [^\circ\text{C}]$
1000	20,000	76.2	74.2	78.9	79.7
1000	10,000	126.2	124.2	129.0	129.7
1000	5000	226.3	224.3	229.1	229.8
2000	20,000	76.4	74.4	82.2	83.8
2000	10,000	126.5	124.6	132.3	134.0
2000	5000	226.7	224.7	232.5	234.1

Table 8 Temperatures for $400 \mu\text{m}$ hot-spot diameter for various heat flux and cooling conditions

$q''_{\text{spot}} [\text{W/cm}^2]$	$h [\text{W/m}^2 \text{ K}]$	$T_{\text{chip}} [^\circ\text{C}]$	$T_{\text{conv}} [^\circ\text{C}]$	$T_{\text{spot}} [^\circ\text{C}]$	$T_{\text{spot_max}} [^\circ\text{C}]$
1000	20,000	7.75	75.6	89.1	92.2
1000	10,000	128.2	126.4	140.2	143.3
1000	5000	229.4	227.6	241.6	244.7
2000	20,000	79.2	77.5	103.7	110.3
2000	10,000	130.7	129.0	155.9	162.5
2000	5000	233.2	231.6	259.0	265.6

is established by the superposition of these two effects, any reduction in the average chip temperature has a salutary effect on the peak chip temperature, as well.

Thus, as also revealed in Fig. 66, micro-gap coolers, along with effective thermal spreading in the chip, appears to offer the potential for successfully limiting the chip and hot-spot temperature rise to acceptable levels for a wide range of operating conditions. Most significantly, a heat-transfer coefficient of 20 kW/m² K, which is thought to be attainable in a microgap cooler [132] could be used to effectively cool a most challenging 2 kW/cm², 400 μ m hot spot, along with maintaining an acceptable average temperature, for a 100 W/cm² chip.

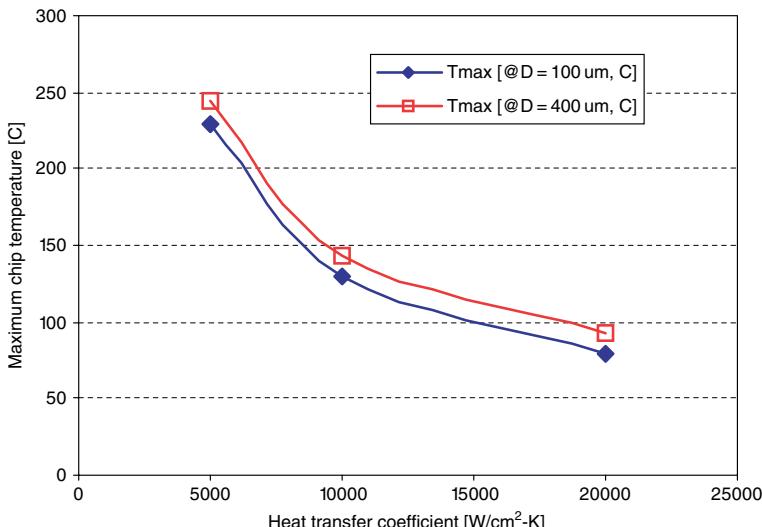


Fig. 66 Effect of convective coefficient on the hot-spot temperature for $q''_{hs} = 1000 \text{ W/cm}^2$

Interestingly, the on-chip temperature rise – of the hot-spot center relative to the chip average – can be seen to vary almost directly with the product of the heat flux and diameter, yielding a ninefold increase from 3.6°C for a 100 μ m, 1 kW/cm² hot spot to 32.4°C for a 400 μ m, 2 kW/cm² hot spot. Due to the superposition of the convective and conductive effects, it may also be noted that while a large change (~50%) in the maximum excess temperature results from increasing the micro-gap heat-transfer coefficient from 5 to 10 kW/m² K, further increases to 20 kW/m² K only reduces the maximum temperature rise by approximately 30%.

5 Conclusions

The preceding chapter addresses on-chip, hot-spot cooling, which has become one of the most active and challenging domains in the thermal management of nano-electronic devices and packages. Following a brief discussion of several passive

and active high heat flux thermal management techniques, attention was turned to the physical phenomena underpinning the most promising on-chip thermal management approaches, including thin-film and miniaturized thermoelectric coolers, orthotropic TIMs/heat spreaders, and phase-change microgap coolers, and their use for remediation of these hot spots. It was shown that, with proper thermal optimization, mini-contact enhanced miniaturized “bulk” thermoelectric coolers can yield hot-spot temperature reductions in excess of 15°C for near millimeter-sized hot spots with kW/cm²-level heat fluxes, but to be vulnerable to the deleterious effects of thermal contact resistance. Micro-scaled thin-film silicon thermoelectric coolers, monolithically grown, or fashioned, on the back of silicon chips, were similarly found to provide effective thermal management of high heat flux spots, and to be capable of neutralizing the local temperature rise for a large variety of sub-millimeter hot spots. Orthotropic TIM/spreaders with high in-plane thermal conductivities were also shown to offer significant temperature reduction capability even for large, high-flux spots, when used in conjunction with a more conventional, very high heat-transfer coefficient thermal management approaches. Initial research results for microgap coolers, relying on the boiling and evaporation of a dielectric liquid flowing in a miniature gap at the back of the chip, strongly suggest that this cooling technique could provide both the local and global heat-transfer coefficients needed to meet many of the most demanding nanoelectronic cooling challenges, including severe on-chip hot spots.

References

1. ITRS, The International Technology Roadmap for Semiconductors, 2004, Semiconductor Industry Association, <http://www.ITRSnemi.org>.
2. Shelling P., Li S., and Goodson K. E. Managing heat for electronics, Materials Today, 2005; 8: 30–35.
3. Mahajan R., Chiu C., and Chrysler G. Cooling a microprocessor chip, Proceedings of IEEE, 2006; 94: 1476–1486.
4. Pedram M., and Nazarian S. Thermal modeling, analysis, and management in VLSI circuits: principles and methods, Proceedings of the IEEE, 2006; 9: 1487–1501.
5. Bar-Cohen A., Arik M., and Ohadi M. Direct liquid cooling of high flux micro and nano electronic components, Proceedings of the IEEE, 2006; 94: 1549–1570.
6. Mudawar I. Assessment of high-heat-flux thermal management schemes, IEEE Transactions on Components and Packaging Technologies, Part A: Packaging Technologies, 2001; 24: 122–141.
7. Garimella S. V. Advances in mesoscale thermal management technologies for microelectronics, Microelectronics Journal, 2006; 37: 1165–1185.
8. Sinha S., and Goodson K. E. Thermal conduction in sub-100 nm transistors, Microelectronics Journal, 2006; 37: 1148–1157.
9. NEMI, Electronics Manufacturing Initiative Technology Roadmap, 2006, <http://www.nemi.org>.
10. Jannadham K., Watkins T. R., and Dinwiddie R. B. Novel heat spreader coatings for high power electronic devices, Journal Materials Sciences, 2002; 37: 1363–1376.
11. Dahlgren S. High-pressure polycrystalline diamond as a cost effective heat spreader, Proceedings of Thermal and Thermomechanical Phenomena in Electronic Systems, (ITHERM 2000), 2000; 1: 23–26.

12. Goodson K. E., and Ju Y. S. Heat conduction in novel electronic films, *Annual Review of Materials Science*, 1999; 29: 261–293.
13. Le Berre M., Pandraud G., Morfouli P., and Lallemand M. The performance of micro heat pipes measured by integrated sensors, *Journal of Micromechanics and Microengineering*, 2006; 16: 1047–1050.
14. Peterson G. P., Duncan A. B., and Weichold M. H. Experimental investigation of micro heat pipes fabricated in silicon wafers, *Journal of Heat Transfer*, 1993; 115: 750–756.
15. Peterson G. P., Duncan A. B., and Weichold M. H. Experimental investigation of micro heat pipes fabricated in silicon wafers, *ASME Journal of Heat Transfer*, 1993; 115: 751–756.
16. Mallik A. K., Peterson G. P., and Weichold M. H. Fabrication of vapor deposited micro heat pipes arrays as an integral part of semiconductor devices, *IEEE Journal of Microelectromechanical System*, 1995; 4: 119–131.
17. Karimi G., and Culham J. R. Review and assessment of pulsating heat pipe mechanism for high heat flux electronic cooling, *Proceedings of IITHERM*, 2004; 2: 52–59.
18. Benson D. A., Adkins D. R., Peterson G. P., Mitchell R. T., Tuck M. R., and Palmer D. W., Turning silicon substrates into diamond: micro machining heat pipes, in *Proc. Adv. Design Mater. Process*, Apr. 1996, pp. 19–21.
19. Suman B. Modeling, experiment, and fabrication of micro-grooved heat pipes: an update, *Applied Mechanics Reviews*, 2007; 60: 107–119.
20. Akachi H., Structure of a heat pipe, U.S. Patent #4,921,041, 1990.
21. Zuo Z. J., North M. T., and Wert K. L. High heat flux heat pipe mechanism for cooling of electronics, *IEEE Transactions on Components and Packaging Technologies*, 2001; 24: 220–225.
22. Lin C., Ponnappan R., and Leland J. High performance miniature heat pipe, *International Journal of Heat and Mass Transfer*, 2002; 45: 3131–3142.
23. Lee M., Wong M., and Zohar Y. Integrated micro-heat-pipe fabrication technology, *Journal of Microelectromechanical Systems*, 2003; 12: 138–146.
24. Lee M., Wong M., and Zohar Y. Characterization of an integrated micro heat pipe, *Journal of Micromechanics and Microengineering*, 2003; 13: 58–64.
25. Khrustalev D., and Faghri A. Thermal characteristics of conventional and flat miniature axially-grooved heat pipes, *Journal of Heat Transfer*, 1995; 117: 740–747.
26. Ma H. B., Lofgreen K. P., and Peterson G. P. An experimental investigation of a high flux heat pipe heat sink, *Journal of Electronic Packaging*, 2006; 128: 18–22.
27. Gillot G., Avenas Y., Cezac N., Poupon G., Schaeffer C., and Fournier E. Silicon heat pipes used as thermal spreaders, *IEEE Transactions on Components and Packaging Technologies*, 2003; 26: 332–339.
28. Lin C., Ponnappan R., and Leland J. High performance miniature heat pipe, *International Journal of Heat and Mass Transfer*, 2002; 45: 3131–3142.
29. Tuckerman D. B., and Pease R. F. W. High-performance heat sinking for VLSI, *IEEE Electron Device Letters*, 1981; EDL-2: 143–150.
30. Prasher R., Chang J., Sauciu I., Narasimhan S., Chau D., Chrysler G., Myers A., Prstic A., and Hu C. Nano and micro technology-based next-generation package level cooling solutions, *Intel Journal of Technology*, 2005; 9: 285–296.
31. Bowers M., and Mudawar I. High-flux boiling in low-flow rate, low-pressure drop mini-channel and microchannel heat sinks, *International Journal of Heat and Mass Transfer*, 1994; 37: 321–332.
32. Mudawar I., and Maddox D. E. Enhancement of critical heat flux from high power micro-electronic heat sources in a flow channel, *Journal of Electronic Packaging*, 1990; 112: 241–248.
33. Bar-Cohen A., Arik M., and Ohadi M. Direct liquid cooling of high flux micro and nano electronic components, *Proceedings of the IEEE*, 2006; 94: 1549–1570.
34. Garimella S. V., Singhal V., and Liu D. On-chip thermal management with microchannel heat sinks and integrated micropumps, *Proceedings of the IEEE*, 2006; 94: 1534–1548.

35. Prasher R. Thermal interface materials: historical perspective, status and future directions, Proceedings of the IEEE, 2006; 94: 1571–1586.
36. Zhang L., Wang E. N., Koo J. M., Goodson K. E., Santiago J. G., and Kenny T. W. Microscale liquid jet impingement, Proceedings of AMSE IMECE, 2001; Vol.2: Paper No. MEME-23820.
37. Wolf D. H., Incropera F. P., and Viskanta R. Local jet impingement boiling heat transfer, International Journal of Heat and Mass Transfer, 1996; 39: 1395–1406.
38. Kiper A. M. Impinging water jet cooling of VLSI circuits, International Communications in Heat and Mass Transfer, 1984; 11: 126–129.
39. Harman T. C., Taylor P. J., Walsh M. P., and LaForge B. E. Quantum dot superlattice thermoelectric materials and devices, Science, 2002; 297: 2229–2232.
40. Venkatasubramanian R., Siivola E., Colpitts T., and O’Quinn B. Thin-film thermoelectric devices with high room-temperature figures of merit, Nature (London), 2001; 413: 597–602.
41. Fan X., Zeng G., LaBounty C., Bowers J., Croke E., Ahn C., Huxtable S., and Majumdar A. SiGeC/Si superlattice micro-coolers, Applied Physics Letters, 2001; 78: 1580–1600.
42. Chen C., Yang B., and Liu W. L. Engineering nanostructures for energy conversion, Heat Transfer and Fluid Flow in Microscale and Nanoscale Structures, Faghri, M. and Sundén, B., Eds., Southampton, UK: WIT Press, 2004.
43. Yang B., Liu W. L., Wang K. L., and Chen G. Simultaneous measurements of Seebeck coefficient and thermal conductivity across superlattice, Applied Physics Letters, 2002; 80: 1758–1760.
44. Shakouri A., and Zhang Y. On-chip solid-state cooling for integrated circuits using thin-film microrefrigerators, IEEE Transactions on Components and Packaging Technologies, 2005; 28: 65–69.
45. Zhang Y., Zeng G. H., Piprek J., Bar-Cohen A., and Shakouri A. Superlattice microrefrigerators fusion bonded with optoelectronic devices, IEEE Transactions on Components and Packaging Technologies, 2005; 28: 658–666.
46. Simons R. E., Ellsworth M. J., and Chu R. C. An assessment of module cooling enhancement with thermoelectric coolers, Journal of Heat Transfer, 2005; 127: 76–84.
47. Yeh L., and Chu C. Thermal Management of Microelectronic Equipment, New York: ASME Press, 2002.
48. Kraus A. D., and Bar-Cohen A. Thermal Analysis and Control of Electronic Equipment, New York, USA: Hemisphere Publishing Corporation, 1983.
49. Fan X., Ph. D. Thesis, University of California at Santa Barbara, March 2002.
50. Fleuriol J. -P., Borshchevsky A., Ryan M. A., Phillips W., Kolawa E., Kacisch K., and Ewell R., Thermoelectric microcoolers for thermal management applications, Proceedings of 16th International Conference on Thermoelectrics, 1997; 641–645.
51. Venkatasubramanian R., Siivola E., Colpitts T., and O’Quinn B. Thin-film thermoelectric devices with high room-temperature figures of merit, Nature (London), 2001; 413: 597–602.
52. Harman, T.C., Taylor, P. J., Walsh, M. P. and LaForge, B. E., Quantum Dot Superlattice Thermoelectric Materials and Devices, *Science*, Vol. 297, No. 2229, 2002.
53. Fan, X., Zeng, G., Croke, E., LaBounty, C., Shakouri, A., and Bowers, J. E., Integrated SiGeC/Si Micro Cooler, *Applied Physics Letters*, Vol. 78, No.11, 12 March 2001.
54. Semenyuk V., Thermoelectric micro modules for spot cooling of high density heat sources, Proceedings of the 20th International Conference on Thermoelectrics, 2001; 391–396.
55. Semenyuk V., Cascade Thermoelectric micro modules for spot cooling high power electronic components, Proceedings of the 21st International Conference on Thermoelectrics, 2002; 531–534.
56. Semenyuk V. Thermoelectric cooling of electro-optic components, Thermoelectrics Handbook: Macro to Nano, Rowe, D. M., Ed., Boca Raton, FL: CRC Press, 2006.
57. www.thermion-company.com.

58. Ioffe A. F. *Semiconductor Thermoelements and Thermoelectric Cooling*, London, UK: Infosearch Ltd., 1957.
59. Ettenberg M. H., Jesser M. A., and Rosi E. D., A new n-type and improved p-type pseudo-ternary $(\text{Bi}_2\text{Te}_3)(\text{Sb}_2\text{Te}_3)(\text{Sb}_2\text{Se}_3)$ alloy for Peltier cooling, *Proceedings of the 15th International Conference on Thermoelectrics*, 1996; 52–56.
60. Yamashita O., and Tomiyoshi S. Effect of annealing on thermoelectric properties of bismuth telluride compounds, *Japan Journal of Applied Physics*, 2003; 42: 492–500.
61. Yamashita O., Tomiyoshi S., and Makita K. Bismuth telluride compounds with high thermoelectric figures of merit, *Journal of Applied Physics*, 2003; 93: 368–374.
62. Yamashita O., and Tomiyoshi S. High performance n-type bismuth telluride with highly stable thermoelectric figure of merit, *Journal of Applied Physics*, 2004; 95: 6277–6283.
63. Shakouri A., and Zhang Y. On-chip solid-state cooling for integrated circuits using thin-film microrefrigerators, *IEEE Transactions on Components and Packaging Technologies*, 2005; 28: 65–69.
64. Pandey R. K., Sahu S. N., and Chandra S., *Handbook of Semiconductor Deposition*, Ed., New York: Marcel Dekker, 1996.
65. Snyder G. J., Lim J. R., Huang C., and Fleuriol J.-P. Thermoelectric microdevice fabricated by a MEMS-like electrochemical process, *Nature Materials*, 2003; 2: 528–531.
66. da Silva L. W., Kaviani M., and Uher C. Thermoelectric performance of films in the bismuth-tellurium and antimony-tellurium systems, *Journal of Applied Physics*, 2005; 97: 114903.
67. Böttner H., Nurnus J., Gavrikov A., Kühner G., Jägle M., Künzel C., Eberhard D., Plescher G., Schubert A., and Schlereth K. New thermoelectric components using microsystem technologies, *Journal of Microelectromechanical Systems*, 2004; 13: 414–420.
68. Bottner H., Micropelt miniaturized thermoelectric devices: small size, high cooling power densities, short response time, *Proceedings of the 24th International Conference on Thermoelectrics*, 2005; 1–8.
69. Zhou H., Rowe D. M., and Williams S. Peltier effect in a co-evaporated $\text{Sb}_2\text{Te}_3(\text{P})-\text{Bi}_2\text{Te}_3(\text{N})$ thin film thermocouple, *Thin Solid Films*, 2002; 408: 270–274.
70. Semenyuk V., Miniature thermoelectric modules with increased cooling power, *Proceedings of the 25th International Conference on Thermoelectrics*, 2006; 322–326.
71. Semenyuk V., and Ph. D.. Dissertation, Odessa Technological Institute of Food and Refrigerating Engineering, Odessa, USSR, 1967 (in Russian).
72. Hicks L. D., and Dresselhaus M. S. Thermoelectric figure of merit of a one-dimensional conductor, *Physics Review B*, 1993; 47: 16631–16634.
73. Balandin A., and Wang K. L. Effect of phonon confinement on the thermoelectric figure of merit of quantum wells, *Journal of Applied Physics*, 1998; 84: 6149–6153.
74. Balandin A., and Lazarenko O. L. Mechanism for thermoelectric figure-of-merit enhancement in regimented quantum dot superlattices, *Applied Physics Letters*, 2003; 82: 415–417.
75. Yang B., and Chen G. *Thermal Conductivity: Theory, Properties and Application*, Tritt, T. M., Ed., New York: Kluwer Press, 2005.
76. Harman T. C., Taylor P. J., Walsh M. P., and LaForge B. E. Quantum dot superlattice thermoelectric materials and devices, *Science*, 2002; 297: 2229–2232.
77. Venkatasubramanian R., Silvana E., Colpitts T., and O’Quinn B. Thin-film thermoelectric devices with high room-temperature figures of merit, *Nature*, 2001; 413: 597–602.
78. Mahan G. D., and Woods L. M. Multilayer thermionic refrigeration, *Physical Review Letter*, 1998; 80: 4016–4019.
79. Shakouri A., LaBounty C., Piprek J., Abraham P., and Bowers J. E. Thermionic emission cooling in single barrier heterostructures, *Applied Physics Letters*, 1999; 74: 88–89.
80. Rowe D. M. *Thermoelectrics Handbook Macro to Nano*, Boca Raton, FL: CRC Press, 2005.
81. Venkatasubramanian R., Colpitts T., Liu S., El-Masry N., and Lamvik M. Low-temperature organometallic epitaxy and its application to superlattice structures in thermoelectrics, *Applied Physics Letters*, 1999; 75: 1104–1106.

82. Zhang Y., Zeng G., Bar-Cohen A., and Shakouri A. Is ZT the main performance factor for hot spot cooling using 3D microrefrigerators?, IMAPS on Thermal Management, 2005, Oct. 26th–28th, Palo Alto, CA.
83. Zeng G., Shakouri A., LaBounty C., Robinson G., Croke E., Abraham P., Fan X., Reese H., and Bowers J. E. SiGe micro-cooler, Electronics Letter, 1999; 35: 2146–2147.
84. Zeng G., Fan X., LaBounty C., Croke E., Zhang Y., Christofferson J., Vashaee D., Shakouri A., and Bowers J. E. Cooling power density of SiGe/Si superlattice micro refrigerators, Proceedings of Thermoelectric Materials Research and Applications, 2003; 793: 43–49.
85. Fan X., Zeng G., LaBounty C., Vashaee D., Christofferson J., Shakouri A., and Bowers J. E., Integrated cooling for Si-based microelectronics: Proceedings of 20th International Conference on Thermoelectrics, 2001; 405–408.
86. Fan X. SiGeC/Si superlattice microcoolers, Applied Physics Letters, 2001; 78: 1580–1582.
87. Fan X., Zeng G., LaBounty C., Croke E., Vashaee D., Shakouri A., Ahn C., and Bowers J. E. High cooling power density SiGe/Si micro coolers, Electronics Letter, 2001; 37: 126–127.
88. Herwaarden A. W., and Sarro P. M. Thermal sensors based on the Seebeck effect, Sensors and Actuators, 1986; 10: 321–346.
89. Geballe T. H., and Hull G. W. Seebeck effect in silicon, Physical Review, 1955; 98: 940–970.
90. Rowe D. M. CRC Handbook of Thermoelectrics, Roca Raton, FL: CRC Press, 1995.
91. Zhang Y., Shakouri A., and Zeng G. High-power-density spot cooling using bulk thermoelectrics, Applied Physics Letters, 2004; 85: 2977–2979.
92. Nieveld G. D. Thermopiles fabricated using silicon planar technology, Sensors and Actuators, 1983; 3: 179–183.
93. Chapman P. W., Tfte O. N., Zook J. D., and Long D. Electrical properties of heavily doped silicon, Journal of Applied Physics, 1963; 34: 3291–3295.
94. Wang P., Bar-Cohen A., and Yang B. Analytical modeling of silicon thermoelectric microcooler, Journal of Applied Physics, 2006; 100: 14501.
95. Wang P., and Bar-Cohen A. On-chip hot spot cooling using silicon-based thermoelectric microcooler, Journal of Applied Physics, 2007; 102: 034503.
96. Solbrekken G. L., Zhang Y., Bar-Cohen A., and Shakouri A., Use of superlattice thermionic emission for “Hotspot” reduction in convectively-cooled chip, Proceedings of 9th IITHERM’s;04, 2004; 610–616.
97. Wang P., Bar-Cohen A., Yang B., Solbrekken G. L., Zhang Y., and Shakouri A., Thermoelectric microcooler for hotspot thermal management, Proceedings of InterPACK’s;05, 2005; Paper No: 2005-7324.
98. Lide D. R., CRC Handbook of Chemistry and Physics, 75th edition, Boca Raton, USA: CRC Press, 1994.
99. Kraus A. D., and Bar-Cohen A. Thermal Analysis and Control of Electronic Equipment, New York, USA: Hemisphere Publishing Corporation 1983.
100. Geballe T. H., and Hull G. W. Seebeck effect in silicon, Physical Review, 1955; 98: 940–970.
101. Herwaarden A. W., and Sarro P. M. Thermal sensors based on the Seebeck effect, Sensors and Actuators, 1986; 10: 321–346.
102. Horn F. H. Densitometric and electrical investigation of boron in silicon, Physical Review, 1955; 97: 1521–1525.
103. Fritzsche H. A General expression for the thermoelectric power, Solid State Communication, 1971; 9: 1813–1815.
104. Chang C. Y., Fang Y. K., and Sze S. M. Specific contact resistance of metal-semiconductor barriers, Solid State Electronics, 1971; 14: 541–550.
105. Fan X., Silicon M., Ph. D. Thesis, University of California at Santa Barbara, 2002.
106. Yang B., Wang P., and Bar-Cohen A. Mini-contact enhanced thermoelectric cooling of hot spot in high power devices, IEEE Transactions on Components and Packaging Technologies, Part A, 2007; 30: 432–438.

107. Narasimhan S., Lofgreen K., Chau D., and Chrysler G., Thin film thermoelectric cooler thermal validation and product thermal performance estimation, Proceedings of 10th Intersociety Conference on Thermal and Thermo-mechanical Phenomena in Electronics Systems, San Diego, CA, May 30–June 2, 2006.
108. Gwinn J. P., and Webb R. L. Performance and testing of thermal interface materials, *Microelectronics Journal*, 2003; 34: 215–222.
109. Singhal V., Siegmund T., and Garimella S. V. Optimization of thermal interface materials for electronics cooling applications, *IEEE Transactions on Components and Packaging Technologies*, 2004; 27: 244–252.
110. Labudovic M., and Li J. Modeling of TE cooling of pump lasers, *IEEE Transactions on Components and Packaging Technologies*, 2004; 27: 724–730.
111. So W. W., and Lee C. C. Fluxless process of fabricating In-Au joints on copper substrates, *IEEE Transactions on Components and Packaging Technologies*, 2000; 23: 377–382.
112. www.thermion-company.com.
113. Semenyuk V. Thermoelectric cooling of electro-optic components, *Thermoelectrics Handbook: Macro to Nano*, Rowe, D. M. , Ed., Boca Raton, FL: CRC Press, 2006.
114. Wijngaards D. D. L., de Graaf G., and Wolffenbuttel R. F. Single-chip micro-thermostat applying both active heating and active cooling, *Sensors and Actuators A*, 2004; 110: 187–195.
115. Corrèges P., Bognard E., Millerin C., Masiero A., Andrivet J. P., Bloc A., and Dunant Y. A simple, low-cost and fast Peltier thermoregulation set-up for electrophysiology, *Journal of Neuroscience Methods*, 1998; 83: 177–184.
116. McKinney C. J., and Nader M. W. A Peltier thermal cycling unit for radiopharmaceutical synthesis, *Applied Radiation and Isotopes*, 2001; 54: 97–100.
117. Elsgaard L., and Jørgensen L. W. A sandwich-designed temperature gradient incubator for studies of microbial temperature responses, *Journal of Microbiological Methods*, 2002; 49: 19–29.
118. Reid G., Amuzescu B., Zech E., and Flonta M. L. A system for applying rapid warming or cooling stimuli to cells during patch clamp recording or ion imaging, *Journal of Neuroscience Methods*, 2001; 111: 1–8.
119. Hodgson J. Gene sequencing's industrial revolution, *IEEE Spectrum*, 2000; 37: 36–42.
120. Maltezos G., Johnston M., and Scherer A. Thermal management in microfluidics using micro-Peltier junctions, *Applied Physics Letters*, 2005; 87: 154105.
121. Bachmann C., and Bar-Cohen A., Hotspot remediation with anisotropic thermal interface materials, *Proceedings of ITERM 2008*, 2008; 238–247.
122. Muzychka Y. S., Culham J. R., and Yovanovich M. M. Thermal spreading resistance of eccentric heat sources on rectangular flux channels, *Journal of Electronic Packaging*, 2003; 125: 178–185.
123. Smalc M., Thermal performance of natural graphite heat spreaders, *Proceedings IPACK2005*, San Francisco, California, USA, July 17–22, PaperNumber: IPACK2005-73073.
124. Bar-Cohen A., Arik M., and Ohadi M. Direct liquid cooling of high flux micro and nano electronic components, *Proceedings of the IEEE*, 2006; 94: 1549–1570.
125. Montesano M., Annealed pyrolytic graphite, *Advanced Materials and Processes*, 2006; June:1–3.
126. Viswanath R., Wakarkar V., Watwe A., and Lebonheur V. Thermal performance challenges from silicon to systems, *Intel Technology Journal*, 2000; Quarter 3: 1–16.
127. Xong Y., Smalc M. et al., Thermal tests and analysis of thin graphite heat spreader for hot spot reduction in handheld devices, *Proceedings of ITERM 2008*, 2008; 583–590.
128. Bergles A. E., and Bar-Cohen A. Immersion cooling of digital computers, *Cooling of Electronic Systems*, Kakac, S., Yuncu, H. and Hijikata, K., Eds., Boston, MA: Kluwer Academic Publishers, 1994, 539–621
129. Bergles A. E., and Bar-Cohen A. Direct liquid cooling of microelectronic components, *Advances in Thermal Modeling of Electronic Components and Systems*, A. Bar-Cohen and A. D. Kraus, Eds., New York: ASME, 1990; 2: 241–250.

130. Kim D., Rahim E., Bar-Cohen A., and Han B., Thermofluid characteristics of two-phase flow in micro-gap channels, Proceedings of ITERM 2008, 2008; 979–992.
131. Kim D. W., Ph.D. Thesis, University of Maryland, 2007.
132. Bar-Cohen A., and Rahim E. Modeling and prediction of two-phase microgap channel heat transfer characteristics, Heat Transfer Engineering, 2009; 30: 601–625.

Some Aspects of Microchannel Heat Transfer

Y. Joshi, X. Wei, B. Dang, and K. Kota

Abstract Systems for energy conversion, heat rejection, and sensing and control often incorporate heat exchange devices. Recent developments in microfabrication and assembly methods have led to significant miniaturization of these systems. Miniaturized heat exchange devices have commonly utilized microchannel flow passages. This chapter reviews the fundamental flow and heat-transfer phenomena in microchannels, the most common numerical and experimental characterization techniques, microfabrication methods, and the application of microchannels in thermal management of high heat flux electronics. The range of channel hydraulic diameters covered in this chapter is from a few micrometers to a few millimeters where the larger diameter transport characteristics become valid.

Keywords Microchannel flows · Heat transfer · Two-phase flows · Numerical techniques · Experimental techniques · Stacked microchannels · Micro-PIV · Electronics cooling · Two-phase flow regimes · Flow boiling · Flow condensation · Microfabrication · Laser-induced fluorescence · Total internal reflection fluorescence

1 Fundamentals of Microchannel Pressure Drop and Heat Transfer

Microchannels may utilize single-phase flow or phase change. It is now generally believed that for microchannels of few microns or larger in hydraulic diameter involving single-phase flow, the fluid flow and heat transport are adequately characterized by the same pressure drop and heat-transfer correlations as larger sized channels. On the other hand, two-phase flows are significantly affected by microchannel sizes. We summarize the key aspects of pressure drop and heat transfer in single-phase and two-phase heat transfer.

Y. Joshi (✉)

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: yogendra.joshi@me.gatech.edu

1.1 Single-Phase Flows

If the length of the microchannel is significantly longer than the hydraulic entrance length, $X_{\text{fd,h}}$ and thermal entrance length, $X_{\text{fd,th}}$, as defined in equations (1) and (2), respectively, the flow and heat transfer are considered fully developed. In equations (1) and (2), Re and Pr are the Reynolds number and the Prandtl number for the fluid, respectively, and d_h is the hydraulic diameter.

$$X_{\text{fd,h}} \approx 0.05Re \cdot d_h \quad (1)$$

$$X_{\text{fd,th}} \approx 0.05Re \cdot d_h \cdot Pr \quad (2)$$

As a result, the nondimensional Nusselt number, Nu , is a constant for particular channel geometry and boundary conditions. Thus the heat-transfer coefficient, h , is inversely proportional to the hydraulic diameter, as shown in equation (3), where k is the thermal conductivity of the fluid.

$$Nu = \frac{h \cdot d_h}{k} = \text{const} \Rightarrow h \propto \frac{1}{d_h} \quad (3)$$

Following this equation, heat-transfer coefficient in microchannels is orders of magnitude higher when compared to channels of macroscale. This benefit is demonstrated in Tuckerman and Pease (1981), who reported a microchannel heat sink with thermal resistance 0.09 K cm²/W. This is three times or more lower than conventional heat sinks. Typical microchannel heat sinks as described in this work consist of parallel microflow passages 50 μm wide and 300 μm deep.

Following Tuckerman and Pease [1], several early studies [2–10] indicated that the flow and heat-transfer parameters deviate from the classical theory developed for macrosize channels and the transition from laminar to turbulent flow occurs at a considerably smaller critical Reynolds number. However, there is no consensus on the trend of this deviation among different researchers. Some of the reported deviations may be attributed to experiment conditions such as surface roughness and flow maldistribution. In contrast, results from more recent studies [11–16] appear to indicate that the pressure drop and transport characteristics for macroscale channels also hold for the microchannels for various applications with channel hydraulic diameters of 10 μm or larger, with gases or liquids, such as in microelectronics cooling.

1.1.1 Simplified Model for Single-Phase Microchannel Heat Sink

A simplified model of a microchannel heat sink incorporating a stack of microchannels seen in Fig. 1 [17] is described in this section. At each layer, a number of parallel microchannels are machined in the surface of a substrate, e.g., copper, silicon, or diamond. These layers are then stacked and bonded to form a micro heat sink. The benefits of using stacked microchannels are the significant reduction

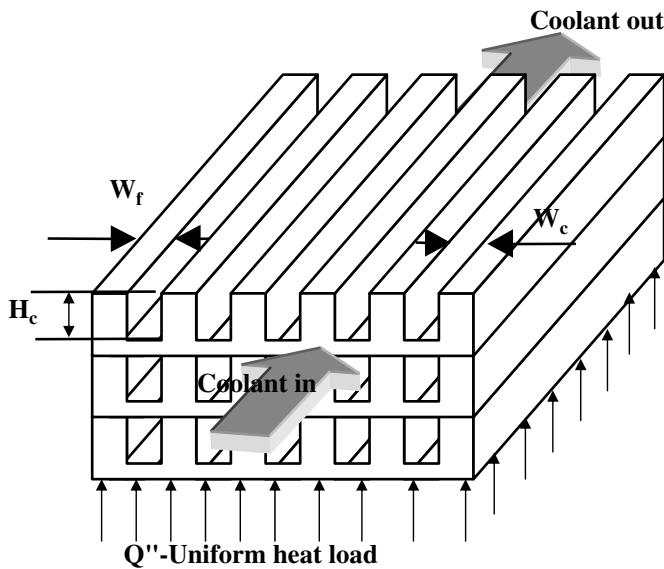


Fig. 1 Three-dimensional stack of micro-channels (Reprinted with permission from Wei and Joshi [17])

in pressure drop and temperature gradient, when compared with a single-layered microchannel heat sink. It also allows tailored design of the fluid passage for potential hot-spot cooling for highly nonuniform heat density applications. Although the focus of this section is on stacked microchannel, the same preliminary modeling procedure can be followed for a single-layered microchannel heat sink, since it is simply a special case of the stacked microchannel heat sink.

1.1.2 Correlations for Friction Factor and Nusselt Number (Wei and Joshi, 2003)

The pressure drops associated with the stack structure include the contraction and expansion pressure drops in the inlet and outlet, respectively, the pressure drop due to 90° bend and the flow friction loss. As pointed out by many researchers [1, 18], the friction losses dominate the pressure drop in laminar flow in rectangular duct. Therefore, only the friction loss is included here.

To consider the friction factor (f_{app}) for developing laminar flow in a rectangular duct, a Churchill–Usagi asymptotic type of model shown in equations (4), (5), (6) is used [19].

$$f_{app} \operatorname{Re}_{\sqrt{A_c}} = \left[\left(\frac{3.44}{\sqrt{y^+}} \right)^2 + (8\sqrt{\pi} G(\alpha))^2 \right]^{0.5} \quad (4)$$

$$G(\alpha) = \left[1.086957^{1-\alpha} \left(\sqrt{\alpha} - \alpha^{1.5} \right) + \alpha \right]^{-1} \quad (5)$$

$$y^+ = \frac{y}{\text{Re}_{\sqrt{A_c}} \sqrt{A_c}} \quad (6)$$

In this equation, y^+ is the nondimensional length scale for friction factor calculation, α is the aspect ratio or the inverse of the aspect ratio of the microchannel that is always smaller than 1. It is noted here that the length scale for the Reynolds number (Re) and the dimensionless length is the square root of the channel cross-sectional area (A_c). The above equations allow the calculation of pressure drop across a microchannel.

Similarly, the Nusselt number for thermally developing condition in a rectangular duct is given as [20],

$$Nu_{\sqrt{A_c}}(y^*) = \left[\left(C1 \cdot C2 \left(\frac{8\sqrt{\pi}G(\alpha)}{y^*} \right)^{\frac{1}{3}} \right)^5 + \left(C3 \left(\frac{G(\alpha)}{\alpha^\gamma} \right) \right)^5 \right]^{0.2} \quad (7)$$

$$y^* = \frac{y}{\text{Re}_{\sqrt{A_c}} \cdot \text{Pr} \cdot \sqrt{A_c}} \quad (8)$$

$C1 = 1$ for isoflux condition, $C2 = 0.501$, $C3 = 3.66$, $\gamma = 0.1$, y^* is the nondimensional length scale for Nusselt number calculation, and $G(\alpha)$ is the aspect ratio function. The heat-transfer coefficient can be calculated from the above Nusselt number relations.

1.1.3 Thermal Resistance Network Analysis

A thermal resistance analysis based on existing correlations for heat-transfer coefficients and friction factor in laminar channel flow [17] is described below. Heat spreading and constriction resistances for small-sized heat sources on larger substrates, adapted from [18], are also incorporated into the resistance network.

The thermal resistance network for a three-layered microchannel stack is illustrated in Fig. 2. All component resistances except R_2 and R_4 are listed in Table 1. To calculate these resistances, one-dimensional conduction is assumed. This assumption is reasonable as long as the aspect ratio is sufficiently large. In the base area, however, the aspect ratio is close to one. Two-dimensional heat conduction in that region is taken into account by considering constriction resistance for area contraction and spreading resistance for area expansion.

where

$$R_{\text{cond}} = \frac{t}{k_s WL} \quad (9)$$

$$R_{\text{cont}} = \frac{L \cdot n \left[1 / \sin \left(\frac{\pi}{2} \frac{1}{1+W_c/W_f} \right) \right]}{k_s WL} \quad (10)$$

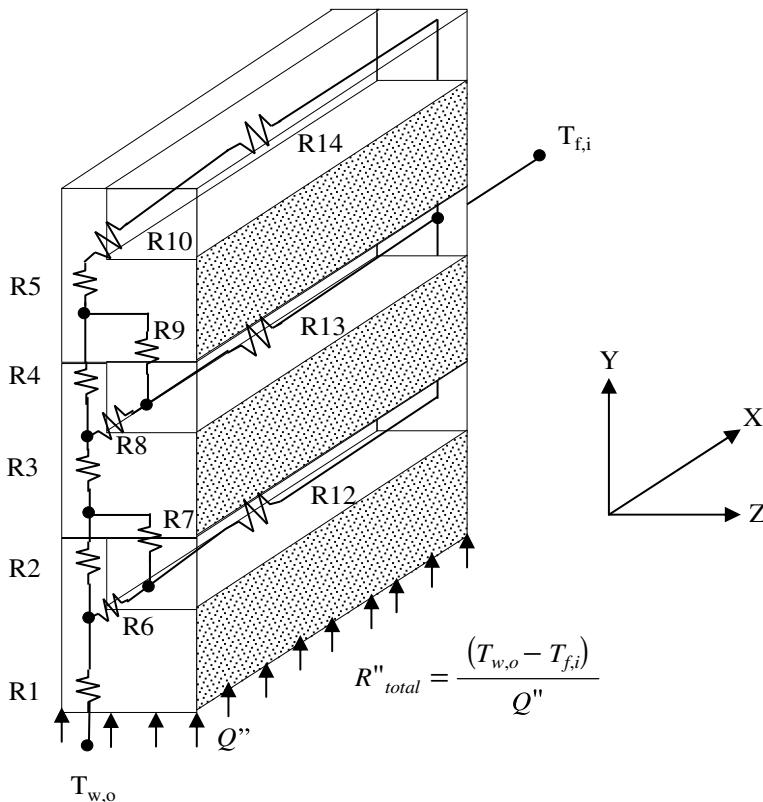


Fig. 2 Thermal resistance network for a three-layered micro-channel stack (Reprinted with permission from Wei and Joshi [17])

Table 1 Component resistances for the network

$R_1 = R_{\text{cond}} + R_{\text{cont}}$	$R_6 = R_{\text{conv,fin}}$	$R_7 = R_{\text{conv,c}}$	$R_{12} = R_{\text{bulk}}$
$R_3 = R_{\text{cond}} + R_{\text{cont}} + R_{\text{spread}}$	$R_8 = R_{\text{conv,fin}}$	$R_9 = R_{\text{conv,c}}$	$R_{13} = R_{\text{bulk}}$
$R_5 = R_{\text{cond}} + R_{\text{cont}} + R_{\text{spread}}$	$R_{10} = R_{\text{conv,fin}}$		$R_{14} = R_{\text{bulk}}$

$$R_{\text{spread}} = \frac{L \cdot n \left[1 / \sin \left(\frac{\pi}{2} \frac{1}{1 + W_c/W_f} \right) \right]}{k_s WL} \quad (11)$$

$$R_{\text{Conv,fin}} = \frac{1}{h \cdot n \cdot L \cdot W_c + 2 \cdot h \cdot \eta_f \cdot n \cdot L \cdot H_c} \quad (12)$$

$$R_{\text{Conv,c}} = \frac{1}{h \cdot n \cdot L \cdot W_c} \quad (13)$$

$$R_{\text{bulk}} = \frac{1}{\rho \cdot C_p \cdot v_m \cdot n \cdot H_c \cdot W_c} \quad (14)$$

where t is the base thickness for microchannel, k_s the solid thermal conductivity, k_f the fluid thermal conductivity, W_c the channel width, W_f the fin width, L the channel length, n the number of channels in each layer, h the heat-transfer coefficient, H_c the channel depth, η_f the fin efficiency, ρ the fluid density, C_p the specific heat of the fluid, and v_m the mean fluid velocity.

The fin equation shown in equation (15) can be solved to get the two conduction resistances R_2 and R_4 .

$$\frac{d^2\theta}{dx^2} = m^2 \cdot \theta, 0 < x < H_c \quad (15)$$

$$\text{where } \theta = T - T_f, m = \sqrt{\frac{2 \cdot h}{k_s \cdot W_f}}$$

Since the boundary conditions for the fin equation are part of the solution, the solution procedure is iterative in nature. The simple model described in this section can be readily applied in an optimization algorithm to achieve optimum designs of microchannel heat sinks [17].

1.2 Two-Phase (Liquid–Vapor) Flows

In general, flow boiling regime maps and heat-transfer and pressure drop correlations developed for macroscale channels may show significant errors when the channel sizes go down below 1–2 mm. Within the past 5 years, flow regime maps for microchannels flows have been characterized. Evidence for the transition from macroscale to microscale in the form of flow regime changes and the associated parameters is described below.

1.2.1 Two-Phase Flow Regimes for Microchannels

1.2.1.1 Flow Regimes

Revellin and Thome [21] proposed a diabatic flow regime map for microchannels based on: isolated bubble (IB) flow, coalescing bubble (CB) flow, annular (A) flow, and dryout (DO) flow. The transition from IB to CB was found to be governed by the boiling number (Bo_i), the liquid Reynolds number (Re_L) and the vapor Weber number (We_G). The following correlation predicted the transition vapor quality.

$$x_{IB/CB} = 0.763 (Re_L Bo_i / We_G)^{0.41} = 0.763 \left(q r G s / u_L h_{LG} m^2 \right)^{0.41} \quad (16)$$

where

$$Re_L = G D / \mu_L \quad (17)$$

$$Bo_i = q / (G h_{LG}) \quad (18)$$

$$We_G = G^2 D / (\rho_G \sigma) \quad (19)$$

Only two channel diameters were employed in this correlation, and hence no diameter dependence was discerned.

The CB/A transition occurs when the bubble frequency reaches zero, or at the end of distinct vapor bubbles and liquid slugs. This is predicted by

$$x_{CB/A} = 1.4 \times 10^{-3} Re_L^{1.47} We_L^{-1.23} \quad (20)$$

where the liquid Weber number is given by

$$We_L = G^2 D / (\rho_L \sigma) \quad (21)$$

The A/DO correlation was obtained by combining the critical heat flux correlation of Wojtan, Revillin, and Thome [22] and an expression for critical vapor quality, x_{crit} . The resulting expression is

$$x_{crit} \frac{q_c}{G(h_{LV} + h_{sub})} \frac{4L_h}{D} \quad (22)$$

with G as the mass flux, L as the channel length, D is the diameter, h_{LV} is the latent heat of vaporization, h_{sub} is the subcooling enthalpy, q is the heat flux, σ is the surface tension, μ is the kinematic viscosity, and q_c is the critical heat flux with subscripts L and G signifying liquid and vapor phases, respectively.

1.2.1.2 Flow Transitions

In microscale two-phase horizontal flows, the effects of gravity are insignificant, resulting in the disappearance of stratified wavy and fully stratified flows, found in larger diameter channels. Microscale flows provide a uniform film thickness in annular and slug flows due to the small diameters. In vertical tube flows for diameters below 2 mm, the observed flow patterns in R-134a were found to be very different from macroscale flows [23]. Void fractions also seem to show deviations from homogeneous void fraction laws, as sizes are reduced. This has been observed for nitrogen in water, ethanol, and water–ethanol mixtures in channels from 0.25 to 0.05 mm. Thus, size effects in two-phase flows are seen for considerably larger sizes (up to ~ 2 mm) than in single-phase flows.

1.2.1.3 Transition Criteria

Kew and Cornwall [24] found that as the passage diameter is reduced, influence of gravity is suppressed in comparison to surface tension for boiling number, $Bo_i < 4$. This results in a threshold diameter d_{th} below which, the macroscale methods are not suitable for predicting heat-transfer coefficients or flow patterns. The threshold to confined bubble flow to the channel diameter is covered by the equation

below, where Co is the confinement number (>0.5) and σ , ρ_L , ρ_G , and g are the surface tension, densities of liquid and vapor phases, and gravitational acceleration, respectively.

$$d_{th} = \frac{\sqrt{\sigma/(\rho_L - \rho_G)g}}{Co} \quad (23)$$

This ignores flow forces, i.e., no superficial velocities and shear effects on bubbles and annular films. Further refinements have introduced additional dimensional groups to define the threshold [25].

Jacobi and Thome [26] assumed that no stratification exists at the microscale. During bubble growth, the bubble diameter reaches tube diameter, prior to departure. The subsequent growth then takes place only along the tube length in the flow direction. Bubble growth relations from nucleate pool boiling have typically been used.

Li and Wang [28] performed an analysis to determine the transition from symmetric to asymmetric flow during condensation in tubes and is described in a later section on condensation in microchannels. These criteria could also be applied to adiabatic flows and flow boiling as well.

Computations of slug flows of elongated bubbles have been carried out for R134a at a saturation pressure of 7.74 bar ($T_{sat} = 30^\circ\text{C}$). The computations have compared the liquid film thickness at the top of the bubble to that at the bottom of the bubble, as a function of the channel diameter and six nondimensional parameters. The transition criteria suggest the transition is in the range of 0.1–0.5 mm channel sizes, comparable to Li and Wang [28].

1.2.2 Pressure Drop Correlations and Models

The available evidence for pressure drop seems to be counter-intuitive, suggesting that in the range of 0.1–3 mm diameter the pressure drops estimated from several macroscale models, including homogeneous flow, predicted microscale data quite well, even for flows that would fall in slug and annular regimes provided the flow patterns are properly matched.

The pressure drop is calculated using the following equation for two-phase flow in microchannels:

$$\frac{dp_{tp}}{dz} = \frac{dp_f}{dz} + \frac{dp_m}{dz} + \frac{dp_s}{dz} \quad (24)$$

where the terms on the right-hand side of the equation represent the losses due to friction, acceleration of the fluid, and the static head due to gravitational forces (which is usually negligible).

Though a generic model is not available to predict pressure drops in two-phase flows, most of the proposed pressure drop correlations in the literature are based on one of the two classical two-phase flow models. For modeling two-phase pressure drops, the two commonly used models are the homogeneous model and the

separated flow model. In the former, the two-phase fluid is assumed to behave as a single-phase fluid with pseudo-properties. The effective density, viscosity, and specific heat are evaluated based on vapor quality. In the latter approach, the two phases are segregated into separate streams, which interact through an interface.

1.2.2.1 Homogeneous Model

In the homogeneous model, the two phases are assumed to be constituents of a single fluid and the properties of the phases are calculated based on the vapor mass quality (x). Both fluids are assumed to have the same velocity with a slip ratio of 1. The density of the mixture is based on the mass quality and is defined as in equation (25).

$$\rho_m = \left(\frac{x}{\rho_G} + \frac{1-x}{\rho_L} \right)^{-1} \quad (25)$$

Different empirical correlations are proposed in literature for the effective viscosity of the mixture. Few of the references are McAdams et al. [28], Cicchitti et al. [29], Dukler et al. [30], and Awad and Muzychka [31]. The most common definition of viscosity is by McAdams et al. [28] as shown in equation (26).

$$\mu_m = \left(\frac{x}{\mu_G} + \frac{1-x}{\mu_L} \right)^{-1} \quad (26)$$

Several researchers have proposed new definitions for viscosity and have compared their predictions with the published data for frictional pressure drop based on the homogeneous model. The homogeneous model is found to be more appropriate to predict pressure drops of bubbly flows.

1.2.2.2 Separated Flow Model

In the separated flow model, the frictional pressure drop is expressed in terms of frictional multipliers. It is defined as the ratio of the frictional pressure gradient to that for one or the other phase flow alone or for a flow at the same total mass flux having gas or liquid properties.

$$\frac{dp_f}{dz} = \phi_L^2 \frac{dp_L}{dz} = \phi_G^2 \frac{dp_G}{dz} \quad (27)$$

where ϕ_L, ϕ_G are two-phase multipliers.

The pressure gradient due to acceleration or deceleration effects is given by

$$\frac{dp_m}{dz} = G^2 \frac{d}{dz} \left[\frac{(1-x)^2}{(1-\alpha)\rho_L} + \frac{x^2}{\alpha\rho_G} \right] \quad (28)$$

where G is the mass flux, z is the length variable in flow direction, and α is the void fraction defined as the ratio of gas flow cross-sectional area to the total cross-sectional area. This model is more suitable for predictions in slug flows.

1.2.3 Heat-Transfer Correlations and Models

1.2.3.1 Heat-Transfer Coefficient

Some of the experimental evidence for R134a suggests an increase in the average heat-transfer coefficient in transitioning from macroscale to microscale (~ 2 mm). Other studies indicate a weaker or no dependence.

Agostini and Thome [32] compared flow boiling heat-transfer coefficient data in microchannels from 13 studies. Wide variations were found in the level and trends due to differences in the conditions of the experiments. For example, the flow maldistribution resulting in mass flow rate variations within a multichannel evaporator fed from an inlet plenum could be addressed by having orifices at the inlet of each channel. The following general conclusions could be drawn from the studies.

At very low vapor qualities ($x < 0.05$), the heat-transfer coefficient increases with heat flux, and either increases with, or is insensitive to vapor quality.

At low to medium quality ($0.05 < x < 0.5$), the heat-transfer coefficient increases with heat flux, and decreases with or is relatively insensitive to vapor quality.

For $x > 0.5$, the heat-transfer coefficient decreases sharply with vapor quality and does not depend on heat flux or mass velocity.

The heat-transfer mechanisms depend on the flow regime during boiling. In bubbly flow nucleate boiling, liquid convection can be assumed to be dominant. In slug flow, thin film evaporation is the dominant mode. Liquid convection and vapor convection may also be important; the latter if there is a dry zone. In annular flow, thin liquid film evaporation is the dominant mode. In mist flow, vapor heat-transfer processes along with droplet impingement are important.

Kandlikar and coworkers (e.g., Kandlikar and Balasubramanian [33] and Peters and Kandlikar [34]) have proposed the determination of two-phase heat-transfer coefficient for the nucleate boiling dominant (NBD) and convective boiling dominant (CBD) regions:

$$h_{tp,NBD} = 0.6683Co^{-0.2}(1-x)^{0.8}f_2(Fr_{lo})h_{lo} + 1058Bo_i^{0.7}(1-x)^{0.8}F_{fl}h_{lo} \quad (29)$$

$$h_{tp,NBD} = 1.136Co^{-0.9}(1-x)^{0.8}f_2(Fr_{lo})h_{lo} + 667.2Bo_i^{0.7}(1-x)^{0.8}F_{fl}h_{lo} \quad (30)$$

The first term in each equation is referred to as the convective boiling term, and the second as the nucleate boiling term. The authors propose the use of these two equations for calculating the heat-transfer coefficient during flow boiling, based on the flow regime and microchannel configuration. The larger of the two is to be used as the two-phase heat-transfer coefficient, h_{tp} . The h_{lo} in the above equations is calculated as a single-phase heat-transfer coefficient appropriate for the flow condition, assuming that the entire mass flux is liquid. Suggested correlations are as follows.

Turbulent flow ($Re_{lo} \geq 3000$):

$3000 \leq Re_{lo} \leq 10^4$: Gnielinski correlation

$$h_{lo} = \frac{(Re_{lo} - 1000) Pr_l(f/2) (K_1/D)}{1 + 12.7 \left(Pr_l^{2/3} - 1 \right) (f/2)^{0.5}} \quad (31)$$

$10^4 \leq Re_{lo} \leq 5 \times 10^6$: Petukhov and Popov correlation

$$h_{lo} = \frac{(Re_{lo} - 1000) Pr_l(f/2) (K_1/D)}{1 + 12.7 \left(Pr_l^{2/3} - 1 \right) (f/2)^{0.5}} \quad (32)$$

$$\text{where } f = [1.58 \ln (Re_{lo}) - 3.28]^{-2} \quad (33)$$

Laminar flow ($Re_{lo} < 1600$):

$$h_{lo} = \frac{Nu \cdot k}{D_h} \quad (34)$$

where Nu is chosen based on the channel geometry and the wall boundary condition.

Transition flow:

h_{lo} is taken as a linear interpolation between the laminar value at $Re_{lo} = 1600$ and Gnielinski correlation value at $Re_{lo} = 3000$.

Peters and Kandlikar [34] propose the following modifications for $Re_{lo} < 410$:
For $100 < Re_{lo} < 410$:

$$h_{tp} = h_{tp,NBD} = 0.6683 Co^{-0.2} (1-x)^{0.8} h_{lo} + 1058 Bo_i^{0.7} (1-x)^{0.8} F_{fl} h_{lo} \quad (35)$$

For $Re_{lo} < 100$, the convective boiling contribution is negligible, resulting in,

$$h_{tp} = 1058 Bo_i^{0.7} (1-x)^{0.8} F_{fl} h_{lo} \quad (36)$$

where the parameter F_{fl} accounts for the effects of the fluid–surface interactions in nucleate boiling. For copper and brass surface, $F_{fl} = 1$ for water, and $F_{fl} = 1.63$ for R-134a.

Jacobi and Thome [26] proposed an elongated bubble (slug) flow boiling model to predict the heat-transfer coefficient. They accounted for evaporation of the thin liquid film trapped between the bubble and the channel wall, as well as the convection by the liquid slugs between bubbles. The former was significantly larger than the latter. Their thin film evaporation model predicted $h_{tp} \sim q^n$, where the exponent n depends on elongated bubble frequency and initial liquid film thickness resulting from the passing bubble.

1.2.3.2 Critical Heat Flux (CHF)

Critical heat flux is an extremely important operational limit in two-phase flow-based cooling devices. It signifies the highest heat flux, which can be sustained for a particular set of conditions. Exceeding this value results in a dryout at the wall, and an associated large increase in the wall temperature. Qu and Mudawar [35] reported water CHF data in microtubes from 1 to 3 mm diameter. These show mean absolute error of $\sim 17\%$ and an error band of $\pm 40\%$ compared to Katto and Ohno correlation for macrotubes and channels.

Bar-Cohen and coworkers [36] have studied the effect of flow patterns on CHF by considering over 4000 data points from the literature. The data have been categorized in terms of the Taitel and Duckler [37] flow regime maps, in terms of annular, intermittent, and bubbly flows. They suggest that a similar CHF mechanism as reported by Qu and Mudawar [35] would apply within any given flow regime.

Qu and Mudawar [35] considered a multi-microchannel heat sink made up of 21 parallel microchannels, 0.215 mm wide and 0.821 mm high. Instabilities in the form of vapor backflow into the inlet plenum were observed as the CHF was approached. The backflow resulted in negating any advantage of inlet subcooling on CHF. They recommend the following CHF correlation for saturated boiling. This correlation predicted their data for water and R-113 within 4%.

$$q_{\text{CHF}} = 33.43 \left[h_{\text{LG}} (\rho_G / \rho_L)^{1.11} G \right] / \left[We^{0.21} (L/d_h)^{0.36} \right] \quad (37)$$

Wojtan, Revillin, and Thome [22] investigated CHF in stainless-steel tubes of diameter 0.509 and 0.790 mm for refrigerants R-134a and R-245fa. A mass velocity range of 400–1600 kg/m²s was examined, as well as four heated lengths (between 20 and 70 mm), two saturation temperatures (30–35°C), and subcooling in the range 2–15 K. The following correlation was proposed.

$$q_{\text{cr}} = 0.437 (\rho_V / \rho_L)^{0.073} \cdot We^{-0.24} (L_h/D)^{-0.72} \cdot G h_{\text{LV}} \quad (38)$$

More recent data by Agostini et al. [38] support this correlation. The same research group estimates that a 2×20 mm silicon chip may dissipate 112–250 W/cm² in flow boiling, based on this correlation.

Based on the hypothesis that CHF at a microchannel location is achieved when during evaporation in annular flow; the height of the interfacial waves reaches the mean thickness of the annular film, Revellin and Thome [39] proposed a CHF model. The governing equations for mass, momentum, and energy assuming annular flow, ignoring interfacial wave formation, were first solved to determine the annular liquid film thickness along the channel length. The wave height was modeled using the Kelvin–Helmholtz critical wavelength criterion and the slip ratio. They compared their predicted results with several data sets, both from their group, as well as from others, with a resulting agreement within $\pm 20\%$ for over 90% of the data.

Reillin et al. [40] extended the calculations for a local hot spot on a microprocessor chip. The question investigated was whether the presence of the hot spot may

trigger CHF during forced convection boiling. A 0.4 mm hot spot along the periphery of the entire channel of 0.5 mm diameter and 20 mm length was considered for R-134a. The saturation temperature was 30°C and the mass flux 500 kg/m²s, with no inlet subcooling. It was found that at a uniform heat flux of 218 kW/m², the hot spot should have a heat flux of 3000 kW/m² for CHF to occur. The value of CHF without the hot spot for the same conditions was 396 kW/m². The location of hot spots near the exit of the channel was found to increase the likelihood of getting CHF.

1.2.4 Condensation in Microchannels

Two-phase thermal management devices most commonly utilize a closed-loop arrangement, where the working fluid collects the heat in the evaporator, and rejects it to the ambient, or another working fluid in a condenser. As such, the performance of miniature condensation devices is of great importance. Most of the prior work on condensation flow regime maps was for tube diameters greater than 5 mm (Breber et al. [41] and Soliman [42]). The major nondimensional parameters considered in these studies are the liquid–vapor volume ratio, Froude number and Weber number. These maps primarily focused on the hydraulic diameters between ~ 5 and ~ 25 mm, and are therefore meant for determining the transitions between gravity-dominated stratified wavy flows and shear-dominated annular flows. For small diameter tubes, the tube diameter affects the transition boundaries significantly, and the traditional flow regime maps fail to predict the flow behavior accurately. This is because of the increased influence of surface tension relative to gravity. This effect particularly in noncircular channels is usually termed as “Gregorius effect” and was found to occur with hydraulic diameters around 1 mm [43]. Because of this, the condensate surface curvature changes, which in conjunction with surface tension, leads to pressure gradients in the condensate film and enhanced heat transfer. Some of the prior efforts that focused on small-sized channels include Coleman and Garimella [44] and Yang and Shieh [45]. Dimensional parameters such as liquid and vapor superficial velocities or mass flux and quality were used to characterize the different flow regimes.

Coleman and Garimella [44] observed annular, wavy, intermittent, and dispersed flows. In addition, they found that high-pressure refrigerant flow regimes and transitions are significantly different from those for air–water mixtures. Hence, flow regime maps for air–water data cannot be directly applied to refrigerants. In agreement with Coleman and Garimella [44], Chen and Cheng [46] reported results of visualization studies of condensation of steam in trapezoidal silicon microchannels with hydraulic diameters of 75 μm and 80 mm. They observed that droplet condensation took place near the inlet of the microchannels while an intermittent flow of vapor and condensate was observed downstream of the channels. The traditional annular flow, wavy flow, and dispersed flow were not observed in microchannels.

Bandhauer et al. [47] and Nema [48] derived a set of nondimensional parameters to characterize flow regimes for condensation in practical microchannels. They found that the important flow transitions include annular or wavy regime to

intermittent, discrete- to disperse-wave, intermittent or wavy to annular film flow, intermittent to dispersed flow, and transition to mist flow. They modeled the transition to the intermittent regime (consisting of slug and plug flow patterns) from the wavy or annular regimes using the vapor-phase Weber number (We_G) and the Martinelli parameter (X_{tt}). A critical Bond number ($Bo_{Critical}$) determined whether the transition occurs from the wavy or the annular regime.

$$Bo_{Critical} = \frac{(\rho_L - \rho_G) D_h^2 g}{\sigma} = \frac{1}{\left(\frac{\rho_L}{(\rho_L - \rho_G)} - \frac{\pi}{4} \right)} \quad (39)$$

Below the critical Bond number, they found that annular flow transitions directly to intermittent flow, suppressing the wavy flow regime. This is because of the increased prominence of surface tension over gravity forces as the Bond number or tube diameter is decreased. This transition criterion predicted 85% of data in the correct regimes. The corresponding transition liquid volume fraction was modeled as follows:

$$X_{tt_{slug}} = X_{tt_0} \text{ if } Bo \leq Bo_{Critical}, \text{ where } X_{tt_0} = 0.3521 \quad (40)$$

$$X_{tt_{slug}} = X_{tt_0} + \frac{X_{tt_1} (Bo - Bo_{Critical})}{Bo - Bo_{Critical} + e} \text{ if } Bo > Bo_{Critical} \quad (41)$$

where $X_{tt1} = 1.6 - X_{tt_0}$ and $e = 5.5$.

They also found that a constant value of the modified Froude number could be used to model this transition and this criterion predicted 93% of the data correctly. The annular flow data were divided into three categories; annular film, annular film with mist, and mist flow. To model the transitions, annular film and annular film with mist were considered as one region (annular film), while mist flow was treated as the other region. The transition to the annular film regime from intermittent flow was established based on the vapor-phase Weber number and Martinelli parameter as described above. It was also found that most of the mist flow data existed only for liquid volume fractions (represented by the Martinelli parameter) less than a critical value. This transition criterion using constant vapor-phase Weber number and Martinelli parameter predicted 67% of the data accurately.

The sustenance of annular flow was modeled using the vapor-phase Weber number and Bond number as follows:

For $Bo \leq Bo_{Critical}$, annular film flow exists if $We_G > 6$.

For $Bo > Bo_{Critical}$, annular film flow exists if $We_G > 6 + c(Bo - Bo_{Critical})^d$, where the constants c and d depend on the size of the flow channel.

Based on the Young–Laplace equation, Li and Wang [27] propose the following critical and threshold values of tube diameter.

$$d_{cr} = 0.224 \sqrt{\sigma / (\rho_L - \rho_G) g} \quad (42)$$

$$d_{th} = 7.81 d_{cr} \quad (43)$$

The following regimes for condensation were defined based on these values:

- (a) When $d_{cr} \leq d \leq d_{th}$, the flow condensation will be asymmetrical and the corresponding flow regimes would be annular, asymmetrical annular, lengthened bubble, and bubble flow.
- (b) When $d \leq d_{cr}$, the flow patterns will be annular, lengthened bubble, and bubble flow.
- (c) When $d \geq d_{th}$, the flow regimes will be similar to those for macro-sized channels/tubes, except for the decreasing effect of surface tension with increasing tube size.

In accordance with lack of sufficient and reasonably accurate flow regime maps for condensation in microchannels, there is a dearth of definitive correlations for predicting the pressure drop and heat-transfer performance. Wang and Rose [43, 49] developed a theoretical model for heat transfer during condensation in square and triangular section horizontal microchannels including the effects of interfacial shear stress, gravity, and transverse surface tension force on the motion of the condensate film. Results are presented for fluids of R134a, R22, and R410A in square and triangular section channels of sides in the range 0.5–5 mm and for mass velocities in the range 100–1300 kg/m² s. The results demonstrated significant heat-transfer enhancement by surface tension toward the channel entrance. For smaller channels, the initial enhancement is higher but was found to fall off after a shorter distance along the channel; higher mass velocities were found to result in increased length of enhanced heat transfer.

As the channel sizes become small, different mixture viscosity models as against conventional ones need to be used to make homogeneous flow models work. This is attributed to the lower mixing losses due to the weak momentum coupling between the phases in the smaller channels. Garimella et al. [50] represented the total pressure drop as the summation of the frictional pressure drop in the slug and bubble regions and the pressure drop associated with the transitions between these regions.

Baird et al. [51] conducted an experimental investigation to determine the local heat-transfer coefficient during condensation in 0.92 and 1.95 mm internal diameter tubes. The data showed a strong influence of mass flux and local quality on the heat-transfer coefficient and a relatively weaker influence of system pressure. The observed heat-transfer coefficient was found to increase with mass flux. To predict the heat-transfer coefficients, they proposed an approach similar to that developed by Moser et al. [52] based on a core annular shear-driven gas–liquid flow in which the gas–liquid interface is assumed to be smooth and the liquid film is turbulent, with modifications to the film thickness parameter.

Bandhauer et al. (2006) [53] developed a model based on boundary layer analyses including the shear–stress estimation from the pressure drop models of Garimella et al. [54] developed specifically for microchannels. Their model also indirectly accounts for surface tension through a surface tension parameter in the pressure drop used for the shear–stress calculation.

Agarwal [55] found that both pressure drop and heat transfer increased with increasing vapor quality, increasing mass flux, and decreasing saturation temperature. The flow velocities and other parameters determined for the pressure drop model were used as inputs for the heat-transfer models. The slug and bubble regions were again analyzed separately to determine the slug and film heat-transfer coefficients. A time-averaged refrigerant heat-transfer coefficient was determined by combining the slug and film heat-transfer coefficients according to their transit times through the channel. The proposed pressure drop (equation (44)) and heat-transfer models (equation (45)) predict 95%, and 94% of the data, respectively, within $\pm 25\%$. As the channel hydraulic diameter decreases and the flow velocity increases, the pressure drop and heat-transfer coefficients were found to increase due to a decrease in film thickness and channel diameter. As the aspect ratio increases, the pressure drop and heat-transfer coefficients increased due to an increased occurrence of slugs.

The total pressure drop according to their model is given by

$$\Delta P_{\text{fric, mod el}} = N_{\text{UC}} \cdot \Delta P_{\text{transition}} + \Delta P_{\text{fric,only}} \quad (44)$$

where N_{UC} depends on the microchannel dimensions and aspect ratio, Reynolds number, flow quality and density of phases, and $\Delta P_{\text{transition}}$ and $\Delta P_{\text{fric,only}}$ (which mainly depends on the lengths of slug and bubble flow regimes) are given in Agarwal [55].

On a similar note, they predicted the heat-transfer coefficient for a refrigerant as follows:

$$h_{\text{refg,model}} = h_{\text{slug}} \cdot \left(\frac{l_{\text{slug}}}{l_{\text{slug}} + l_{\text{bubble}}} \right) + h_{\text{film}} \cdot \left(1 - \frac{l_{\text{slug}}}{l_{\text{slug}} + l_{\text{bubble}}} \right) \quad (45)$$

where the method to obtain h_{slug} and h_{film} is detailed in Agarwal (2006) and l_{slug} and l_{bubble} are the slug and bubbly flow regime lengths, respectively. The above equation shows that the heat-transfer coefficient also depends on the lengths of the slug and bubbly flow regimes.

1.3 Rotating Flows

Rotating flows in microchannels can occur in microfluidics applications for flow control, like in electro-osmosis and mixing as in a lab-on-a-chip system. Owing to the nature of the applications, mostly single-phase flow characteristics were focused upon in literature on radially rotating microchannels with centrifugal force-driven flow. Conventional pressure drop estimations seem to work well for rotating microchannels also for slow rotational speeds. But at high flow and rotational Reynolds numbers, Coriolis force dominates the centrifugal force component and needs to be included in estimating the frictional pressure drop [56]. This is found to be true also in the case of microchannels rotating about their own axis [57].

2 Numerical Techniques

For numerical modeling of flow and heat transfer in microchannels for hydraulic diameters of $1\text{ }\mu\text{m}$ and beyond, continuum-based mass, momentum, and energy transport equations are usually sufficient. For nanoscale channels, molecular techniques have been employed in modeling of transport phenomena. The following section briefly describes some of these techniques.

2.1 Continuum Models

Single-phase liquid flows are usually modeled using continuum forms of the momentum and energy conservation equations, with appropriate boundary conditions. These can include various boundary conditions and fluid enhancements such as microsized particle suspension flows [58–61]. These equations for fluids are simultaneously coupled with heat conduction in microchannel walls. Conventional computational methods to predict the flow and temperature fields provide reasonable accuracy when compared with experiments.

For modeling two-phase flows, flow boiling models are usually preferred, which must take into account the incipience of boiling, predict heat-transfer coefficient and pressure drop, and identify the critical heat flux. Modeling in general involves finding the point of incipience of boiling and then proceeding with using the conservation equations depending on the model employed. Boiling heat-transfer models are mostly based on specific flow pattern observations from experiments. For example, in the microlayer model of Jacobi and Thome [26], the critical nucleation radius is experimentally estimated and is used to solve the coupled algebraic-differential equation set based on momentum and energy balance. The annular model of Qu and Mudawar [62] treats the liquid and vapor zones as separate and conservation equations for individual phases are solved. Sarangi et al. [63] used both the homogeneous model and the annular model and used a known exit pressure condition to march backward in the microchannel rather than assuming the pressure at the boiling incipience point based on the assumed inlet conditions. Their results were found to either lie above or below (based on the model type) the experimental results of Qu and Mudawar [62]. The major drawback of most of the two-phase flow models is their applicability to specific flow conditions.

2.2 Molecular Models

For conditions in microchannels, where continuum approaches fail (Knudsen number >0.1) or in flows where capturing of particle–particle and particle–fluid interactions is of significance (e.g., estimating thermophysical properties such as thermal conductivity of nanoparticle suspensions), molecular approaches such as molecular dynamics simulations (MDS), and direct simulation Monte Carlo (DSMC) are

used. This subsection provides the basics of lattice Boltzmann method (LBM), a macromolecular model applied to suspension flows.

The lattice Boltzmann model most widely uses the lattice Bhatnagar–Gross–Krook (LBGK) single-relaxation-time model. This model is based on a discrete Boltzmann equation that is usually solved on a regular lattice. Rather than considering a large number of individual molecules like in the molecular dynamics approach, LBM method considers a much smaller number of fluid “particles” (a large group of molecules) which is much larger than a molecule but still considerably smaller than the smallest length scale of the simulation. This reduces the amount of data which need to be stored for simulation. The particles are allowed to interact through hard-sphere, inelastic, and frictional collisions. The dynamics of the method is given by the LB equation,

$$f_i(\vec{r}, \vec{c}_i, t + 1) = f_i(\vec{r}, t) + \frac{1}{\tau} [f_i^{\text{eq}}(\vec{r}, t) - f_i(\vec{r}, t)] \quad (46)$$

in which f_i is the distribution function of the fluid particles moving in the \vec{c}_i direction, and $f_i^{\text{eq}}(\vec{r}, t)$ is the equilibrium distribution toward which the distribution functions are relaxed. The density (ρ) and velocity (\vec{u}) of the fluid are determined at each lattice node from the first and second velocity moments of the distribution function f_i , respectively. One of the advantages of this method is the easy implementation of the fluid–solid no-slip boundary condition by using the heuristic bounce-back rule. Ladd [64, 65] generalized the bounce-back rule for moving boundaries,

$$f_i(\vec{r}, t + 1) = f_{-i}(\vec{r}, t) + \frac{2\rho t_i}{c_s^2} (\vec{u}_w \cdot \vec{c}_i) \quad (47)$$

Here, $-i$ denotes the link opposite to i in the lattice, \vec{u}_w is the local velocity of the moving boundary, t_i is a weight factor related to i direction, and c_s is the speed of sound. This generalization makes the simulation of moving particles straightforward. It does not only describe the action of the particles on the fluid, but can also be used to compute the hydrodynamic forces and torques action on the particles at a very small computational cost. When equation (46) is solved, the velocities, angular velocities, and positions of the particles can be computed by using methods similar to molecular dynamics. This simulation will provide the time-averaged particle concentration, fluid and particle velocities, fluid–particle interaction force, shear and normal stresses due to particle–particle interaction, and fluid-phase viscous stresses across a microchannel. From this simulation, a rheological constitutive model can be obtained that describes shear and normal stresses in neutrally buoyant suspensions as functions of apparent shear rate and volume fraction of solids. LBM has also been applied for simulating multiphase flows (e.g., Chang [66]).

3 Experimental Techniques

Characterizing the transport phenomena inside microchannels is of vital importance to the design of these devices. Existing research has often focused on measuring bulk

transport characteristics such as friction factors or mean thermal resistances for idealized geometries of microtubes (Judy et al. [67]), rectangular microchannels (Harns et al. (1999) [68]), and trapezoidal microchannels (Wu and Cheng [69]). Such bulk transport characteristics are often influenced by factors, which are otherwise negligible for macroscale channels such as flow rate distribution, surface roughness, and flow inlets and outlets. A direct measurement of the fluid flow and heat transfer inside the microchannel becomes necessary. To this end, this section discusses some nonintrusive optical measurement techniques for flow and/or temperature field inside microchannels. Most of the techniques involve some types of seeding particles that are used to trace the flow and temperature field through tracking the position and intensity, respectively. Section 3.1 gives a detailed introduction to the micro-particle image velocimetry (μ PIV) system for velocity measurement. Section 3.2 discusses various temperature measurement techniques including μ PIV using Brownian motion, two-color laser-induced fluorescence (LIF) for fluid temperature, single quantum dot tracking for near-wall temperature and velocity measurement.

3.1 Measuring Full Field Flow Inside Microchannels Using Micro-PIV

Particle image velocimetry (PIV) is a well-established technique for measuring flow in macroscale channels. In a typical PIV implementation [70] the flow field to be measured is seeded with flow-tracing particles, which are illuminated with either a pulsed light source or a continuous light source gated by the camera shutter. The images of the particles are taken at two known instants and the displacement of the particle images is then statistically determined by correlation techniques. Significant success has been achieved in developing microresolution PIV (micro-PIV) systems. The first implementation of micro-PIV was reported in Santiago et al. [71], where an epi-fluorescent microscope was coupled with a continuous Hg-arc lamp and a CCD camera to record the flow around a nominally 30×10^{-6} m elliptical cylinder in a Hele-Shaw flow cell. The resolution of this system was reported to be 6.9×10^{-6} m \times 6.9×10^{-6} m \times 1.5×10^{-6} m. As the time delay between image exposures was controlled through the camera shutter, application of such a system is limited to very slow flow. A micro-PIV system capable of measuring over a broader range of velocity was described in Meinhart et al. [72]. Velocity distribution in a glass microchannel, 30×10^{-6} m deep and 300×10^{-6} m wide, were measured. A 5×10^{-9} s (5 ns) pulsed Nd: YAG laser was used as the light source to illuminate the 200×10^{-9} m diameter fluorescent particles. The images of the flow field were recorded using a $1300 \times 1030 \times 12$ -bit CCD camera. The streamwise velocity profile estimated from the PIV measurements on one plane agreed within 2% with analytical results.

Recently, the effects of out-of-focus particles were examined in [73–76] considering the effects of out-of-plane motions and Brownian motions. These work laid the foundation for applying the micro-PIV technique to characterize flow inside microchannels [77]. Park et al. [78] demonstrated that confocal laser scanning

microscopy (CLSM) can be used to improve the out-of-plane resolution. Mielnik and Sætran [79] showed that the out-of-focus resolution can be further improved using a selective-seeding technique to reduce the background noise caused by out-of-focus particles associated with volume seeding technique.

In a recent work by Wei and Joshi [16], velocity distributions inside slightly tapered microchannels were directly measured using the nonintrusive micro-PIV technique. An image parity exchange technique developed by Tsuei and Savas [80] for macroscale PIV was adapted to improve the accuracy in velocity measurement near the wall–fluid interface. The measured results agree well with theoretical predictions. It was found that the location of the maximum velocity deviates significantly from the mid-plane in the depth direction for the slightly tapered microchannel.

3.1.1 Typical Micro-PIV System for Fluid Flow

A typical micro-PIV system is illustrated in Fig. 3. The imaging system includes an Epi-fluorescent inverted microscope (Nikon Eclipse TE2000-S), dual Nd: YAG lasers (Gemini PIV 200) and a cooled interline CCD camera. The epi-fluorescent filter cube consists of an excitation filter, a dichromatic mirror, and an emission filter. The pulsed green light from the dual lasers expands through a beam expander, and passes through the diffuser, the excitation filter, and the dichromatic mirror to

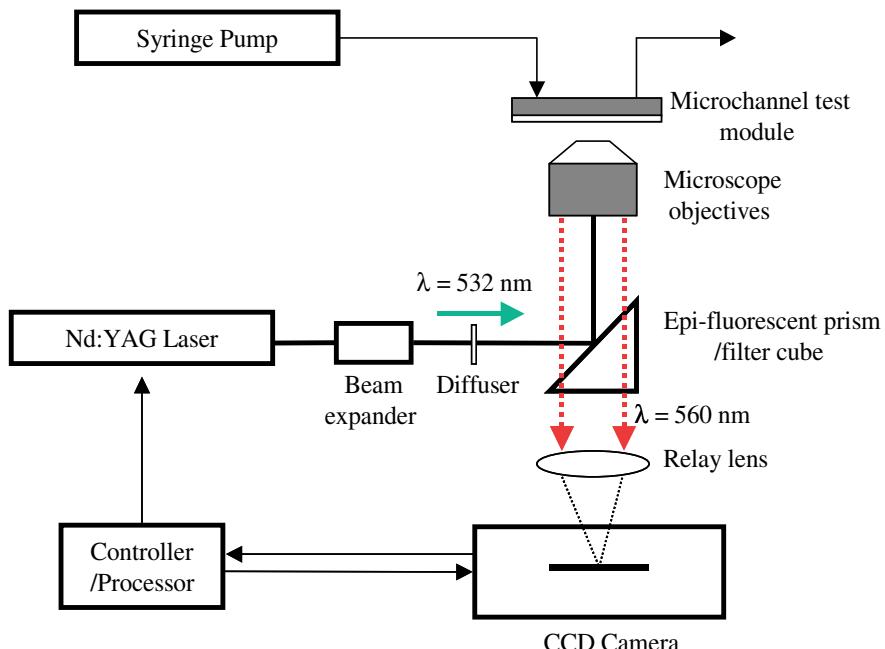


Fig. 3 Schematic of a micro-PIV test setup

the microscope objectives. Fluorescent particles (e.g., Nile Red, Molecular Probe) excited by this incident green light emit red light peaked around 560×10^{-9} m (560 nm). The emitted red light passes through the long pass emission filter to reach the CCD camera, which operates in the double-frame single-exposure mode. Scattered green light from the background and ambient is blocked at the emission filter.

For a particular optical system, a scale factor needs to be determined as the ratio of the known physical size of the channel width to the image size in the CCD camera array. This scale factor is used to obtain the actual displacement between two laser pulses with known time intervals. The fine focus knob of the microscope objectives can be used to traverse the measurement focus plane, two-dimensional velocity measurement is thus obtained at different depths of the channel.

It is noted here that the displacement of the objectives (L_o) is different from the displacement of the objective planes. If an air-immersion lens is used to record the image and the objective plane is located in water, which has a different refractive index, from Fig. 4, equations (48) and (49) can be derived based on the geometry and the Snell's law of refraction. The ratio of the measurement plane displacement to that of objective lens displacement is equal to 1.333, which is the ratio of the refractive index of water (n_w) to air (n_a). A near optical axis condition is assumed in deriving equations (48) and (49).

$$L_o = (s_o - L_i) \frac{n_w}{n_a} \quad (48)$$

$$\Delta L_o = -\Delta L_i \frac{n_w}{n_a} \quad (49)$$

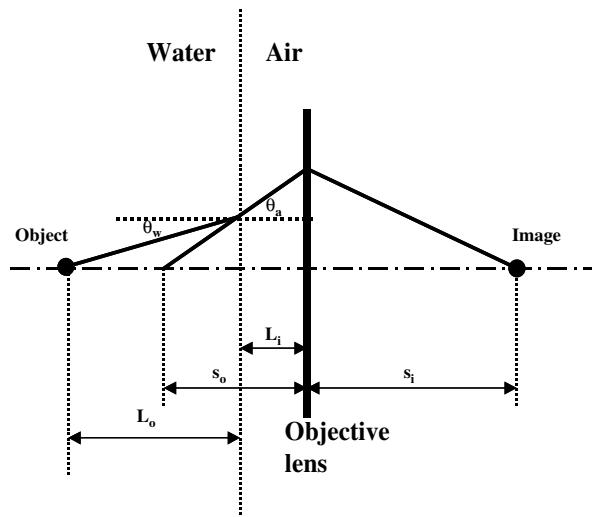


Fig. 4 Effects of the refractive index on the object plane displacement

One important parameter for PIV measurement is the time interval between the two laser pulses. The displacement between the two interrogation pairs is the product of the time interval and the velocity. An insufficient time interval will make it difficult to detect a slow velocity whereas an excessive time interval will cause severe loss of pairs as a lot of particles will travel out of the second interrogation area during the interval. As a general rule, the time interval for cross-correlation is selected so that the particles will travel no larger than one-fourth of the interrogation window width. For typical experimental conditions, the time interval in the range 5–100 μs can be used as a starting point.

Compared to a traditional PIV, micro-PIV is different in that volume illumination is used instead of the light sheet method. In addition, due to the small channel size, it is important to carefully select the seeding particles and apply proper image-processing techniques to improve the signal-to-noise ratio (SNR).

3.1.2 Seeding Particles

In selecting particle sizes, consideration should be given to tracking fidelity, visibility, spatial resolution, and Brownian motion effects. A convenient measure of particle-tracking fidelity is the relaxation time, which basically describes how long it takes for the particle to attain velocity equilibrium with the fluid. For Stokes flow, the relaxation time is proportional to the square of the particle diameter. Small particles should be used in order to track the fluid flow faithfully. For the particles used in the current study, the relaxation time is in the order of 10^{-8} s. However, particles should also be large enough to be visible by the imaging optics.

Another issue closely related to the particle size is the Brownian motion effects. Brownian motion results from the random collisions between the fluid molecules and the suspended microparticles. The relative error in measuring velocity caused by the Brownian motion is estimated to be

$$\varepsilon_B = \frac{1}{u} \sqrt{\frac{2D}{\delta t}} \quad (50)$$

where u is the fluid velocity, δt is the time interval, and D is the Brownian motion diffusion coefficient given by

$$D = \frac{\kappa T}{3\pi \mu d_p} \quad (51)$$

In this equation κ is the Boltzmann's constant, T is absolute temperature, μ is the dynamic viscosity, and d_p is the particle diameter. Clearly, smaller particles will cause larger uncertainties in the velocity measurement due to Brownian motion. For this study, the relative error due to Brownian motion is typically less than 2%. The error due to Brownian motion is unbiased and can be further reduced by ensemble averaging over repeated measurements. This diffusive error is inversely proportional to the square root of the size of the ensemble.

Particle concentration is another important parameter to be considered in a micro-PIV experiment. Too many particles will increase the background glow and thus reduce the signal-to-noise ratio (SNR). On the other hand, insufficient seeding will result in a weak signal. In general at least 2–3 particles are required for each interrogation area. The SNR can be enhanced by the ensemble averaging technique as described in Meinhart et al. [81]. It was shown that ensemble averaging of the correlation function before the peak detection is superior to averaging the image or the vector map to enhancing the SNR.

For typical applications, the volumetric concentration of the fluorescent particles should be maintained in the range 0.05–0.07%. The size of the ensemble is typically around 40 to sufficiently enhance the SNR and reduce the uncertainties caused by Brownian motion.

3.1.3 Spatial Resolution and Depth of Measurement

The spatial resolution of any PIV measurement is defined by the size of the interrogation region and ultimately limited by the effective particle image size. The image of a finite-diameter particle is the convolution of the diffraction limited point response function with the geometric image of the particle. In a study performed at Georgia Tech, a long working distance objective with magnification $M = 60$ and numerical aperture NA = 0.7 was used to resolve the 0.5×10^{-6} m particles. The effective image diameter was about 52×10^{-6} m, and it corresponds to 0.87×10^{-6} m when projected back to the flow field. The interrogation area used was 64 pixels in streamwise direction and 32 pixels in the transverse direction. This corresponds to spatial resolution of $7.3 \times 10^{-6} \times 3.6 \times 10^{-6}$ m. When 50% overlap was used for the interrogation, about 20 vectors were measured across the channel width of 34×10^{-6} m.

One important parameter in interpreting the velocity data is the depth of measurement, since the out-of-focus particles also contribute to the correlation function. Olsen and Adrian [75] derived the depth of correlation based on the assumption that the particle image intensity follows Gaussian distribution and the illumination over the interrogation area is uniform. The depth of correlation is given as

$$\delta z_{\text{corr}} = 2 \left[\frac{(1 - \sqrt{\varepsilon})}{\sqrt{\varepsilon}} \left(f^{\#2} d_p^2 + \frac{5.95(M + 1)^2 \lambda^2 f^{\#4}}{M^2} \right) \right]^{1/2} \quad (52)$$

where ε is a threshold defined as the ratio of the weighting function for the out-of-focus particles to that for the in-focus particles, $f^{\#}$ is the f -number of the optics, λ is the recording wavelength, d_p is the particle diameter, and M is the magnification. Typically, ε is chosen as 0.01 [75]. For the current imaging system settings, equation (52) gives a measurement depth of about 2.65 μm .

3.1.4 Image Processing and Correlation Analysis

In PIV, the measured velocity is defined by the particle image displacement divided by the product of the magnification and the time intervals. The measured velocity is essentially an average over the interrogation region and is assigned to the centroid of the interrogation area. Effectively, the velocity is thus given by [75]

$$u_m = \frac{\int u(X,t) W(X) dX}{\int W(X) dX} \quad (53)$$

where $W(X)$ is the weighting function, which determines the relative contribution to the correlation function at different positions of the interrogation area. If the velocity is fully developed, it is only dependent on the y -coordinates as shown in Fig. 5. The velocity within the interrogation area can be expanded into Taylor series around the centroid of the interrogation area as

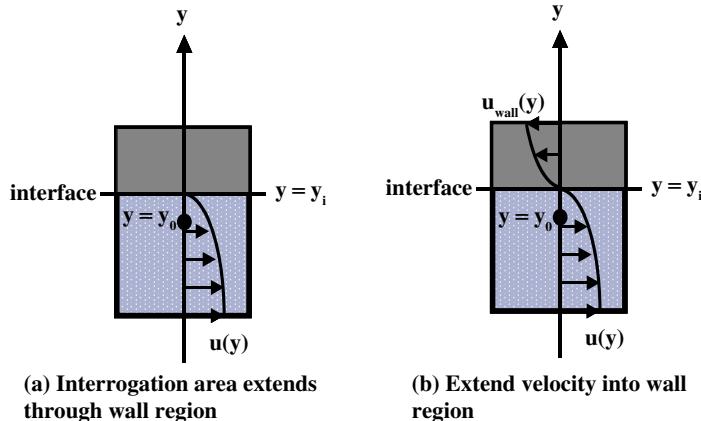


Fig. 5 Effects of the presence of wall boundary inside interrogation area on the velocity measurement

$$u(y) = u_0 + \frac{\partial u}{\partial y} \Big|_{y_0} (y - y_0) + \frac{1}{2} \frac{\partial^2 u}{\partial y^2} \Big|_{y_0} (y - y_0)^2 + O(y - y_0)^3 \quad (54)$$

where u_0 is the true velocity at the centroid. Clearly, the velocity is biased when there is a velocity gradient within the interrogation area. If a uniform weighting function is used, the measurement bias is the second order as the first-order term is terminated during the integration across the interrogation area.

If the interrogation area extends beyond the wall boundary, however, additional large bias in the velocities will occur. The largest bias occurs when the centroid of the interrogation area is located on the solid–liquid interface. As the velocity should be zero for the no-slip boundary condition, a nonzero velocity will be measured as

only half of the interrogation area contributes to the correlation function. This issue could be addressed if the velocity field can be extended to the wall area according to equation (49), which is proposed by Tsuei and Savas [80] for macroscale PIV.

$$u_{\text{wall}}(y) = 2u(y_i) - u(2y_i - y) \quad (55)$$

where y is the coordinate orthogonal to the flow direction, as shown in Fig. 5, y_i is the location of the interface. A Taylor series expansion around the centroid reveals that the extension in equation (51) is accurate to the first order; consequently using equation (51) improves the accuracy near the boundary region to the same degree for the bulk fluid region. Implementation of equation (51) is straightforward. With reference to Fig. 5, point y in the wall region and point $(2y_i - y)$ inside the channel are symmetric about the interface located at y_i . Since the interface velocity $u(y_i)$ is zero for the no-slip boundary condition, equation (51) indicates a reflection and a reverse of the velocity field inside the channel about the interface. This is achieved through the image process steps illustrated in Fig. 6 such that images for both frames are reflected about the interface. The wall region of the second frame is replaced with the reflected image of the fluid region in the first frame and vice versa. This process is referred to as image parity exchange (IPX) in Tsuei and Savas [80] and will be used hereafter. In micro-PIV measurement, the wall boundary is usually hard to define as it blurs as a result of defocusing. In the present study, a Matlab (commercial technical computing software) function which searches for the maximum image intensity gradient is used to locate the channel boundary.

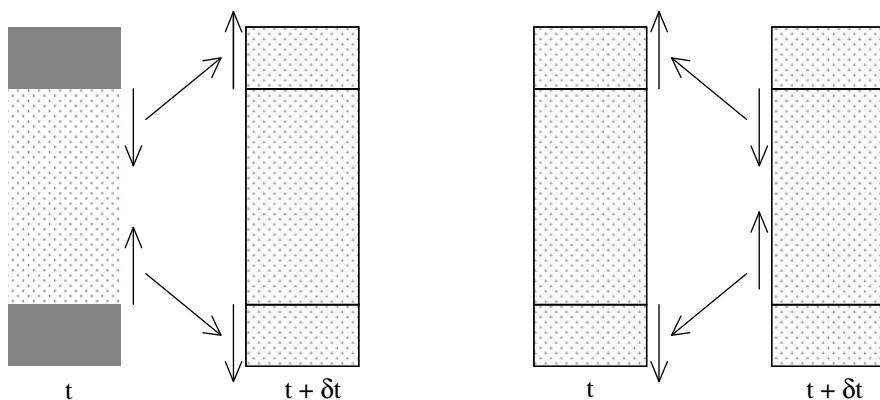


Fig. 6 Implementation of equation (55) to extend the velocity field to the wall region

It is noted here that before the IPX process, mean pixel intensity value is subtracted from the raw image maps to reduce the background noise. To identify the effects of IPX, correlation analysis is also performed on the image maps where only mean intensity is subtracted. These results are compared with predicted velocity profiles from the series solution for rectangular channel [82] and numerical solution

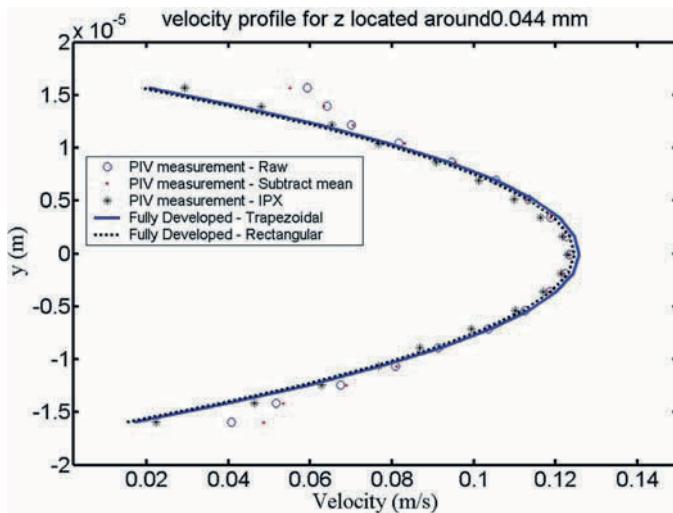


Fig. 7 Velocity profile using different image processing techniques

for trapezoidal channel. As can be seen from Fig. 7 applying IPX before the correlation significantly improves the accuracy of the velocity profile prediction near the wall. This is particularly useful for measuring velocity for very narrow channels where boundary effects are significant such as the channel sizes in the current study.

To determine the velocity vectors, adaptive correlation can be first conducted to the image maps. The average velocity map obtained can then be used as the offset for the ensemble average correlation. This offset is applied to reduce the loss-of-pair issue successfully.

3.1.5 Velocity Profiles for Fully Developed Flow

Figure 8 plots the velocity profile for a channel of $34 \mu\text{m}$ wide and $118 \mu\text{m}$ deep at different depth positions for the flow rate of $2.78 \times 10^{-10} \text{ m}^3/\text{s}$ (1 ml/h). Predicted results based on analytical solutions for an effective rectangular channel and numerical simulations for a trapezoidal channel are plotted for comparison. As can be clearly seen, the measured velocity profile matches very well with results predicted from laminar theory for all the planes measured. Even though the actual channel is not perfectly trapezoidal, the trapezoidal model results generally agree better with experimental measurements than the rectangular model, particularly for the measurement plane located away from the geometric symmetry plane. These comparisons clearly prove the validity of the laminar theory for microchannels. These results also confirm the feasibility of using micro-PIV as a diagnostic tool for microfluidics for channels of these sizes.

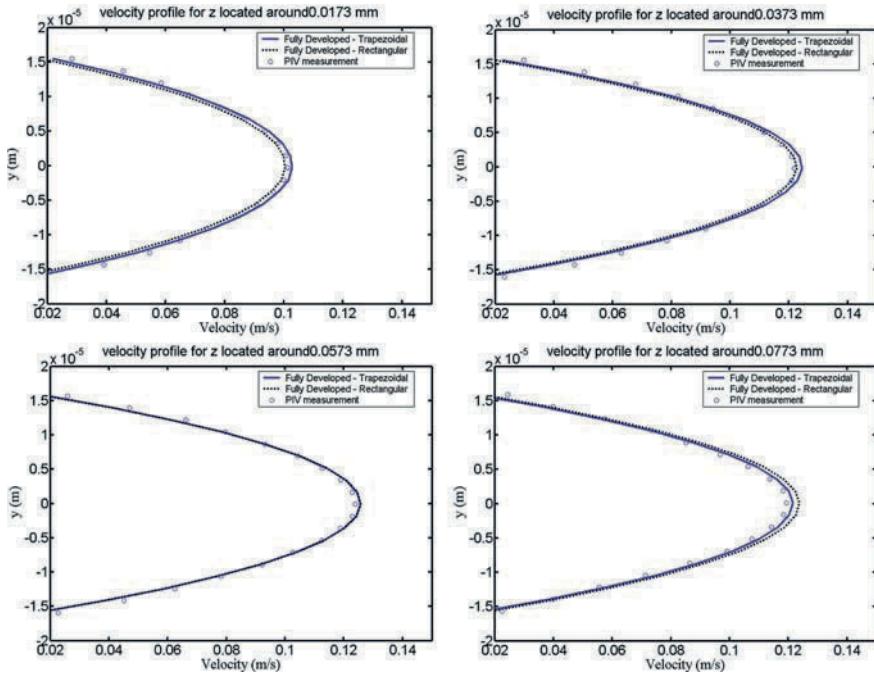


Fig. 8 Velocity profile at different focus planes (flow rate $2.78 \times 10^{-10} \text{ m}^3/\text{s}$ or mean velocity 0.07 m/s)

3.2 Measuring Fluid Temperature Inside Microchannels

Measuring temperature field inside microchannels is very challenging due to the small length scale involved. Optical nonintrusive techniques are probably the only feasible way for this type of work. This section discusses various temperature measurement techniques including micro-PIV using Brownian motion, two-color laser-induced fluorescence (LIF) for fluid temperature, single quantum dot tracking for near-wall temperature and velocity measurement.

3.2.1 Temperature Measurement Using Micro-PIV

As discussed in the previous section, velocity measurement using micro-PIV is often affected by the Brownian motion of the seeding particles. Derived from the balance of the Stokes drag force and the diffusion force, the diffusion coefficient is a function of the absolute temperature and the viscosity as indicated in equation (51), which is rewritten here for convenience.

$$D = \frac{\kappa T}{3\pi \mu d_p} \quad (56)$$

where κ is the Boltzmann constant, μ is the viscosity of the fluid and d_p is the particle size, and T is the absolute temperature.

The RMS displacement x_{rms} by the particles within the time window of Δt is given by

$$x_{\text{rms}} = \sqrt{2D\Delta t} \quad (57)$$

Clearly, this displacement is proportional to the square root of the ratio of temperature to viscosity. In the presence of the Brownian motion, according to Hohreiter et al. [83], the probability function for the particles follows a Gaussian distribution, which is shown here in equation (58).

$$f = (4\pi D\Delta t)^{-3/2} \text{EXP}\left(-(x - x_0 - u\Delta t)^2 / 4D\Delta t\right) \quad (58)$$

where x is the position of the particle, x_0 is the original position of the particle, and u is the mean velocity. The net effects are the broadening of the peak and decrease in the peak height in cross-correlation. The broadening in the peak width can be quantified by the difference in the correlation peak areas between the auto-correlation and the cross-correlations, as given by [83],

$$\Delta A = 4M^2\kappa/3d_p(T/\mu)\Delta t \quad (59)$$

where M is the magnification of the objective lens.

This technique is implemented in Hohreiter et al. [83] to measure the water temperature inside a microchannel with no net flow. Images were taken sequentially at a fixed time interval. The peak width area of the auto-correlation of the first frame is used as the reference. The first frame is then correlated with the second and subsequent frames. As shown in equation (59), the increase in the peak area is linearly proportional to time. In the plot of the peak area and time, the slope represents the temperature multiplied by a constant. This relationship holds true as long as the heat transfer is in steady state such that the temperature and viscosity does not change with time. It is noted that the viscosity and particle size are needed as the input to calculate the temperature.

The measurement uncertainty of the temperature field depends on the accurate account of the peak area, particle diameter, and time interval. Due to the temperature dependency of the viscosity, the measurement uncertainty is also depending on the temperature. For the system described, the uncertainty is $\pm 3^\circ\text{C}$.

Simultaneous measurement of the temperature and the velocity is limited by the proper time intervals between the image frames. On the one hand, the time interval should be sufficiently long to allow measurable peak area increase. On the other, the time delay has to be short to avoid losing pairs between the cross-correlation pairs. Hohreiter et al. [83] reported the measurable velocity for water with temperature higher than 20°C is less than 8 mm/s for their system. Such difficulties stem from the fact that the same tracer particles are used for both Brownian motion and bulk

flow. The following section introduces a technique with dual tracing particles where the temperature measurement and velocity can be separated.

3.2.2 Two-Color Laser-Induced Fluorescence (LIF) for Temperature Measurement

Temperature-sensitive fluorescent dyes have been used for temperature measurement in fluid systems [84]. The fluorescent intensity absorbed by the imaging camera is proportional to the quantum efficiency, incident light intensity, dye concentration, and absorptivity of the dye. The fluid field temperature can be mapped if the incident light intensity is uniform and constant and the dye concentration is constant. However, the requirement for uniform and constant incident light is very difficult to achieve and the imaging system itself adds uncertain noises to the signal. One elegant way to address this is to introduce a second dye, which is nonsensitive to the temperature. The ratio of the intensity between the two should help to reduce the dependency on the incident light and the noises by the cameras and optics. This is shown in equation (60) for the system of Rhodamine-B and Rhodamine-110 by Kim et al. [85]. The ratio of the absorption spectral intensity $\frac{I_{0\text{rhb}}}{I_{0\text{rh110}}}$ is fixed for single light source. The fluorescent intensity ratio $\frac{I_{\text{rhb}}}{I_{\text{rh110}}}$ is only dependant on the quantum efficiency ratio $\frac{Q_{\text{rhb}}}{Q_{\text{rh110}}}$ since the absorptivity ratio $\frac{\varepsilon_{\text{rhb}}}{\varepsilon_{\text{rh110}}}$ and the concentration ratio $\frac{C_{\text{rhb}}}{C_{\text{rh110}}}$ can be fixed.

$$\frac{I_{\text{rhb}}}{I_{\text{rh110}}} = \frac{I_{0\text{rhb}}\varepsilon_{\text{rhb}}C_{\text{rhb}}Q_{\text{rhb}}}{I_{0\text{rh110}}\varepsilon_{\text{rh110}}C_{\text{rh110}}Q_{\text{rh110}}} \quad (60)$$

If the difference in the quantum efficiency between the two dyes are significant enough, the temperature can be mapped based on the intensity ratios extracted from the images. The dual dye system of Rhodamine-B and Rhodamine-110 [85] fits this requirement. Rhodamine-B has a temperature sensitivity of 2% per K while Rhodamine-110 has a temperature sensitivity of 0.13% per K.

A typical optical system [86] is illustrated in Fig. 9, which includes an Ar-ion laser (wavelength = 488 nm). A scanning mirror sweeps vertically to form the light sheet to illuminate the test field of interest. The fluorescent light emitted from the particles was splitted by a beam splitter, which transmits the wavelengths $\lambda_T \leq 520$ nm and reflects wavelengths $\lambda_R \geq 570$ nm. As can be seen from Table 2, this splitter transmits light emission from Rhodamine-110, but reflects majority of Rhodamine-B. Since the background light has wavelength of 488 nm, another filter rejects wavelengths below 495 nm for Rhodamine-110. An optional filter helps eliminate wavelengths higher than 570 nm for Rhodamine-B. The CCD cameras are arranged to detect the Rhodamine-110 and Rhodamine-B, respectively.

To accurately calculate the intensity ratios at every location, it is very important to align the two cameras with each other. A calibration plate with equally spaced grid pattern can be used to calibrate the image coordinates against the physical coordinates. A simple correlation can be performed to give the coefficients for the transfer

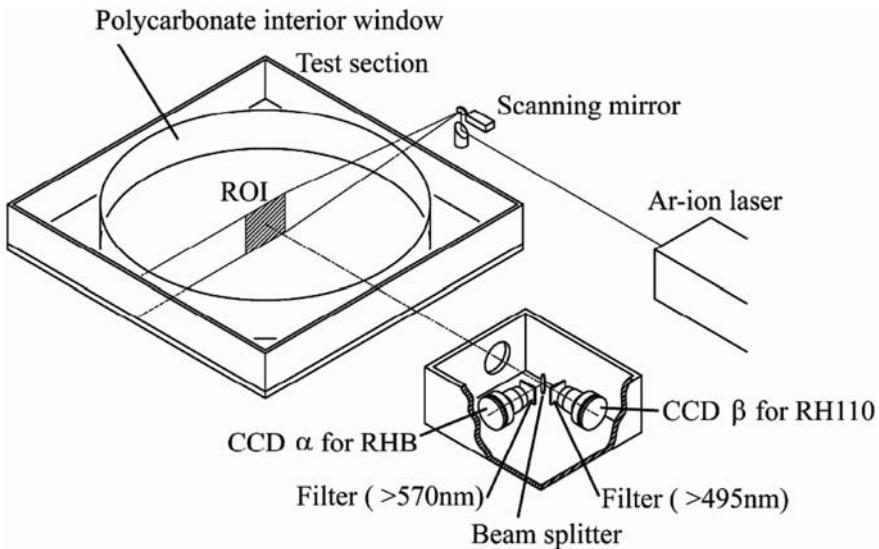


Fig. 9 Measurement of temperature field of a Rayleigh–Benard convection using two-color laser-induced fluorescence (Reprinted with permission from Sakakibara and Adrian [86])

Table 2 Characteristics of Rhodamine-B and Rhodamine-110

Dye	Molecular weight	Peak absorption wavelength	Peak emission wavelength	Quantum yield	Absorptivity at 488 nm
Rhodamine-B	479.02	554	575	0.31	4.4
Rhodamine-110	366.8	496	520	0.8	34

function from the physical coordinates to the image coordinates. The dual camera system allows the simultaneous measurement of the image intensity for both dyes. However, it is found in Sakakibara and Adrian [86] that intensity ratio measured by the two cameras is spatially dependant even when the temperature field is uniform. This may have been caused by the differences in object lens and the astigmatic aberrations by the inclined beam splitter. To overcome this problem, a location by location correction has to be performed through Fourier transformation [86].

A simplified version of the test system is used in Kim et al. [85], where the beam splitter is removed. To split between the Rhodamine-B and Rhodamine-110, an alternating filter is used before the emission from the particle reaches the single camera. Since only one camera is used, there is no need to do a pixel by pixel correction of the image coordinates.

Before actual temperature field measurement, the system is calibrated for a constant fluid temperature case. The intensity ratios are correlated with the known temperature. Such an example is given in Fig. 10. As can be clearly seen, the temperature sensitivity for Rhodamine-B is significantly higher than that of

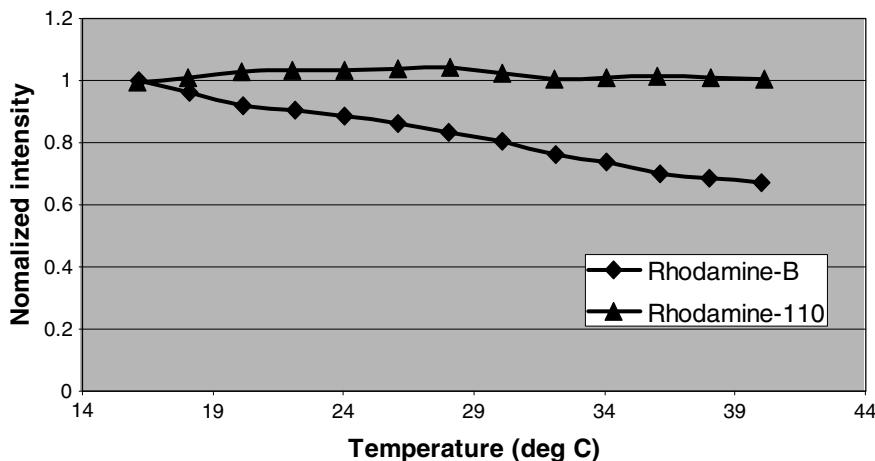


Fig. 10 Comparison of fluorescence intensities of Rhodamine-B and Rhodamine-110, normalized by the corresponding intensities at 14°C (Reprinted with permission from Kim et al. [85])

Rhodamine-110. The calibrated system can then be used to measure the actual field temperature. Kim et al. [85] reported the uncertainty is 2.3°C for spatial resolution of 19 μm.

One limitation of the system reported here is that a light sheet has to be employed. Similar to micro-PIV techniques, a volumetric illumination system can also be implemented where the emissions from the dual dyes are received by a single camera sequentially through an alternating camera. However, such a system is subject to the optical influences of the out-of-focus particles much like the micro-PIV. Furthermore, since the temperature measurement is based on the image intensity, the background noise by the out-of-focus particles is even more unpredictable. To address this issue, a total internal reflection fluorescence (TIRF) technique can be used [87].

3.2.3 TIRF for Simultaneous Measurement of the Temperature and Velocity for Near Regions

Total reflection happens when the incident light angle is larger than the critical angle determined by the refractive index for both media. An evanescent wave is generated when total reflection happens. The evanescent electromagnetic field decays exponentially from the interface as shown in equation (61).

$$J(z) = J_0 \exp(-z/d) \quad (61)$$

where $J(z)$ is the intensity distribution, J_0 is the wall intensity, z is the distance from the wall, and d is the penetration depth. Equation (62) is used to determine the penetration depth.

$$d = \frac{\lambda_0}{4\pi} \left[n_1^2 \sin^2 \theta - n_2^2 \right]^{-1/2} \quad (62)$$

where λ_0 and θ are the incident wavelength and angle, respectively, n_1 and n_2 are the refractive index for the medium, respectively.

Typical penetration depth is about 100–200 nm into the fluid field and thus allows selective illumination of the fluid field. Unlike the flood illumination used in micro-PIV, contributions from the bulk of the fluid are minimized.

Using this technique, Guasto and Breuer [88] conducted simultaneous measurement of temperature and velocity with single quantum dot tracking. Quantum dots (QDs) are nanometer-sized tracer particles made out of semiconductor nanocrystals. These QDs are single fluorophores with significant quantum efficiency and resistance to photo-bleaching. The small size (5–25 nm in diameter) allows QDs to be used for true nano-scale measurements. They are typically coated with surface coatings to achieve solubility in either organic or inorganic solvents. In the system reported in Guasto and Breuer [88], CdSe/ZnS QDs are used together with the temperature-insensitive Rhodamine-110 in a borate buffer. QDs exhibit significant temperature-dependant quantum efficiency. Similar to LIF, Rhodamine-110 is used as the background dye for normalization purposes since its quantum efficiency is insensitive to the temperature. A schematic illustration of the optical system used in Guasto et al. [87], is reproduced in Fig. 11. The PDMS channel dimension is $340 \times 30 \times 10$ mm. Dual wavelength images produced by the fluorescence emission of the QDs and Rhodamine-110 are separated by wavelength using an image splitter through a dichronic mirror with cutoff wavelength at 560 nm. The images are filtered before recombination and projection onto the ICCD. A silicon reference grid is used for image coordinates calibration purposes. The images of the QDs on the camera are defined by the diffraction limited spot since the diameter is significantly smaller than the light wavelength. QDs are detected by thresholding the maximum pixel value of an image until a possible particle is found. A set of the brightest particles are taken to achieve sufficient signal-to-noise-ratio (SNR). Roughly, 25 QDs are detected for an effective area of 22×22 μm .

To measure the temperature, the peak intensity of the single QD is normalized by the single pixel intensity from the Rhodamine-110 reference images. To account for the temporal fluctuations, the Rhodamine-110 reference intensity is averaged over time.

To determine the temperature sensitivity of the QDs and Rhodamine-110, verification experiments are done first using bulk solutions with flood illumination. The result indicates significantly higher temperature sensitivity for QDs than for Rhodamine-110. The calibrated curve for intensity and temperature is then used for temperature measurement of the microchannel flow shown in the figure. The temperature of the flow is controlled by the controller attached to the microscope stage. Although the measured intensity does not follow a perfect Gaussian distribution, the mean value when normalized by the reference intensity shows good agreement with the temperature measured independently. The uncertainty is within 5°C although better accuracy is expected since some transient effects may have not been completely removed in the experiments.

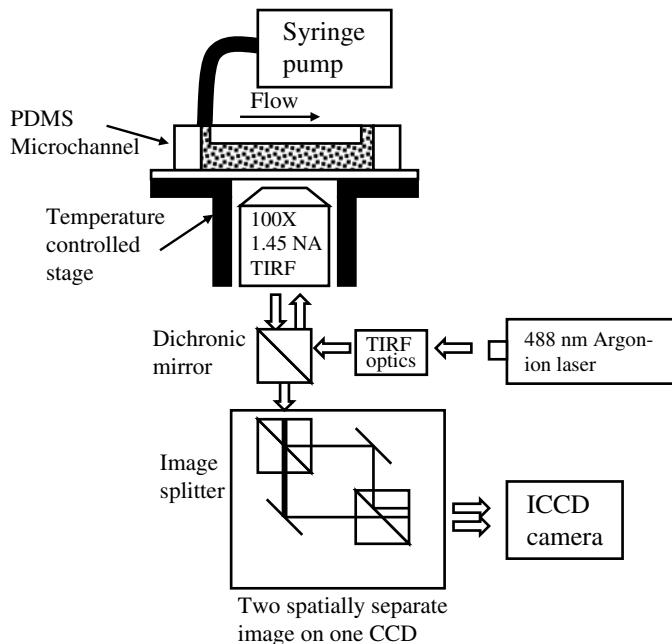


Fig. 11 Setup for simultaneous imaging of solutions of Rhodamine-110 and QDs using TIRF (Reprinted with permission from Guasto et al. [87])

After the QDs locations are detected in the image, the displacement of the QDs can be used to determine the stream and cross-stream velocity. Due to the extremely small diameter of the QDs, the Brownian motion effects are significant when measuring the velocity. The mean displacement is inversely proportional to the square root of the diameter. The net effect of the Brownian motion is the broadening of the velocity distribution. To minimize this effect, ensemble average is performed over 2000 samples, which reduced the standard deviation significantly.

It is noted here that QDs are typically subject to bleach and photostability issues. Although the shelf life is typically 6 months, QDs degrade rapidly once diluted. It is therefore important to carry out experiments within a short period of time (hours). QD delays due to QD blinking phenomena could further impose illumination time limits on the test.

In summary, this section provided introduction of the measurement techniques inside microchannel for velocity and temperature. Details of the micro-PIV technique for velocity measurement are provided. While velocity measurement techniques have seen significant progress in micro-PIV, temperature measurement has remained a challenge. A brief introduction is given to three representative techniques in micro-PIV with Brownian motion, two-color LIF, and TIRF using QDs. Significant development is needed to improve the measurement accuracy in the future.

4 Microfabrication and Assembly

Currently, most microchannels used in applications fall into the range of 30–300 μm [89]. Microchannels can be fabricated using micromachining, molding, or embossing in a variety of materials, such as Si, glass, metals, or polymers [90–92]. The most important feature of the microchannels is their large surface-to-volume ratio, which leads to a high rate of heat transfer (or heat-transfer coefficient). In addition, microchannels can be integrated close to the Si chips to enable compact heat sinks.

Microchannel heat sinking was first proposed for very large-scale integrated (VLSI) electronic components in 1980 s. As shown in Fig. 12, Tuckerman and Pease [1] studied chip cooling with an on-chip microfluidic heat sink and demonstrated that a heat flux of $\sim 790 \text{ W/cm}^2$ can be removed with a temperature rise of 71°C . This translates into a thermal resistance as low as 0.09°C/W . The deep microtrenches were first formed through micromachining in the backside of a Si substrate with integrated circuits. A cover plate was then bonded on top of the microtrenches to form enclosed microchannels. Following this landmark work, numerous studies have been carried out to implement on-chip microchannel cooling [93–95]. The most common bonding technique for high-quality bonding of Si wafers is field-assist bonding (anodic bonding) process. Such bonding technique usually requires ultra smooth/flat/clean bonding interfaces, as well as high pressure,

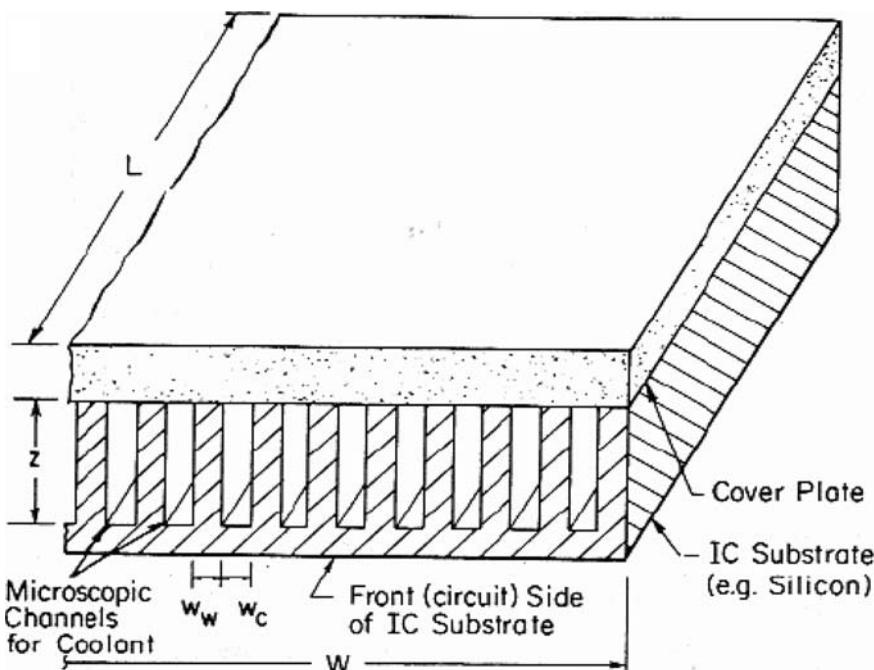


Fig. 12 Microchannel heat sink integrated in IC substrate

high temperature (up to $\sim 1000^{\circ}\text{C}$), and high voltage (500–1000 V). These processing conditions are not very friendly to back-end-of-the-line (BEOL) wafers, where CMOS devices and on-chip interconnects are completed. Therefore, CMOS compatible processes are necessary for fabrication of on-chip microchannels following BEOL processing.

Several low-temperature CMOS-compatible microchannel fabrication approaches have been studied in the past decades. Electroplating method to form metallic microchannels is shown in Fig. 13 [92]. In this approach, a photoresist was used to define the microchannels and metallic shells were then deposited through electroplating. After the photoresist is dissolved in acetone bath, hollow microchannels can be formed. As shown in Fig. 14, another example is the buried microchannels fabricated in Si substrate through two-step electrochemical process, which combines formation of porous Si and electropolishing [93].

Dang et al. [97, 98] also studied a monolithic method to fabricate polymer-capped microchannels in Si substrate using a sacrificial polymer at low temperature, as illustrated in Fig. 15. Instead of wafer bonding, the deep trenches are first filled with a sacrificial polymer by spin coating. After surface planarization through mechanical polishing, the trenches are overcoated with a layered structure. The sacrificial polymer is then decomposed by heat treatment. The layered overcoat consists of a SiO_2 thin film ($1\text{--}2 \mu\text{m}$) deposited with PECVD at low temperature ($\sim 150^{\circ}\text{C}$) and a thicker layer of Avatrel 2195P polymer that is spin coated and cured as the structural polymer ($\sim 25 \mu\text{m}$). Since the SiO_2 film is porous, it allows permeation of the gaseous products from decomposition of the sacrificial polymer. The SiO_2 also plays an important role in maintaining the structural integrity of the overcoat by preventing “sagging” of the above structural polymer. In fact, decomposition of the sacrificial polymer and curing of the structural polymer occur simultaneously during a heat treatment in a program-controlled oven. To further increase the mechanical

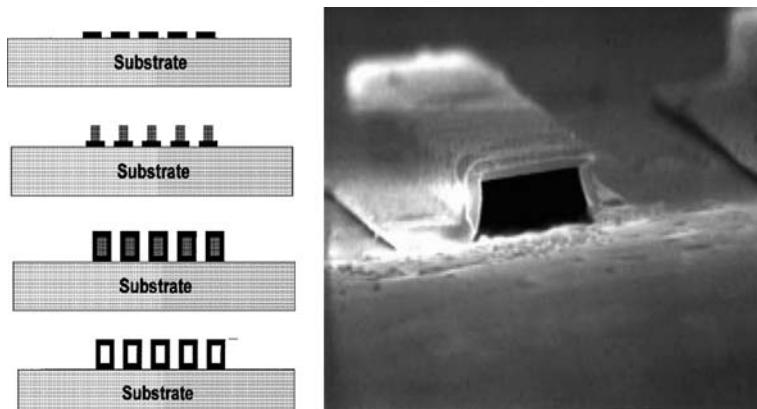


Fig. 13 Metallic microchannels formed by electroplating around sacrificial photoresist lines (Papautsky et al. [92])

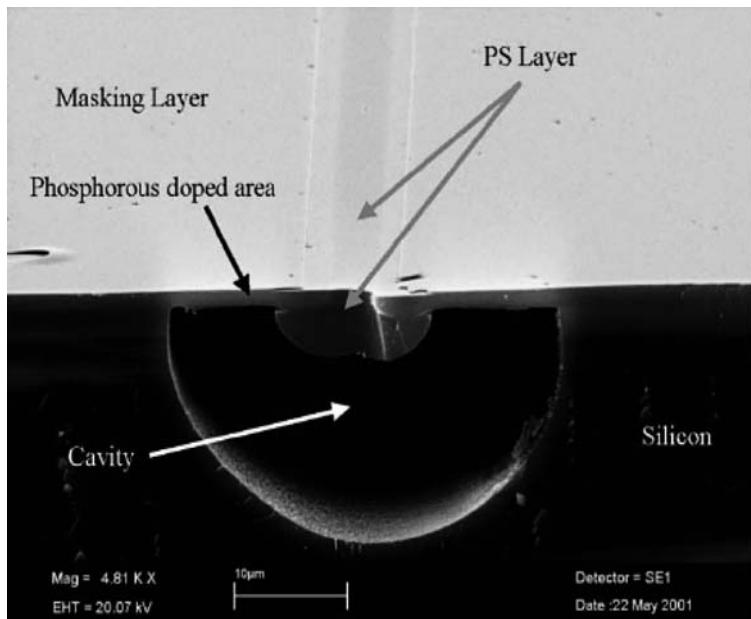


Fig. 14 Planar buried microchannels in Si substrate (Kaltsas et al. [96])

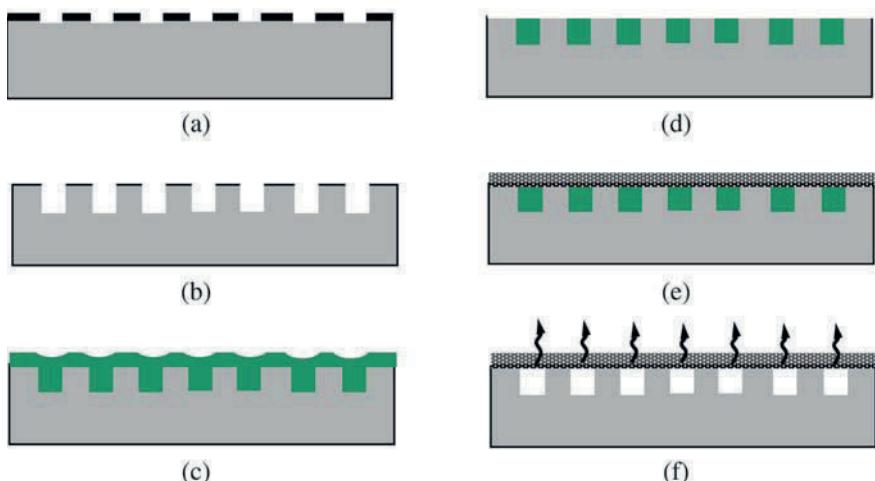
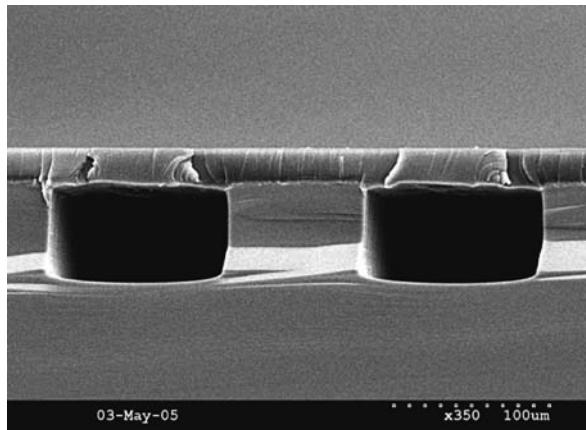


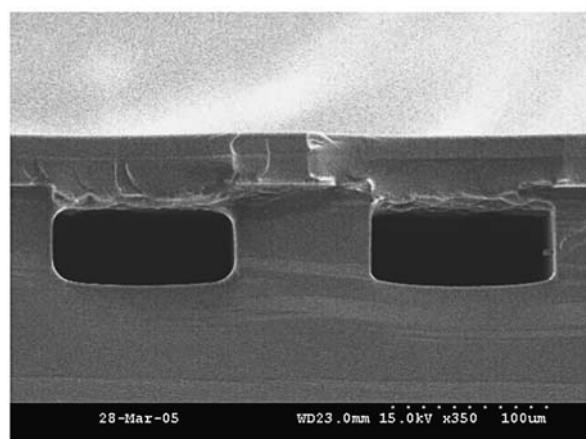
Fig. 15 Schematic illustration of low-temperature monolithic process of microchannel fabrication in Si wafer through a sacrificial polymer and an overcoat. (a) Trench patterning, (b) trench etching and surface cleaning, (c) trench refilling by spin coat, (d) polishing and descum, (e) overcoating with a layered structure (a porous SiO_2 layer and a polymer layer), and (f) sacrificial polymer decomposition and overcoat curing [101]

strength and hermeticity of the microchannels, an additional overcoat layer can be easily applied, as shown in Fig. 16 (a, b). In addition to the dielectric polymer, many other materials may also be used as water-proof coating, such as metal layers deposited by DC sputtering and electro/electroless plating, or spin-on glass.

Fig. 16 Cross-sectional SEM photographs of completed microfluidic channels enclosed by a layered overcoat after decomposition and curing. (a) Single overcoat layer and (b) double overcoat layer [101]



(a)



(b)

High-pressure fluidic testing confirmed good adhesion, hermeticity and high mechanical strength of the polymer overcoat. The polymer overcoat of only 30 μm in thickness can withstand a DI water pressure of 2.5 atm (35 psi) during the continuous testing. No delamination or cracking was observed. In addition, the adhesion between the overcoat polymer and the Si substrate did not degrade and no cracking or delamination occurred during thermal cycling test based on JEDEC Standard (Condition J of JESD 22-A104-B, 0–100°C, 3 cycles per hour).

In addition to microchannel fabrication, fluidic interconnection and assembly raise another challenge in on-chip microfluidic cooling system. Conventional microchannel interconnection method is illustrated in Fig. 17 [99, 100]. Inlet and outlet of the microchannel heat sink are located at the back side of a chip and are connected to the external fluidic circulation system.

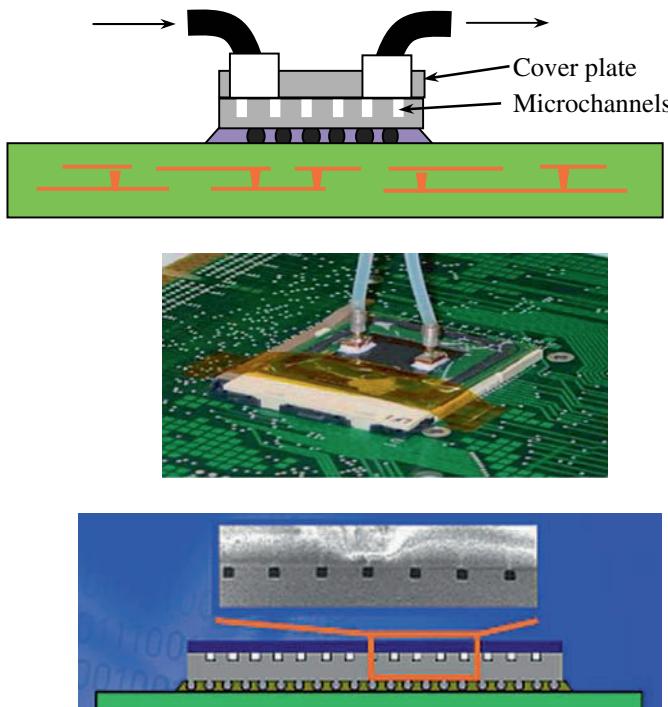


Fig. 17 Schematic of typical microchannel heat sinks in literature (the microchannel heat sink is integrated as part of a Si chip) [99, 100]

Dang et al. [98] proposed a compact microelectronic system with board-level fluidic delivery for on-chip microchannel heat sink using thermal-fluidic I/O interconnection. The basic concept of the thermal-fluidic I/O assembly is illustrated in Fig. 18 [101]. Micropipes are fabricated and integrated with C4 (Controlled Collapse Chip Connection) solder bumps to provide simultaneous electrical and thermal-fluidic I/Os, as shown in Fig. 19.

As shown in Fig. 20, with a precision flip-chip bonder, a compression force is applied after the chip and the package substrate are aligned. Since the polymer pipes are usually taller than the solder bumps, a portion of the polymer pipes will be inserted into the orifices, when the solder balls come into contact with the metal pad on the substrate. Once reflowed, metallurgical joining will be formed by the solder bumps and a chip is then mechanically bonded with the board substrate.

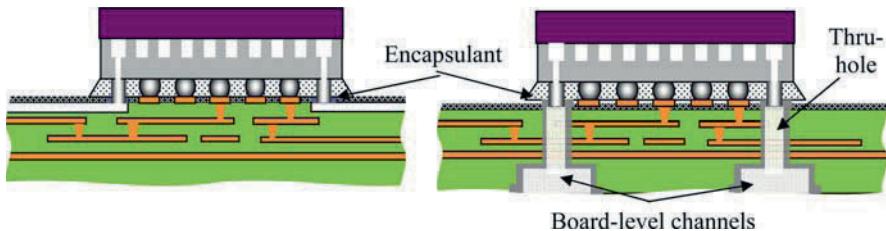


Fig. 18 Integrated thermal-fluidic I/O interconnection configurations [101]

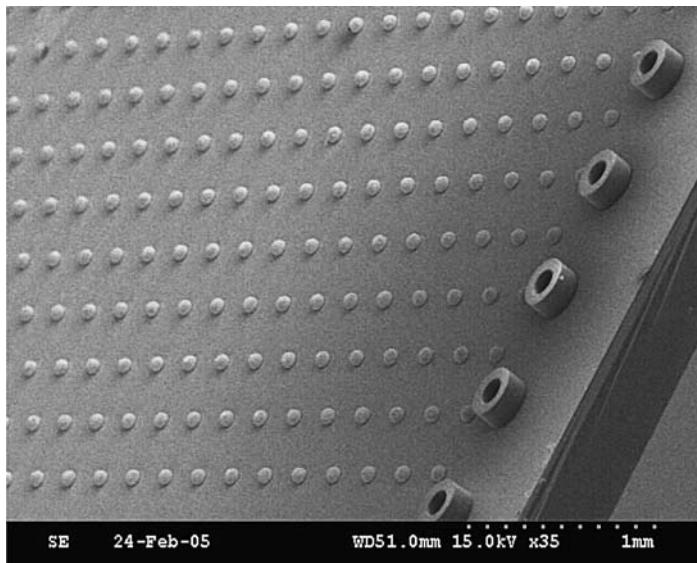


Fig. 19 SEM photographs of a row of micro-polymer pipes as the fluidic interconnects and the area-array solder bumps for as high-density electrical interconnects [98]

The micropipes are designed to have the same diameter as that of the orifices. To ensure hermetic sealing of the micropipes, a void-free polymer should be applied at the chip-edge after assembly, which is the same as the conventional underfilling process.

Figure 21 shows a “microfluidic flip chip” assembled onto a substrate with fluidic path for DI water circulation. For the purpose of simplicity, the microchannels on the test thermal chip were enclosed by glue bonding of a Si cover plate. The height of the microchannels is approximately 200 μm and their width is around 100 μm . After underfilling and curing, both fluidic continuity and electrical continuity were achieved successfully.

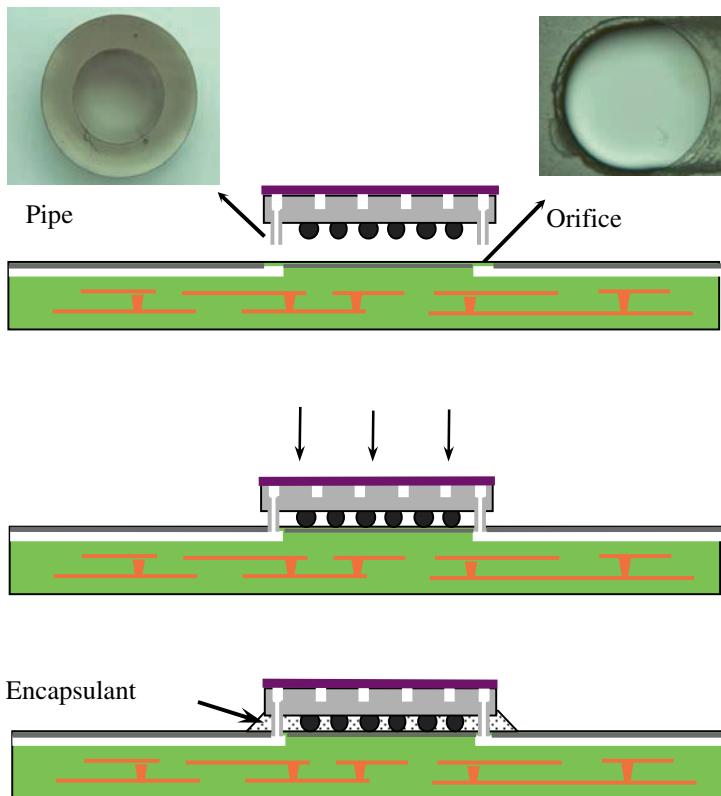


Fig. 20 Schematic of the assembly process of a “microfluidic flip chip” (a) alignment, (b) bonding through compression and heating, and (c) encapsulation with underfill [101]

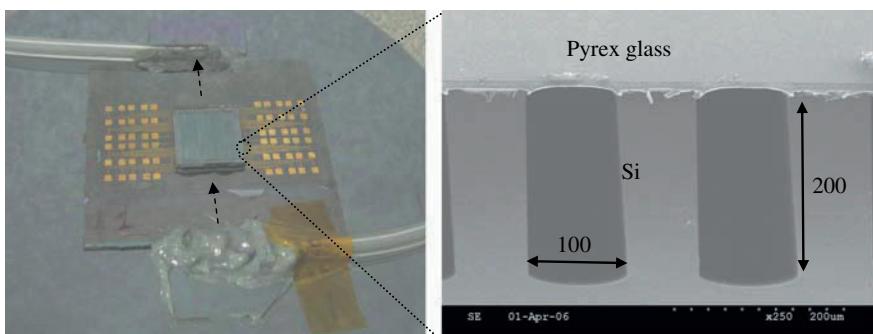


Fig. 21 A test thermal chip bonded on a substrate with through-vias connecting with fluidic tubes (51 microchannels distributed across a chip area of 1 cm^2) [101]

5 Application in Electronics Thermal Management

Electronics are finding newer applications in the areas of nano- and biotechnologies, which lead to high-power densities. Microchannel cooling is very attractive owing to its simplicity in implementation and facilitate for both single-phase and two-phase cooling with high heat flux removal capabilities (usually, 100–1000 W/cm²). With microchannels (and their progressive variations) proven over the last two decades to be suitable for high heat flux removal with narrow range of operating temperatures, they can be used to provide highly effective and competitive cooling solutions and as high-precision temperature control systems for nano- and bioelectronic devices in future.

Different variations of microchannels were successfully implemented in recent years for cooling of electronics. Stacked microchannels [17] and manifold microchannels [102] are two major examples. While stacked microchannels were discussed before in this chapter, manifold microchannels (MMCs) are primarily designed to significantly reduce the pressure drop by having a normal flow. Pressure drop can be a major impediment to implementing the microchannel technology for electronics especially for sub-hundred micron channel sizes and tailored stacked and manifold microchannels offer an attractive solution under such scenarios. An array of microchannels having a fraction of length of the total heat sink constitutes an MMC heat sink and has been demonstrated successfully [103].

In addition to variations in microchannel designs, efforts have also been made to improve thermophysical properties of the working fluids in microchannels. Examples include nanoparticle loaded fluids [104], micro-encapsulated and nano-encapsulated solid–liquid-phase change material slurries [61] and emulsions, which ameliorate thermal conductivity and/or heat capacity of the working fluid for better heat penetration and absorption compared to pure fluids. Studies found an optimum loading of thermophysical property boosting media, inlet temperature, and fluid and loaded particle type are all important to achieve the best thermal performance. The advantage of high latent heat associated with liquid–vapor phase change of the working fluids has also been studied in microchannels and has been reviewed under Section 1 of this chapter.

Recent efforts [105–107] studied the effectiveness of passive heat-transfer enhancement surface features like etched dimples/concavities to improve heat transfer between the working fluid and the microchannel walls for cooling microelectronics. They found that such surface features help in increasing the heat transfer by even up to ~11% in the laminar flow regime, which is the case with most microchannel flows. Figure 22 shows a bottom wall dimpled microchannel geometry [105] as an example.

While most of the prior efforts by various research groups focused on reducing the fluid or solid–fluid interface thermal resistance, the largest thermal resistance in most typical microchannel cooling implementation for electronics such as microprocessors occurs at the junction between the heat sink and the electronic component, which is usually filled with thermal conductive grease. Hence, integrated microchannel cooling has gained prominence in recent years, wherein the microchannels are

Fig. 22 3D representation of bottom-dimpled wall microchannels (Reprinted with permission from Wei et al. [105])



formed directly in one or more of the components that constitute the package. This technology is particularly useful when the electronic structures are stacked to achieve small form factors and high-power density.

Sekar et al. [108] presented a technology that involves 3D devices with integrated microchannel cooling. They propose the idea of fabricating through-hole silicon vias (TSVs) on a wafer using a two-step lithography process involving chip fabrication and etching to form the fluidic TSVs and microchannels. The device is then spin coated and polished with a sacrificial polymer material, which is spun-on, patterned, and cured to form a cover for the TSVs and microchannels. Cooling fluid was suggested to be delivered to the 3D stack either using tubes on the back side of the 3D stack or using fluidic channels on the substrate. Using this technology, their study showed a junction-to-ambient thermal resistance of $0.24^{\circ}\text{C}/\text{W}$ for a two-chip 3D stack.

Thermal management is a crucial facilitating technology in the development of modern electronics. As the power trends in electronics increase, the problem of thermal control becomes critical. This has led to the development and implementation of advanced thermal management solutions that are reliable, simple, and effective. Of them, microchannel cooling is established as one of the forefront technologies that can enable many of the Moore's law advances in micro, nano, and bioelectronics in future. With recent developments in microfabrication methods and materials, microchannels with advanced technologies such as using tailored single-phase and two-phase fluids, active/passive heat-transfer-enhancement structures, and substrate integration could be realized and will play a prominent role in thermal management of future high-power density nano- and bioelectronics.

References

1. Tuckerman D.B., R.F.W. Pease. High-performance heat sinking for VLSI. *IEEE Electron Device Letters* 1981; 2:126–129.
2. Wu P., W.A. Little. Measurement of friction factors for the flow of gases in very fine channels used for microminiature Joule–Thompson refrigerators. *Cryogenics* 1983; 23:273–277.
3. Pfahler J., J. Harley, H.H Bau, J. Zemel. *Gas and Liquid Flow in Small Channels. Micromechanical Sensors, Actuators and Systems*. ASME: New York, 1991.

4. Peng X.F., G.P. Peterson, B.X. Wang. Frictional flow characteristics of water flowing through rectangular microchannels. *Experimental Heat Transfer* 1994; 7:249–264.
5. Peng X.F., G.P. Peterson, B.X. Wang. Heat transfer characteristics of water flowing through rectangular microchannels. *Experimental Heat Transfer* 1994; 7:265–283.
6. Papautsky I., J. Brazile, T. Ammel, A.B. Frazier. Laminar fluid behavior in microchannels using micropolar fluid theory. *Sensors and Actuators* 1999; 73:101–108.
7. Harms T.M., M.J. Kazmierczak F.M. Gerner. Developing convective heat transfer in deep rectangular microchannels. *International Journal of Heat and Fluid Flow* 1999; 210: 149–157.
8. Mala G.M., D. Li. Flow characteristics of water in microtubes. *International Journal of Heat and Fluid Flow* 1999; 20:142–148.
9. Mala G.M., D. Li, C. Werner, H.J. Jacobasch, Y.B. Ning. Flow characteristics of water through microchannels between two parallel plates with electrokinetic effects. *International Journal of Heat and Fluid Flow* 1997; 18:489–496.
10. Tso C. P., S.P Mahulikar. The use of the Brinkman number for single phase forced convective heat transfer in microchannels. *International Journal of Heat and Mass Transfer* 1998; 41:1759–1769.
11. Xu B., K.T. Ooi, N.T. Wong. Experimental investigation of flow friction for liquid flow in microchannels. *International Communications in Heat and Mass Transfer* 2000; 27: 1165–1176.
12. Qu W., I. Mudawar. Flow boiling heat transfer in two-phase micro-channel heat sinks – II. Annular two-phase flow model. *International Journal of Heat and Mass Transfer* 2002; 46:2773–2784.
13. Liu D., S.V. Garimella. Investigation of liquid flow in microchannels. Eighth AIAA/ASME Joint Thermophysics and Heat Transfer Conference 2002.
14. Lee P., S.V. Garimella. Experimental investigation of heat transfer in microchannels. ASME Summer Heat Transfer Conference 2003.
15. Kohl M.J., S.I. Abdel-Khalik, S.M. Jeter, D.L. Sadowski. An experimental investigation of microchannel flow with internal pressure measurements. *International Journal of Heat and Mass Transfer* 2005; 48:15518–1533.
16. Wei X., Y. Joshi. Experimental and numerical study of sidewall profile effects on flow and heat transfer inside microchannels. *International Journal of Heat and Mass Transfer* 2007; 50:4640–4651.
17. Wei X., Y. Joshi. Optimization study of stacked micro-channel heat sinks for micro-electronic cooling. *IEEE Transactions on Components and Packaging Technologies* 2003; 26:55–61.
18. Phillips R.J. *Micro-channel Heat Sinks: Advances in Thermal Modeling of Electronic Components and Systems*. ASME: New York, 1990.
19. Muzychka Y.S., M.M. Yovanovich. Modeling friction factors in non-circular ducts for developing laminar flow. 2nd AIAA Theoretical Fluid Mechanics Meeting 1998.
20. Muzychka Y.S., M.M. Yovanovich. Modeling Nusselt numbers for thermally developing laminar flow in non-circular ducts. Seventh AIAA/ASME Joint Thermophysics and Heat Transfer Conference 1998.
21. Revellin R., J.R Thome. A new type of diabatic flow pattern map for boiling heat transfer in microchannels. *Journal of Micromechanics and Microengineering* 2007; 17:788–796.
22. Wojtan L., R. Revellin, J.R. Thome. Investigation of critical heat flux in single, uniformly heated microchannels. *Experimental Thermal and Fluid Science* 2007; 30:765–774.
23. Chen L., Y.S. Tian, T.G. Karayiannis. The effect of tube diameter on vertical two-phase flow regimes in small tubes. *International Journal of Heat and Mass Transfer* 2006; 49: 4220–4230.
24. Kew P., K. Cornwall. Correlations for the prediction of boiling heat transfer in small diameter channels. *Applied Thermal Engineering* 1997; 17: 705–715.

25. Kawaji M., M.Y. Chung. Unique characteristics of adiabatic gas–liquid flows in microchannels: Diameter and shape effects on flow patterns, void fraction and pressure drop. First ASME International Conference on Microchannels and Minichannels 2003.
26. Jacobi A., J.R. Thome. Heat transfer model for evaporation of elongated bubble flows in microchannels. *Journal of Heat Transfer* 2002; 124:1131–1136.
27. Li J., B. Wang. Size effect on two-phase regime for condensation in micro/mini tubes, *Heat Transfer - Asian Research* 2003; 32:65–71.
28. McAdams W.H., W.K. Woods, L.C. Heroman. Vaporization inside horizontal tubes II-benzene-oil mixtures. *Transactions of ASME* 1942; 64:193–200.
29. Cicchitti A., C. Lombaradi, M. Silversti, G. Soldaini, R. Zavattarli. Two-phase cooling experiments – pressure drop heat transfer burnout measurements. *Energia Nucleare* 1960; 7:407–425.
30. Dukler A.E., W. Moye, R.G. Cleveland. Frictional pressure drop in two-phase flow. Part A: a comparison of existing correlations for pressure loss and holdup, and part B: an approach through similarity analysis. *AIChE Journal* 1964; 10:38–51.
31. Awad M.M., Y.S. Muzychka. Effective property models for homogeneous two-phase flows. *Experimental Thermal and Fluid Science* 2008; 33:106–113.
32. Agostini B., J.R. Thome. Comparison of an extended database of flow boiling heat transfer coefficient in multi-microchannel elements with the three-zone model. ECI International Conference on Heat Transfer and Fluid Flow in Microscale 2005.
33. Kandlikar S.G., P. Balasubramanian. Extending the applicability of the flow boiling correlation at low Reynolds number flows in microchannels. First ASME International Conference on Microchannels and Minichannels 2003.
34. Peters J.V.S., S.G. Kandlikar. Further evaluation of a flow boiling correlation for microchannels and minichannels. Proceedings of the Fifth International Conference on Nanochannels, Microchannels and Minichannels 2007.
35. Qu W., I. Mudawar. Measurement and correlation of critical heat flux in two-phase microchannel heat sinks. *International Journal of Heat and Mass Transfer* 2004; 47: 2045–2059.
36. Pribyl D.J., A. Bar-Cohen, A.E. Bergles. An investigation of critical heat flux and two-phase flow regimes for upward steam and water flow. Proceedings of the Fifth International Conference on Boiling Heat Transfer 2003.
37. Taitel T., A.E. Dukler. A model for flow regime transition in horizontal and near horizontal gas–liquid flows. *AIChE Journal* 1975; 19:47–55.
38. Agostini B., J.R. Thome, M. Fabbri, B. Michel, D. Calmi, U. Kloster. High heat flux flow boiling in silicon multi-microchannels – Part II: Heat transfer characteristics of refrigerant R245fa. *International Journal of Heat and Mass Transfer* 2008; 51:5415–5425.
39. Revellin R., J.R. Thome. A theoretical model for the prediction of the critical heat flux in heated microchannels. *International Journal of Heat and Mass Transfer* 2008; 51: 1216–1225.
40. Revellin R., J.M. Quiben, J. Bonjour, J.R. Thome. Effect of local hot spots on the maximum dissipation rates during flow boiling in a microchannel. *IEEE Transactions on Components and Packaging Technologies* 2008; 31:407–416.
41. Breber G., J.W. Palen, J. Taborek. Prediction of horizontal tubeside condensation of pure components using flow regime criteria. *Journal of Heat Transfer* 1980; 102:471–476.
42. Soliman H.M. Correlation of mist-to-annular transition during condensation. *Canadian Journal of Chemical Engineering* 1983; 61:178–182.
43. Wang H.S., J.W. Rose. A theory of film condensation in horizontal noncircular section microchannels. *Journal of Heat Transfer* 2005; 127:1096–1105.
44. Coleman J.W., S. Garimella. Two-phase flow regimes in round, square and rectangular tubes during condensation of refrigerant R134a. *International Journal of Refrigeration* 2003; 26:117–128.

45. Yang C.-Y., C.-C. Shieh. Flow pattern of air-water and two-phase R-134a in small circular tubes. *International Journal of Multiphase Flow* 2001; 27:1163–1177.
46. Chen Y.P., P. Cheng. Condensation of steam in silicon microchannels. *International Communications in Heat and Mass Transfer* 2005; 32:175–183.
47. Bandhauer T.M., A. Agarwal, S. Garimella. Condensation pressure drop in circular microchannels. *Heat Transfer Engineering* 2005; 26:28–35.
48. Nema G. Flow Regime Transitions during Condensation in Microchannels. The George W. Woodruff School of Mechanical Engineering. Georgia Institute of Technology: Atlanta, 2008.
49. Wang H.S., J.W. Rose. Film condensation in horizontal microchannels: Effect of channel shape. *International Journal of Thermal Sciences* 2006; 45:1205–1212.
50. Garimella S., J.D. Killion, J.W. Coleman. An experimentally validated model for two-phase pressure drop in the intermittent flow regime for circular microchannels. *Journal of Fluids Engineering* 2002; 124:205–214.
51. Baird J. R., D.F. Fletcher, B.S. Haynes. Local condensation heat transfer rates in fine passages. *International Journal of Heat and Mass Transfer* 2003; 46:4453–4466.
52. Moser K.W., R.L. Webb, B. Na. A new equivalent Reynolds number model for condensation in smooth tubes. *Journal of Heat Transfer* 1998; 120:410–417.
53. Bandhauer T.M., A. Agarwal, S. Garimella. Measurement and modeling of condensation heat transfer coefficients in circular microchannels. *Journal of Heat Transfer* 2006; 128:1050–1059.
54. Garimella S., A. Agarwal J.D. Killion. Condensation pressure drop in circular microchannels. *Heat Transfer Engineering* 2005; 26:28–35.
55. Agarwal A. Heat Transfer and Pressure Drop during Condensation of Refrigerants in Microchannels. The George W. Woodruff School of Mechanical Engineering. Georgia Institute of Technology: Atlanta, 2006.
56. Maruyama T., T. Maeuchi. Centrifugal-force driven flow in cylindrical micro-channel. *Chemical Engineering Science* 2008; 63:153–156.
57. Ducre J., S. Haeberle, T. Brenner, T. Glatzel, R. Zengerle. Patterning of flow and mixing in rotating radial microchannels. *Microfluid Nanofluid* 2006; 2:97–105.
58. Hao Y.L., Y.-X. Tao. A numerical model for phase-change suspension flow in microchannels. *Numerical Heat Transfer A* 2004; 46:44–77.
59. Xing K.Q., Y.X. Tao, Y.L. Hao. Performance evaluation of liquid flow with PCM particles in microchannels. *Journal of Heat Transfer* 2005; 127: 931–940.
60. Sabbah R., M.M. Farid, S. Al-Hallaj. Micro-channel heat sink with slurry of water with micro-encapsulated phase change material: 3D-numerical study. *Applied Thermal Engineering* 2009; 29:445–454.
61. Kuravi S., K. Kota, J. Du, L. Chow. Numerical investigation of flow and heat transfer performance of nano-encapsulated phase change material slurry in microchannels. *Journal of Heat Transfer* 2009; 131:062901 (1–10).
62. Qu W., I. Mudawar. Experimental and numerical study of pressure drop and heat transfer in a single-phase micro-channel heat sink. *International Journal of Heat and Mass Transfer* 2002; 45:2549–2565.
63. Sarangi R.K., A. Bhattacharya, R.S. Prasher. Numerical modeling of boiling heat transfer in microchannels. *Applied Thermal Engineering* 2009; 29:300–309.
64. Ladd A.J.C. Numerical simulation of particulate suspensions via a discretized Boltzmann equation, part I. theoretical foundations. *Journal of Fluid Mechanics* 1994; 271: 285–309.
65. Ladd A.J.C. Numerical simulation of particulate suspensions via a discretized Boltzmann equation, part II. Numerical results. *Journal of Fluid Mechanics* 1994; 271:311–339.
66. Chang Q. Lattice Boltzmann Method for Thermal Multiphase Fluid Dynamics. Mechanical and Aerospace Engineering. Case Western Reserve University: Cleveland, 2005.

67. Judy J., D. Maynes, B.W. Webb. Characterization of frictional pressure drop for liquid flows through microchannels. *International Journal of Heat and Mass Transfer* 2002; 45: 3477–3489.
68. Harms T.M., M.J. Kazmierczak, F.M. Gerner. Developing convective heat transfer in deep rectangular microchannels. *International Journal of Heat and Fluid Flow* 1999; 20: 149–157.
69. Wu H.Y., P. Cheng. Friction factors in smooth trapezoidal silicon microchannels with different aspect ratios. *International Journal of Heat and Mass Transfer* 2003; 46:2519–2525.
70. Adrian R.J. Particle imaging techniques for experimental fluid mechanics. *Annual Review of Fluid Mechanics* 1991; 23: 261–304.
71. Santiago J. G., S.T. Wereley, C.D. Meinhart, D.J. Beebe, R.J. Adrian. A particle image velocimetry system for microfluidics. *Experiments in Fluids* 1998; 25:316–319.
72. Meinhart C.D., S.T. Wereley, J.G. Santiago. PIV measurements of a microchannel flow. *Experiments in Fluids* 1999; 27:414–419.
73. Meinhart C.D., S.T. Wereley, M.H.B. Gray. Volume illumination for two-dimensional particle image velocimetry. *Measurements Science and Technology* 2000; 11:809–814.
74. Olsen M.G., R.J. Adrian. Brownian motion and correlation in particle image velocimetry. *Optics Laser Technology* 2000; 32:621–627.
75. Olsen M.G., R.J. Adrian. Out-of-focus effects on particle image visibility and correlation in microscopic particle image velocimetry. *Experiments in Fluids* 2000; 29:166–174.
76. Olsen M.G., C.J. Bourdon. Out-of-plane motion effects in microscopic particle image velocimetry. *Journal of Fluids Engineering* 2003; 125:895–901.
77. Liu D., S.V. Garimella, S.T. Wereley. Infrared micro-particle image velocimetry of fluid flow in silicon-based microdevices. *ASME Heat Transfer/Fluids Engineering Summer Conference* 2004.
78. Park J.S., C.K. Choi, K.D. Kihm. Optically sliced micro-PIV using confocal laser scanning microscopy (CLSM). *Experiments in Fluids* 2004; 37:105–119.
79. Miernik M.M., L.R. Saetran. Improvement of micro-PIV resolution by selective seeding. *Proceedings of 4th International Conference on Nanochannels, Microchannels and Minichannels* 2006.
80. Tsuei L., O. Savas. Treatment of interfaces in particle image velocimetry. *Experiments in Fluids* 2000; 29:203–214.
81. Meinhart C.D., S.T. Wereley, J.G. Santiago. A PIV algorithm for estimating time-averaged velocity fields. *Journal of Fluids Engineering* 2000; 122:285–289.
82. Shah R.K., A.L. London. *Laminar Flow Forced Convection in Ducts*. Academic Press: New York, 1978.
83. Hohreiter V., S.T. Wereley, M.G. Olsen, J.N. Chung. Cross-correlation analysis for temperature measurement. *Measurements Science and Technology* 2003; 13:1072–1078.
84. Sakakibara J., K. Hishida M. Maeda. Measurements of thermally stratified pipe flow using image processing techniques. *Experiments in Fluids* 1993; 16:82–96.
85. Kim H.J., K.D. Kihm, J.S. Allen. Examination of ratiometric laser induced fluorescence thermometry for microscale spatial measurement resolution. *International Journal of Heat and Mass Transfer* 2003; 46:3967–3974.
86. Sakakibara J., R.J. Adrian. Measurement of temperature field of a Rayleigh–Benard convection using two-color laser induced fluorescence. *Experiments in Fluids* 2004; 37: 331–340.
87. Guasto J.S., P. Huang, K.S. Breuer. Statistical particle tracking velocimetry using molecular and quantum dot tracer particles. *Experiments in Fluids* 2006; 41:869–880.
88. Guasto J.S., K.S. Breuer. Simultaneous, ensemble-average measurement of near-wall temperature and velocity in steady state micro-flows using single quantum dot tracking. *Experiments in Fluids* 2008; 45:157–166.
89. Gad-el-Hak M. *MEMS Handbook*. CRC/Taylor & Francis: Florida, 2006.

90. de Boer M. J., R. W. Tjerkstra, J.W. Berenschot, H.V. Jansen, G.J. Burger, J.G.E Gardeniers, M. Elwenspoek, A. van den Berg. Micromachining of buried micro channels in silicon. *Journal of Microelectromechanical Systems* 2000; 9:94–103.
91. Jayachandran J. P., H.A. Reed, H. Zhen, L.F. Rhodes, C.L. Henderson, S.A.B. Allen, P.A. Kohl. Air-channel fabrication for microelectromechanical systems via sacrificial photosensitive polycarbonates. *Journal of Microelectromechanical Systems* 2003; 12:147–159.
92. Papautsky I., J. Brazzle, H. Swerdlow, A.B. Frazier. A low-temperature IC-compatible process for fabricating surface-micromachined metallic microchannels. *Journal of Microelectromechanical Systems* 1998; 7:267–273.
93. Pijnenburg R. H. W., R. Dekker, C.C.S. Nicole, A. Aubry, E.H.E.C. Eummelen. Integrated micro-channel cooling in silicon. *Proceedings of the 34th European Solid-State Device Research Conference* 2004.
94. Oprins H., C.C.S. Nicole, J.C. Baret, G. Van der Veken, C. Lasance, M. Baelmans. On-chip liquid cooling with integrated pump technology. *Proceedings of 21st Annual IEEE Semiconductor Thermal Measurement and Management Symposium* 2005.
95. Chang J.Y., R.S. Prasher, D. Chau, A. Myers, J. Dirner, S. Prstic, D. He. Convective performance of package based single phase microchannel heat exchanger. *Proceedings of ASME InterPACK 2005*.
96. Kaltsas G., D.N. Pagonis, A.G. Nassiopoulou. Planar CMOS compatible process for the fabrication of buried microchannels in silicon, using porous-silicon technology. *Journal of Microelectromechanical Systems* 2003; 12:863–872.
97. Dang B., P.J. Joseph, M.S. Bakir, T. Spencer, P. Kohl, J. Meindl. Wafer-level microfluidic cooling interconnects for GSI. *Proceedings of IEEE International Interconnect Technology Conference* 2005.
98. Dang B., M. Bakir, J. Meindl. Integrated thermal-fluidic I/O interconnects for an on-chip microchannel heat sink. *IEEE Electron Device Letters* 2006; 27:117–119.
99. Kandlikar S.G., H.R. Upadhye. Extending the heat flux limit with enhanced microchannels in direct single-phase cooling of computer chips. *Proceedings of 21st Annual IEEE Semiconductor Thermal Measurement and Management Symposium* 2005.
100. Pal A., Y.K. Joshi, M.H. Beitelmal, C.D. Patel, T.M. Wenger. Design and performance evaluation of a compact thermosyphon. *IEEE Transaction on Components and Packaging Technologies* 2002; 25:601–607.
101. Dang B. Integrated Input/Output Interconnection and Packaging for GSI, Ph.D. Thesis, The School of Electrical and Computer Engineering, Georgia Institute of Technology: Atlanta, 2006.
102. Copeland D., M. Behnia, W. Nakayama. Manifold microchannel heat sinks: Isothermal analysis. *IEEE Transactions on Components, Packaging and Manufacturing Technology, Part A* 1997; 20:96–102.
103. Kim Y., W. Chun, J. Kim, B. Pak, B. Baek. Forced air cooling by using manifold microchannel heat sinks. *Journal of Mechanical Science and Technology* 1998; 12:709–718.
104. Xu D., L. Pan. Numerical study of nanofluid flow and heat transfer in microchannels. *International Journal of Nanoscience* 2006; 5:747–752.
105. Wei X.J., Y.K. Joshi, P.M. Ligrani. Numerical simulation of laminar flow and heat transfer inside a microchannel with one dimpled surface. *Journal of Electronic Packaging* 2007; 129:63–70.
106. Silva C., E. Marotta, L. Fletcher. Flow structure and enhanced heat transfer in channel flow with dimpled surfaces: Application to heat sinks in microelectronic cooling. *Journal of Electronics Packaging* 2007; 129:157–166.
107. Silva C., D. Park, E. Marotta, L. Fletcher. Optimization of fin performance in a laminar channel flow through dimpled surfaces. *Journal of Heat Transfer* 2009; 131:021702(1–9).
108. Sekar D., C. King, B. Dang, T. Spencer, H. Thacker, P. Joseph, M. Bakir, J. Meindl. A 3D-IC technology with integrated microchannel cooling. *Proceedings of IEEE International Interconnect Technology Conference* 2008.

Nanoprobes for Live-Cell Gene Detection

Gang Bao, Won Jong Rhee, and Andrew Tsourkas

Abstract The ability to visualize in real time the expression dynamics and localization of specific RNAs *in vivo* offers tremendous opportunities for biological and disease studies including cancer detection. However, quantitative methods such as real-time polymerase chain reaction (PCR) and DNA microarrays rely on the use of cell lysates thus not able to obtain important spatial and temporal information. Fluorescence proteins and other reporter systems cannot image endogenous RNA in living cells. Fluorescence *in situ* hybridization (FISH) assays require washing to achieve specificity, therefore can only be used with fixed cells. Here we review the recent development of nanostructured probes for living cell RNA detection and discuss the biological and engineering issues and challenges of quantifying gene expression *in vivo*. In particular, we describe methods that use oligonucleotide probes, combined with novel delivery strategies, to image the relative level, localization, and dynamics of RNA in live cells. Examples of detecting endogenous messenger RNAs, as well as imaging their subcellular localization are given to illustrate the biological applications, and issues in probe design, delivery, and target accessibility are discussed. The nanostructured probes promise to open new and exciting opportunities in sensitive gene detection for a wide range of biological and medical applications.

Keywords Hairpin probe · Oligonucleotide probe · RNA detection · Live cell · Molecular beacon · Fluorescence resonance energy transfer

1 Introduction

The ability to image specific ribonucleic acid (RNA) in living cells in real time can provide essential information on RNA synthesis, processing, transport, and

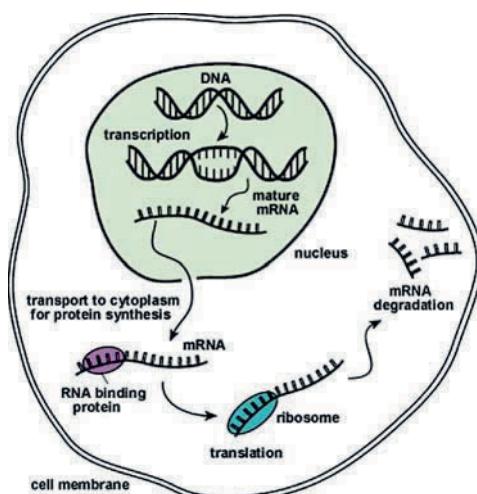
G. Bao (✉)

Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA, USA

e-mail: gang.bao@bme.gatech.edu

localization, and on the dynamics of RNA expression and localization in response to external stimuli; it will offer unprecedented opportunities for advancement in molecular biology, disease pathophysiology, drug discovery, and medical diagnostics. Over the last decade or so, there is increasing evidence to suggest that RNA molecules have a wide range of functions in living cells, from physically conveying and interpreting genetic information, to essential catalytic roles, to providing structural support for molecular machines, and to gene silencing. These functions are realized through control of the expression level and stability, both temporally and spatially, of specific RNAs in a cell. Therefore, determining the dynamics and localization of RNA molecules in living cells will significantly impact on the molecular biology and medicine.

Of particular interest is the fluorescence imaging of specific messenger RNAs (mRNAs), both their expression level and subcellular localization, in living cells. As shown schematically in Box 1, for eukaryotic cells a pre-mRNA molecule is synthesized in cell nucleus. After processing (including splicing and polyadenylation) the mature mRNAs are transported from cell nucleus to cytoplasm, and often localized at specific sites. The mRNAs are then being translated by ribosome to make specific proteins, and degraded by RNases after a certain amount of time. The limited lifetime of mRNA enables a cell to alter protein synthesis rapidly in response to its changing needs. During the entire life cycle of an mRNA, it is always complexed with RNA-binding proteins to form a ribonucleoprotein (RNP). This has significant implications to the live-cell imaging of mRNAs.



Messenger RNA (mRNA) that encodes the chemical “blueprint” for a protein is synthesized (transcribed) from a DNA template, and the pre-mRNA is processed (spliced) to have mature mRNA, which is transported to specific locations in the cell cytoplasm. The coding information carried by mRNA is being used by ribosomes to produce proteins (translation). After a certain amount of time, the message is degraded. mRNAs are almost always complexed with RNA-binding proteins to form ribonucleoprotein (RNP) molecules.

Box 1 mRNA life cycle

Many *in vitro* methods have been developed to provide a relative (mostly semi-quantitative) measure of gene expression level within a cell population using purified DNA or RNA obtained from cell lysate. These methods include polymerase chain reaction (PCR) [1], Northern hybridization (or Northern blotting) [2], expressed sequence tag (EST) [3], serial analysis of gene expression (SAGE) [4], differential display [5], and DNA microarrays [6]. These technologies, combined with the rapidly increasing availability of genomic data for numerous biological entities, present exciting possibilities for understanding human health and disease. For example, pathogenic and carcinogenic sequences are increasingly being used as clinical markers for diseased states. However, using *in vitro* methods to detect and identify foreign or mutated nucleic acids is often difficult in a clinical setting due to the low abundance of diseased cells in blood, sputum, and stool samples. Further, these methods cannot reveal the spatial and temporal variation of RNA within a single cell.

Labeled linear oligonucleotide (ODN) probes have been used to study intracellular mRNA via *in situ* hybridization (ISH) [7] in which cells are fixed and permeabilized to increase the probe delivery efficiency. Unbound probes are removed by washing to reduce background and achieve specificity [8]. To enhance the signal level, multiple probes targeting the same mRNA can be used [7]. However, fixation agents and other supporting chemicals can have considerable effect on signal level [9] and possibly on the integrity of certain organelles such as mitochondria. Thus, fixation of cells, by either cross-linking or denaturing agents, and the use of proteases in ISH assays may prevent from obtaining an accurate description of intracellular mRNA localization. It is also difficult to obtain a dynamic picture of gene expression in cells using ISH methods.

In order to detect RNA molecules in living cells with high specificity, sensitivity, and signal-to-background ratio, especially for low-abundance genes and clinical samples containing a small number of diseased cells, the probes need to recognize RNA targets with high specificity, convert target recognition *directly* into a measurable signal, and differentiate between true and false-positive signals. It is important for the probes to quantify low gene expression levels with high accuracy, and have fast kinetics in tracking alterations in gene expression in real time. For detecting genetic alterations such as mutations, insertions, and deletions, the ability to recognize single-nucleotide polymorphisms (SNPs) is essential. To achieve this optimal performance, it is necessary to have a good understanding of the structure–function relationship of the probes, probe stability, and RNA target accessibility in living cells. It is also necessary to achieve efficient cellular delivery of probes with minimal probe degradation.

In the remaining sections, we review commonly used fluorescent probes for RNA detection, discuss the critical issues in living cell RNA detection, including probe design, target accessibility, cellular delivery of probes, and detection sensitivity, specificity, and signal-to-background ratio. Emphasis is placed on the design and application of molecular beacons, although some of the issues are common to other oligonucleotide probes.

2 Fluorescent Probes for Live Cell RNA Detection

Several classes of molecular probes have been developed for RNA detection in living cells, including (1) tagged linear oligonucleotide (ODN) probes; (2) oligonucleotide hairpin probes; (3) probes using fluorescent proteins as reporter. Although probes composed of full-length RNAs (mRNA or nuclear RNA) tagged with a fluorescent or radioactive reporter have been used to study the intracellular localization of RNA [10–12], these probes are not discussed here since they cannot be used to measure the expression level of specific RNAs in living cells.

2.1 Tagged Linear ODN Probes

Single fluorescently labeled linear oligonucleotide probes have been developed for RNA tracking and localization studies in living cells [13–15]. In many cases, the ODN probe has a DNA backbone. A DNA molecule is a long polymer made from repeating nucleotides including adenine (A), quanine (G), cytosine (C), and thymine (T). For RNA, thymine is replaced by uracil (U). Typically G pairs with C, and A pairs with T (or U), according to Watson–Crick base pairing. Although these probes may recognize specific endogenous RNA transcripts in living cells via Watson–Crick base pairing and reveal subcellular RNA localization, this approach lacks the ability to distinguish background from true signal, since both bound probes (i.e., probes hybridized to RNA target) and unbound probes give fluorescence signal. It may also lack detection specificity since partial match between the probe and target sequences could induce probe hybridization to RNA molecules of multiple genes. A novel way to increase signal-to-noise ratio and improve detection specificity is to use two linear probes with a fluorescence resonance energy transfer (FRET) pair of (donor and acceptor) fluorophores [13]. However, the dual linear-probe approach may still have a high background signal due to direct excitation of the acceptor and emission detection of the donor fluorescence. Further, it is difficult for linear probes to distinguish targets that differ by a few bases since the difference in free energy of the two hybrids (with and without mismatch) is typically rather small. This limits the application of linear ODN probes in biological and disease studies.

2.2 ODN Hairpin Probes

Hairpin nucleic acid probes have the potential to be highly sensitive and specific in live-cell RNA detection. As shown in Fig. 1a, b one class of such probes is molecular beacons, which are dual-labeled oligonucleotide probes with a fluorophore at one end and a quencher at the other end [16]. They are designed to form a stem-loop hairpin structure in the absence of a complementary target so that fluorescence of the fluorophore is quenched by the quencher molecule (typically through FRET

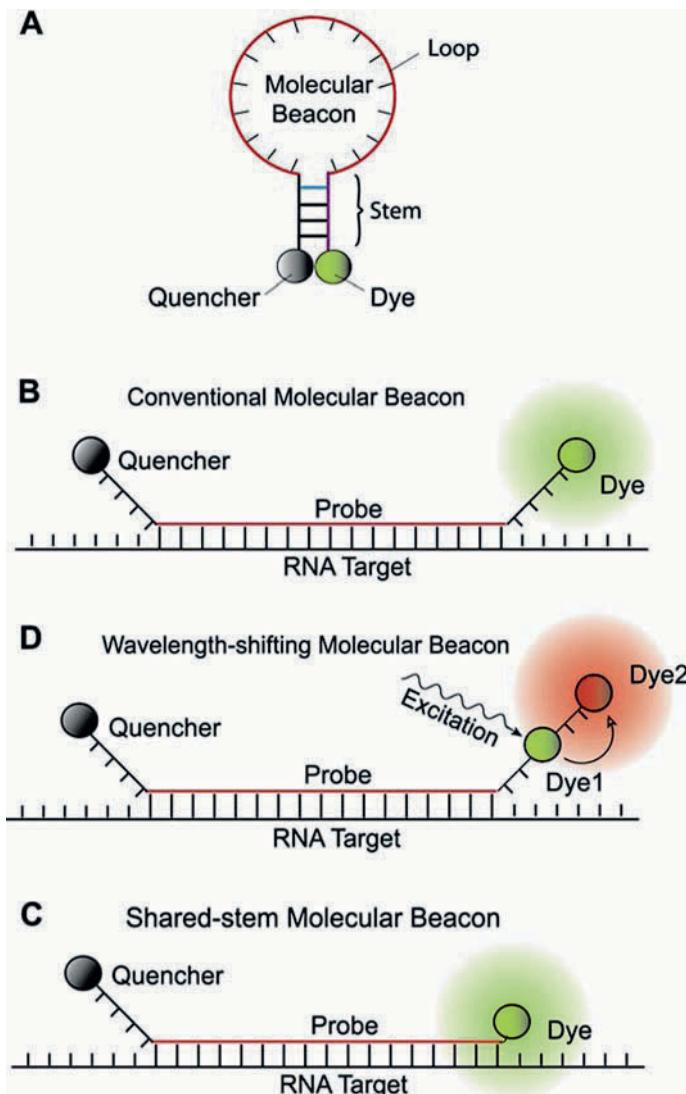


Fig. 1 Illustrations of molecular beacons. (a) Molecular beacons are stem-loop hairpin oligonucleotide probes labeled with a reporter fluorophore at one end and a quencher molecule at the other end. (b) Conventional molecular beacons are designed such that the short complementary arms of the stem are independent of the target sequence. (c) Wavelength-shifting molecular beacons containing two fluorophores, one absorbs in the wavelength range of the monochromatic light source, and the other emits light at the desired emission wavelength due to FRET. (d) Shared-stem molecular beacons are designed such that one arm of the stem participates in both stem formation and target hybridization

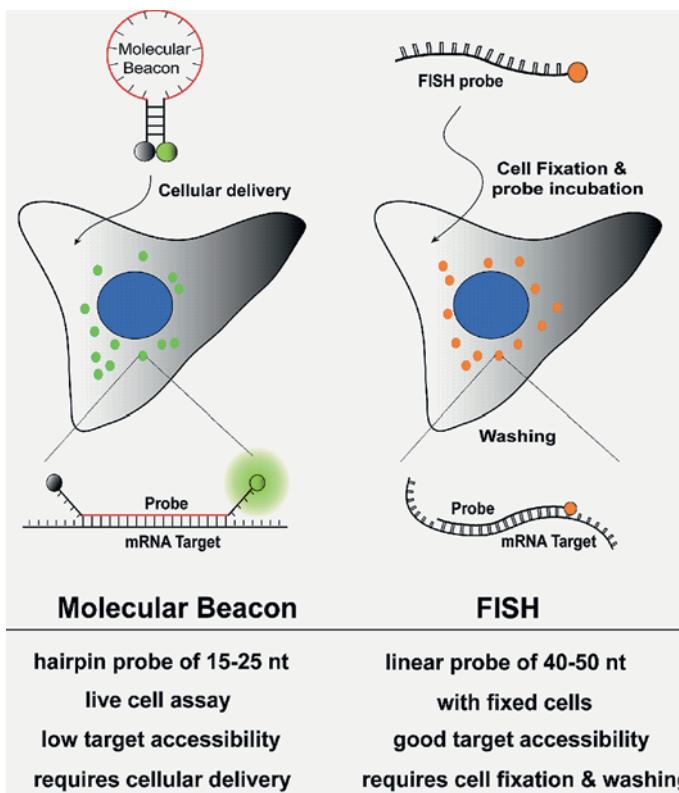
between fluorophore and quencher, and contact quenching). Hybridization with the target nucleic acid opens the hairpin and physically separates the fluorophore from quencher, allowing a fluorescence signal to be emitted upon excitation (Fig. 1b). Under optimal conditions, the fluorescence intensity of molecular beacons can increase by >200-fold upon binding to their targets [16]. This enables a molecular beacon to function as a sensitive probe with a high signal-to-background ratio. The stem-loop hairpin structure provides an adjustable energy penalty for hairpin opening which improves probe specificity [17, 18]. The ability to transduce target recognition *directly* into a fluorescence signal with high signal-to-background ratio, coupled with an improved specificity, has allowed molecular beacons to enjoy a wide range of biological and biomedical applications, including multiple analyte detection, real-time enzymatic cleavage assaying, cancer cell detection, real-time monitoring of PCR, genotyping and mutation detection, viral infection studies, and mRNA detection in living cells [14, 19–32].

As illustrated in Fig. 1a, a conventional molecular beacon has four essential components: loop, stem, fluorophore, and quencher. The loop, which is the sensing (targeting) domain of the probe, usually consists of 15–25 nucleotides and is selected to have a unique target sequence and proper melting temperature (defined as the temperature at which half of the beacons are hybridized to the target). The stem, which services as the ‘controller’ of the probe and formed by the base pairing of two complementary short-arm sequences, is to close the loop domain therefore having a hairpin structure. The stem is typically 4–6 bases long and chosen to be independent of the target sequence (Fig. 1a). The fluorophore is a reporter of the probe that emits a fluorescent signal when excited. The quencher, typically a small molecule such as Dabcyl, is to suppress the fluorescence of the fluorophore partially or completely through FRET between the fluorophore and quencher as well as contact quenching.

A novel design of hairpin probes is the wavelength-shifting molecular beacons that can fluoresce in a variety of different colors [33]. As shown in Fig. 1c, in this design, a molecular beacon contains two fluorophores: a first fluorophore that absorbs strongly in the wavelength range of the monochromatic light source, and a second fluorophore that emits at the desired emission wavelength due to fluorescence resonance energy transfer from the first fluorophore to the second fluorophore. It has been demonstrated that wavelength-shifting molecular beacons are substantially brighter than conventional molecular beacons that contain a fluorophore that cannot efficiently absorb energy from the available monochromatic light source.

One major advantage of the stem-loop hairpin probes is that they can recognize their targets with higher specificity than linear ODN probes. Solution studies suggested that [17, 18], using molecular beacons, it is possible to discriminate between targets that differ by a single nucleotide. In contrast to current techniques for detecting single nucleotide polymorphism (SNP), which are often labor-intensive and time-consuming, molecular beacons may provide a simple and promising tool for detecting SNPs in disease diagnosis.

Box 2 compares the basic features of molecular beacon versus fluorescence in-situ hybridization (FISH). Specifically, molecular beacons are dual-labeled hairpin probes of 15–25 nucleotides (nt), while FISH probes are dye-labeled linear



Box 2. Comparison of molecular beacon and FISH approaches

oligonucleotides of 40–50 nt. The molecular beacon-based approach has the advantage of detecting RNA in live cells without the need of cell fixation and washing. However, it requires cellular delivery of the probes and has low target accessibility (discussed in later sections). The advantage of FISH assays is the ease of probe design due to better target accessibility. Although FISH assays can be used to image the localization of mRNA in fixed cells, they rely on stringent washing to achieve signal specificity and do not have the ability to image the dynamics of gene expression in living cells.

In the conventional molecular beacon design, the stem sequence is typically independent of the target sequence (Fig. 1b), although sometimes two end bases of the probe sequence, each adjacent to one arm sequence of the stem, could be complementary with each other, thus forming part of the stem (light blue base of the stem shown in Fig. 1a). Molecular beacons can also be designed such that all the bases of one arm of the stem (to which a fluorophore is conjugated) are complementary to the target sequence, thus participating in both stem formation and target hybridization (shared-stem molecular beacons) [34] (Fig. 1d). The advantage of this shared-stem design is to help fix the position of the fluorophore that attached to the stem arm,

limiting its degree-of-freedom of motion, and increasing the fluorescence resonance energy transfer (FRET) in the dual FRET molecular beacon design, as discussed below.

A dual FRET molecular beacons approach was developed [26–28] to overcome the difficulty that, in live-cell RNA detection, molecular beacons are often degraded by nucleases or open due to nonspecific interaction with hairpin-binding proteins, causing a significant amount of false-positive signal. In this dual probe design, a pair of molecular beacons labeled with a donor and an acceptor fluorophore, respectively are employed (Fig. 2). The probe sequences are chosen such that this pair of molecular beacons hybridizes to adjacent regions on a single RNA target (Fig. 2). Since FRET is very sensitive to the distance between donor and acceptor fluorophores and typically occurs when the donor and acceptor fluorophores are within ~10 nm, FRET signal is generated by the donor and acceptor beacons only if both probes are bound to the same RNA target. Thus, the sensitized emission of the acceptor fluorophore upon donor excitation serves as a positive signal in the FRET-based detection assay, which is differentiable from non-FRET false-positive signals due to probe degradation and nonspecific probe opening. This approach combines the lowbackground signal and high specificity of molecular beacons with the ability of FRET assays in differentiating between true target recognition and false-positive signals, leading to an enhanced ability to quantify RNA expression in living cells [28].

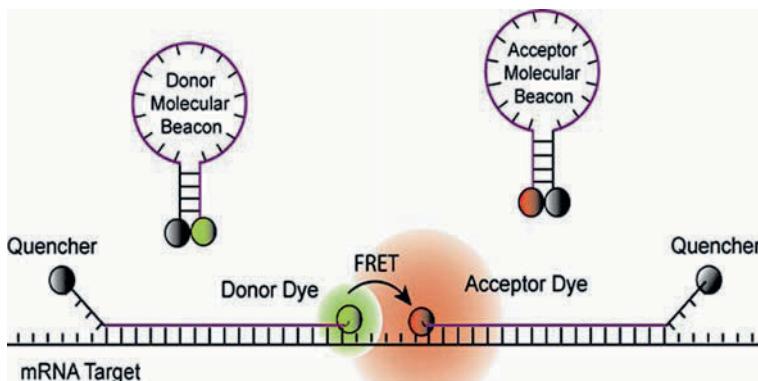


Fig. 2 A schematic showing the concept of dual FRET molecular beacons. Hybridization of donor and acceptor molecular beacons to adjacent regions on the same mRNA target results in FRET between donor and acceptor fluorophores upon donor excitation. By detecting FRET signal, fluorescence signals due to probe–target binding can be readily distinguished from that due to molecular beacon degradation and nonspecific interactions

2.3 Fluorescent Protein-Based Probes

In addition to oligonucleotide probes, tagged RNA-binding proteins such as those with green fluorescent protein (GFP) tags have been used to detect mRNA in live

cells [35]. One limitation is that it requires the identification of a unique protein, which only binds to the specific mRNA of interest. To address this issue, a coat protein of the RNA bacteriophage MS2 was tagged with GFP and a RNA sequence corresponding to several MS2-binding sites was introduced to the mRNA of interest, which allowed for the specific targeting of the nanos mRNA in live *Drosophila* eggs [36]. The GFP-MS2 approach has been used to track the localization and dynamics of RNA in living cells with single molecule sensitivity [37, 38]. However, since unbound GFP-tagged MS2 proteins also give fluorescence signal, the background signal in the GFP-MS2 approach could be high, leading to a low signal-to-background ratio in live-cell imaging of RNA.

An interesting fluorescent-protein-based approach that overcomes this problem is to utilize the fluorescent protein complementation [39, 40]. In this approach (split-GFP), a RNA-binding protein is dissected into two fragments, which are respectively fused to the split fragments of a fluorescent protein. Binding of the two tagged fragments of the RNA-binding protein to adjacent sites on the same mRNA molecule (or two parts of an aptamer sequence inserted to the mRNA sequence) brings the two halves of the fluorescent protein together, reconstituting the fluorescent protein and restoring fluorescence [40]. Alternatively, two RNA-binding proteins that bind specifically to adjacent sites on the same mRNA molecule can be tagged with the split fragments of a fluorescent protein, and their binding to the target mRNA results in the restoration of fluorescence [39]. The advantage of this novel approach is that background signal is low – no fluorescence signal unless the RNA-binding proteins (or protein fragments) are bound to the target mRNA. The split-GFP method, however, may have difficulties in tracking the dynamics of RNA expression in real time, since the reconstitution of the fluorescent protein from the split fragments typically takes 2–4 h, during which the RNA expression level may change. Transfection efficiency could also be a major concern in the GFP-based approaches in that usually only a few percent of the cells express the fluorescent proteins following transfection. This limits the application of the split-GFP methods in detecting diseased cells using mRNA as a biomarker for the disease.

3 Probe Design and Structure–Function Relations

3.1 Target Specificity

There are three major design issues of molecular beacons: probe sequence, hairpin structure, and fluorophore–quencher selection. In general, the probe sequence is selected to ensure specificity, and to have good target accessibility. The hairpin structure as well as the probe and stem sequences are determined to have the proper melting temperature and the fluorophore–quencher pair should give high signal-to-background ratio. To ensure specificity, for each gene to target, one can use the Basic Local Alignment Search Tool (BLAST) developed by the National Center for

Biotechnology Information (NCBI) [41] or similar software to select multiple target sequences of 15–25 bases that are unique for the target RNA. Since the melting temperature of a molecular beacon affects both the signal-to-background ratio and detection specificity, especially for mutation detection, it is often necessary to select the target sequence with a balanced G–C content (the percentage of G–C pairs), and to adjust the loop and stem lengths and the stem sequence of the molecular beacon to realize the optimal melting temperature. In particular, it is necessary to understand the effect of molecular beacon design on melting temperature so that, at 37°C, single-base mismatches in target mRNAs can be differentiated. This is also a general issue for detection specificity in that, for any specific probe sequence selected, there might be multiple genes in the mammalian genome that have sequences that differ from the probe sequence by only a few bases. Therefore, it is important to design the molecular beacons so that only the specific target RNA would give a strong signal.

Several approaches can be taken to validate the signal specificity. For example, one could upregulate or downregulate the expression level of a specific RNA, quantify the level using reverse transcriptase PCR (RT-PCR), and comparing the PCR result with that of molecular beacon-based imaging of the same RNA in living cells. However, complications may arise when the approach used to change the RNA expression level in living cells has an effect on multiple genes, leading to some ambiguity even when the PCR and beacon results match. Perhaps the best way to downregulate the level of a specific mRNA in live cells is to use small interfering RNA (siRNA) treatment, which typically leads to >80% reduction of the specific mRNA level. Since the effect of siRNA treatment varies depending on the specific siRNA probes used, the siRNA delivery method, and the cell type, optimization of the protocol (probe design and delivery method/conditions) is often needed.

3.2 Molecular Beacon Structure–Function Relations

The loop, stem lengths, and sequences are critical design parameters for molecular beacons, since at any given temperature they largely control the fraction of molecular beacons that are bound to the target [17, 18]. In many applications, the choices of the probe sequence are limited by target-specific considerations, such as the sequence surrounding a single nucleotide polymorphism (SNP) of interest. However, the probe and stem lengths, and stem sequence, can be adjusted to optimize the performance (i.e., specificity, hybridization rate, and signal-to-background ratio) of a molecular beacon for a specific application [17, 34].

To demonstrate the effect of molecular beacon structure on its melting behavior, the melting temperature for molecular beacons with various stem-loop structures is displayed in Fig. 3a. In general, it was found that the melting temperature increased with probe length but appeared to plateau at a length of ~20 nucleotides. It was also found that the stem length of the molecular beacon could strongly influence the melting temperature of molecular beacon–target duplexes. If the melting temperature (T_m) is too high, the beacon becomes very stable, resulting in very slow hybridization kinetics and loss of the ability to discriminate mismatches between

probe and target at 37°C. On the other hand, if the T_m is too low, molecular beacons are too each to open, thus increasing the background signal. Therefore, the melting temperature of molecular beacons is an important parameter that has a large influence on signal-to-background ratio, hybridization kinetics, and detection specificity.

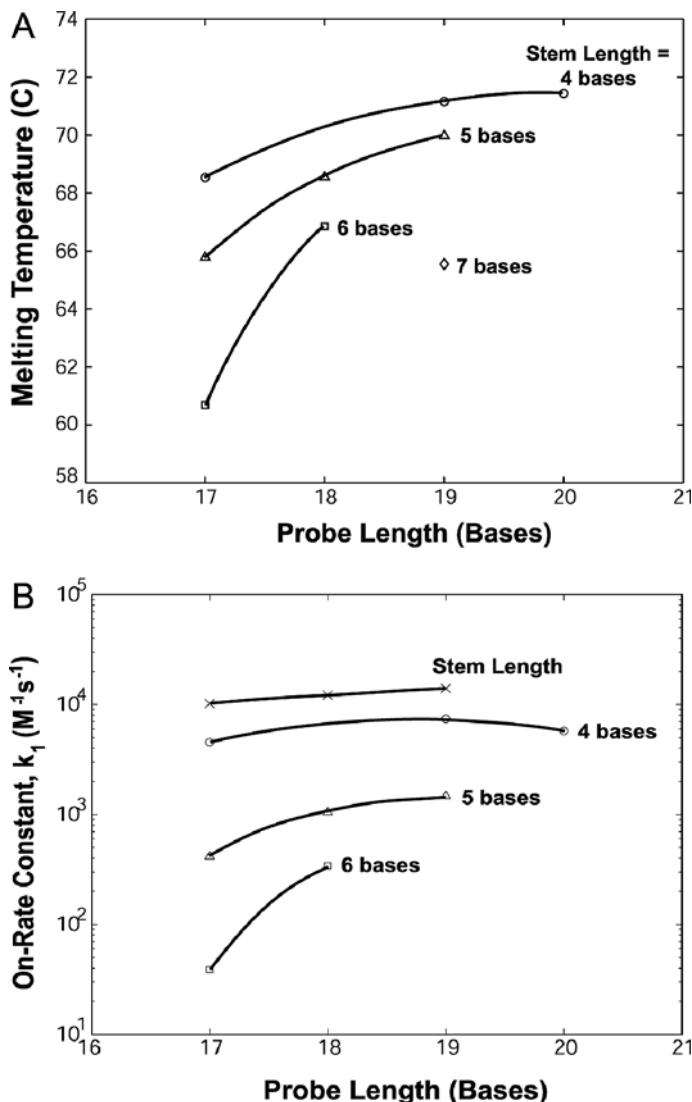


Fig. 3 Structure–function relations of molecular beacons. **(a)** Melting temperatures for molecular beacons with different structures in the presence of target. **(b)** The on-rate constant of hybridization k_1 for molecular beacons with various probe and stem lengths hybridized to their complementary targets

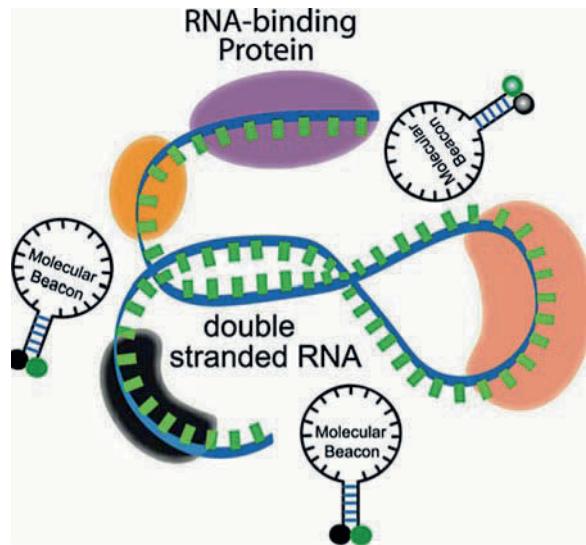
While both the stability of the hairpin probe and its ability to discriminate targets over a wider range of temperatures increase with increasing stem length, it is accompanied by a decrease in hybridization on-rate constant, as shown in Fig. 3b. For example, molecular beacons with a 4-base stem had an on-rate constant up to 100 times greater than molecular beacons with a 6-base stem. Changing the probe length of a molecular beacon may also influence the rate of hybridization, as demonstrated by Fig. 3b.

From the thermodynamic and kinetic studies, it was found that, if the stem length is too large, it will be difficult for the beacon to open upon hybridization. On the other hand, if the stem length is too small, a large fraction of beacons may open due to the thermal force. Similarly, relative to the stem length, a longer probe may lead to a lower dissociation constant; however, it may also reduce the specificity, since the relative free energy change due to one base mismatch would be smaller. A long probe length may also lead to coiled conformations of the beacons, resulting in reduced kinetic rates. Therefore, the stem and probe lengths need be carefully chosen in order to optimize both hybridization kinetics and MB specificity [17, 34]. In general, it has been found that molecular beacons with longer stem lengths have an improved ability to discriminate between wild-type and mutant targets in solution over a broader range of temperatures. This can be attributed to the enhanced stability of the molecular beacon stem-loop structure and the resulting smaller free energy difference between closed (unbound) molecular beacons and molecular beacon–target duplexes, which generates a condition where a single-base mismatch reduces the energetic preference of probe–target binding. Longer stem lengths, however, are accompanied by a decreased probe–target hybridization kinetic rate. Similarly, molecular beacons with short stems have faster hybridization kinetics but suffer from lower signal-to-background ratios compared with molecular beacons with longer stems.

3.3 Target Accessibility

A critical issue in molecular beacon design is target accessibility, as is the case for most oligonucleotide probes for live-cell RNA detection. It is well known that a functional mRNA molecule in a living cell always has RNA-binding proteins on it, forming a ribonucleoprotein (RNP). Further, an mRNA molecule often has double stranded portions and forms secondary (folded) structures (Fig. 4). Therefore, in designing a molecular beacon, it is necessary to avoid targeting mRNA sequences that are double-stranded, or occupied by RNA-binding proteins, for otherwise the probe has to compete off the RNA strand or the RNA-binding protein in order to hybridize to the target. Indeed, molecular beacons designed for targeting a specific mRNA often show no signal when delivered to living cells. One difficulty in the molecular beacon design is that, although predictions of mRNA secondary structure can be made using software such as *Beacon Designer* (www.premierbiosoft.com) and *mfold* (<http://www.bioinfo.rpi.edu/applications/mfold/old/dna/>), they may be inaccurate due to limitations of the biophysical models used, and the limited understanding of protein–RNA interaction. Therefore, for each gene to target, it may be necessary to select multiple unique sequences along the target RNA and

Fig. 4 A schematic illustration of a segment of the target mRNA with a double-stranded portion and RNA-binding proteins. A molecular beacon has to compete off an mRNA strand or RNA-binding protein(s) in order to hybridize to the target



have corresponding molecular beacons designed, synthesized, and tested in living cells to select the best target sequence.

To uncover the possible molecular beacon design rules, the accessibility of BMP-4 mRNA was studied using different beacon designs [42]. Specifically, molecular beacons were designed to target the start codon and termination codon regions, the siRNA and antisense oligonucleotide probe sites identified previously, and sites that were randomly chosen. All the target sequences are unique to BMP-4 mRNA. Of the eight molecular beacons designed to target BMP-4 mRNA, it was found that only two beacons gave strong signal, one targets the start codon region, and the other targets the termination codon region. It was also found that, even for a molecular beacon that works well, shifting its targeting sequence by just a few bases toward the 3' or 5' ends would significantly reduce the fluorescence signal from beacons in a live-cell assay, indicating that the target accessibility is quite sensitive to the location of the targeting sequence. These results, together with molecular beacons validated previously, suggest that the start codon and termination codon regions and the exon-exon junctions are more accessible than other locations in an mRNA.

3.4 Fluorophores and Quenchers

With proper backbone synthesis and fluorophore-quencher conjugation, in theory, a molecular beacon can be labeled with any desired reporter-quencher pair. However, proper selection of the reporter and quencher could improve the signal-to-background ratio and multiplexing capabilities. Selecting a fluorophore label for a molecular beacon as the reporter is usually not as critical as the hairpin probe design since many conventional dyes can yield satisfactory results. However, proper selection could yield additional benefits such as an improved signal-to-background ratio

and multiplexing capabilities. Since each molecular beacon utilizes only one fluorophore it is possible to use multiple molecular beacons in the same assay, assuming that the fluorophores are chosen with minimal emission overlap [19]. Molecular beacons can even be labeled simultaneously with two fluorophores, i.e., “wavelength shifting” reporter dyes (Fig. 1c), allowing multiple reporter dye sets to be excited by the same monochromatic light source yet fluorescing in a variety of colors [33]. Clearly, multicolor fluorescence detection of different beacon/target duplexes can become a powerful tool for the simultaneous detection of multiple genes.

For dual FRET (fluorescence resonance energy transfer) molecular beacons (Fig. 2), the donor fluorophores typically emit at shorter wavelengths compared with that of acceptor. Energy transfer occurs as a result of long-range dipole–dipole interactions between the donor and the acceptor. The efficiency of energy transfer depends upon the extent of the spectral overlap of the emission spectrum of the donor with the absorption spectrum of the acceptor, the quantum yield of the donor, the relative orientation of the donor and acceptor transition dipoles [43], and the distance between the donor and acceptor molecules (usually 4–5 bases). In selecting donor and acceptor fluorophores, in order to have high signal-to-background ratio, it is important to optimize the above parameters, and to avoid direct excitation of the acceptor fluorophore at the donor excitation wavelength, as well as minimizing donor emission detection at the acceptor emission detection wavelength. Examples of FRET dye pairs include Cy3 (donor) and Cy5 (acceptor), TMR (donor) and Texas Red (acceptor), fluorescein (FAM) (donor), and Cy3 (acceptor).

It is relatively straightforward to select the quencher molecules. Organic quencher molecules such as dabcyl, BHQ2 (blackhole quencher II) (Biosearch Tech), BHQ3 (Biosearch Tech), and Iowa Black (IDT) can all effectively quench a wide range of fluorophores by both fluorescence resonance energy transfer (FRET) and the formation of an exciton complex between the fluorophore and the quencher [44].

4 Cellular Delivery of Nanoprobes

One of the most critical aspects of measuring the intracellular level of RNA molecules using synthetic probes is the ability to deliver these probes into cells through the plasma membrane, which is quite lipophilic and restricts the transport of large and charged molecules. Therefore, it is a very robust barrier to polyanionic molecules such as hairpin oligonucleotides. Further, even if the probes enter the cells successfully, the efficiency of delivery in an imaging assay should be defined not only by how many probes enter the cell or how many cells have probes internalized, but also how many probes remain functioning inside cells. This is different from both antisense and gene delivery applications where the reduction in level of protein expression is the final metric used to define efficiency or success. For measuring RNA molecules (including mRNA and rRNA) in the cytoplasm, a large amount of probes should remain in the cytoplasm.

Existing cellular delivery techniques can be divided into two categories: endocytic and nonendocytic methods. Endocytic delivery typically employs cationic and polycationic molecules such as liposomes and dendrimers, while nonendocytic methods include microinjection and the use of cell-penetrating peptides (CPP) or streptolysin *O* (SLO). Probe delivery via the endocytic pathway typically takes 2–4 h. It has been reported that ODN probes internalized via endocytosis are predominately trapped inside endosomes and often lysosomes, and being degraded there due to cytoplasmic nucleases [45]. Consequently, only 0.01–10% of the probes remain functioning after escaping from endosomes and lysosomes [46].

Oligonucleotide probes (including molecular beacons) have been delivered into cells via microinjection [47]. In most of the cases, the ODNs exhibited a fast accumulation in the cell nucleus, preventing the probes from targeting mRNAs in the cell cytoplasm. Depletion of intracellular ATP or lowering the temperature from 37 to 4°C did not have a significant effect on ODN nuclear accumulation, ruling out active, motor-protein-driven transport [47]. It is unclear if the rapid transport of ODN probes to nucleus is due to electrostatic interaction, or driven by microinjection-induced flow, or the triggering of some signaling pathway. There is no fundamental biological reason why ODN probes should accumulate in the cell nucleus. To prevent nuclear accumulation, streptavidin (60 kDa) molecules were conjugated to linear ODN probes via biotin [13]. After microinjected into cells, dual FRET linear probes could hybridize to the same mRNA target in the cytoplasm, resulting in a FRET signal. More recently, it was demonstrated that when tRNA transcripts are attached to molecular beacons with 2'-*O*-methyl backbone and injected into the nucleus of HeLa cells, the probes are exported into the cytoplasm. When these constructs are introduced into the cytoplasm, they remain cytoplasmic [48]. However, even without the problem of unwanted nuclear accumulation, microinjection is inefficient in delivering probes into a large number of cells.

Another nonendocytic delivery method is toxin-based cell membrane permeabilization. For example, streptolysin *O* (SLO) is a pore-forming bacterial toxin that has been used as a simple and rapid means of introducing oligonucleotides into eukaryotic cells [49–51]. SLO binds as a monomer to cholesterol and oligomerizes into a ring-shaped structure to form pores of approximately 25–30 nm in diameter, allowing the influx of both ions and macromolecules. It was found that SLO-based permeabilization could achieve an intracellular concentration of ODNs of approximately 10 times that of electroporation and liposomal-based delivery. Since cholesterol composition varies with cell types, the permeabilization protocol needs to be optimized for each cell type by varying temperature, incubation time, cell number, and SLO concentration. An essential feature of this technique is that the toxin-based permeabilization is reversible. This can be achieved by introducing oligonucleotides with SLO under serum-free conditions and then removing the mixture and adding normal media with serum [50, 52].

Cell penetrating peptides (CPP) have been used to introduce proteins, nucleic acids, and other biomolecules into living cells [53–55]. Among the family of peptides with membrane translocating activity are antennapedia, HSV-1 VP22, and the

HIV-1 Tat peptide. To date the most widely used peptide are HIV-1 Tat peptide and its derivatives due to their small size and high delivery efficiency. The Tat peptide is rich in cationic amino acids especially arginines, which is very common in many of the cell penetrating peptides. However, the exact mechanism for CPP-induced membrane translocation remains elusive.

A wide variety of cargos have been delivered to living cells both in cell culture and in tissue using cell penetrating peptides [56, 57]. For example, Allinquant et al. [58] linked antennapedia peptide to the 5' end of DNA oligonucleotides (with biotin on the 3' end) and incubated both peptide-linked ODNs and ODNs alone with cells. By detecting biotin using streptavidin-alkaline phosphatase amplification, it was found that the peptide-linked ODNs were internalized very efficiently into all cell compartments compared with control ODNs. No indication of endocytosis was found. Similar results were obtained by Troy et al. [59] with a 100-fold increase in antisense delivery efficiency when ODNs were linked to antennapedia peptides. Recently, Tat peptides were conjugated to molecular beacons using different linkages (Fig. 5); the resulting peptide-linked molecular beacons were delivered into living cells to target GAPDH and survivin mRNAs [29]. It was demonstrated that, at relatively low concentrations, peptide-linked molecular beacons were internalized into living cells within 30 min with nearly 100% efficiency. Further, peptide-based delivery did not interfere with either specific targeting by or hybridization-induced fluorescence of the probes, and the peptide-linked molecular beacons could have self-delivery, targeting, and reporting functions. In contrast, liposome-based (Oligofectamine) or dendrimer-based (Superfect) delivery of molecular beacons required 3–4 h and resulted in a punctate fluorescence signal in the cytoplasmic vesicles and a high background in both cytoplasm and nucleus of cells [29]. It was clearly demonstrated that cellular delivery of molecular beacons using the peptide-based approach has far better performance compared with conventional transfection methods.

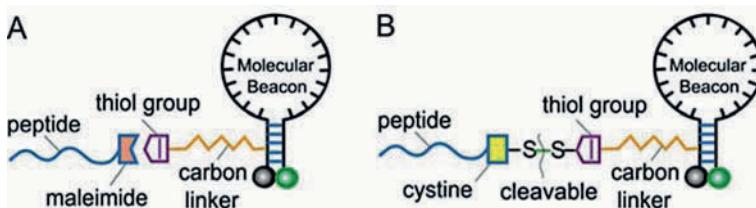


Fig. 5 A schematic of peptide-linked molecular beacons. (A) A peptide-linked molecular beacon using the thiol-maleimide linkage in which the quencher arm of the molecular beacon stem is modified by adding a thiol group, which can react with a maleimide group placed to the C terminus of the peptide to form a direct, stable linkage. (B) A peptide-linked molecular beacon with a cleavable disulfide bridge in which the peptide is modified by adding a cysteine residue at the C terminus, which forms a disulfide bridge with the thiol-modified molecular beacon. This disulfide bridge design allows the peptide to be cleaved from the molecular beacon by the reducing environment of the cytoplasm

5 Living Cell RNA Detection Using Molecular Beacons

Sensitive gene detection in living cells presents a significant challenge. In addition to issues with target accessibility, detection specificity, and probe delivery as discussed above, achieving high detection sensitivity and signal-to-background ratio requires not only careful design of the probes and advanced fluorescence microscopy imaging, but also a better understanding of RNA biology and probe–target interactions. It is likely that different applications have different requirements on the properties of probes. For example, rapid determination of RNA expression level and localization requires fast probe–target hybridization kinetics, while long-time monitoring of gene expression dynamics requires probes having high intracellular stability.

To demonstrate the capability of molecular beacons in sensitive detection of specific endogenous mRNAs in living cells, dual FRET molecular beacons were designed to detect K-ras and survivin mRNAs in HDF and MIAPaCa-2 cells, respectively [28]. K-ras is one of the most frequently mutated genes in human cancers [60]. A member of the *G*-protein family, K-ras is involved in transducing growth-promoting signals from the cell surface. Survivin, one of the inhibitor of apoptosis proteins (IAPs), is normally expressed during fetal development but not in most normal adult tissues [61], thus can be used as a tumor biomarker for several types of cancers. Each FRET probe pair consisted of two molecular beacons, one labeled with a donor fluorophore (Cy3, donor beacon) and a second labeled with an acceptor fluorophore (Cy5, acceptor beacon). These molecular beacons were designed to hybridize to adjacent regions on an mRNA target so that the two fluorophores lie within the FRET range (~6 nm) when probe–target hybridization occurs for both beacons. BHQ-2 and BHQ-3 were used as quenchers for the donor and acceptor molecular beacons, respectively. One pair of molecular beacons targets a segment of the wild-type K-ras gene whose codon 12 mutations are involved in the pathogenesis of many cancers. A negative control dual FRET molecular beacon pair was also designed (“random beacon pair”) whose specific 16-base target sequence was selected using random walk, thus having no exact match in the mammalian genome. It was found that detection of the FRET signal significantly reduced false positives, leading to sensitive imaging of K-ras and survivin mRNAs in live HDF and MIAPaCa-2 cells. For example, FRET detection gave a ratio of 2.25 of K-ras mRNA expression in stimulated versus unstimulated HDF, comparable to the ratio of 1.95 using RT-PCR, and in contrast to single-beacon result of 1.2. The detection of survivin mRNA also indicated that, compared with the single-beacon approach, dual FRET molecular beacons gave lower background signal, thus having a higher signal-to-background ratio [28].

5.1 Biological Significance

An intriguing discovery in detecting K-ras and survivin mRNAs using dual FRET molecular beacons is the clear and detailed mRNA localization in living cells [28]. To demonstrate, in Fig. 6a, a fluorescence image of K-ras mRNA in stimulated HDF

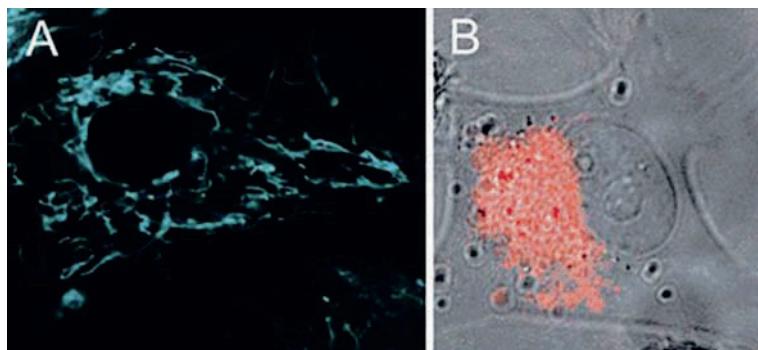


Fig. 6 mRNA localization in HDF and MIAPaCa-2 cells. **(a)** Fluorescence images of K-ras mRNA in stimulated HDF cells. Note the filamentous K-ras mRNA localization pattern. **(b)** A fluorescence image of survivin mRNA localization in MIAPaCa-2 cells. Note that survivin mRNAs often localized to one side of the nucleus of the MIAPaCa-2 cells

cells is shown, indicating an intriguing filamentous localization pattern. The localization pattern of K-ras mRNA was further studied and found to be colocalized with mitochondria inside live HDF cells [62]. Since K-ras proteins interact with proteins such as Bcl-2 in mitochondria to mediate both antiapoptotic and proapoptotic pathways, it seems that cells localize certain mRNAs where the corresponding proteins can easily bind to their partners.

The survivin mRNA, however, is localized in MIAPaCa-2 cell very differently. As shown in Fig. 6b in which the fluorescence image was superimposed with a white-light image of the cells, survivin mRNAs seemed to localize in a nonsymmetrical pattern within MIAPaCa-2 cells, often to one side of the nucleus of the cell. These mRNA localization patterns raise many interesting biological questions. For example, how mRNAs are transported to their destination and how the destination is recognized? To what subcellular organelle might the mRNAs be colocalized? What is the biological implication of mRNA localization? Although mRNA localization in living cells is believed to be closely related to post-transcriptional regulation of gene expression, much remains to be seen if such localization indeed targets a protein to its site of function by producing the protein “right on the spot.”

The transport and localization of oskar mRNA in *Drosophila melanogaster* oocytes has also been visualized [26]. In this work, molecular beacons with 2'-O-methyl backbone were delivered into cells using microinjection and the migration of oskar mRNAs were tracked in real time, from the nurse cells where it is produced to the posterior cortex of the oocyte where it is localized. Clearly, the direct visualization of specific mRNAs in living cells with molecular beacons will provide important insight into the intracellular trafficking and localization of RNA molecules.

As another example of targeting specific genes in living cells, molecular beacons were used to detect the viral genome and characterize the spreading of bovine respiratory syncytial virus (bRSV) in living cells [63]. It was found that molecular beacon

signal could be detected in single living cells infected by RSV with high detection sensitivity, and the signal revealed a connected, highly three-dimensional, amorphous inclusion body structure not seen in fixed cells. Figure 7 shows the molecular beacon signal indicating the spreading of viral infection at days 1, 3, 5, and 7 post-infection (PI), which demonstrates the ability of molecular beacons to monitor and quantify in real time the viral infection process. Molecular beacons were further used to image the viral genomic RNA (vRNA) of human RSV (hRSV) in live Vero cells, revealing the dynamics of filamentous virion egress, and providing insight into how viral filaments bud from the plasma membrane of the host cell [64].

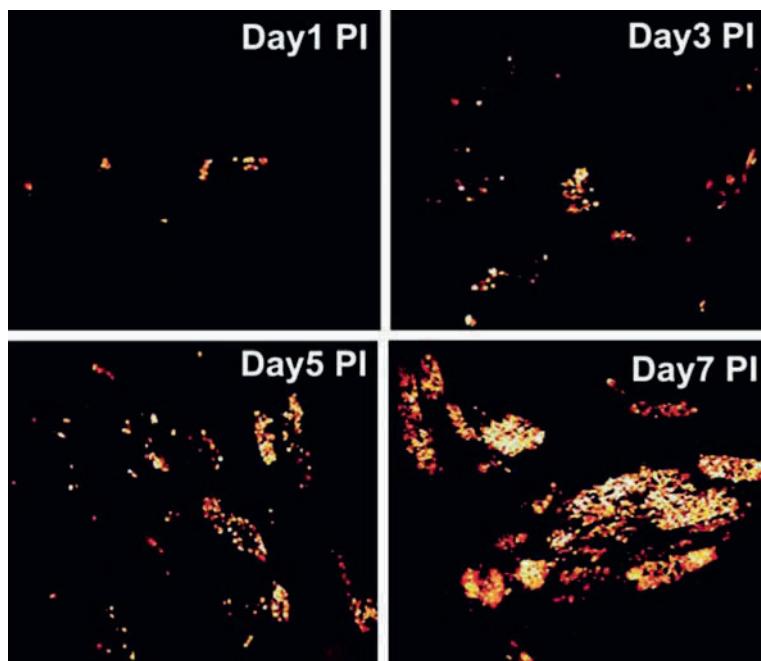


Fig. 7 Live-cell fluorescence imaging of the genome of bovine respiratory syncytial virus (bRSV) using molecular beacons shows the spreading of infection in host cells at days 1, 3, 5, and 7 post-infection (PI). Primary bovine turbinate cells were infected by a clinical isolate of bRSV, CA-1, with a viral titer of $2 \times 10^{3.6}$ TCID₅₀/ml. Molecular beacons were designed to target several repeated sequences of the gene-end-intergenic-gene-start signal within the bRSV genome, with a signal-to-noise ratio of 50–200

6 Engineering Challenges

Nanostructured molecular probes such as molecular beacons have the potential to enjoy a wide range of applications that require sensitive detection of genomic sequences. For example, molecular beacons are used as a tool for the

detection of single-stranded nucleic acids in homogeneous in vitro assays [65, 66]. Surface-immobilized molecular beacons used in microarray assays allow for the high-throughput parallel detection of nucleic acid targets while avoiding the difficulties associated with PCR-based labeling [65, 67]. Another novel application of molecular beacons is the detection of double-stranded DNA targets using PNA “openers” that form triplexes with the DNA strands [68]. Further, proteins can be detected by synthesizing “aptamer molecular beacon” [69, 70] which, upon binding to a protein, undergoes a conformational change that results in the restoration of fluorescence.

The most exciting application of nanostructured oligonucleotide probes, however, is living cell gene detection. As demonstrated, molecular beacons can detect endogenous mRNA in living cells with high specificity, sensitivity, and signal-to-background ratio, thus having the potential to provide a powerful tool for laboratory and clinical studies of gene expression *in vivo*. For example, molecular beacons can be used in high-throughput cell-based assays to quantify and monitor the dose-dependent changes of specific mRNA expression in response to different drug leads. The ability of molecular beacons to detect and quantify the expression of specific genes in living cells will also facilitate disease studies, such as viral infection detection and cancer diagnosis.

There are a number of challenges in detecting and quantifying RNA expression in living cells. In addition to issues of probe design and target accessibility, quantifying gene expression in living cells in terms of mRNA copy number per cell poses a significant challenge. For instance, it is necessary to distinguish true and background signals, determine the fraction of mRNA molecules hybridized with probes, and quantify the possible self-quenching effect of the reporter, especially when mRNA is highly localized. Since the fluorescence intensity of the reporter may be altered by the intracellular environment, it is also necessary to create an internal control by, for example, injecting fluorescently labeled oligonucleotides with known quantity into the same cells and obtaining the corresponding fluorescence intensity. Further, unlike in RT-PCR studies where the mRNA expression is averaged over a large number of cells (usually over 1 million), in optical imaging of mRNA expression in living cells, only a relatively small number of cells (typically less than 1000) are observed. Therefore, the average copy number per cell may change with the total number of cells observed due to the (often large) cell-to-cell variation of mRNA expression.

Another issue in living cell gene detection using hairpin ODN probes is the possible effect of probes on normal cell function, including protein expression. As revealed in the antisense therapy research, complementary pairing of a short segment of an exogenous oligonucleotide to mRNA can have a profound impact on protein expression levels and even cell fate. For example, tight binding of the probe to the translation start site may block mRNA translation. Binding of a DNA probe to mRNA can also trigger RNase H-mediated mRNA degradation. However, the probability of eliciting antisense effects with hairpin probes may be very low when low concentrations of probes (<200 nM) are used for mRNA detection, in contrast to

the high concentrations (typically 20 μM ; [51]) employed in antisense experiments. Further, it generally takes 4 h before any noticeable antisense effect occurs, whereas visualization of mRNA with hairpin probes requires less than 2 h after delivery. However, it is important to carry out a systematic study of the possible antisense effects, especially for molecular beacons with 2'-*O*-methyl backbone, which may also trigger unwanted RNA interference.

As a new approach for *in vivo* gene detection, the nanostructured probes can be further developed to have enhanced sensitivity and a wider range of applications. For example, it is likely that hairpin ODN probes with quantum dot as the fluorophore will have a better ability to track the transport of individual mRNAs from the cell nucleus to the cytoplasm. Hairpin ODN probes with near-infrared (NIR) dye as the reporter, combined with peptide-based delivery have the potential to detect specific RNAs in tissue samples, animals, or even humans. It is also possible to use lanthanide chelate as the donor in a dual FRET probe assay and perform time-resolved measurements to dramatically increase the signal-to-noise ratio, thus achieving high sensitivity in detecting low-abundance genes. Although very challenging, the development of these and other nanostructured ODN probes will significantly enhance our ability to image, track, and quantify gene expression *in vivo* and provide a powerful tool for basic and clinical studies of human health and disease.

There are many possibilities for nanostructured ODN probes to become clinical tools for disease detection and diagnosis. For example, molecular beacon could be used to perform cell-based early cancer detection using clinical samples including blood, saliva, and other bodily fluid. In this case, cells in the clinical sample are separated and molecular beacons designed to target specific cancer genes are delivered to the cytoplasm for detecting mRNAs of the cancer biomarker genes. Cancer cells having a high level of the target mRNAs (such as survivin) or mRNAs with specific mutations that cause cancer (such as K-ras codon 12 mutations) would show high level of fluorescence signal, while normal cells would show just low background signal. In this approach, the target mRNAs would not be diluted compared with the approaches using cell lysate. Thus, molecular beacons-based assay has the potential to positively identify cancer cells in a clinical sample with high specificity and sensitivity. It may also be possible to detect cancer cells *in vivo* by using NIR-dye-labeled molecular beacons in combination with endoscopy. Although there remain significant challenges, imaging methods using nanostructured probes have a great potential in becoming a powerful clinical tool for disease detection and diagnosis.

Acknowledgments This work was supported by the National Heart Lung and Blood Institute of the NIH as a Program of Excellence in Nanotechnology (HL80711), by the National Cancer Institute of the NIH as a Center of Cancer Nanotechnology Excellence (CA119338), and by the NIH Roadmap Initiative in Nanomedicine through a Nanomedicine Development Center award (PN2EY018244).

Table 1 Descriptions of key biological and technical terminologies

Deoxyribonucleic acid (DNA)	A DNA molecule is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms and some viruses. It is a long polymer made from repeating nucleotides including adenine (A), quanine (G), cytosine (C), and thymine (T). Typically G pairs with C and A pairs with T, according to Watson–Crick base pairing
Ribonucleic acid (RNA)	RNA is a biologically important type of molecule that consists of a long chain of nucleotide units. Each nucleotide contains a ribose sugar, with carbons numbered 1' through 5'. A base is attached to the 1' position, generally adenine (A), cytosine (C), quanine (G), or uracil (U). Typically G pairs with C and A pairs with U
Oligonucleotide (ODN)	An oligonucleotide is a short nucleic acid polymer, typically with 20 or fewer bases. Because oligonucleotides readily bind to their respective complementary nucleotides, they are often used as probes for detecting DNA or RNA
Polymerase chain reaction (PCR)	PCR is a technique to amplify a single or few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence. Almost all PCR applications employ a heat-stable DNA polymerase which enzymatically assembles a new DNA strand from DNA building blocks, the nucleotides, by using single-stranded DNA as a template and DNA oligonucleotides (DNA primers) which are required for initiation of DNA synthesis
Fluorescence in situ hybridization (FISH)	FISH is a cytogenetic technique used to detect and localize the presence or absence of specific DNA sequences on chromosomes. FISH uses fluorescent probes that bind to only those parts of the DNA that show a high degree of sequence similarity. FISH is often used for finding specific features in DNA for use in genetic counseling, medicine, and species identification. FISH can also be used to detect and localize specific mRNAs within tissue samples
Ribonucleoprotein (RNP)	RNP is a nucleoprotein that contains RNA and protein, i.e., it is an association that combines ribonucleic acid and protein together.
Fluorescence resonance energy transfer (FRET)	FRET is the non-radiative transfer of energy from one molecule to another based on the proximity of the two molecules and their spectral overlap. It is a widely applied tool for optical sensing and studies of biomolecular interaction
Single-nucleotide polymorphism (SNP)	A single-nucleotide polymorphism (SNP) is a DNA sequence variation occurring when a single nucleotide—A, T, C or G,—in the genome (or other shared sequence) differs between members of a species (or between paired chromosomes in an individual)
Fluorophore	A fluorophore, in analogy to a chromophore, is a component of a molecule which causes a molecule to be fluorescent. It is typically a functional group in a molecule which will absorb energy of a specific wavelength and re-emit energy at a different (but equally specific) wavelength
Quencher	A quencher is a substance that absorbs excitation energy from a fluorophore; a dark quencher dissipates the energy as heat thus eliminates a fluorophore's ability to fluoresce
Codon	Codon is a sequence of three adjacent nucleotides constituting the genetic code that specifies the insertion of an amino acid in a specific structural position in a polypeptide chain during the synthesis of proteins

References

1. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350–1354.
2. Alwine, J.C., Kemp, D.J., Parker, B.A., Reiser, J., Renart, J., Stark, G.R. and Wahl, G.M. (1979) Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper. *Methods Enzymol*, **68**, 220–242.
3. Adams, M.D., Dubnick, M., Kerlavage, A.R., Moreno, R., Kelley, J.M., Utterback, T.R., Nagle, J.W., Fields, C. and Venter, J.C. (1992) Sequence identification of 2,375 human brain genes. *Nature*, **355**, 632–634.
4. Velculescu, V.E., Zhang, L., Vogelstein, B. and Kinzler, K.W. (1995) Serial analysis of gene expression. *Science*, **270**, 484–487.
5. Liang, P. and Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, **257**, 967–971.
6. Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467–470.
7. Bassell, G.J., Powers, C.M., Taneja, K.L. and Singer, R.H. (1994) Single mRNAs visualized by ultrastructural *in situ* hybridization are principally localized at actin filament intersections in fibroblasts. *J Cell Biol*, **126**, 863–876.
8. Buongiorno-Nardelli, M. and Amaldi, F. (1970) Autoradiographic detection of molecular hybrids between RNA and DNA in tissue sections. *Nature*, **225**, 946–948.
9. Behrens, S., Fuchs, B.M., Mueller, F. and Amann, R. (2003) Is the *in situ* accessibility of the 16S rRNA of *Escherichia coli* for Cy3-labeled oligonucleotide probes predicted by a three-dimensional structure model of the 30S ribosomal subunit? *Appl Environ Microbiol*, **69**, 4935–4941.
10. Huang, Q. and Pederson, T. (1999) A human U2 RNA mutant stalled in 3' end processing is impaired in nuclear import. *Nucleic Acids Res*, **27**, 1025–1031.
11. Glotzer, J.B., Saffrich, R., Glotzer, M. and Ephrussi, A. (1997) Cytoplasmic flows localize injected oskar RNA in Drosophila oocytes. *Curr Biol*, **7**, 326–337.
12. Jacobson, M.R. and Pederson, T. (1998) Localization of signal recognition particle RNA in the nucleolus of mammalian cells. *Proc Natl Acad Sci USA*, **95**, 7981–7986.
13. Tsuji, A., Koshimoto, H., Sato, Y., Hirano, M., Sei-Iida, Y., Kondo, S. and Ishibashi, K. (2000) Direct observation of specific messenger RNA in a single living cell under a fluorescence microscope. *Biophys J*, **78**, 3260–3274.
14. Dirks, R.W., Molenaar, C. and Tanke, H.J. (2001) Methods for visualizing RNA processing and transport pathways in living cells. *Histochem Cell Biol*, **115**, 3–11.
15. Molenaar, C., Abdulle, A., Gena, A., Tanke, H.J. and Dirks, R.W. (2004) Poly(A)+ RNAs roam the cell nucleus and pass through speckle domains in transcriptionally active and inactive cells. *J Cell Biol*, **165**, 191–202.
16. Tyagi, S. and Kramer, F.R. (1996) Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol*, **14**, 303–308.
17. Tsourkas, A., Behlke, M.A., Rose, S.D. and Bao, G. (2003) Hybridization kinetics and thermodynamics of molecular beacons. *Nucleic Acids Res*, **31**, 1319–1330.
18. Bonnet, G., Tyagi, S., Libchaber, A. and Kramer, F.R. (1999) Thermodynamic basis of the enhanced specificity of structured DNA probes. *Proc Natl Acad Sci USA*, **96**, 6171–6176.
19. Tyagi, S., Bratu, D.P. and Kramer, F.R. (1998) Multicolor molecular beacons for allele discrimination. *Nat Biotechnol*, **16**, 49–53.
20. Li, J.J., R. Geyer, W. Tan. (2000) Using molecular beacons as a sensitive fluorescence assay for enzymatic cleavage of single-stranded DNA. *Nucleic Acids Res*, **28**, E52.
21. Molenaar, C., Marras, S.A., Slats, J.C., Truffert, J.C., Lemaitre, M., Raap, A.K., Dirks, R.W. and Tanke, H.J. (2001) Linear 2' O-Methyl RNA probes for the visualization of RNA in living cells. *Nucleic Acids Res*, **29**, E89.

22. Sokol, D.L., Zhang, X., Lu, P. and Gewirtz, A.M. (1998) Real time detection of DNA.RNA hybridization in living cells. *Proc Natl Acad Sci USA*, **95**, 11538–11543.
23. Vet, J.A., Majithia, A.R., Marras, S.A., Tyagi, S., Dube, S., Poiesz, B.J. and Kramer, F.R. (1999) Multiplex detection of four pathogenic retroviruses using molecular beacons. *Proc Natl Acad Sci USA*, **96**, 6394–6399.
24. Kostrikis, L.G., Tyagi, S., Mhlanga, M.M., Ho, D.D. and Kramer, F.R. (1998) Spectral genotyping of human alleles. *Science*, **279**, 1228–1229.
25. Piatek, A.S., Tyagi, S., Pol, A.C., Telenti, A., Miller, L.P., Kramer, F.R., Alland, D. (1998) Molecular beacon sequence analysis for detecting drug resistance in *Mycobacterium tuberculosis*. *Nat Biotech*, **16**, 359–363.
26. Bratu, D.P., Cha, B.J., Mhlanga, M.M., Kramer, F.R. and Tyagi, S. (2003) Visualizing the distribution and transport of mRNAs in living cells. *Proc Natl Acad Sci USA*, **100**, 13308–13313.
27. Tsourkas, A., Behlke, M.A., Xu, Y. and Bao, G. (2003) Spectroscopic features of dual fluorescence/luminescence resonance energy-transfer molecular beacons. *Anal Chem*, **75**, 3697–3703.
28. Santangelo, P.J., Nix, B., Tsourkas, A. and Bao, G. (2004) Dual FRET molecular beacons for mRNA detection in living cells. *Nucleic Acids Res*, **32**, e57.
29. Nitin, N., Santangelo, P.J., Kim, G., Nie, S. and Bao, G. (2004) Peptide-linked molecular beacons for efficient delivery and rapid mRNA detection in living cells. *Nucleic Acids Res*, **32**, e58.
30. Tyagi, S., Alsmadi, O. (2004) Imaging native beta-actin mRNA in motile fibroblasts. *Biophys J*, **87**, 4153–4162.
31. Peng, X.H., Cao, Z. H., Xia, J. T., Carlson, G. W., Lewis, M. M., Wood, W. C., Yang, L. (2005) Real-time detection of gene expression in cancer cells using molecular beacon imaging: new strategies for cancer research. *Cancer Res*, **65**, 1909–1917.
32. Medley, C.D., Drake, T.J., Tomasini, J.M., Rogers, R.J., Tan, W. (2005) Simultaneous monitoring of the expression of multiple genes inside of single breast carcinoma cells. *Anal Chem*, **77**, 4713–4718.
33. Tyagi, S., Marras, S.A. and Kramer, F.R. (2000) Wavelength-shifting molecular beacons. *Nat Biotechnol*, **18**, 1191–1196.
34. Tsourkas, A., Behlke, M.A. and Bao, G. (2002) Structure–function relationships of shared-stem and conventional molecular beacons. *Nucleic Acids Res*, **30**, 4208–4215.
35. Brodsky, A.S. and Silver, P.A. (2002) Identifying proteins that affect mRNA localization in living cells. *Methods*, **26**, 151–155.
36. Forrest, K.M. and Gavis, E.R. (2003) Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in *Drosophila*. *Curr Biol*, **13**, 1159–1168.
37. Shav-Tal, Y., Darzacq, X., Shenoy, S.M., Fusco, D., Janicki, S.M., Spector, D.L., and Singer, R.H. (2004) Dynamics of single mRNPs in nuclei of living cells. *Science*, **304**, 1797–1800.
38. Haim, L., Zipor, G., Aronov, S. & Gerst, J. E. (2007) A genomic integration method to visualize localization of endogenous mRNAs in living yeast. *Nat Methods*, **4**, 409–412.
39. Ozawa, T., Natori, Y., Sato, M. & Umezawa, Y. (2007) Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat Methods*, **4**, 413–419.
40. Valencia-Burton, M., McCullough, R. M., Cantor, C. R. & Broude, N. E. (2007) RNA visualization in live bacterial cells using fluorescent protein complementation. *Nat Methods*, **4**, 421–427.
41. States, D.J., Gish, W. and Altschul, S.F. (1991) Improved sensitivity of nucleic acid database searches using application-specific scoring matrices. *Methods*, **3**, 66–70.
42. Rhee, W.J., Santangelo, P.J., Jo, H., Bao, G. (2007) Target accessibility and signal specificity in live-cell detection of BMP-4 mRNA using molecular beacons. *Nucleic Acids Res*, submitted.

43. Lakowicz, J.R. (1999) *Principles of Fluorescence Spectroscopy*. 2nd ed. Springer-Verlag, New York.
44. Marras, S.A., Kramer, F.R. and Tyagi, S. (2002) Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res*, **30**, e122.
45. Price, N.C. and Stevens, L. (1999) *Fundamentals of Enzymology: The Cell and Molecular Biology of Catalytic Proteins*. 3rd ed. Oxford University Press, New York.
46. Dokka, S. and Rojanasakul, Y. (2000) Novel non-endocytic delivery of antisense oligonucleotides. *Adv Drug Deliv Rev*, **44**, 35–49.
47. Leonetti, J.P., Mechti, N., Degols, G., Gagnor, C. and Lebleu, B. (1991) Intracellular distribution of microinjected antisense oligonucleotides. *Proc Natl Acad Sci USA*, **88**, 2702–2706.
48. Mhlanga, M.M., Vargas, D.Y., Fung, C.W., Kramer, F.R., Tyagi, S. (2005) tRNA-linked molecular beacons for imaging mRNAs in the cytoplasm of living cells. *Nucleic Acids Res*, **33**, 1902–1912.
49. Giles, R.V., Ruddell, C.J., Spiller, D.G., Green, J.A. and Tidd, D.M. (1995) Single base discrimination for ribonuclease H-dependent antisense effects within intact human leukaemia cells. *Nucleic Acids Res*, **23**, 954–961.
50. Barry, M.A. and Eastman, A. (1993) Identification of deoxyribonuclease II as an endonuclease involved in apoptosis. *Arch Biochem Biophys*, **300**, 440–450.
51. Giles, R.V., Spiller, D.G., Grzybowski, J., Clark, R.E., Nicklin, P. and Tidd, D.M. (1998) Selecting optimal oligonucleotide composition for maximal antisense effect following streptolysin O-mediated delivery into human leukaemia cells. *Nucleic Acids Res*, **26**, 1567–1575.
52. Walev, I., Bhakdi, S.C., Hofmann, F., Djondjer, N., Valeva, A., Aktories, K. and Bhakdi, S. (2001) Delivery of proteins into living cells by reversible membrane permeabilization with streptolysin-O. *Proc Natl Acad Sci USA*, **98**, 3185–3190.
53. Snyder, E.L., Dowdy, S.F. (2001) Protein/peptide transduction domains: potential to deliver large DNA molecules into cells. *Curr Opin Mol Ther*, **3**, 147–152.
54. Wadia, J.S., Dowdy, S.F. (2002) Protein transduction technology. *Curr Opin Biotechnol*, **13**, 52–56.
55. Becker-Hapak, M., McAllister, S.S., Dowdy, S.F. (2001) TAT-mediated protein transduction into mammalian cells. *Methods*, **24**, 247–256.
56. Wadia, J.S., Dowdy, S.F. (2005) Transmembrane delivery of protein and peptide drugs by TAT-mediated transduction in the treatment of cancer. *Adv Drug Deliv Rev*, **57**, 579–596.
57. Brooks, H., Lebleu, B., Vives, E. (2005) Tat peptide-mediated cellular delivery: back to basics. *Adv Drug Deliv Rev*, **57**, 559–577.
58. Allinquant, B., Hantraye, P., Mailleux, P., Moya, K., Bouillot, C. and Prochiantz, A. (1995) Downregulation of amyloid precursor protein inhibits neurite outgrowth in vitro. *J Cell Biol*, **128**, 919–927.
59. Troy, C.M., Derossi, D., Prochiantz, A., Greene, L.A. and Shelanski, M.L. (1996) Downregulation of Cu/Zn superoxide dismutase leads to cell death via the nitric oxide-peroxynitrite pathway. *J Neurosci*, **16**, 253–261.
60. Minamoto, T., Mai, M., Ronai, Z. (2000) K-ras mutation: early detection in molecular diagnosis and risk assessment of colorectal, pancreas, and lung cancers—a review. *Cancer Detect Prev*, **24**, 1–12.
61. Altieri, D.C., and Marchisio, P.C. (1999) Survivin apoptosis: an interloper between cell death and cell proliferation in cancer. *Lab Invest*, **79**, 1327–1333.
62. Santangelo, P.J., Nitin, N. and Bao, G. (2005) Direct visualization of mRNA colocalization with mitochondria in living cells using molecular beacons. *J Biomed Opt*, **10**, 44025.
63. Santangelo, P., Nitin, N., LaConte, L., Woolums, A., Bao, G. (2006) Live-cell characterization and analysis of a clinical isolate of bovine respiratory syncytial virus, using molecular beacons. *J Virol*, **80**, 682–688.

64. Santangelo, P.J., Bao, G. (2007) Dynamics of filamentous viral RNPs prior to egress. *Nucleic Acids Res*, **35**, 3602–3611.
65. Liu, X. and Tan, W. (1999) A fiber-optic evanescent wave DNA biosensor based on novel molecular beacons. *Anal Chem*, **71**, 5054–5059.
66. Kambhampati, D., Nielsen, P.E. and Knoll, W. (2001) Investigating the kinetics of DNA–DNA and PNA–DNA interactions using surface plasmon resonance-enhanced fluorescence spectroscopy. *Biosens Bioelectron*, **16**, 1109–1118.
67. Steemers, F.J., Ferguson, J.A. and Walt, D.R. (2000) Screening unlabeled DNA targets with randomly ordered fiber-optic gene arrays. *Nat Biotechnol*, **18**, 91–94.
68. Kuhn, H., Demidov, V.V., Coull, J.M., Fiandaca, M.J., Gildea, B.D. and Frank-Kamenetskii, M.D. (2002) Hybridization of DNA and PNA molecular beacons to single-stranded and double-stranded DNA targets. *J Am Chem Soc*, **124**, 1097–1103.
69. Hamaguchi, N., Ellington, A. and Stanton, M. (2001) Aptamer beacons for the direct detection of proteins. *Anal Biochem*, **294**, 126–131.
70. Yamamoto, R., Baba, T. and Kumar, P.K. (2000) Molecular beacon aptamer fluoresces in the presence of Tat protein of HIV-1. *Genes Cells*, **5**, 389–396.

Packaging for Bio-micro-electro-mechanical Systems (BioMEMS) and Microfluidic Chips

Edward S. Park, Jan Krajniak, and Hang Lu

Abstract In the last two decades, fundamental and application-driven research on microfluidics and bio-micro-electro-mechanical systems (BioMEMS) has flourished in academia and industries and has begun to make impact on medicine and biosciences. Packaging of these systems is an integral if not critical part of the device/system design and function. Because the applications and the designs of the chips are wide ranging, it is difficult to achieve a universal packaging scheme that meets the requirements of all applications. Instead, research and manufacturing practices of each type of biochip have come up with specialty techniques. This chapter will review these techniques in the specific contexts of the chip applications, as well as materials requirements. In addition, we will highlight common and advanced practices and point out research needs in these areas.

Keywords Packaging · Lab-on-a-chip · BioMEMS · Microfluidics · Medical diagnostics · Chip-to-world connection · Biocompatibility · Surface modification · Interface · Integration · Fluidic interconnect · Electrical interconnect · Optical transparency · Disposable · Point-of-care diagnostic · Sample preparation · Fluid handling · Soft lithography · Large-scale integration · Microfabrication

1 Introduction

In recent years, applications of lab-on-a-chip, microfluidics, and bioMEMS devices have ranged from basic research on macromolecules (e.g., DNAs and proteins), cells, tissues, and organisms to clinical applications (e.g., blood analysis, disease diagnosis, and drug delivery). The device design and the final packaging of the devices demand two very important aspects of compatibility. First, the fabrication processes and sequences of all components (including device or reagents) need to be compatible with each other. Second, the devices and their packaging need to be

Hang Lu (✉)

School of Chemical Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA, USA

e-mail: hang.lu@chbe.gatech.edu

biocompatible, whether it is referring to minimizing sample contamination, optimizing sample interaction with device/packaging materials, or eliminating immune responses to implanted devices. Therefore, in designing these biochips, one has to consider not only the functionalities but also the packaging scheme with the final product and the environment in which it will be used.

This chapter is organized to address packaging issues in a variety of applications, packaging schemes that are commonly used, the types of interfaces one needs to think about (e.g., interface with the outside environment and interface between devices), and the materials of choice and their modifications. In designing a good packaging scheme, one needs to strategize in terms of constraints and a series of trade-offs. No packaging scheme is universal, and there exist many options. Depending on how the devices are designed and how they are intended to be used, the packaging scheme can be very different, and aspects that have smaller design space (miniaturization) need to be considered first. For example, if an implantable device needs to be packaged, the most important aspect is the biocompatibility; next, one needs to consider where the device is to be implanted and then match the mechanical properties of the tissues and consider all the geometrical constraints. More than often, the infrastructure that is available to manufacture the devices and the packaging will also put constraints on selecting the processes. For industrial applications, overall cost is likely an important parameter.

This chapter aims to point out some standard solutions and research opportunities in bioMEMS packaging and will refer to many primary research articles for detailed information on existing technologies.

2 Packaging Schemes Based on Application

2.1 Portable and Point-of-Care (POC) Diagnostics and Analysis

Owing to their small size, bioMEMS are ideal platforms upon which to design portable and point-of-care (POC) diagnostics and analytical systems. Portability is an essential consideration in the design and packaging of systems intended for use in resource-poor or similarly challenging environments. These environments lack the infrastructure of typical “developed-world” clinical laboratories, which provide stable electrical power, cleanliness, refrigeration, and highly trained personnel. Interest in developing bioMEMS-based portable analytical systems originated in the 1990s from the US Defense Advanced Research Projects Agency (DARPA), seeking to enhance chemical and biological weapons detection and to improve battlefield medical care. More recently, interest from both governmental and charitable organizations has led to significant investments in developing portable diagnostics to avert epidemics in poor and remote geographies [1, 2].

The design and packaging of portable diagnostic systems must satisfy a challenging set of requirements. Such requirements include (i) small size and weight, (ii) full integration of necessary functions (unit operations), (iii) low power consumption, (iv) easy operation by potentially non-skilled personnel, (v) functionality in wide

temperature ranges, (vi) ruggedness and mechanical protection, (vii) protection from dust or other contaminants, and (viii) low cost.

The packaging schemes that have emerged to address the above requirements can be grouped into two broad categories. The first category focuses upon integrating as many functions as possible onto a single chip in order to maximize portability. The second category consists of systems that are partitioned into two components: (i) a disposable chip (or cartridge) and (ii) a readout machine (or scanner). The cartridge is inserted into the scanner, read and analyzed, and then discarded. Both the integrated and partitioned categories are discussed in the following sections.

2.1.1 Integrated BioMEMS Packaging Schemes

Integrating functions onto a single bioMEMS device is the central aim for packaging diagnostic systems when size and weight are the most important criteria. Basic functions in a typical diagnostic system include fluid handling (i.e., pumping and mixing), reagent storage, species separation and concentration, and detection. Today, diagnostic chips that integrate all necessary functions are rare. In most cases, the devices are designed with a specific function in mind.

The major challenges in portable chip-scale integration involve fluid handling, detection, and electrical power. Fluid pumping for analytical purposes has traditionally been carried out using macroscale machines (i.e., peristaltic, syringe pump). Likewise, analyte detection is normally accomplished using complex optical or spectroscopic techniques that require bench-top equipment, which are usually expensive and bulky. In such forms, both functions require power sources that exceed the latest capabilities of portable batteries. To circumvent these challenges, integrated biochips must often rely upon passive fluid transport mechanisms (e.g., capillary action) and simplified “yes/no” detection readouts for portable diagnostics.

Immunochromatographic test strips. The classic example of integrated, portable diagnostic devices is immunochromatographic strip (ICS) tests. ICS tests are packaged with a strip of fibrous material that is divided into a sample pad, conjugation/labeling region, and detection zone. The strip is contained in a plastic casing with access windows over the sample pad and detection zone. Fluid samples are transported along the strip via capillary action through the pores and channels in the fibrous material and are bound to an antigen (or by an antibody) as they pass through the labeling region. Typically, a dye is labeled to the analyte, allowing for easy visual readout by the operator. ICS tests have been successfully deployed “in the field” for over 10 years to diagnose infectious agents like diphtheria toxin [3] and various sexually transmitted diseases [4–6]. The ICS packaging scheme is advantageous because it does not include any active power source. In addition, materials and reagents are off-the-shelf, inexpensive, and compatible with high-volume production methods.

The primary drawback of ICS design is its relatively crude “yes/no” readout mechanism (although visual readouts can also be an advantage if the operator is non-skilled). Paper-based microfluidic systems are being developed that provide colorimetric readout of digital images [7–9]. Such systems are comprised of a

paper sheet that is embedded with a photoresist pattern that guides sample fluid into different detection zones. Each zone contains different labels for parallel diagnoses. Paper-based systems exhibit the same packaging advantages as ICS systems. However, the image acquisition is performed using a digital camera, which reduces the degree of integration but increases the precision of the measurement due its semi-quantitative nature. For both ICS and paper-based microfluidics, this simple packaging scheme compatible with the economic requirement of the applications.

Electrical detection. In order for portable diagnostics to provide fully quantitative measurements, researchers are investigating ways to miniaturize sophisticated detection components and package them on-chip. The most actively researched techniques are electrical, mechanical, and optical. To achieve electrical detection on-chip, conductive electrodes are integrated into the flow path of the sample fluid. The electrodes are typically metallic and are sometimes functionalized with molecular probes or modified with a selective membrane. The simplest packaging configuration for electrical detection is to pattern electrodes on a planar substrate, which is interfaced with another substrate that contains the sample flow path (Fig. 1). Several examples of this configuration exist, most notably for glucose and urea detections [10, 11], where a current or a potential is measured as a result of redox (reduction and oxidation) reactions on the substrate. A similar packaging configuration enables ICS test strips to give fully quantitative outputs using screen-printed, disposable electrodes that couple to quantum dot labels [12]. Also, highly sensitive and reversible electrical detection of antibodies has been demonstrated using silicon nanowires grown on an oxidized silicon surface and coupled to a polydimethylsiloxane (PDMS) flow channel [13, 14]. The challenge is that often sensitive off-chip electronics are needed to detect minute changes in current or potential.

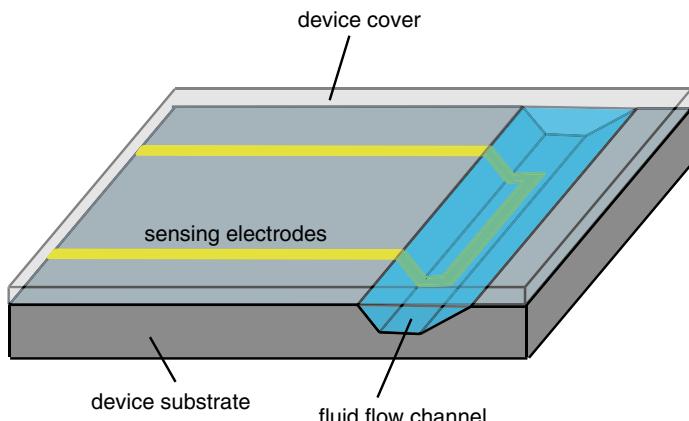


Fig. 1 Electrodes used for sensing are deposited on the surface of the device, with the sensing portion protruding into the fluidic components. Packaging materials compatible with the manufacturing requirements have to be chosen. For example, if the device is to be polymer based, deposition of metal electrodes will be complicated and will require modification of the polymer

Mechanical detection. On-chip mechanical detection has been achieved using a two-substrate packaging scheme similar to electrical detection. Mechanical detection has been demonstrated using micro- and nanoscale cantilever beams that are machined from bulk silicon using MEMS fabrication techniques and functionalized with the appropriate capture probes [15, 16]. A change in the surface stress or resonance frequency of the cantilever indicates binding of biological species. The cantilever's mechanical output is then measured in two ways: (i) optically, by tracking the change in location of a laser beam reflected off the cantilever or (ii) electrically, by coupling a piezoresistive element to the edge of the beam. As a result, cantilever detectors must be packaged in a manner that can interface to optical components or electrical leads, respectively. Specifically, optical detection would require a viewing aperture with a transparent substrate (and laser beam alignment and position detection), while electrical detection would require attachment of leads to bond pad arrays (Fig. 2).

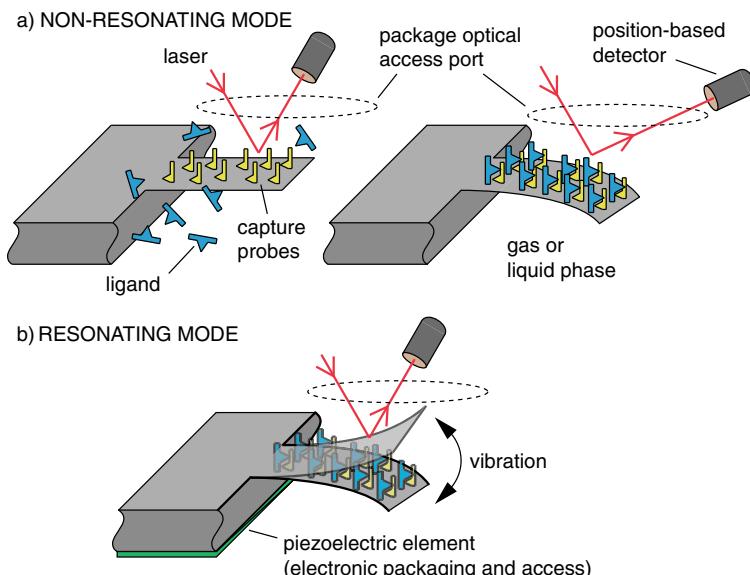


Fig. 2 Mechanically based bio-detection via cantilever beams requires optical and electrical access in packaging. **(a)** A stress-based cantilever functionalized with capture probes is exposed to fluid containing a ligand of interest; attachment of the ligand causes the beam to deflect, changing the angle of reflected light. DNA hybridization and protein–protein interactions have been reported using stress-based methods [17, 18]. **(b)** A resonance-based cantilever responds by changes in resonance frequency due to changes in mass as ligand attaches to capture probes [19, 20]; piezoelectric actuation requires electrical leads and interconnect in packaging

Because of such additional components, portable and fully integrated cantilever-based diagnostic systems have yet to be realized. However, a major advantage of cantilever-based detection, which is the ability to detect molecules without attaching optical (e.g., fluorescent) labels to the molecules themselves, motivates continuing

research. Label-free, stress-based detection has been demonstrated with DNA and proteins [17, 18, 21, 22]. In addition, resonance-based detection has been demonstrated with cells [20] and viruses [19]. A particularly novel resonator is configured so that the sample flows through a microchannel in a hollow cantilever, which enables sub-femtogram measurement resolution [23, 24]. High sensitivity is made possible by packaging the cantilever and drive electrode under high vacuum in a hermetically sealed cartridge. Such innovation in both fabrication and packaging is a good example of collaboration between academia (MIT) and industry (Innovative Micro Technology, Inc. and Affinity Biosensors). Although some progress has been made in this area in miniaturizing and integrating the detection modules, most mechanical detection chips still require some off-chip components (i.e., optics). This area therefore remains an active research area for miniaturization.

Optical detection. The miniaturization and integration of optical detection on-chip is also an active area of research. Packaging schemes vary dramatically, depending on which optical components are deemed critical for integration. The most basic approach is to attach a separate module containing a miniaturized excitation source and photodetector directly to a chip containing the appropriate flow network. Although it is not true chip-scale integration, the modular approach confers advantages with respect to flexibility and convenience, while maintaining acceptable (though not ideal) size and weight. Such an approach has been successfully employed to couple a glass capillary electrophoresis chip to a light-induced fluorescence (LIF) module to detect protein biotoxins (i.e., ricin, botulinum toxin) without preconcentration [25].

A similar modular approach assembles a miniature diode laser, sample well plate, and detection chip. The use of a self-contained detection chip in lieu of a full-size camera makes the assembly better suited for field use, and diffractive optics split the excitation beam into 16 channels, resulting in parallel analysis of 16 corresponding wells. The chip-to-plate approach is shown to detect low concentrations of *Bacillus globigii*, a surrogate marker for *Bacillus anthracis*, utilizing ELISA signal amplification to compensate for the detection chip's limited sensitivity [26].

True on-chip integration of optical components has been achieved on a limited basis. The concept of a "microscope-on-a-chip," integrating miniature lenses and complementary metal oxide semiconductor (CMOS) detectors on-chip, has been investigated [27]. Optical fibers micromachined with angled grooves have been embedded underneath microfluidic chips to guide excitation beams into multiple flow channels simultaneously [28]. Similarly, optical fibers have also been directly inserted into waveguides patterned in photoresist (SU-8, a multifunctional epoxy negative photoresist, manufactured by MicroChem Corp.) [29].

Using soft lithography, the fabrication of several optical components, such as beam splitters, lenses, and prisms, has been shown [30]. Moreover, integrated circuit (IC) fabrication techniques have been utilized to produce monolithic light-emitted diodes (LEDs), photodetectors, and optical fibers [31]. From a packaging standpoint, this represents a high degree of integration, uniting the excitation source, waveguides, and detectors onto a single silicon chip that is attached to a PDMS microfluidic flow channel. In a similar display of high integration, optical

waveguides in photoresist, an optically pumped dye laser, and photodiodes embedded in silicon have been integrated on-chip [32]. On-chip optical integration is a large and diverse field [12, 33]. Nevertheless, the field is still in its infancy, providing few portable packages for diagnostic applications today.

Other components. The portability of diagnostic devices has been improved by the integration of other components, such as vapor barriers to prevent the evaporation of fluids from microfluidic systems constructed from gas-permeable elastomers [34]. Also, on-chip diaphragm-based pumps and valves, which rely upon thermal actuation from resistive losses from adjacent electrodes, have been integrated into a portable polymerase chain reaction-capillary electrophoresis (PCR-CE) system [35]. Furthermore, the storage and release of reagents on-chip has been addressed by the integration of torque-actuated screw valves, which eliminates the necessity for electrical power when closing or opening microfluidic flow paths [36].

Integration outlook. Although the full integration of all essential functions onto a single chip has not been realized, some devices have come very close to being total analysis system “on a chip” [37–39]. As previously discussed, fluid handling, detection (i.e., electrical, mechanical, optical), and electrical power are the major challenges in developing a fully integrated device. In addition, stored reagents could also add to the bulk of a diagnostic chip. The trade-off one makes as more functions are integrated on-chip is that detection sensitivity typically decreases. In addition, the cost per chip increases. Because of these limitations, the idea of inserting a simple and disposable diagnostic chip into a separate scanner box could provide a workable compromise between portability and performance.

2.1.2 Portable BioMEMS Chip in Tandem with a Readout Box

By partitioning a portable diagnostic system into a disposable chip and a scanner, not only can high performance be provided, but also cleaning and per-test costs are reduced. From the design standpoint, a major task is to decide what mix of functions will be located on the disposable chip versus on the reader. Careful consideration will result in functions that are located in a manner that maximizes the combined benefit of detection sensitivity, ease of use (during testing and maintenance), and low cost. From there, packaging can be developed that properly interfaces the chip to the reader and the reader to the operator. Figure 3 illustrates the concept of the partitioned approach.

A champion of the partitioned design concept is the Yager Research Group at the University of Washington. Its development of a disposable diagnostic card for saliva testing is an insightful case study [2, 40]. Most active functions, such as fluid pumping and surface plasmon resonance [41] excitation and detection, are located on the scanner. Functions located on the chip, which include sample preconditioning, mixing, and assaying, are specifically designed to be passive in nature. For instance, the mixing region relies upon chaotic mixing that is induced by herringbone structures as fluid flows through the channel. Also, the diffusion of small molecules across

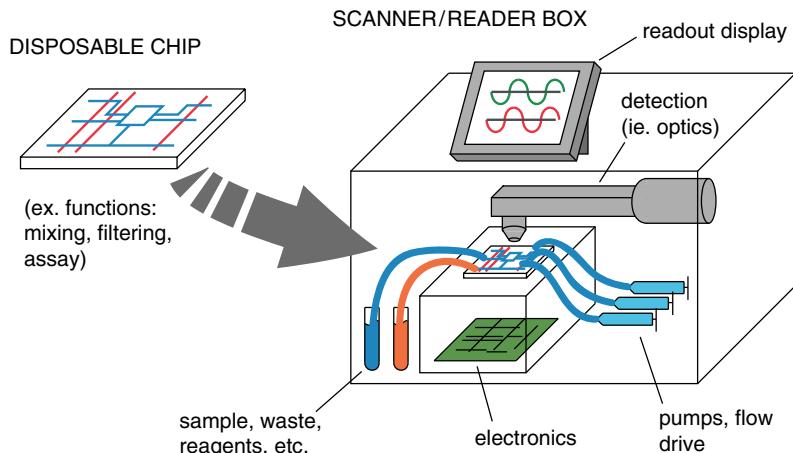


Fig. 3 A bioMEMS chip used in tandem with a scanner. Functions are partitioned between the disposable chip and scanner box, thereby optimizing critical factors (i.e., cost, detection sensitivity, portability) [2]. Interconnect between the two elements, as well as scanner maintenance, is an additional packaging consideration

the interface of two laminar flow streams (buffer and sample) is utilized to precondition the sample before assay. Such a preconditioning approach cleverly removes the necessity for less portable preconditioning methods (e.g., centrifugation). The chip (roughly 4 inches by 2 inches in size) consists of a PDMS slab containing the flow channel network that is bonded to a glass slide, which is gold-coated and functionalized for antibody capture. Rubber gaskets (O-rings) are embedded into the fluidic inlet/outlet ports to facilitate connection with off-chip pumps and other active elements.

A similar partitioned package, consisting of a microfluidic chip whose channels are loaded with polyacrylamide gel, has been used to perform electrophoretic immunoassays to measure the concentration of the protein MMP-8 from a saliva sample [42]. The glass chip is inserted into a portable scanner containing a user interface, miniature high-voltage power supply, data acquisition, and software control. Although some sample preparation and fluorescence imaging was performed with standard laboratory equipment, the system proves the principle of a modular, point-of-care microfluidic assay for a clinically relevant disease marker. The choice of glass also minimizes work that has to be performed for surface modification and ensures optical compatibility for the readout.

The partitioned design strategy is being pursued in other contexts also [43]. A capillary electrophoresis (CE) chip for the separation and detection of organophosphate nerve agents has been demonstrated as a chemical weapon detector [44]. The glass chip contains a CE channel coupled to a screen-printed carbon electrode for amperometric detection. The scanner houses the electrochemical detector, reference/counter electrodes, and buffer/sample reservoirs. A portable,

highly parallel SPR system has also been developed for chemical weapons detection [45]. The system is based on the Texas Instruments' Spreeta sensing chip, which contains three LEDs and a diode array detector. The sensing chips are inserted into a control box that houses the power supply, LCD touch-screen display, custom electronic board, and fluidic elements, such as pumps and valves. Each chip is replaced via a snap-in mechanism that mates it to a silicone flow cell. The combination of eight chips, each having three channels, enables the simultaneous monitoring of up to 24 analytes.

Yet another embodiment of the tandem strategy utilizes a cartridge in which cells are cultured and stimulated for use as a cell-based biosensor [46]. The cartridge is made of a CMOS silicon chip bonded to a PDMS slab patterned with cell chambers. Circuitry on the CMOS chip enables temperature control and electrophysiological sensing with microelectrodes. The cartridge interfaces with a handheld electronic reader that monitors the action potential (AP) activity of cardiomyocytes within the cell chambers. The system has successfully monitored the AP response of cells to different biochemical stimulation under conditions outside the laboratory (desert field testing).

2.1.3 Outlook for Portable Diagnostics

Portable diagnostic and analysis systems have the potential to dramatically improve the health of people in the poorest, most remote areas in the world, especially with respect to infectious diseases. Moreover, the benefits of portability will undoubtedly contribute to the ease and decentralization of disease management in the developed world as well. One could imagine patients wearing portable diagnostic devices that monitor their response to drugs for screening or dose titration purposes, which would be a critical step toward realizing the concept of “personalized medicine” [47, 48].

The major challenges facing packaging, as previously discussed, are function integration and partitioning between disposable and non-disposable components. Another challenge is cost. Virtually all prototype devices today have been fabricated out of relatively expensive base materials, such as glass, silicon, or PDMS. Furthermore, their fabrication processes may not be ideal for mass production, unlike thermoplastics that can be hot-embossed or injection-molded, which makes them disposable. For example, a microfluidic immunosensor made of embossed plastic has recently been applied to monitor cardiac inflammation markers from human serum [49]. On the other hand, some applications have stringent materials requirements that disqualify some plastic materials; hence, glass and PDMS continue to serve in many areas as the materials of choice.

Besides the device components, the packaging of reagents presents an additional challenge to system portability. To reduce bulk and transportation costs, reagents would ideally be packaged in dry form and reconstituted into solution at the point of care. In situations where the reagent can neither be dried nor reconstituted with local solvents to maintain activity (e.g., antibodies), the reagent solution must typically be packaged off-chip due to inadequate space within the minuscule volume of the

on-chip microfluidic network. To address this, the use of plastic clinical tubing as a reagent cartridge has been investigated [50]. Plugs of different reagents are sequentially loaded into a tube and spaced by air. Similar packaging may enable reliable transport of valuable, ready-to-use solutions in portable diagnostic systems.

2.2 *Implantable Devices*

In the recent years, MEMS technology has become popular in the medical field due to its many advantages over traditional technologies. BioMEMS chips are highly functional, versatile, and power efficient and can be fabricated much smaller than other implantable systems. Furthermore, since bioMEMS fabrication processes originate in the IC industry, the infrastructures of the fabrication are well established, and these techniques are reasonably well developed and standardized [51]. Therefore, it is no surprise that many implantable devices that primarily use IC fabrication technologies, such as wireless pressure sensors [52], subcutaneous drug delivery chips [53], and glucose sensors [54, 55], have been and are being developed and commercialized.

The major packaging challenges for implantable MEMS devices include but are not limited to biocompatibility and size considerations. Another consideration is matching the mechanical properties and sometimes optical properties of the surrounding tissue. Implantation is an invasive process and triggers a healing and immune response from the body. Not only do implants have to be designed to minimize the immediate and long-term immune response, but they are also required to be non-toxic to the surrounding tissue. For this purpose, various packaging methods, materials, and protective coatings are available.

These designs must be developed with the miniaturization requirement in mind. They need to add as little volume or length to the device as possible and have to be adaptable to the specific tissue environment and their applications. For example, ocular implants need to be small enough for implantation through hypodermic needles, flexible enough to not cause irritation to the patients, and to match the mechanical properties of the surrounding tissues, but mechanically sturdy enough to survive the implantation process. The latter is particularly critical because inadequate protection during implantation can cause a complete loss of device functionality [56]. This section will review some successful strategies for implantable devices with these considerations in mind.

2.2.1 *Drug Delivery Devices*

Packaging for drug delivery devices is highly dependent on the duration of the intended delivery applications. Short- or intermediate-term delivery can be achieved through oral or subcutaneous administration, whose packaging requirement is minimal, typically involving encapsulating drugs in microbeads or delivery through various types of microneedles [57–60]. However, for long-term controlled delivery, bioMEMS devices seem to be most suitable: these devices can be designed to be

highly functional; in addition, if proper packaging is designed, the drug reservoirs for controlled delivery can be made refillable [61].

A highly efficient and reliable controlled-release biochip design for implantation was introduced by Santini *et al.* [62]. The release of drug occurs from an array of wells covered by a gold thin-film membrane, which is electrochemically dissolved during the time of release (Fig. 4a). The reservoirs are fabricated using standard wet-etch MEMS technology in silicon wafers. The gold membrane covering the wells serves simultaneously as an electrical component and as a protective layer for the content of the reservoirs. The gold membrane is chemically stable and mechanically able to resist pressures exerted by the tissue and interstitial fluids [62]. Applying proper voltage to the anodic membrane leads to the dissolution of the membrane. This process is dependent on the presence of chlorine ions (Cl^-) in bodily fluids.

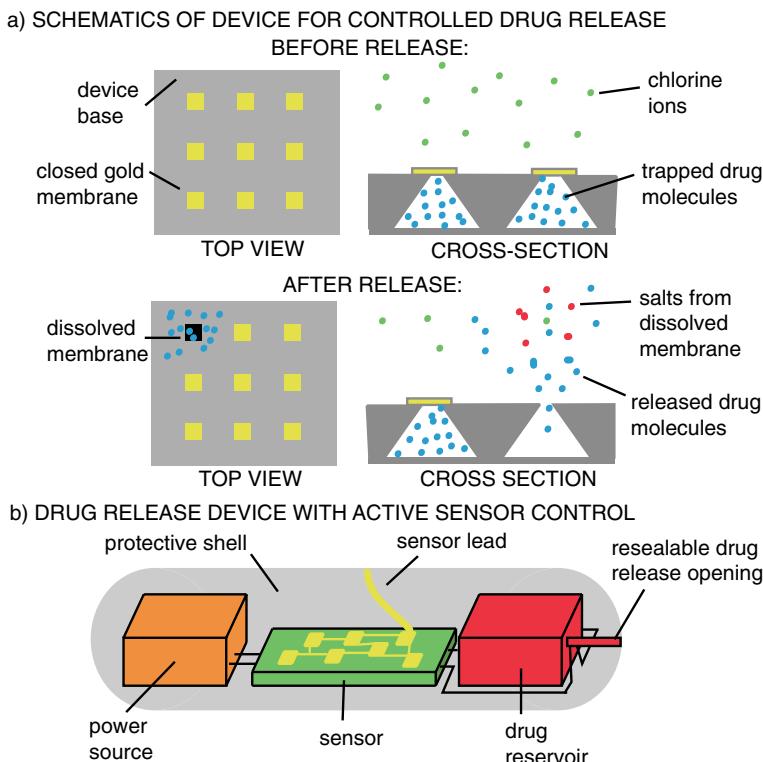


Fig. 4 Implantable drug release devices. (a) Basic design schematic of the device design by Santini *et al.* Prior to release, gold membranes form protective caps for drug compartments etched into the device. Once current is applied to the membrane, the gold forms soluble salts with Cl^- ions from the environment, dissolving the membrane and releasing the drug. Coating for protection and passivation can be applied. (b) Concept drawing of a drug release system combining active sensing technology and controlled drug release, contained within an implantable protective shell. Concept adapted from research by MicroCHIPS and ChipRx.

Soluble salts of gold form due to the applied potential and the presence of Cl^- , and instead of which dissolve into the surrounding fluid. Santini *et al.* established that these salts are biocompatible and are present in very low concentrations after the dissolution of the gold metal [63]. Because silicon is the structural material for the chip, the device overall is suitable for implantation. To protect the metallic components of the chip further from the interstitial fluids, SiO_2 is deposited as a protective coating for the device.

This technology is being further developed for commercial application by MicroCHIPS [64]. The aim is to couple the release of drug to sensor devices for improved release timing and schemes. As such, a microchip packaged with a sensor array and a power source would regulate the release of drugs from the reservoirs [53] (Fig. 4b). A modified version of the approach, which relies on re-sealable compartments, is being investigated by ChipRx [64]. This technique could prove useful in limiting the size requirement for larger amounts of drugs through the use a single refillable compartment. While silicon is a chemically and mechanically robust material and is usually easy to be rendered biocompatible, one drawback of silicon-based devices is that these devices may not be suitable for other drug delivery formats that require soft and flexible surface materials, such as ocular drug delivery. New approaches and new materials need to be considered.

2.2.2 Ocular Implants

Li *et al.* [61, 65] developed a drug delivery device enclosed in silicone rubber for ocular drug delivery (Fig. 5a). The soft material protects a two-part chip – a reservoir and a delivery chip; the refillable drug reservoir is bound to the delivery chip during the packaging process. The packaging is comprised of a silicone rubber outer surface molded to fit the curvature of the eye, which is filled with paraffin wax and encloses the delivery chip, provided that the chip is small compared to the eye. The advantage of this approach is that the device is completely enclosed in the flexible and biocompatible material [61, 66]. Furthermore, since the enclosure is made of flexible material, the reservoir can be punctured by a needle and refilled with drug, after which the silicone rubber seals itself [61]. This already efficient packaging could be further improved through various surface modification technologies. Integration of traditional IC or MEMS components is also made possible with flexible polymeric packaging [67–69]. The chip device can thus be fabricated from the required or desired material and packaged for flexibility. Another option is to design a hard-shell packaging scheme, which must minimally interfere with the tissue in the eye [70, 71].

Another type of ocular implant is the retinal stimulation device (Fig. 5b). The purpose of this implant is to stimulate the ocular nerve in diseases caused by photoreceptor degeneration. These implants incorporate image sensors and electrode stimulators to facilitate image capture and nerve stimulation, which renders the chip more complex [69]. Because of the electrical components, CMOS fabrication technology is used to produce the device components. Typically all of the

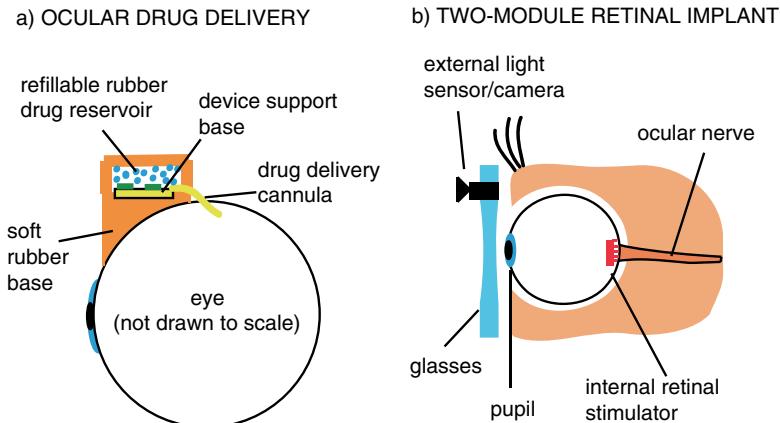


Fig. 5 Examples of devices for ocular implantation. (a) Example of an ocular drug delivery device fabricated with a soft silicone rubber shell for minimal eye irritation designed by Li et al. [61, 65]. The base is molded to conform to the curvature of the eye. The reservoir can be refilled by puncturing the rubber reservoir cap, which re-seals itself after the procedure. The protective silicone rubber is a soft and biocompatible material and thus suitable for implantable devices. (b) The light sensor is separated from the retinal stimulator in this ocular implant. Only the stimulation chip is implanted in the eye, while the light sensor or camera is worn on the outside. This allows for minimal invasiveness, especially if the implanted chip is made from soft and flexible biocompatible materials

components are fabricated on a highly flexible silicon substrate. The silicon substrate between these components is then thinned, leading to an overall flexible device [69].

Monolithic integration of all parts for ocular implant devices can be advantageous in terms of simpler design because of the electrical connectivity and signal transduction among the sensing, stimulation, and electronics components. However, this monolithic scheme comes with the price of having larger devices. The alternative is then to reduce the number of components in the system to be implanted. In the case of ocular implants, image sensors/CMOS cameras can be worn on the outside, while only a signal receiver and the stimulation chip (with its electrodes) need to be implanted. In one packaging scheme, the implanted components are then coated with Parylene-C to form a biocompatible protective layer [67, 68] where the device is fabricated on a flexible (but not necessarily biocompatible) substrate, such as polyimide [68].

Hard-shell protection for the implanted device is also an option. Since the shell is not flexible, its shape and size should be minimally invasive to the eye. Tube-like capsules can be fabricated from glass, ceramics, metals, or silicon, all of which must be hermetically sealed to the device [71]. This method has been pursued by BIONTM implants. These implants incorporate circuitry, a power source, and other functional parts within a hermetically sealed tantalum tube [70]. The rugged yet biocompatible tube is small enough to be implanted through a needle.

2.2.3 Neural Interface Implants

Some of the most common neural interfaces include pacemakers, neural probes, brain stimulators, and cochlear implants [72, 73]. The connection to the nervous tissue usually occurs through electrodes implanted into the brain tissue and in the vicinity of neurons and their processes. These electrodes are commonly organized into arrays such as the Michigan probe array and Utah array [74–76]. Both of these are variations of arrays of long silicon multi-electrode probes (Fig. 6) and are commonly coated with silicon nitride (SiN), silicon dioxide (SiO₂), or gold [73] to improve biocompatibility.

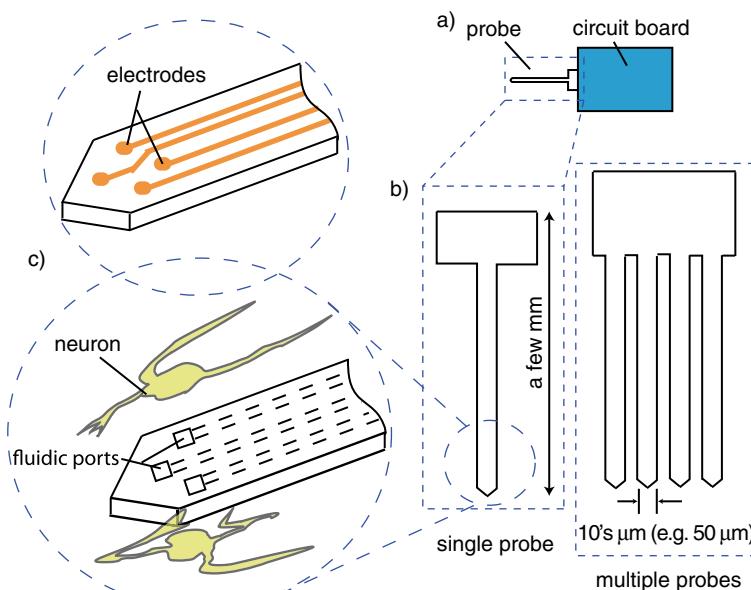


Fig. 6 Device details and packing of typical neural probes. (a) Microfabricated neural probes are usually mechanically mounted on and electrically connected to a circuit board. The probes are inserted into brain tissues for recording of neural activities. (b) These microfabricated probes can include either a single probe head or multiple parallel probe heads. (c) The tip of the probes can incorporate both metal electrode leads and fluid delivery ports. Materials to render the devices biocompatible or to include additional drug delivery capabilities (not drawn)

Besides silicon, which is mechanically rigid and easy for insertion during surgery, other materials are used to create flexible arrays. Common choices include polyimide, benzocyclobutene and Parylene. Some of these devices also include microfluidic channels for delivery of drugs to discourage scar tissue buildup for long-term recording and stimulation of the electrodes. In most arrays (stiff or soft), it is common to coat the surfaces with (i) poly(ethylene glycol) (PEG) films to minimize protein adsorption and (ii) drug eluting coatings [73]. The latter is achieved by using amphiphilic hydrogel particles that are capable of releasing drugs over time. One challenge of incorporating these particles into the coating is the choice of

polymers, which need to be stable and biocompatible over time and also giving the desired drug release profile. Modification of the surfaces to increase porosity may also be used to reduce cell adhesion, which may be especially useful in preventing isolation and preserving functionality of probes [73].

Tokuda *et al.* [77] developed a general neural interface based on a multi-chip system similar to the segmented ocular implant described above. Several interconnected unit chips equipped with stimulating and recording electrodes can be implanted as a thin sheet. Due to the small size of each chip and flexible interconnects, the array can be bent and extended. However, for added mechanical protection, the chips are anchored on a polyimide film and covered with epoxy resin, which protects the body of the chip while leaving the electrodes exposed [77].

As was the case with ocular implants, hard-shell enclosures may be used with neural interface implants. Biocompatible materials such as titanium could be used to manufacture the shell. The shell has to be hermetically sealed, which can be achieved through the use of medical-grade epoxy, which also provides additional shock absorption for the circuitry [78].

2.2.4 Cardiovascular Implants

Stents. Though stents are not typically considered bioMEMS devices, packaging technologies for stents can be applied to MEMS devices designed to interact with the cardiovascular system. The cardiovascular system is very sensitive to disturbances, and the presence of foreign objects can lead to problems such as intimal hyperplasia, or thickening of the blood vessel wall due to injury, and blood coagulation.

Stents often suffer from restenosis or the re-narrowing of the blood vessel wall, rendering them not functional. To prevent restenosis, various protective coatings have been developed, with the goal to provide a biocompatible interface, which passively or actively limits vessel wall thickening or fouling. TiNO_x films can be deposited by physical vapor deposition (PVD); tuning the ratios of nitrogen to oxide allows for control of electrical properties of the film [79]. Polymeric coatings are also often used to provide a protective function. Drugs to prevent restenosis and hyperplasia are often contained in the polymer matrix and eluted from the matrix after the implantation [80, 81]. Although the type of drug may vary, the coating procedure generally relies on the application of a thin layer of polymer–drug mixture in solvents (e.g., by dip coating). Some of common polymers used for this purpose include methacrylate and ethylene-based polymers [81].

Stents can also be implanted in combination with a pressure sensor to monitor the success of the implantation procedure [52]. CardioMEMS Inc. has developed a capacitance-based sensor enclosed in flexible polymers (Fig. 7). The capacitance change of the sensor can be correlated to deflection of the capacitor component from which pressure data can be calculated. The sensor is constructed from ultra-flexible alloys for additional flexibility. Pressure-related data are obtained wirelessly from an external readout machine. Communication between the devices is achieved through an energizing radio frequency (RF) signal sent by the readout machine. The signal is coupled to the sensor via magnetic coupling and induces a current in the

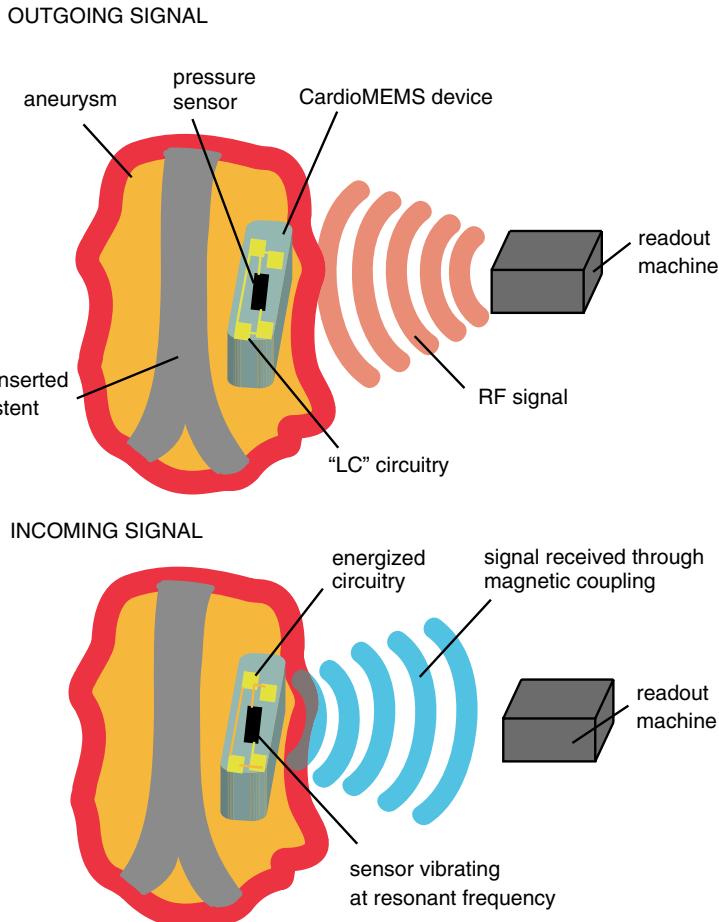


Fig. 7 CardioMEMS wireless pressure sensor. The sensor is implanted inside an aneurism along with the stent. Pressure measurements are obtained using an external readout module. This module sends out RF signals, which energize the LC circuit on the implanted device, including a pressure sensor. The sensors resonant frequency is determined through a series of RF signals with different frequencies; the module receives data through magnetic coupling. Once the resonant frequency has been determined, the pressure reading can be calculated. Packaging for this device not only needs to be biocompatible but also to pose minimal interference to the incoming and outgoing signals

sensor, which vibrates at a specific frequency depending on the applied pressure. The readout machine receives the response from the sensor through magnetic coupling and determines pressure based on the frequency of the sensor [82]. In this case, the packaging materials should not interfere with the RF signals.

Pacemakers. Pacemakers are electrical devices that interact with the cardiovascular system and the nervous system. Although typically not considered as MEMS devices, they are sometimes interfaced with MEMS devices and have successful packaging schemes that implantable MEMS devices can borrow. Packaging

schemes for different components of pacemaker devices face the challenge of different requirements. Typically the power source and control mechanism can be housed in traditional enclosures manufactured from titanium [83], PDMS, to a myriad of other materials. However, packaging of the pacemaker leads has to be treated separately.

Pacemakers are packaged to minimize the required pacing current and the polarization of the lead, to protect the lead from damage, and to reduce the inflammatory response around the lead. The packaging schemes are similar to that of stents; for example, protective polymeric sheaths, such as polyether polyurethane sheaths, are further covered by silicone rubber sheaths [84–86]. As was the case with stents, these sheaths may elute drugs, such as steroids, to prevent inflammatory response [84, 85].

2.2.5 Implantable Biosensors

Glucose sensors. Continuous glucose measurement is critical in monitoring the well-being of diabetes patients. A myriad of glucose measurement techniques is available for subcutaneously implantable glucose sensors. In these techniques, blood sugar can be measured through changes in fluorescence of a glucose sensitive system [87], output current in electrochemistry [54], or changes in viscosity [55]. Since each of these methods relies on different chemical and/or physical principles, packaging schemes must be fitted to reflect the different demands.

Encapsulation of fluorescently sensitive systems, whether they are housed on-chip or in micro particles, can be achieved through coating by PEG hydrogels [87]. To obtain a flexible sensor, which can be worn or implanted, biocompatible polymers can be used [54]. Flexible polymers are particularly useful if bendable electrodes are used in a current-sensing-based device. Bioinert polymers, such as PDMS and dimethylacrylamide (DMA), may be combined with other types of polymers for additional functionality, such as increased permeability of small molecules [54]. Biocompatible rigid sensor enclosures can also be manufactured from materials such as silicon, glass, or medical-grade epoxy [55] to reduce the risk of immune response.

Pressure sensors. MEMS-based minimally invasive pressure sensors can be used to monitor intraocular, intracranial, and cardiovascular pressures [52], each being of great interest to parts of the medical community [88]. Although the designs may differ, these sensors are generally fabricated using CMOS technology and enclosed in silicon [89–91]. The overall device may be further coated with deposited silicon nitride or silicon oxide layers.

As with other implantable devices, pressure sensors can also be coated with a polymer film for protection during the implantation process and normal operation. One such example is the use of Parylene-C as a protective membrane for ocular implants, which can be deposited through CVD processes at room temperature [92]. The polymer package may be further treated to reduce bubble formation on the surface through oxygen plasma roughening [92].

Communication between the sensors and the outside world must be considered in the packaging scheme. For example, ocular pressure sensors can be observed directly through the tissue of the eye, since this tissue is mostly transparent. Choosing transparent packaging such as biocompatible polymers will thus enable readouts by direct observation [93]. However, this option is not available for sensors implanted in non-transparent tissue, such as the CardioMEMS sensor described previously. For deep tissue implants, other readout methods are required, such as the RF-based method.

2.3 BioMEMS Packaging for Clinical Applications

BioMEMS technology is rapidly gaining attention in the clinical field due to its high versatility and broad range of functions. The most common bioMEMS uses can be divided into two broad categories: (i) analytical tools and (ii) components of medical devices, such as endoscopes or catheters. Section 2.2 reviews many implantable devices. This section will include additional devices used in medical diagnostics. Analytical devices used in the laboratory environment share designs common with point-of-care diagnostic devices, with some key differences in packaging schemes. The major difference is rooted in the availability of off-chip analytical technology and a clean and protected processing environment in the laboratory. Many of these devices are a direct miniaturization of existing conventional macroscopic assays. Even though all sample modification and processing occur on chip, analysis of the results can be performed off-chip [94] without exposure to an interfering environment. Hence, the challenge for packaging schemes lies in providing ease of interfacing to external instruments and primarily a high level of re-usability of bioMEMS devices.

Enzyme-linked immunosorbent assay (ELISA) devices are a prime example of the common device design. ELISA chips utilize antibody–antigen interactions to detect the presence of specific molecules in the analyte fluid. The sample is pumped through the chip resulting in (i) bonding of specific analyte to a localized protein or enzyme and (ii) changes in chemiluminescence, from which presence of the desired analyte can be inferred. The change is observed by an external camera [94] or on a microscope and processed on a computer. Analogous designs may include a sensor array and wireless transmitters to gather signals and relay them to the final analysis system [95]. Packaging schemes for these types of devices do not face biocompatibility issues, but must allow for interfacing the chip to the outside world. This most commonly occurs in the fluidic, optical, and electrical domains, since analytes are introduced in solution and signal gathering occurs through imaging or microelectrode sensing.

Polymerase chain reaction (PCR) chips are also heavily used in the medical field when genetical analysis is required. PCR chips can be manufactured using typical MEMS materials such as silicon, glass, epoxy, SU-8 (an epoxy negative photoresist, MicroChem Inc.), PDMS, poly(methyl methacrylate) (PMMA), poly(ethylene terephthalate) (PET or PETE), and common fabrication processes [96, 97]. These

devices operate on fluids, making use of standard microfluidic approaches for pumping, mixing, and flow control. As such, their packaging must allow for easy sample introduction and removal of product. However, reuse of PCR chips (or similar biological or medical analytics devices) is complicated due to risk of contamination from previous samples [96]. This risk is minimized through the use of various cleaning and packaging schemes. Prakash *et al.* use a removable silane coating in a glass-based device to prevent cross-contamination. A silane layer is first applied to improve the hydrophilicity of internal channels for PCR. After the reactions, the silane is stripped and re-applied for a new run [96]. This method or its analogs for different coatings may improve the re-usability of bioMEMS devices used in the clinical environment. PCR chip packaging is further discussed in Section 2.4.1.

Microfluidic-based bioMEMS chips are also used for analysis of human blood [98, 99]. The purpose of these devices is to separate specific molecules or cells from the sample for analysis or transfusion. For example, microfluidic devices can be used to remove leukocytes from blood before it is received by the patient to minimize the immune response. Micronics, Inc. has designed a microfluidic device to work in tandem with a readout machine for the purpose of blood (and even urine) analysis [99]. The analysis occurs on a microfluidic card designed to perform all functions related to sample manipulation. The card is placed into a manifold, which allows the card to interact with a computerized pumping and control mechanism. The system can then automatically perform an analysis of the sample included in the card. The packaging for this system must ensure that the card interacts with the control device. Thus, proper precautions have to be taken, especially at the card–manifold interface, where leaking can occur.

Unlike stand-alone bioMEMS devices, packaging schemes for bioMEMS components of medical tools, such as endoscopes [100], are heavily influenced by the tool design and its application. Packaging has to allow for incorporation into the overall apparatus without inhibiting function. Since there are no general device types, no general packaging schemes exist for this application. Rather, each device or chip must be integrated into the complete system with specific considerations for that system.

2.4 General Research for the Life Sciences

Whether a research laboratory is academic or industrial, it is undoubtedly in need of data of higher quality, greater throughput, and potentially newer forms. To address these unmet needs, bioMEMS-based research devices are being developed to offer (i) greater sensitivity and lower noise, (ii) higher throughput, automation, and standardization, and (iii) new functional capabilities.

The packaging schemes for research bioMEMS are diverse and non-standardized. This diversity spans both academic and industrial systems and is due in part to the relative youth of the bioMEMS field. The following subsections will highlight general packaging strategies and considerations for four of the

most actively researched applications: PCR–CE analysis, microarrays, microfluidic large-scale integration (MLSI), and cell culture.

2.4.1 Genetic Analysis via PCR and CE

The analysis of DNA and RNA is fundamental to the life sciences. One of the most widely used methods for such analysis is the polymerase chain reaction (PCR). The PCR process amplifies minute amounts of nucleic acids by subjecting a sample solution to a sequence of temperature cycles. PCR is one of the first applications for which biochips have been designed and where integration of different functionalities is demonstrated and therefore is a good example for packaging considerations (Fig. 8). The advantages of miniaturized PCR include shorter cycle times, reduced reagent consumption, lower fabrication costs, and reduced contamination.

Materials. One of the most important factors to consider in the packaging of PCR chips is material selection. First, the choice of material will dictate its thermal conductivity, which in turn affects the device's temperature cycling rate and overall throughput. Second, optical transparency must be considered, as less transparent materials may limit the device's utility for real-time optical detection of amplification products. Third, the binding affinity of sample molecules to the material will determine whether the inner surface of the reaction chamber requires a passivation coating. Finally, adequate chemical resistance and dimensional stability are required for cleaning steps and temperature cycles, respectively. One should bear in mind that the factors just discussed are by no means an exhaustive list. For example, cost and manufacturability may also be important.

PCR biochips have been manufactured in silicon, glass, and various polymers. A thorough review of PCR chip technology is by Zhang *et al.* [97], which among other things includes a detailed list of materials used in PCR chips. In recent years, polymers have gained the most interest due to the potential for very low cost for both raw materials and fabrication. Notable polymers are embossed polycarbonate [101, 102], embossed PMMA [103], and compression-molded poly(cyclic olefin) [104]. Also, hybrid PCR chips constructed out of two base materials have also been investigated. Silicon–glass [105, 106], polymer–silicon [107], and polymer–glass [108, 109] chips have also been realized for various applications.

Temperature cycling and heating. Another factor in PCR chip packaging is the choice of heating elements. Thin-film electrodes have been integrated on-chip out of platinum [110], other metals, and doped polysilicon [111]. Indium tin oxide (ITO) thin film has also been used due to its optical transparency [112]. The advantage of thin film electrodes is the small thermal mass of the integrated chip and rapid heating. This comes at the cost of a more complex fabrication process and the risk of sample contamination due to electrode degradation.

External heating strategies have also been employed. For example, Peltier-based thermoelectric heating plates have been attached to PCR chips [113–115]. To achieve intimate thermal contact, a layer of material with high thermal conductivity, such as mineral oil [113, 116] or metal [117], is placed at the interface between

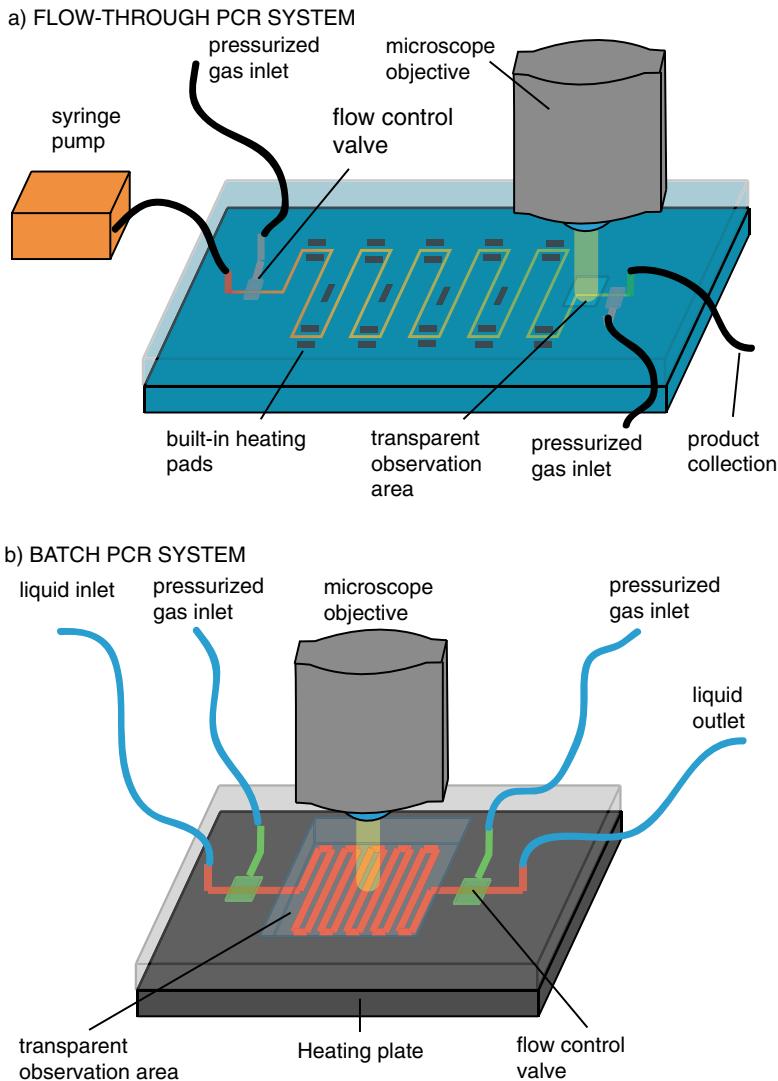


Fig. 8 Examples of different types of PCR systems. **(a)** In a flow-though system, a syringe pump/injector drives the fluid through the device. The fluid is heated in a cyclic fashion for the PCR to occur, and the result, such as an increase in fluorescence, is observed at the end. Packaging for flow-through systems has to allow for connecting the syringe pump, and including inputs for valve control and a source of heat to run the PCR, and a transparent area for microscopy. **(b)** Batch PCR systems do not require an input for a syringe pump, but all other components of the packaging scheme need to be present, such as a heating pad, flow control, and transparent packaging for imaging purposes

the Peltier heater and PCR chip. Peltier heaters are a reliable and modular temperature cycling method. However, their large thermal mass makes it difficult for a single Peltier heater to achieve an adequate ramp rate; as a result, two or more heaters must sometimes be packaged above and below the chip [113, 116, 118]. Other external heaters, such as commercial thin-film resistors [104] and resistive heater coils [119, 120], have been utilized.

PCR heating is also achieved without direct physical contact between the chip and the heat source. This scheme makes the fabrication process simpler. Examples of non-contact heat sources are hot/cold air streams [121, 122], infrared (IR) radiation [123–125], and lasers [126]. When non-contact heating methods are employed, the package must provide good contact and thermal conductivity to the chip. In addition, thermal insulation is also important in IR heating, where heat should not be easily lost to the surroundings.

Flow control. In flow-through PCR systems, where the sample undergoes temperature cycles by flowing past a series of heating zones, a fluid pump must be included in the system. Syringe pumps [127] and peristaltic pumps [128] are examples of flow drivers. In such cases, the pump and PCR chip are separate modules within the entire package. In a few cases, the flow driver has been integrated into the chip itself. Miniaturized peristaltic pumps have been integrated [129]. Also, electrokinetic pumping driven by on-chip electrodes has been used [37]. In these flow delivery schemes, sealing of fluids and interfaces between the chip and the macro-world need to be considered.

CE and integration of other functions. Capillary electrophoresis (CE) benefits from miniaturization through dramatically increased separation efficiency. One noteworthy innovation in CE packaging is a radial array CE system, which utilizes a laser-excited rotary confocal scanner with four color detection channels, enabling the simultaneous analysis of 96 samples [130]. Moreover, PCR and CE are commonly integrated onto a single chip [37, 43, 104, 108, 113, 116, 131, 132]. By packaging PCR and CE together on-chip, sample handling and risk of contamination are minimized, sensitivity increases, and large-scale parallelization becomes possible. In addition, fluids can be driven by electrokinetic means, thereby obviating the need for an external pump.

The integration of sample preparation and PCR–CE onto a single chip is a less developed endeavor. Sample preparation includes cell separation, isolation, washing, and lysis. Attempts at integrating sample preparation steps include on-chip cell capture and lysis using immunomagnetic beads [38], as well as cell separation using dielectrophoresis [100, 133], or DEP coupled with field flow fractionation (DEP-FFF) [134]. Despite such efforts, sample preparation largely remains an off-chip operation; therefore, major opportunities exist for higher speed and lower sample consumption if sample preparation is integrated into a single-chip package.

The integration of detection systems on-chip is in a similarly early stage of development. PCR products are typically detected optically using laser-induced or other fluorescence methods. Optical excitation sources and emission detectors (e.g., mercury lamps and CCD cameras, respectively) are sophisticated and bulky instruments

usually located off-chip. Using these systems, an integrated PCR–CE chip must be optically transparent and have an unobstructed observation path. Efforts to integrate detection on-chip are limited. For instance, photodiodes have been coupled to PCR chambers via integrated optical fibers, with the excitation source remaining off-chip [135].

2.4.2 Microarrays

Microarray technology is considered one of the first realizations of truly high-throughput biological analysis [136]. A microarray is a grid-like arrangement of micrometer-scale spots on a planar substrate. Each spot is a surface-deposited cluster of molecules known as capture probes. The two most common types of microarrays are DNA and protein arrays. The capture probes on DNA microarrays are either single-stranded oligonucleotides or complementary DNA (cDNA), whereas the capture probes on protein microarrays could by definition be one of many different types of proteins, including enzymes, antibodies, peptides, or protein complexes. When a sample solution containing DNA, RNA, or proteins (a.k.a. the target) is incubated with the array, the target molecules bind to the capture probes of the spots on the array to varying degrees. Binding and capturing are facilitated (i) by hybridization between complementary DNA–DNA or DNA–RNA sequences or (ii) by protein–protein or protein–DNA interactions. Fluorescent labeling of the target probes enables the researcher to optically detect the presence and/or interactions of thousands of different target probes simultaneously on a single array.

The emergence of microarrays can be attributed to a serendipitous merger of precision robotics, microelectronics fabrication, biology, and chemistry over a decade ago. Today, a wide variety of microarrays is commercially available from several manufacturers (e.g., Affymetrix Inc. and Nanogen Inc.). Microarrays are used for both basic and applied purposes, including gene expression analysis, mutation analysis, protein function studies, drug development, diagnostics, and forensics. Recent reviews offer a detailed discussion of current and future applications [137–141].

General packaging and fabrication. The typical packaging scheme for a commercial microarray consists of a glass or quartz substrate that is encased by a card-like plastic cartridge. The substrate acts as a rigid support upon which the capture probe array is patterned and subsequently read. To facilitate the attachment of capture probes, the surface of the substrate must be pre-treated or otherwise modified (not discussed here) [136]. Furthermore, the substrate in the optical path is required to be transparent and minimally autofluorescent so as not to interfere with the probes' fluorescence signals.

The capture probe array is patterned using one of many techniques, including an *in situ* photolithographic synthesis or an electric field-mediated attachment (for DNA arrays), as well as robotic spotting or microstamping (for protein arrays) [136, 142].

Cartridge. The cartridge serves a number of purposes: (i) protecting the substrate from mechanical damage, (ii) forming an enclosed reaction chamber and flow path, and (iii) interfacing the substrate with external supporting equipment. An enclosed

reaction chamber is beneficial because it lowers the usage of sample/wash buffers and prevents the evaporation to the external environment; the enclosure also permits straightforward connection to external fluid ports (which in turn link with pumps, sample, and waste) and eases handling by the operator. Rubber septa/gaskets are embedded at the inlets to the cartridge to interface with external fluid connections, forming a leak-free re-sealable fluid port. Also, the cartridge must have a viewing window for optical interrogation of the array in addition to a shape that facilitates easy loading and automated processing in a scanning system.

Supporting infrastructure. The microarray package is supported by a substantial collection of bench-scale equipment. This includes a fluidic module (i.e., pumps and reservoirs for samples, buffers, and waste), scanning system and housing, computer, and equipment for sample preparation (e.g., PCR, fluorescent labeling) [143]. The scanning system contains the excitation source, such as laser or mercury lamp, and the emission detector, such as photomultiplier tube (PMT) or charge-coupled device (CCD).

Integrated packages. The footprint, complexity, and sample usage of the total microarray system can be reduced by miniaturizing and packaging elements of the supporting infrastructure into the cartridge. In one instance, sample preparation, including PCR and target labeling, was integrated with the array on a single chip that is less than the size of a credit card [144]. The chip interfaced with 10 buffer connections, three Peltier heater-coolers, and over 100 gas pressure lines to actuate on-chip valves. Functionality of the chip was demonstrated by detecting mutations from a low copy sample of RNA. Another device was demonstrated to have even greater upstream integration by including cell capture and lysis, as well as PCR and DNA hybridization, on a single chip [38]. In this setup, no external pressure sources were required, as fluid pumping was provided by thermoelectrically actuated on-chip valves. In addition, the hybridization was detected using an integrated electrochemical sensor. Although the chip's resolution and sensitivity are less favorable in comparison to that of optical arrays, the simplicity to detect the presence of bacteria from a sample of whole blood was an advantage.

The integration of new functions can also make dramatic improvements in performance. A notable example is the desire to speed up hybridization. Under normal circumstances, hybridization is driven solely by molecular diffusion of targets to the capture probes, which could take many hours. In one device, the target solution was oscillated by an on-chip pump, thereby introducing convection to the fluid and increasing the hybridization rate many times over [145]. Cavitation microstreaming, which mixes fluids by vibration, has also been integrated on-chip and shows a fivefold increase in hybridization [38, 146]. Because DNA molecules are charged, it is also possible to use electrical field to increase the mass transfer of the target molecules to the surface. The packaging of any of these flow devices would require somewhat more complex schemes, where a pump, a piezo-actuator, or electrical leads need to be accommodated.

Integration outlook. Naturally, higher integration will lead to more interconnect-related challenges in the package, particularly for electrical and fluidic interconnects. In addition, the increasingly commercial nature of the microarray field will

demand fabrication processes and raw materials of lower cost. The use of polymer substrates that are amenable to high-volume manufacturing (i.e., embossing, injection-molding) may bring the end results closer to both goals [147–149]. The integration of new functions and materials, while introducing short-term challenges, will likely lead to higher performance and more economical microarray packages [150].

2.4.3 Microfluidic Large-Scale Integration

Microfluidic large-scale integration (MLSI) is a term describing microchips that possess a dense arrangement of fully integrated valves, channels, and chambers [151]. Invented by researchers in Stephen Quake's group, the enabling technology for MLSI is a miniaturized, elastomeric valve that is fabricated monolithically into the microfluidic network of the device using soft lithography [152] and thermal bonding [153]. Such valves are used to perform fluid handling operations, such as pumping and mixing, and to isolate chambers (or sections) of a device from fluid flow.

Out of all bioMEMS platforms, the layout and packaging of MLSI devices is most analogous to that of microelectronic integrated circuits (ICs). In MLSI, multiple layers of channels, chambers, and valves are stacked upon each other in a grid-like configuration that is reminiscent of the multilayer architecture comprised of metal lines, transistors, and gate electrodes found in ICs. Moreover, fluidic interconnections between MLSI inlets (or outlets) and external tubes is accomplished by rows of densely packed, hollow metal pins positioned in a manner that resemble the bond pads and vias at the edge of ICs (Fig. 9).

The applications of MLSI are diverse. The detection of mRNA from single cells [154], synthesis of precious reagents [155], and realization of a non-fouling bacterial chemostat [156] are a few questions that have been addressed by exploiting the highly integrated nature of MLSI. In addition, commercial MLSI systems are currently used as tools for protein crystallization screening and genetic analysis [157–160].

Packaging and infrastructure. Although MLSI systems are employed in diverse applications, their design layout and packaging are similar. All MLSI devices are fabricated out of the elastomeric polydimethylsiloxane (PDMS) or a fluorinated elastomer that behaves similarly to PDMS [161], which is both amenable to replica molding and exhibits plastic deformability that is essential for valve operation. The features in MLSI devices are typically configured in an array of modules that share common bussing channels, pumps, or specific unit operations. Also, MLSI systems are almost universally packaged to interface with external optical detection instruments. Therefore, glass slides are used as the device's support substrate and inlets/outlets (with accompanying interconnect pins) are located near the edge of the chip to allow for unobstructed viewing.

The off-chip infrastructure consists of a computer-controlled battery of solenoid valves that stand between the chip and a pressure source (i.e., compressed gas tank). Selective actuation of solenoid valves actuates corresponding on-chip valves via

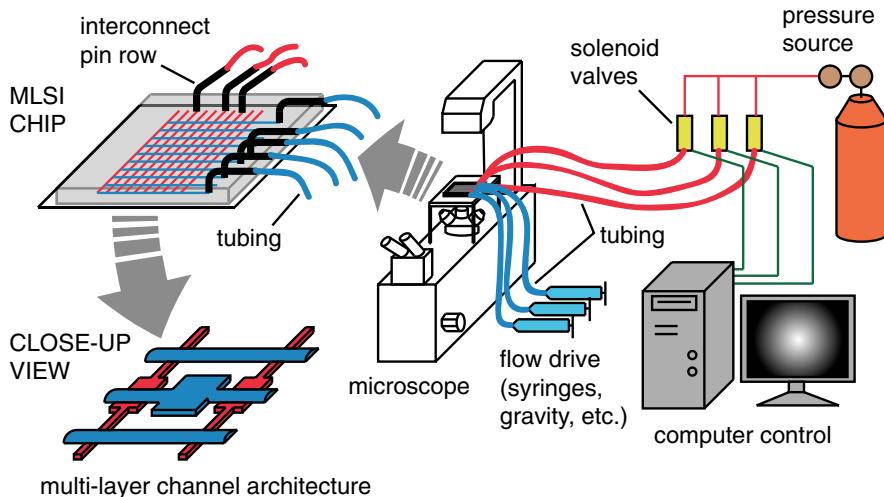


Fig. 9 An illustration of a MLSI system. The MLSI chip is mounted to a microscope stage for optical interrogation. The chip is comprised of an elastomeric block, patterned with a microfluidic network, which is bonded to a glass substrate. Rows of hollow L-shaped pins connect the on-chip network to external tubing. A magnified view shows a portion of the multilayer architecture that enables the construction of monolithic valves and dense arrays. MLSI systems typically require extensive off-chip infrastructure, including solenoid valves, a pressure source, and computer control

pressurization or depressurization. Furthermore, reagents are commonly stored off-chip in vials or syringes.

Outlook. The dense packing and computer control of fluidic components have led to unprecedented parallelization and automation, respectively. However, the extensive off-chip infrastructure, as well as the lack of standardized fabrication and interconnect methods are aspects of ongoing research, which will potentially enable wide applications of MLSI. Strategies to minimize the number of off-chip solenoid valves for a given set of on-chip operations, as well as general design rules, are being advocated [146, 162]. The development of alternative valve actuation techniques could ease implementation by eliminating the need for solenoid valves and bulky compressed gas tanks. For example, bioMEMS chips have been mounted to Braille displays, whose piezoelectric pins slide up and down to actuate on-chip valves [163–165]. Similar to pneumatic monolithic valves, when Braille pins slide upward, they deform the PDMS membrane and seal the flow channel directly above them. The tandem package of a chip and Braille display could substantially improve MLSI device portability; however, it may restrict the optical access of the chip. In addition, standardized interconnect strategies are likely being pursued by commercial manufacturers on a proprietary basis [166, 167]. This could lead to the development and acceptance of industry-wide standards in the near future.

2.4.4 Cell Culture and Assay

One of the most promising applications of bioMEMS and microfluidics is the culture and assay of living cells. Microtechnologies possess unique capabilities that can dramatically increase the resolution and content of data from cellular experiments. In addition, experimental parallelization and throughput can be greatly improved [168–170].

The maintenance of living entities brings with it additional requirements, which in turn affect the packaging of devices. One requirement is the control of temperature at physiological (or experimentally perturbed) levels. Also, the transport of nutrients to the cells must be provided on a continuous or semi-continuous basis to match the cells' metabolic rate. Added to that must be the means to monitor the levels of critical nutrients or other culture conditions, such as pH and dissolved gas concentration. Moreover, biocompatibility of the device interior must be assured through appropriate selection of base material, polymer coatings, or adsorption/functionalization with specific biomolecules.

Integration of functions. One way to accommodate for the additional requirements of cell-based systems is to integrate more functions on-chip. Specifically, many requirements are met by integrating electrical components into the device. Temperature control has been demonstrated by the use of microfabricated indium tin oxide (ITO) electrodes [171, 172]. In addition to being conductive, ITO has the added benefit of being transparent, making it an ideal choice when optical interrogation is required. Also, the on-chip electrochemical generation [173, 174] and measurement [65, 173, 175] of dissolved gas have been demonstrated. Similarly, integrated measurement of pH has been achieved [176]. Moreover, the manipulation and positioning of cells into specific locations within a device has been investigated using dielectrophoresis [177–179]. As a consequence of integrating electrical functions on-chip, the packaging of bioMEMS is complicated by the attachment (or bonding) of electrical leads to bond pads at the edge of the chip. External power supplies and driver electronics are also required.

Integration of other functions can be achieved through MLSI or arraying approaches. The culture of stem cells has been demonstrated in an array of cell chambers, each of which were semi-continuously perfused with the necessary growth medium [180]. Fluid handling within the device is performed by integrated valves and pumps, and the corresponding fluid interconnect is achieved via rows of densely packed hollow pins. The pins are in turn connected to a computer-controlled row of solenoid valves that lead to a compressed gas pressure source. Other cell-based bioMEMS with innovative linear [181, 182] and radial [183] arrays avoid reliance upon integrated pumps, which reduces the on-chip complexity. However, this necessitates the use of additional off-chip equipment, such as syringe pumps or rotary motors. These trade-offs need to be evaluated according to the specific needs of the applications.

Optical interface. One off-chip function that has a great impact upon cell-based device packaging is optical detection. Many assays require very high magnification microscope systems. For instance, the forces exerted by cells as they crawl along

a substrate have been measured by observing minute changes in the location of dots embedded in an elastomeric substrate [184] and bending of elastomeric pillars [185]. To manipulate and position cells optically, high-power laser beams are necessary using optical tweezing techniques [186, 187]. BioMEMS for such applications must be fabricated of optically transparent materials, and just as importantly, their thickness must accommodate the short working distance of the high-magnification objectives (Fig. 10). Occasionally when there is thermal stress(induced by laser power adsorption by the medium or the cell), one must also provide effective means to dissipate the heat so as to reduce cellular damage.

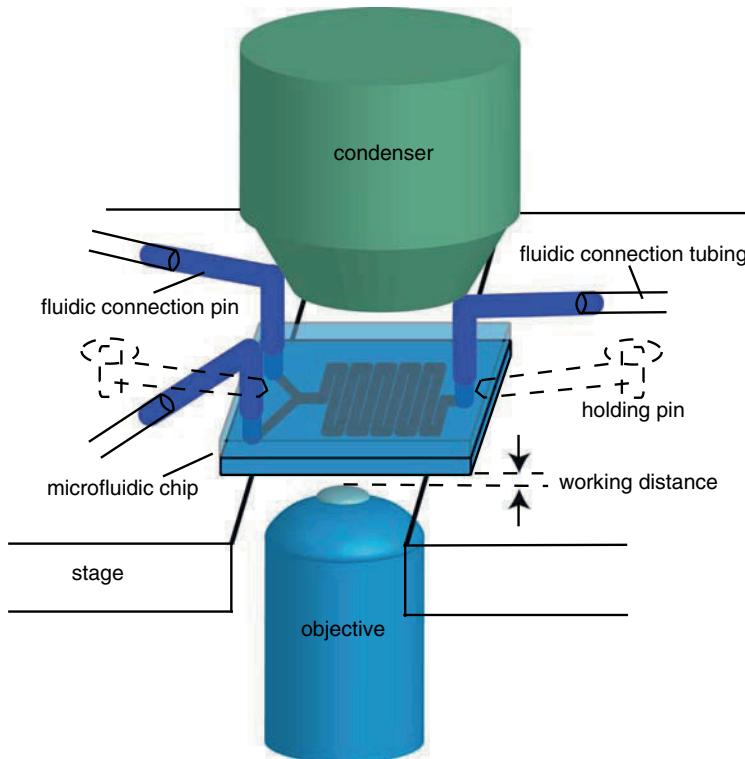


Fig. 10 Optical access to the chip, especially at high magnifications, strongly depends on the chip packing. In transmission mode, both the condenser and the objective need to have access to the chip without hindrance of the fluidic, electrical, and mechanical connectors. The working distance of the lenses is a critical parameter to consider. In fluorescent mode, the objective still needs access to the chip; since no condenser is necessary, the fluidic and electrical connections as well as the mechanical clamping mechanisms can be placed on the open side of the device

3 BioMEMS Chip Interfacing

The previous section covers the specific strategies of packaging in many applications of microtechnologies in biological and medical research and practice. This section reviews two important interfaces in biochip designs: the interface between various components of a microsystem and the chip-to-world interface that is critical to operations and functions of each chip.

3.1 Interfacing On-Chip Components in BioMEMS

The integration of various types of components on a bioMEMS chip can prove challenging if the fabrication methods are incompatible. For example, bioMEMS chips often contain immobilized biomolecules, such as DNA or proteins. The high temperatures required for some of the traditional MEMS fabrication steps easily denature these molecules. For instance, although room temperature deposition has been achieved through ammonia catalysis [188], chemical vapor deposition of SiO₂ and SiN commonly occurs at temperatures in excess of 250°C [189]. Therefore, intermediate protective packaging steps need to be introduced to allow the integration and interfacing of seemingly incompatible components, such as the deposition of biomolecules and CMOS technology.

The order of manufacturing steps becomes part of the packaging scheme if protective packaging or modification of fabrication technology is unavailable or infeasible. The order is especially important in manufacturing of hybrid MEMS chips, such as CMOS/microfluidic chips. CMOS and microfluidic components will generally have to be integrated following separated fabrication steps. Integration of biological functionality such as adhesion of DNA, proteins, or enzymes will have to occur last and in already covered devices.

Lastly, interfacing of different components can be difficult when the materials used for their fabrication are incompatible. For example, common bioMEMS materials, such as PDMS, can be difficult to interface with CMOS technologies. PDMS is a poor substrate for the adhesion of deposited metals due to the low surface energy of PDMS [190, 191]. If PDMS packaging is to be used with CMOS technology, better adhesion between the metallic and polymeric layers must be achieved. Furthermore, due to the elevated temperatures during metal deposition, the PDMS surface can buckle upon cooling and cause uneven surfaces and cracks in the metallic layer. Therefore, sufficient cooling of the PDMS layer during metal deposition is required to minimize this effect [192].

3.1.1 Protecting Biomolecules with Intermediate Packaging Steps and Modification of Manufacturing Methods

Because some fabrication steps can be detrimental to proteins and DNAs, these biomolecules must be protected to allow for their introduction before the last steps

in a fabrication process. Trau *et al.* [193] address this type of complications by passivating the DNA by a protective gold layer. The protective gold layer was deposited by chemical vapor deposition (CVD) while the wafer temperature was not allowed to exceed 45°C, a temperature that is deemed benign to DNA molecules. Parts of the gold layer were etched away to expose desired areas on the wafer for traditional wet etching, metal deposition, and other common techniques. By applying this method, they were able to integrate the deposition process of DNA oligonucleotides and proteins during routine microfabrication. The biomolecules maintained functionality through both wet and dry etching processes [193]. After the microfabrication processes, the molecules were deprotected by a solution of potassium cyanide (KCN) in phosphate buffered saline solution (PBS) [193]. This may have been a successful example of protecting biomolecules; in general, however, it is nontrivial to maintain the bioactivities of molecules such as proteins under microfabrication conditions. Rigorous tests need to be performed to assess the bioactivities of protected and deprotected molecules before a method can be adopted. If the biomolecules can be applied after the completion of the microfabrication, it is perhaps the least complicated method for a reliable deposition/introduction of the biomolecules.

Construction of the final protective enclosure for a device can also cause damage to biomolecules, the choice of which also needs careful considerations. For instance, bonding of a protective casing or membrane to the device may require exposure to high temperatures over a long period of time. Therefore the choice of materials is critical. Novel bonding methods performed at lower temperatures may be better alternatives. One example of such low temperature bonding process is through the use of a UV-curable adhesive as illustrated by Kentch *et al.* [194]. The adhesive is applied in ultra-thin film form on the surface, followed by an alignment of surfaces and curing by UV light. Another formulation of the adhesive, Vitralit, allows for bonding of different combinations of materials, such as SU-8 with glass or cyclic olefin copolymer with poly(methyl methacrylate) (PMMA) [194]. Many other light-curable epoxies, two-part epoxies, and medical-grade epoxies (e.g., surgical glues) also exist for various applications.

Certain obstacles must be overcome before this type of packaging method can be universally applied. For applications where thicker films are required, care must be taken not to cause warping of substrates and thus incomplete adhesion in the longer term. In addition, in some coating applications, due to evaporation of solvent from the adhesive, the solvent may condense inside the structures of a device, such as the channels of a microfluidic device [194]. Therefore, great care must be taken in either designing a device system with high tolerance or selecting adhesives and solvents used. Furthermore, the strength of the material interface can be weakened after prolonged exposure to oxygen or corrosive media and buffer such as PBS [194]. This issue may be overcome by choosing an adhesive appropriate for each particular application.

3.1.2 Order of Manufacturing Steps

Hybrid bioMEMS devices more often require precise sequences of fabrication steps for different types of components. For example, CMOS components may need integration of microfluidic channels that require bulk and surface micromachining steps, which may not be directly and fully compatible with the CMOS processes [195]. Similarly, biomolecules may have to be deposited after the last micromachining step if no protective process is available during the deposition process [193, 196]. This type of requirements calls for an optimized order of fabrication steps and also new technologies that allow for monolithic, sequential (i.e., one component at a time), or a hybrid of two approaches.

Microfluidic channels can be formed on top of a CMOS chip through a direct-write fabrication process [195]. Channel features are patterned by the deposition of an organic ink – a sacrificial material. Following the ink deposition, uncured epoxy is deposited over the channels and then cured. Then the organic ink is removed to form the microfluidic network. The downside of this technique is that the channel dimensions are constrained by the dimensions of the ink deposition nozzle. Commercially available nozzles come with a diameter of 100 μm . While smaller nozzles and more precise manipulation stages are being developed, the designs have to contend with high nozzle fragility and surface tension due to viscosity [195]. Additionally, this process is not easily scaled up for batch processing. Ongoing work may explore other techniques (e.g., using photolithography on photopatternable materials as sacrificial materials) to overcome this drawback.

When interfacing components of different types, such as CMOS and microfluidic systems, interfaces and the different size scales must be considered [197]. CMOS chip sizes tend to be in the millimeter range, while microfluidic chips are generally larger, especially if various functionalities such as pumping and mixing are included on chip. Furthermore, inputs and outputs for fluidic systems require standard (macroscale)-sized pins and tubing, even though on device channels are much smaller. These interfaces with the macro-world usually take up large portions of the chip area.

The connection between the CMOS and microfluidic world may be achieved through effective layout design [197]. The CMOS chip can be selectively exposed to some functional areas on the microfluidic chip, such as a mixing or dilution chamber, or electrical or optical interconnects. All other operations on the fluid occur prior to entry in this area, and pins and tubing for inlets and outlets can remain the standard size. If a barrier between the two components is required for fluidic isolation or electrical insulation, it can be as simple as a layer of Pyrex® glass or a membrane made from one of structural materials, such as PDMS [197].

In some applications, it is also necessary to introduce electrical components to the outside of a chip. Some materials pose adhesion issues. For example, PDMS is commonly used in microfluidics and is the subject of numerous studies, but metal layers do not easily adhere using current processes [191]. This problem can be partially resolved by incorporating carbon black or other charged particles in the polymer

matrix. In addition, PDMS-based composites such as Ag-PDMS, which includes silver microparticles, and C-PDMS, which includes carbon black nanoparticles, may be used to increase conductivity of the matrix such that electrical contacts can be made on the matrix directly [191].

In dealing with biomolecules, such as enzymes, the molecules are either protected [193] or deposited after microfabrication [196], as stated above. Novel methods are yet to be developed to achieve biomolecular functionalization for wafer-level fabrication and high-throughput processes. Zimmermann *et al.* [196] have achieved this by introducing enzymes in a polymer solution into target channels as the last step of the microfabrication process. The polymer was then crosslinked through UV exposure, and the uncrosslinked polymer was washed out. This procedure was performed after the final microfabrication step, after the wafer-level bonding was already completed. Similarly in Herr *et al.* [42], forming gels inside protein analysis chips was also done after the device fabrication was completed. Careful choices of sequences of UV crosslinking, washing, and fluid pumping allowed for a complete integration of sophisticated multistep protein analysis on-chip.

3.2 Interfacing BioMEMS with External Systems

The external interface of a bioMEMS or microfluidic device is where the “chip” ends and the “rest of the world” begins. In many applications, the “chip” is an assembly of two planar substrates, at least one of which containing microfabricated or replica-molded features, that are bonded together to seal a microfluidic flow network. One of the substrates is usually rigid in order to provide mechanical support. All off-chip equipment is characterized as “external” to the chip. The interface between the bioMEMS chip and external systems is also called the “world-to-chip” connections or “interconnects,” which is the equivalent term from the microelectronics field.

Due to similarities in form and fabrication, biochips face similar interconnect issues as traditional ICs and MEMS. For instance, the reliability of bioMEMS is heavily dependent on proper interconnection to external systems; most failure mechanisms are associated with interconnect failure. Also, as chip size continues to decrease and feature density increases, space for interconnect junctions becomes more constrained. This motivates the development of techniques and platforms to smoothly transition from macro- to microscale. The microelectronics industry has well-developed approaches for scaling, such as flip-chip packaging and fully automated assembly. On the other hand, bioMEMS interconnect and scaling approaches are relatively unsophisticated. Possible contributing factors to the lack of sophistication include (i) the bioMEMS field being less well developed as compared to the IC industry, (ii) lack of interest (and funding) in packaging, as the majority of current work is in academic research, (iii) diversity in the applications and hence diversity in the needs of packaging, and (iv) traditional reliance on manual techniques that are prone to variability. As the field grows, more research has been devoted to biochip interconnect, as shown in recent reviews [41, 194, 198].

This section covers five areas: (i) fluidic, (ii) electrical, (iii) optical, (iv) thermal, and (v) mechanical. The following subsections will discuss the basic issues of each regime.

3.2.1 Fluid Interconnect

Fluid interconnect relates to the means by which fluids enter and exit from a biochip. In most cases, the fluids also carry biological entities ranging from proteins (often many kinds of proteins), to DNA and RNA, to cells and even organisms. It is perhaps the greatest differentiating issue between bioMEMS interconnect strategies and those of IC/MEMS. Proper fluid interconnect is critical to averting numerous failure modes, including leakage, infiltration (or nucleation) of gas bubbles, and introduction of contaminants. Achieving hermetic seal and having little or no dead volume are the ultimate goals of fluid interconnect. In addition, robust and rapid assembly of interconnects (i.e., alignment, attachment, and bonding) is an active field of research [199–202]. Fluid interconnect is most commonly achieved by the use of pins and ports; in practice, both custom and commercial techniques are employed.

Interconnect pins. Interconnect pins are hollow metal tubes, usually bent 90° into an “L” shape. One end of the pin is inserted vertically into an inlet port on the chip, while the other end points horizontally toward the exterior of the chip. Plastic tubing is attached to the horizontal end of the pin by friction fit. From the pin, the tubing extends to off-chip components, such as reagent sources, pumps, or waste reservoirs. The vertical end of the pin also fits by friction into the chip’s inlet port. Consequently, the substrate into which the pin is inserted must be elastomeric or otherwise plastically deformable. The inlet port on the chip is slightly higher in gauge (smaller in diameter) to provide for a tight, leak-free fit when the pin is inserted.

Pins are a popular means for interconnect in biochips fabricated in PDMS, given its favorable elastomeric properties. The most striking demonstration of this is with MLSI devices. Rows of over 50 pins, spaced by approximately one millimeter, have been packed onto the edge of MLSI chips less than 3 in. on a side [180]. At such densities, space becomes constrained, and pin placement becomes a critical factor in chip design. Another design constraint is substrate thickness. The need for a stable and secure fit between an inlet and pin requires that the thickness of the elastomer substrate be roughly a few millimeters. Therefore, thinner device designs may preclude pins.

Progress in pin technology will be hindered without the evolution of standardized pin sizes and placement templates. Also, current methods for pin insertion and hole punching (for inlet and outlet ports) are painstakingly labor intensive, causing high variation and low throughput. The development of automated assembly methods may improve reliability and allow for more ambitious chip designs. Such automation might already exist within companies that manufacture PDMS-based biochips.

Connector ports. Another fluid interconnect approach is to use connector ports. A connector port usually comes in the form of a small cylindrical assembly, where one end is bonded to the inlet of the chip and the other end acts as a fitting for external tubing. The bond between the substrate and connector port is achieved using epoxy or similar adhesive.

Connector ports are well suited for attachment to hard substrates, such as glass and silicon. Furthermore, the strong bond between the substrate and port makes it possible for the interconnection point to withstand pressures as high as 100 bars. In contrast, interconnect pins are designed for soft, elastomeric substrates, and the nature of their attachment to the substrate (i.e., friction fit) limits the operating pressure to less than 10 bars.

Connector ports are a flexible interconnect platform due in part to the variety of materials (polymers, metals, etc.) and commercial manufacturers from which they are available (e.g., Upchurch Scientific). However, the assembly process is complex. Of the major procedures (i) holes must first be drilled (or etched) through a hard substrate, (ii) adhesive must be dispensed, (iii) the port must be carefully aligned and attached, and (iv) time must be allotted for the adhesive to cure.

Commercial and “plug-and-play” interconnect. The growth of commercial bioMEMS is giving rise to a new set of approaches for fluid interconnects. Commercial bioMEMS are typically packaged as a cartridge that is inserted into a bench-scale scanner box. Examples of such configurations are the Agilent Lab-on-a-Chip, Affymetrix GeneChip[®], and Fluidigm Topaz[®] system. In each of these systems, rubber gaskets and septa around the chip’s inlet holes are used to seal against a chuck (or frame) when the cartridge is pressed into the scanner box. This reduces or eliminates manual attachment of fluid interconnect leads, resembling a so-called plug-and-play approach.

Others are developing the plug-and-play concept as a means to build custom microfluidic systems [203–205]. Custom plug-and-play kits consist of a set of bioMEMS building blocks, each encasing an individual feature or unit operation, which can then be assembled to perform a desired set of functions. The plug-and-play approach will inevitably demand higher standardization, which must exist either for how the building blocks interface with each other or for how they interface with a generalized assembly template. One idea of an assembly template is the microfluidic “breadboard” [206]. The breadboard contains a pattern of bussing channels, which fluidically link different building blocks together based on where each is placed on the board (Fig. 11.).

Although plug-and-play and breadboarding have been proven as concepts, the standardization of fluid interconnect is still in its infancy. Fundamental challenges to standardization persist, such as (i) the need to uniquely tailor surface chemistries for each application and (ii) the question of how to combine electrical and optical components into assembly templates (i.e., multifunctional “breadboards”). By addressing such challenges, the bioMEMS community will move closer to a level of accessibility and scalability comparable to that which is enjoyed by the microelectronics field today.

3.2.2 Electrical Interconnect

Alongside fluid interconnect, electrical interconnect is the most common assembly issue of bioMEMS. Fortunately, their form and fabrication makes them well suited to the same approaches used in IC/MEMS interconnect. Given the well-developed

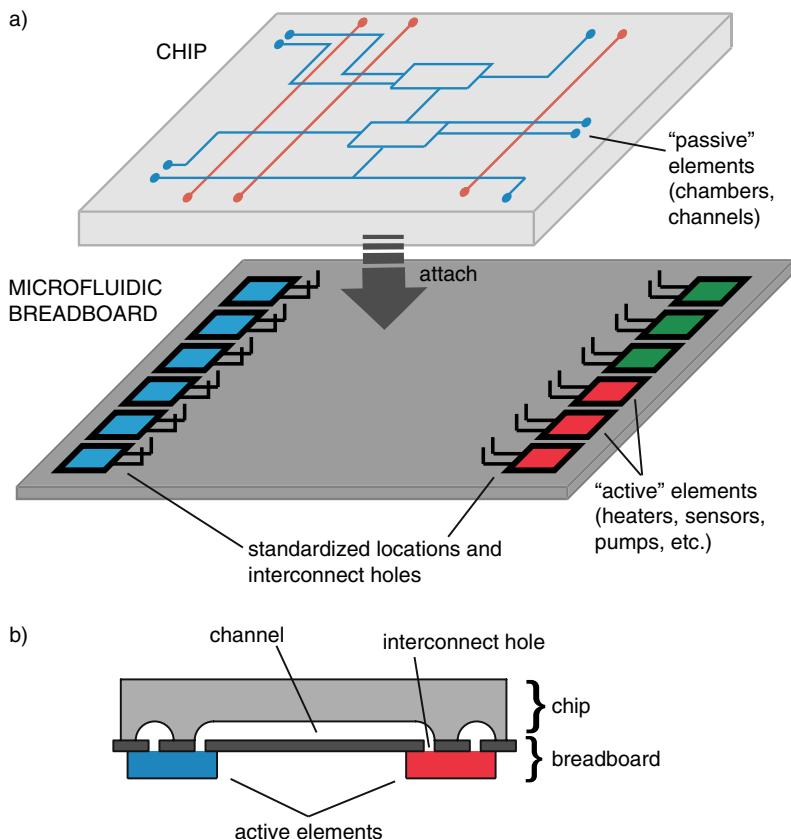


Fig. 11 Scheme of a microfluidic breadboard. **(a)** A microfluidic chip is attached to the breadboard, which possesses an array of active functional elements that are fabricated by standard methods (i.e., CMOS/IC processing) and placed in standardized locations [206]. The on-chip microfluidic network is customized to link with specific elements on the breadboard depending on application. **(b)** Side-view of a channel linking two active elements on a breadboard

state of the microelectronics interconnect field, a thorough discourse will not be provided here, but instead only a few fundamental issues and noteworthy examples will be highlighted. For further information, the reader is encouraged to examine more general reviews [207–209].

In bioMEMS, electrical components are used for a wide variety of purposes, such as transductive readout (electrical or mechanical), resistive heating, electrochemical reactions (providing potential or current), and electrokinetically driven fluid flow. Some applications need simple electrodes and metal wires can serve the purpose. For others, wire bonding to the chip to interface with the on-chip electrodes is used. Similar to ICs and MEMS, the electrodes for bioMEMS “fan out” and expand into bond pads at the edge of the chip. Electrical interconnect is an essential design factor in bioMEMS with a high density of electrical components, such as (i) the

stimulation (and measurement) of action potentials in living neural circuits [210], (ii) the manipulation and sorting of cells by dielectrophoresis [100, 177], and the movement of droplets via electrowetting techniques in digital microfluidics [211]. Each application requires a complex array of electrodes that is coupled to a fluid channel or chamber.

Microelectronic assembly techniques and configurations have been applied to bioMEMS on a limited basis. For instance, flip-chip techniques have been utilized to unite microfluidic to electronic substrates [120, 212, 213]. However, the sensitivity of many bioMEMS to heat might limit the annealing temperature for solder bumps or other contact points. Additionally, some bioMEMS are directly attached to printed circuit boards (PCBs) [214]. The disadvantage of the PCB approach is that by mounting a PCB to the back surface of the chip, it precludes the use of transmissive optical interrogation.

Another noteworthy issue is corrosion. Although interconnect corrosion is a well-known concern for IC/MEMS, the problem is more pronounced in bioMEMS because of the “wetware.” Buffers and reagents used in bioMEMS often contain salts. These fluids are prone to contaminate electrical leads and bond pads while preparing and running experiments. Furthermore, salt solutions commonly flow directly over electrodes within the device’s fluid network. Approaches to passivate electrical components are similar to those of IC/MEMS, although corrosion prevention will also depend heavily on better fluid interconnect technologies.

3.2.3 Optical Interconnect

Optical interconnect is typically achieved through non-contact methods, such as excitation by an off-chip laser or observation through a microscope objective. In such cases, the main concerns are (i) the transparency of the chip’s substrate to the wavelengths of interest, (ii) providing an unobstructed path (or window) for excitation and emission (for fluorescence measurements), (iii) distortion of light as it travels through the chip, and (iv) autofluorescence of the substrate material.

The third and fourth concerns are particularly relevant to PDMS, one of the most popular base materials for biochips. Optical distortion may occur in zones of non-uniform crosslink density. This is in part caused by inadequate mixing of the PDMS prepolymer and crosslinking agent before it is cured. Distortion may also be caused by the insertion of fluidic interconnect pins, which generate non-uniform stress by compression in the polymer. To avoid such distortion, the observation zone must be located a safe distance from the stress fields. Moreover, some polymers (including PDMS and PMMA) are known to be autofluorescent [215], which may disqualify them from fluorescence applications that demand very little background signal.

Contact methods for optical interconnect also exist. For example, on-chip waveguides are often coupled to optical fibers that guide light to off-chip detectors [216]. Also, the concept of an optical backplane, analogous to electrical backplanes (or breadboards), has been investigated [217]. Care must be taken to align and secure the optical fibers to the on-chip components.

3.2.4 Thermal Interconnect

Thermal interconnect is an important issue for many bioMEMS applications, particularly PCR and cell culture. Not only could heat be generated from the application of electrical field (or magnetic field), which needs to be dissipated promptly, but maintaining temperature for biological entities (such as proteins and cells) is critical. Similar to optical interconnect, thermal interconnect can often be achieved by non-contact means. Non-contact heating mechanisms for PCR include infrared radiation [123–125] and convective heating by fans [121, 122]. For cell culture, convective heating via fans or environmental chambers is the common non-contact method.

Heating by contact can be achieved by using Peltier heaters (and coolers) and, if the device is mounted to a microscope, stage and objective heaters. To facilitate intimate thermal contact between a Peltier surface and a biochip, intermediate layers of mineral oil [113, 116] and thin metal [117] have been used.

An important factor that determines the success of the thermal interconnection is the thermal mass of the package. As the chip-heater package becomes smaller, heat can be added and removed more quickly. For applications that require high temporal resolution in heating cycles, packages with minimal thermal mass are the most viable options. For instance, rapid cycling times increase PCR throughput; hence, low-mass IR techniques have been developed [123, 125]. However, when the temperature must be sustained for long periods (i.e., for cell culture), thermal mass is not as important, and perhaps a larger thermal mass is preferred to dampen potential temperature fluctuation.

Other factors to consider for thermal interconnect are (i) material stability (dimensional and chemical) at elevated temperatures and (ii) differences in the coefficient of thermal expansion (CTE) between substrates in hybrid bioMEMS.

Heating via integrated electrodes deserves a brief mention. A key advantage of integrated electrodes is their ability to heat areas with high spatial resolution. This is particularly important for flow-through PCR, where hot and cold zones must be defined in close proximity [37]. Standard heaters cannot provide such precise local heating. An additional advantage is low thermal mass. The thin-film electrodes add no additional mass to the package, which is advantageous for high temporal resolution applications as well.

Although heating and maintaining temperature above room temperature is the main discussion here, integrated cooling systems and Peltier off-chip coolers are some technologies that are used to lower the temperature on-chip for some biological applications [218, 219]. The design considerations are similar to the discussion above on heating schemes.

3.2.5 Mechanical Interconnect

Mechanical interconnect deals with the manner in which a bioMEMS device relates to solid objects around it. First, bioMEMS must be protected from physical breakage. This is especially important for portable bioMEMS, which cannot be handled as carefully in the field as they are in the laboratory. Mechanical protection is provided

either by packaging the chip within a protective enclosure or by fabricating it out of mechanically tough materials. Polymers are the materials of choice due to their resilience, as well as their low cost.

Second, bioMEMS must have the appropriate shape and form to fit into external systems. For instance, bioMEMS are often designed to be attached to standard-sized microscope slides (or cover slips). Therefore, they can be easily mounted to most microscope systems. The same consideration follows for other standardized chucks or frames found on other external systems.

4 Biocompatibility of BioMEMS

Besides the types of interconnects for biochips, an additional consideration is the biocompatibility of the chip and the packaging materials. The criteria for biocompatibility largely depend on the context and the applications the chips are intended for. The major concern for bioanalytical chips is to avoid contamination, so reducing foreign molecules that can get into the analytes or react with analytes is important. On the other hand, for cell culture chips and implant devices, avoiding non-physiological response from the cells or the body is important. To render devices biocompatible, however one defines biocompatibility, the choices of materials and surface modifications are critical. This section aims to review these two important aspects of chip design and packaging schemes for a series of common materials and processes in bioMEMS.

4.1 Biocompatibility of Fabrication and Packaging Materials

The biocompatibility of bulk materials used in bioMEMS and their packaging is essential to clinical and commercial success. One such class of devices is implants and cell culture systems. For implants, the primary concerns are whether the material is toxic or elicits an immune response. In cell culture systems, toxicity is one concern, but additionally it is important to consider whether the material causes an unexpected physiological response that confounds the analysis of experimental data. The lack of biocompatibility with the bulk material could be driven by its inherent chemistry or by the presence of impurities that may leach out of the bulk (usually residues from the manufacturing process). A noteworthy point is that these issues could become exacerbated in bioMEMS cell culture systems because of a substantial increase in the surface-to-volume ratio (SVR) as size is reduced from the macro- to the microscale.

Much of what is known about material biocompatibility has been learned in the medical implant and drug delivery fields. Therefore, a large body of knowledge already exists for certain well-utilized metals, ceramics, glasses, and polymers, which could be applied here. Included below is a brief summary of the biocompatibility of some of the most popular bioMEMS materials.

Glass and silicon. Glass and silicon are the archetypical bulk materials for bioMEMS due to their compatibility with IC/MEMS-based microfabrication and micromachining processes. Although the biocompatibility of glass is well established, that of silicon is still being investigated. In a short-term (<21 day) study, a comparison of the inflammatory and wound healing responses to silicon implanted in rodents showed no difference between implanted and non-implanted cohorts [220]. Silicon has proven itself as a base material for selected bioMEMS applications, including cell culture with multiparametric electrical monitoring [221] and neural prosthetic implantation [222]. In the future, the long-term biocompatibility of silicon may be proven by clinical studies of implanted drug delivery devices [62, 223] and transdermal silicon microneedles [224].

PDMS and SU-8. PDMS is widely used in medical applications [225, 226], and it is emerging as one of the materials of choice for cell culture in bioMEMS. Cells cultured on various compositions/formulations of PDMS have shown growth rates similar to that of polystyrene culture dishes (with some dependency on cell types) [227]. However, the high SVR of bioMEMS culture chambers, combined with the possible leaching of residual molecules (i.e., uncrosslinked oligomer, low molecular unreactive dimethylsiloxane cyclools, solvents, and platinum catalyst) from PDMS, could significantly reduce the viability of mammalian cells in PDMS devices [228]. Fortunately, cell viability can be significantly improved when the devices are pre-washed with several cleaning solvents and sterilized [228, 229]. As for SU-8, its use as a bioMEMS material is far less developed than that of PDMS, although its biocompatibility has been demonstrated in short-term clinical studies [220].

4.2 Surface Modification

When there are limited choices of device materials, surface treatment is most commonly used to render materials more biocompatible and to minimize biofouling, for instance, on implants and in microfluidic channels. This type of surface modifications can be considered part of the packaging scheme since it is an integral part of the post-fabrication process to deliver the functionality and usability of the chips. Additionally, surface modification can serve to protect devices from the environment.

Implantation of foreign material could lead to an immune response from the body, the result of which includes, but is not limited to, inflammation, scarring, fibrous tissue buildup, and atypical cell growth. A properly passivated or functionalized surface can reduce or effectively eliminate unwanted immune response. Furthermore, surface modification can facilitate not only a passive response but also improved adaptation of the device to best suit the environment. For implants, this could lead to a controlled growth of cells around the implant leading to faster healing after insertion and normal tissue formation. Moreover, some proper functionality of the device can be maintained over a longer period of time. In addition, surface treatment can be a useful tool for protecting the device from its corrosive surroundings. Common issues such as fouling and device isolation can thus be addressed.

This following section will provide an overview of existing and emerging techniques which could eventually be used to mitigate the adverse effects of the interaction between the biochip and its environment and enhance the device functionality through active manipulation of that interaction.

4.2.1 Basic Principles of Biological Surface Chemistry – Biorecognition

Although modern bioMEMS devices can be fabricated with micrometer and nanometer features and can be manufactured small enough to fit into a hypodermic needle, the basic device–environment interaction occurs at the molecular level. Thus, for a biological environment, surface treatment schemes have to be designed with biorecognition in mind. Biorecognition is the highly developed ability of biological systems to recognize specially designed features on the molecular scale, whether it is through topographic architecture, chemical architecture, or dynamic properties [230]. This feature of biological systems can determine the rate of adsorption of proteins and other biomolecules onto a surface, the ability of cells to cover a specific surface, and the stability of the interactions.

Furthermore, biorecognition on the nanoscale can affect interactions on much larger length scales [230]. Due to mass transport properties such as diffusivity, the order of arrival of different types of molecules at the device surface is different. The first molecules to arrive are water molecules, followed by proteins and similar types of organic molecules. In the case of implants, the effect of the surface type, along with the arrival order, can influence the ability of cells to cover a device surface [230]. In particular, the initial water shell exhibits different properties based on the type of material with which it interacts. As such, water can form hydration shells, which through interaction with the hydration shells of biomolecules may determine whether proteins denature, how they are oriented, and how effective they can cover the surface [230]. Cells require an extracellular matrix of proteins (ECM) to successfully adhere to a surface and proliferate. If the ECM is altered due to interaction with the surface, the overall cell–surface interaction is also modified. Thus, awareness and understanding of biorecognition becomes imperative for successful design of surfaces.

4.2.2 Surface Treatments for Common BioMEMS Materials

BioMEMS fabrication techniques differ from traditional MEMS microfabrication techniques to various extents, but they share common elements, such as the most typical materials. One such material for IC and CMOS fabrication is silicon; silicon is generally regarded as non-toxic and inert and thus a good material for certain implant and cell culture applications.

Along with silicon, SiO_2 and SiN are used in the fabrication process and often as a protective layer for the device. A layer of the oxide or nitride can be deposited by chemical vapor depositions [189]. These materials do not leech into the environment and under certain conditions are also deemed non-toxic and biocompatible. Other

common, non-polymeric materials include metals such as platinum, titanium oxides (TiO_2), gold, and alloys, as well as ceramics.

Polymer-based devices have become increasingly popular due to their simple and relatively inexpensive manufacturing processes. Polymers can be used not only as the structural material of the bioMEMS device but also as a protective material for silicon-based devices. PDMS, Parylene-C, and PMMA are commonly used polymers as described earlier. They are easy to process, have physical properties (e.g., stiffness) that can be controlled or modified, and are largely non-toxic. Polymer-based photoresists may also be used as structural materials or protective films, but often the concerns are costs and solvents that cannot be completely removed in the process. SU-8, an epoxy-based negative resist, is becoming increasingly popular due to its low toxicity.

It is important to note the limitations of materials' applicability. Many chips are manufactured using a combination of materials, each requiring a different treatment. For example, traditional CMOS components may be combined with polymer-based device components in a non-monolithic process. Therefore, it is important to understand the complexity of a device and choose compatible treatment techniques that will achieve the desired effect.

Optimizing the surface of PDMS. The surface of PDMS is naturally hydrophobic and thus may be unsuitable for certain biological applications. Furthermore, PDMS surfaces can adsorb certain biological molecules, but not in a completely controllable fashion, i.e., the surface property of PDMS may vary depending on the processing conditions. To minimize cell adhesion (e.g., for implant applications for instance) or to better control surface properties, several techniques have been developed to make the surface hydrophilic or to limit adsorption of biomolecules, such as plasma processing, surfactants, polyelectrolyte multilayers, and graft polymerization through radiation exposure or cerium (IV) catalysis [231].

Exposing PDMS to plasma, discharge [231, 232] can alter its surface properties. Through the use of oxygen plasma, the surface of the PDMS in direct contact with the plasma becomes ionized and forms excited species such as ions and radicals. Although the reactions are complex and the mechanisms are not fully understood, the result of the plasma treatment is the formation of a transient silica-like hydrophilic layer of SiO_x with high oxygen content [231]. This renders the surface much more hydrophilic than the original surface.

Generating reactive elements on the surface of PDMS is also possible through exposure to radiation [233]. Free radical formation on the surface as a result of exposure to UV light provides the reactive component required to graft polymerize acrylic acid and other monomers, such as acrylamide (AM), hydroxyethyl acrylate (HEA), PEG, and DMA, onto the surface. A layer of polymer can thus be grafted in the exposed area leading to the desired functionality. Graft polymerization can also be achieved through cerium (IV) catalytic reactions, although the process condition can be more complex [231, 234].

Another simple surface modification to make it hydrophilic and charged is through exposure to surfactants below the critical micelle concentrations [231]. The concentration must be low enough to prevent micelle formation yet sufficiently high

enough to ensure coating of the PDMS surface. The process facilitates the hydrophobic tail of the surfactant attaching to the surface, leaving the hydrophilic end exposed to the environment. Various surfactants have been used, such as sodium dodecyl sulfate [235] and 2-morpholinoethanesulfonic acid [236]. Because surfactants could be damaging to cell membranes, one needs to be judicious in choosing and using surfactants when the chips are used for cell-based experiments.

Exposing a PDMS surface to cationic and anionic polyelectrolyte solutions in an alternating fashion forms a polyelectrolyte multilayer [231]. Although time consuming (because many layers have to be deposited with many rinses in between each deposition), layer-by-layer assembly usually gives robust films. Deposition of a film of this type is also possible with small organic molecules, polymers, natural proteins, inorganic clusters, clay particles, and colloids [237]. The adsorption of polyelectrolytes to a solid surface occurs in solution, in which case substrate size and topology should not play a role in coating efficiency.

Although useful for short-term modifications of surface properties, some of these approaches do not always fully address the issue of long-term stable hydrophilicity and hydrophobic recovery [238]. For example, over time, the silica-like layer on PDMS that forms upon exposure to plasma or radiation may revert back to hydrophobic surfaces due to diffusion of the hydrophilic species from the surface into the bulk, and potentially also crack and deform due to mismatch of physical properties (stiffness and flexibility) from the supporting PDMS layer. Long-term stable films may be achieved through electrostatic layer-by-layer self-assembly of polyelectrolytes with chemical crosslinking [238]. This method uses the same basic principle as the formation of polyelectrolyte multilayers, but introduces crosslinking between the layers for additional stability. It is also possible to covalently bond PEG to the layer-by-layer modified surface for additional protein adsorption resistance [238].

It is important to note that all of the above techniques need to be incorporated in the appropriate order during the manufacturing process. For example, if a device contains biological molecules for analytical purposes, exposure to oxygen plasma without protection may result in unwanted reactions. Thus, the surface of the packaging should be modified either prior to loading with biological molecules or those areas of the chip need to be isolated from the plasma. Exposure to UV light and electrolyte solutions may also render certain components of a device non-functional, and thus they should only be used with proper protection for vulnerable elements.

Modification of silicon surfaces. Using silicon in the biological environment presents several challenges. One of the major issues is biofouling, leading to interference and the suboptimal operation of the device [239]. Silicon surface is hydrophobic and hence potentially attractive to protein deposition. Surface modification provides a convenient way to limit or direct adsorption of biological molecules such as proteins on, and in effect the interactions of cells to, the silicon surface. Other than depositing or growing SiO_2 and SiN_x layers, the most common approach to treating silicon surfaces is to form protective films of other materials.

Plasma polymerization has been used to polymerize a layer of tetraethylene glycol dimethyl ether (tetraglyme) onto a silicon surface [240]. The resulting

protective coating is a PEG-like polymer thin film; silicon surfaces thus treated exhibited less protein adsorption and less fouling. One possible drawback is that coating of high aspect ratio can suffer from non-uniformities [239] and can be as thin as 10 nm [240].

The formation of the protective coating can be achieved by exposing a cleaned silicon surface to tetraglyme vapor in a plasma chamber. The procedure is also applicable to SiO₂ and SiN surfaces which may be desirable for electrical passivation and compatible with photoresist patterned surfaces and may conceivably be used on metallic surfaces [239]. In addition, the film does not interfere with electrical conductivity of metallic electrodes used for neural sensing. As such, it can serve as an optimal protective coating for electrodes used to monitor brain activity and many other applications. However, the formation process is not fully compatible with biological molecules and could destroy biologically active chip components. Hence, as a surface treatment technique, it should only be applied after the device has been completely sealed, preventing vapor access to destructible elements, or prior to the introduction of biologically functional components.

Acrylic acid can also be successfully grafted onto silicon through plasma polymerization [241]. Polymeric acrylic acid coatings lead to higher fluid velocity along the coated surface. Controlling flow and velocity is particularly useful for devices analyzing or interacting with biological fluids. Higher velocity leads to lower biomolecule adsorption, which in the case of blood can lead to less coagulation. Lowering adsorption on the surface can thus lead to less coagulation. Devices aimed at implantation in the cardiovascular system could greatly benefit from understanding this mechanism.

Facilitating controlled cell adhesion and growth is imperative for implantable devices, such as the multi-electrode arrays. Ultrananocrystalline diamond (UNCD) films are emerging as a new type of protective film. This type of films is chemically and electrically inert, has a low friction coefficient and high wear resistance, provides a good substrate for cell attachment, and can be used with a silicon surface as the coating base [242]. Furthermore, the film formation technique is compatible with some common MEMS fabrication processes. The deposition of the film on silicon wafers occurs via microwave plasma chemical vapor deposition, in an argon (99%) and methane (1%) atmosphere [242]. This process leads to C₂ dimer formation, which supplies nucleation species for the entire film. The deposition occurs at a temperature of 400–800°C and is then followed by several cleaning steps before the surface is ready to be dried and used. Compared to pure silicon and platinum surfaces, HeLa, PC12, and MC3T3 cells adhered much better to the UNCD surface. Furthermore, cells attached to this surface exhibited lowest cell rounding and highest cell spreading, which seems to indicate that the UNCD surface is the most biocompatible surface of the three [242].

The drawback of this technique is the complexity of the process and cost, making it unsuitable for single-use applications, for example. Additionally, the high temperature required to form the film and the aggressive cleaning solutions used in the process (e.g., piranha solution) may not be compatible with other processes. The temperature range at which the deposition occurs affects not only

biological molecules or reagents but would also melt some metals commonly used in CMOS processes. Although it is possible to protect different components during the processing, it would lead to additional manufacturing steps and higher cost. Nevertheless, the UNCD approach is perhaps one of the more reliable ways to modify silicon surface and can be useful when processes are well designed.

Control over the thickness of protective films is a process feature desired for every manufacturing process. Atomic layer deposition (ALD) is a process capable of depositing ultra-thin conformal films with atomic-level thickness control [243]. Via this method, materials, such as Al₂O₃, SiC, and biocompatible materials (e.g., TiO₂), may be deposited on devices to form protective films. Similar in principle to polyelectrolyte multilayers, the ALD technique relies on a series of reactions, each of which terminates after one step, between the gaseous precursors and the solid substrate. Once a precursor molecule reacts with the surface, the particular site is unavailable for further reaction. Furthermore, the molecule cannot react with itself, limiting the film thickness to just one layer. The next layer is applied to regenerate the ability of the desired molecule to adsorb. This cycle can be repeated to obtain the desired thickness [243]. The deposition of Al₂O₃ using this technique can be performed at 177°C. Although this temperature is high enough to destroy most biomolecules, it is low enough for the process to be compatible with most CMOS processes. Thus, (i) if insulation of biological chip components can be achieved or (ii) if the process can be applied prior to the introduction of such components, this technique could be used to form protective films of controlled thickness on silicon devices. The advantages of this approach are a low temperatures required (compared to CVD), the ability to form several layers, unlike self-assembling monolayer techniques [243], and most importantly the high degree of conformity of the deposited films [243].

Modification of SU-8 surfaces. Due to its rigidity, chemical stability, low apparent toxicity, and transparency to visible light, SU-8 has been used as a structural component in biochip fabrication [244]. However, since the surface of SU-8 is hydrophobic, it faces similar challenges as other materials for controlling protein deposition and cell interactions.

Modification of SU-8 surfaces has been achieved through methods similar to PDMS surface modification. PEG graft polymerization has been achieved through exposure to radiation [245], and hydrophilicity has been achieved through treatment with ethanolamine [246] or grafting of amine groups to the surface via a CVD treatment [247]. It is also possible to apply the cerium (IV)-based surface graft polymerization process to modify SU-8 surfaces. This can be performed at room temperature and can be used to graft a variety of monomers onto the SU-8 [244]. The method relies on the opening of residual epoxide rings using a mixture of nitric acid and cerium (IV) ammonium nitrate, followed by an incubation of the surface in the same solution containing monomer, after which the grafting reaction is allowed to occur at room temperature. Using this method, polyacrylic acid can be grafted onto the surface to enhance cell attachment [244]. Although this methodology is simple and straightforward, care must be taken to protect vulnerable components of the bioMEMS chip from the nitric acid.

4.2.3 Modification of Surface Topography for Improved and Directed Cell Attachment and Growth

Although the mechanism is not fully understood yet, surface topography can affect cell attachment, growth, and spreading. Thus, topographic features (in conjunction with other surface treatment technologies) could be used to promote or limit cell growth on specific components of a bioMEMS chip, improve integration of implants, and even improve on the functionality of cell-based sensors.

To achieve this goal, a technique for selectively patterning ECM on a surface has been suggested [248]. Microcontact printing techniques based on self-assembling monolayer technology on silicon wafers have been used to fabricate substrates with well-defined regions of ECM [248]. Using this technique and by controlling geometry [248], it is possible to coordinate cell attachment density, at least in the initial phase. Cells, however, are able to migrate onto regions not covered by the ECM after a while, possibly due to protein adsorption from the solution or by secretion of ECM components by the cells themselves [248]. Therefore, it is imperative to passivate and limit the adsorption of protein to areas not covered by the patterned ECM.

Introducing various microtopography features such as ridges of different depth, spacing, and width can influence cell adhesion and spreading [249]. Through photoblation with an excimer laser and the use of a projection lithography/etching technique, well-defined surface features down to submicron scale can be fabricated. The desired pattern is ablated from a polyethylene terephthalate (PET) surface on a fully automated, computer-controlled stage. Although the original experiment tested PET and polyvinyl alcohol (PVA), the methodology is also applicable to mineral or organic surfaces, since it eliminates the need for a photoresist [249]. The laser technology can also be used with temperature-sensitive materials, since heating occurs only in the area targeted by the laser and the process occurs rapidly – 1 s/cm^2 – thus minimizing heat flux to surround areas [250].

Although no theory exists that thoroughly describes the effects of topography on cell growth and orientation, empirical results can be used to determine the type of groove dimensions best suited for the surface treatment [249]. For example, smoother transitions in surface topography appear to promote cell growth [43, 249]. However, this effect was highly dependent on cell type, cell size, and the surface chemistry, such as formation of carbon-rich species at the ablated surface. Nevertheless, the technology provides a valuable blueprint for future research and may soon find practical use.

4.3 Other Considerations

It can be pointed out that miniaturization of devices in some sense may make implantable devices more biocompatible. For example, miniaturized implants can be inserted using needles and catheters instead of using traditional surgical methods [70]. Small wounds elicit smaller immune response. However, there are cases where miniaturization is not the solution. Szarowksi *et al.* [222] tested the response of brain tissue to implanted micromachined silicon devices. Following implantation,

cell density increased in the surrounding area, as many others have previously discovered with implanted devices not free to follow the movement of tissue [251–253]. Furthermore, it was determined that the long-term brain response to devices of various shapes and with different degrees of surface smoothness and corner roundness was effectively the same and led to device encapsulation through formation of dense tissue around the device [222]. For this specific case, since the response was independent of size, techniques other than miniaturization will have to be utilized to improve the device. In general, one would need to consider multiple facets of a packaging scheme in order to determine the optimal strategy to render devices biocompatible.

5 Conclusion

There is no universal packaging scheme that suits all applications for biochips. Many aspects, including materials, geometry, and costs, have to be considered in order to design the best packaging schemes. Most likely, some trade-offs have to be made – sacrificing some flexibility in certain aspects to ensure functionality of the device as a whole. Moreover, continuing research from both academia and industries in the packaging area will further mature this field and advance lab-on-a-chip, microfluidics, and bioMEMS in the long run.

References

1. Chin C.D., V. Linder, S.K. Sia. Lab-on-a-chip devices for global health: Past studies and future opportunities. *Lab on a Chip* 2007;7(1):41–57.
2. Yager P., T. Edwards, E. Fu, K. Helton, K. Nelson, M.R. Tam, B.H. Weigl. Microfluidic diagnostic technologies for global public health. *Nature* 2006;442(7101):412–418.
3. Engler K.H., A. Efstratiou, D. Norn, R.S. Kozlov, I. Selga, T.G. Glushkevich, M. Tam, V.G. Melnikov, I.K. Mazurova, V.E. Kim, G.Y. Tseneva, L.P. Titov, R.C. George. Immunochromatographic strip test for rapid detection of diphtheria toxin: Description and multicenter evaluation in areas of low and high prevalence of diphtheria. *Journal of Clinical Microbiology* 2002;40(1):80–83.
4. Arai H., B. Petchclai, K. Khupulpup, T. Kurimura, K. Takeda. Evaluation of a rapid immunochromatographic test for detection of antibodies to human immunodeficiency virus. *Journal of Clinical Microbiology* 1999;37(2):367–370.
5. Patterson K., B. Olsen, C. Thomas, D. Norn, M. Tam, C. Elkins. Development of a rapid immunodiagnostic test for *Haemophilus ducreyi*. *Journal of Clinical Microbiology* 2002;40(10):3694–3702.
6. Zarakolu P., I. Buchanan, M. Tam, K. Smith, E.W. Hook. Preliminary evaluation of an immunochromatographic strip test for specific *Treponema pallidum* antibodies. *Journal of Clinical Microbiology* 2002;40(8):3064–3065.
7. Martinez A.W., S.T. Phillips, M.J. Butte, G.M. Whitesides. Patterned paper as a platform for inexpensive, low-volume, portable bioassays. *Angewandte Chemie-International Edition* 2007;46(8):1318–1320.
8. Martinez A.W., S.T. Phillips, E. Carrilho, S.W. Thomas, H. Sindi, G.M. Whitesides. Simple telemedicine for developing regions: Camera phones and paper-based microfluidic devices for real-time, off-site diagnosis. *Analytical Chemistry* 2008;80(10):3699–3707.

9. Martinez A.W., S.T. Phillips, G.M. Whitesides. Three-dimensional microfluidic devices fabricated in layered paper and tape. *Proceedings of the National Academy of Sciences* 2008;105(50):19606–19611.
10. Hintsche R., B. Moller, I. Dransfeld, U. Wollenberger, F. Scheller, B. Hoffmann. Chip Biosensors on Thin-Film Metal-Electrodes. *Sensors and Actuators B-Chemical* 1991;4(3–4):287–291.
11. Shulga A.A., A.P. Soldatkin, A.V. Elskaya, S.V. Dzyadevich, S.V. Patskovsky, V.I. Strikha. Thin-Film Conductometric Biosensors for Glucose and Urea Determination. *Biosensors & Bioelectronics* 1994;9(3):217–223.
12. Hunt H.C., J.S. Wilkinson. Optofluidic integration for microanalysis. *Microfluidics and Nanofluidics* 2008;4(1–2):53–79.
13. Cui Y., Q.Q. Wei, H.K. Park, C.M. Lieber. Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. *Science* 2001;293(5533):1289–1292.
14. Zheng G.F., F. Patolsky, Y. Cui, W.U. Wang, C.M. Lieber. Multiplexed electrical detection of cancer markers with nanowire sensor arrays. *Nature Biotechnology* 2005;23(10):1294–1301.
15. Bashir R. BioMEMS: state-of-the-art in detection, opportunities and prospects. *Advanced Drug Delivery Reviews* 2004;56(11):1565–1586.
16. Waggoner P.S., H.G. Craighead. Micro- and nanomechanical sensors for environmental, chemical, and biological detection. *Lab on a Chip* 2007;7(10):1238–1255.
17. Fritz J., M.K. Baller, H.P. Lang, H. Rothuizen, P. Vettiger, E. Meyer, H.J. Guntherodt, C. Gerber, J.K. Gimzewski. Translating biomolecular recognition into nanomechanics. *Science* 2000;288(5464):316–318.
18. Hansen K.M., H.F. Ji, G.H. Wu, R. Datar, R. Cote, A. Majumdar, T. Thundat. Cantilever-based optical deflection assay for discrimination of DNA single-nucleotide mismatches. *Analytical Chemistry* 2001;73(7):1567–1571.
19. Gupta A., D. Akin, R. Bashir. Single virus particle mass detection using microresonators with nanoscale thickness. *Applied Physics Letters* 2004;84(11):1976–1978.
20. Illic B., D. Czaplewski, H.G. Craighead, P. Neuzil, C. Campagnolo, C. Batt. Mechanical resonant immunospecific biological detector. *Applied Physics Letters* 2000;77(3):450–452.
21. McKendry R., J.Y. Zhang, Y. Arntz, T. Strunz, M. Hegner, H.P. Lang, M.K. Baller, U. Certa, E. Meyer, H.J. Guntherodt, C. Gerber. Multiple label-free biodetection and quantitative DNA-binding assays on a nanomechanical cantilever array. *Proceedings of the National Academy of Sciences of the United States of America* 2002;99(15):9783–9788.
22. Wu G.H., R.H. Datar, K.M. Hansen, T. Thundat, R.J. Cote, A. Majumdar. Bioassay of prostate-specific antigen (PSA) using microcantilevers. *Nature Biotechnology* 2001;19(9):856–860.
23. Burg T.P., M. Godin, S.M. Knudsen, W. Shen, G. Carlson, J.S. Foster, K. Babcock, S.R. Manalis. Weighing of biomolecules, single cells and single nanoparticles in fluid. *Nature* 2007;446:1066–1069.
24. Burg T.P., A.R. Mirza, N. Milovic, C.H. Tsau, G.A. Popescu, J.S. Foster, S.R. Manalis. Vacuum-packaged suspended microchannel resonant mass sensor for biomolecular detection. *Journal of Microelectromechanical Systems* 2006;15(6):1466–1476.
25. Fruetel J.A., R.F. Renzi, V.A. VanderNoot, J. Stamps, B.A. Horn, J.A.A. West, S. Ferko, R. Crocker, C.G. Bailey, D. Arnold, B. Wiedenman, W.Y. Choi, D. Yee, I. Shokair, E. Hasselbrink, P. Paul, D. Rakestraw, D. Padgen. Microchip separations of protein biotoxins using an integrated hand-held device. *Electrophoresis* 2005;26(6):1144–1154.
26. Stratist-Cullum D.N., G.D. Griffin, J. Mobley, A.A. Vass, T. Vo-Dinh. A miniature biochip system for detection of aerosolized *Bacillus globigii* spores. *Analytical Chemistry* 2003;75(2):275–280.
27. Psaltis D., S.R. Quake, C.H. Yang. Developing optofluidic technology through the fusion of microfluidics and optics. *Nature* 2006;442(7101):381–386.

28. Irawan R., S.C. Tjin, X.Q. Fang, C.Y. Fu. Integration of optical fiber light guide, fluorescence detection system, and multichannel disposable microfluidic chip. *Biomedical Microdevices* 2007;9(3):413–419.
29. Lin C.H., G.B. Lee, S.H. Chen, G.L. Chang. Micro capillary electrophoresis chips integrated with buried SU-8/SOG optical waveguides for bio-analytical applications. *Sensors and Actuators a-Physical* 2003;107(2):125–131.
30. Kou Q., I. Yesilyurt, V. Studer, M. Belotti, E. Cambril, Y. Chen. On-chip optical components and microfluidic systems. *Microelectronic Engineering* 2004;73–74:876–880.
31. Misiakos K., S.E. Kakabakos, P.S. Petrou, H.H. Ruf. A monolithic silicon optoelectronic transducer as a real-time affinity biosensor. *Analytical Chemistry* 2004;76(5):1366–1373.
32. Balslev S., A.M. Jorgensen, B. Bilenberg, K.B. Mogensen, D. Snakenborg, O. Geschke, J.P. Kutter, A. Kristensen. Lab-on-a-chip with integrated optical transducers. *Lab on a Chip* 2006;6(2):213–217.
33. Dandin M., P. Abshire, E. Smela. Optical filtering technologies for integrated fluorescence sensors. *Lab on a Chip* 2007;7(8):955–977.
34. Prakash A.R., S. Adamia, V. Sieben, P. Pilarski, L.M. Pilarski, C.J. Backhouse. Small volume PCR in PDMS biochips with integrated fluid control and vapour barrier. *Sensors and Actuators B-Chemical* 2006;113(1):398–409.
35. Kaigala G.V., V.N. Hoang, A. Stickel, J. Lauzon, D. Manage, L.M. Pilarski, C.J. Backhouse. An inexpensive and portable microchip-based platform for integrated RT-PCR and capillary electrophoresis. *Analyst* 2008;133(3):331–338.
36. Weibel D.B., M. Kruithof, S. Potenta, S.K. Sia, A. Lee, G.M. Whitesides. Torque-actuated valves for microfluidics. *Analytical Chemistry* 2005;77(15):4726–4733.
37. Burns M.A., B.N. Johnson, S.N. Brahma sandra, K. Handique, J.R. Webster, M. Krishnan, T.S. Sammarco, P.M. Man, D. Jones, D. Heldsinger, C.H. Mastrangelo, D.T. Burke. An integrated nanoliter DNA analysis device. *Science* 1998;282(5388):484–487.
38. Liu R.H., J.N. Yang, R. Lenigk, J. Bonanno, P. Grodzinski. Self-contained, fully integrated biochip for sample preparation, polymerase chain reaction amplification, and DNA microarray detection. *Analytical Chemistry* 2004;76(7):1824–1831.
39. Pal R., M. Yang, R. Lin, B.N. Johnson, N. Srivastava, S.Z. Razzacki, K.J. Chomistek, D.C. Heldsinger, R.M. Haque, V.M. Ugaz, P.K. Thwar, Z. Chen, K. Alfano, M.B. Yim, M. Krishnan, A.O. Fuller, R.G. Larson, D.T. Burke, M.A. Burns. An integrated microfluidic device for influenza and other genetic analyses. *Lab on a Chip* 2005;5(10):1024–1032.
40. Fu E., T. Chinowsky, K. Nelson, K. Johnston, T. Edwards, K. Helton, M. Grow, J.W. Miller, P. Yager. SPR imaging-based salivary diagnostics system for the detection of small molecule analytes. *Oral-Based Diagnostics*. Blackwell Publishing: Oxford, UK, 2007. 335–344.
41. Velten T., H.H. Ruf, D. Barrow, N. Aspragathos, P. Lazarou, E. Jung, C.K. Malek, M. Richter, J. Kruckow. Packaging of bio-MEMS: Strategies, technologies, and applications. *IEEE Transactions on Advanced Packaging* 2005;28(4):533–546.
42. Herr A.E., A.V. Hatch, D.J. Throckmorton, H.M. Tran, J.S. Brennan, W.V. Giannobile, A.K. Singh. Microfluidic immunoassays as rapid saliva-based clinical diagnostics. *Proceedings of the National Academy of Sciences of the United States of America* 2007;104(13):5268–5273.
43. Lagally E.T., J.R. Scherer, R.G. Blazej, N.M. Toriello, B.A. Diep, M. Ramchandani, G.F. Sensabaugh, L.W. Riley, R.A. Mathies. Integrated portable genetic analysis microsystem for pathogen/infectious disease detection. *Analytical Chemistry* 2004;76(11):3162–3170.
44. Wang J., M.P. Chatrathi, A. Mulchandani, W. Chen. Capillary electrophoresis microchips for separation and detection of organophosphate nerve agents. *Analytical Chemistry* 2001;73(8):1804–1808.
45. Chinowsky T.M., S.D. Soelberg, P. Baker, N.R. Swanson, P. Kauffman, A. Mactutis, M.S. Grow, R. Atmar, S.S. Yee, C.E. Furlong. Portable 24-analyte surface plasmon

- resonance instruments for rapid, versatile biodetection. *Biosensors & Bioelectronics* 2007;22(9–10):2268–2275.
- 46. DeBusschere B.D., G.T.A. Kovacs. Portable cell-based biosensor system using integrated CMOS cell-cartridges. *Biosensors & Bioelectronics* 2001;16(7–8):543–556.
 - 47. Hood L., J.R. Heath, M.E. Phelps, B.Y. Lin. Systems biology and new technologies enable predictive and preventative medicine. *Science* 2004;306(5696):640–643.
 - 48. Weston A.D., L. Hood. Systems biology, proteomics, and the future of health care: Toward predictive, preventative, and personalized medicine. *Journal of Proteome Research* 2004;3(2):179–196.
 - 49. Bhattacharyya A., C.M. Klapperich. Design and testing of a disposable microfluidic chemiluminescent immunoassay for disease biomarkers in human serum samples. *Biomedical Microdevices* 2007;9(2):245–251.
 - 50. Linder V., S.K. Sia, G.M. Whitesides. Reagent-loaded cartridges for valveless and automated fluid delivery in microfluidic devices. *Analytical Chemistry* 2005;77(1):64–71.
 - 51. Grayson A.C.R., R.S. Shawgo, A.M. Johnson, N.T. Flynn, Y.W. Li, M.J. Cima, R. Langer. A BioMEMS review: MEMS technology for physiologically integrated devices. *Proceedings of the IEEE* 2004;92(1):6–21.
 - 52. Fonseca M.A.M.D.S.J.W.J.K.; Cardiomems, Inc., assignee. Implantable Wireless Sensor for Pressure Measurement within the Heart. US patent 6855115. 2005 Feb 15
 - 53. Santini J.T.M.J.C.R.S.L.; MIT, assignee. Microchip Drug Delivery Devices. US. 1998 Aug 25
 - 54. Kudo H., T. Sawada, E. Kazawa, H. Yoshida, Y. Iwasaki, K. Mitsubayashi. A flexible and wearable glucose sensor based on functional polymers with Soft-MEMS techniques. *Biosensors & Bioelectronics* 2006;22(4):558–562.
 - 55. Zhao Y.J., S.Q. Li, A. Davidson, B.Z. Yang, Q. Wang, Q. Lin. A MEMS viscometric sensor for continuous glucose monitoring. *Journal of Micromechanics and Microengineering* 2007;17(12):2528–2537.
 - 56. Jauniaux E., A. Watson, O. Ozturk, D. Quick, G. Burton. In-vivo measurement of intrauterine gases and acid-base values early in human pregnancy. *Human Reproduction* 1999;14(11):2901–2904.
 - 57. Prausnitz M.R. Microneedles for transdermal drug delivery. *Advanced Drug Delivery Reviews* 2004;56(5):581–587.
 - 58. Prausnitz M.R., M.G. Allen, I.J. Gujral. Microneedle drug delivery device. US Patent 7, 226, 439; 2007.
 - 59. Gujral I.J., M.G. Allen, M.R. Prausnitz. Microneedle device for extraction and sensing of bodily fluids. US Patent 7,344,499; 2008.
 - 60. McAllister D.V., P.M. Wang, S.P. Davis, J.H. Park, P.J. Canatella, M.G. Allen, M.R. Prausnitz. Microfabricated needles for transdermal delivery of macromolecules and nanoparticles: Fabrication methods and transport studies. *Proceedings of the National Academy of Sciences* 2003;100(24):13755–13760.
 - 61. Li P.Y., J. Shih, R. Lo, S. Saati, R. Agrawal, M.S. Humayun, Y.C. Tai, E. Meng. An electrochemical intraocular drug delivery device. *Sensors and Actuators a-Physical* 2008;143(1):41–48.
 - 62. Santini J.T., M.J. Cima, R. Langer. A controlled-release microchip. *Nature* 1999; 397(6717):335–338.
 - 63. Voskerician G., R.S. Shawgo, P.A. Hiltner, J.M. Anderson, M.J. Cima, R. Langer. In vivo inflammatory and wound healing effects of gold electrode voltammetry for MEMS micro-reservoir drug delivery device. *Ieee Transactions on Biomedical Engineering* 2004;51(4):627–635.
 - 64. Razzacki S.Z., P.K. Thwar, M. Yang, V.M. Ugaz, M.A. Burns. Integrated microsystems for controlled drug delivery. *Advanced Drug Delivery Reviews* 2004;56(2):185–198.
 - 65. Wu C.C., T. Yasukawa, H. Shiku, T. Matsue. Fabrication of miniature Clark oxygen sensor integrated with microstructure. *Sensors and Actuators B-Chemical* 2005;110(2):342–349.

66. Wu H.K., B. Huang, R.N. Zare. Construction of microfluidic chips using polydimethylsiloxane for adhesive bonding. *Lab on a Chip* 2005;5(12):1393–1398.
67. Hungar K., M. Gortz, E. Slavcheva, G. Spanier, C. Weidig, W. Mokwa. Production processes for a flexible retina implant (Eurosensors XVIII, Session C6.6). *Sensors and Actuators a-Physical* 2005;123–24:172–178.
68. Schanze T., L. Hesse, C. Lau, N. Greve, W. Haberer, S. Kammer, T. Doerge, A. Rentzos, T. Stieglitz. An optically powered single-channel stimulation implant as test-system for chronic biocompatibility and biostability of miniaturized retinal vision prostheses. *Ieee Transactions on Biomedical Engineering* 2007;54(6):983–992.
69. Schwarz M., L. Ewe, R. Hauschild, B.J. Hosticka, J. Huppertz, S. Kolnsberg, W. Mokwa, H.K. Trieu. Single chip CMOS imagers and flexible microelectronic stimulators for a retina implant system. *Sensors and Actuators a-Physical* 2000;83(1–3):40–46.
70. Loeb G.E., R.A. Peck, W.H. Moore, K. Hood. BION (TM) system for distributed neural prosthetic interfaces. *Medical Engineering & Physics* 2001;23(1):9–18.
71. Weiland J.D., W.T. Liu, M.S. Humayun. Retinal prosthesis. *Annual Review of Biomedical Engineering* 2005;7:361–401.
72. Schwartz A.B. Cortical Neural Prosthetics. *Annual Review of Neuroscience* 2004;27(1):487–507.
73. Cheung K.C. Implantable microscale neural interfaces. *Biomedical Microdevices* 2007;9(6):923–938.
74. Kipke D.R., R.J. Vetter, J.C. Williams, J.F. Hetke. Silicon-substrate intracortical microelectrode arrays for long-term recording of neuronal spike activity in cerebral cortex. *IEEE Trans on Rehabilitation Engineering* 2003;11(2):151–155.
75. Rutten W.L.C. Selective electrical interfaces with the nervous system. *Annual Review of Biomedical Engineering* 2002;4(1):407–452.
76. Cogan S.F. Neural Stimulation and Recording Electrodes. *Annual Review of Biomedical Engineering* 2008;10(1):275–309.
77. Tokuda T., Y.L. Pan, A. Uehara, K. Kagawa, M. Nunoshita, J. Ohta. Flexible and extendible neural interface device based on cooperative multi-chip CMOS LSI architecture. *Sensors and Actuators a-Physical* 2005;122(1):88–98.
78. Smith B., Z.N. Tang, M.W. Johnson, S. Pourmehdi, M.M. Gazdik, J.R. Buckett, P.H. Peckham. An externally powered, multichannel, implantable stimulator-telemeter for control of paralyzed muscle. *Ieee Transactions on Biomedical Engineering* 1998;45(4):463–475.
79. Windecker S., I. Mayer, G. De Pasquale, W. Maier, O. Dirsch, P. De Groot, Y.P. Wu, G. Noll, B. Leskosek, B. Meier, O.M. Hess, C. Working Grp Novel Surface. Stent coating with titanium-nitride-oxide for reduction of neointimal hyperplasia. *Circulation* 2001;104(8):928–933.
80. Grube E., U. Gerckens, S. Rowold, R. Muller, G. Selbach, J. Stamm, M. Staberock. Inhibition of in-stent restenosis by the Quanam drug eluting polymer stent; Two year follow-up. *Journal of the American College of Cardiology* 2001;37(2):74A–74A.
81. Hiatt B.L., F. Ikeno, A.C. Yeung, A.J. Carter. Drug-eluting stents for the prevention of restenosis: In quest for the holy grail. *Catheterization and Cardiovascular Interventions* 2002;55(3):409–417.
82. Allen M. G M.E.J.K.D.J.M.; CardioMEMS, Inc., assignee. Communication wit an Implanted Wireless Sensor. US. 2007
83. Klose J., E. Rehtanz, C. Rothe, I. Eulitz, V. Guther, W. Beck. Manufacture of titanium implants. *Materialwissenschaft Und Werkstofftechnik* 2008;39(4–5):304–308.
84. Wiegand U.K.H., J. Potratz, F. Luninghamake, G. Taubert, A. Brandes, K.W. Diederich. Electrophysiological characteristics of bipolar membrane carbon leads with and without steroid elution compared with a conventional carbon and a steroid-eluting platinum lead. *Pace-Pacing and Clinical Electrophysiology* 1996;19(8):1155–1161.
85. Wiegand U.K.H., A. Zhdanov, E. Stammwitz, I. Crozier, R.J.J. Claessens, J. Meier, R.J. Bos, F. Bode, J. Potratz. Electrophysiological performance of a bipolar membrane-coated

- titanium nitride electrode: A randomized comparison of steroid and nonsteroid lead designs. *Pace-Pacing and Clinical Electrophysiology* 1999;22(6):935–941.
- 86. Wiggins M.J., B. Wilkoff, J.M. Anderson, A. Hiltner. Biodegradation of polyether polyurethane inner insulation in bipolar pacemaker leads. *Journal of Biomedical Materials Research* 2001;58(3):302–307.
 - 87. Russell R.J., M.V. Pishko, C.C. Gefrides, M.J. McShane, G.L. Cote. A Fluorescence-Based Glucose Biosensor Using Concanavalin A and Dextran Encapsulated in a Poly(ethylene glycol) Hydrogel. *Analytical Chemistry* 1999;71(15):3126–3132.
 - 88. Receveur R.A.M., F.W. Lindemans, N.F. de Rooij. Microsystem technologies for implantable applications. *Journal of Micromechanics and Microengineering* 2007;17(5):R50–R80.
 - 89. Mokwa W., U. Schnakenberg. Micro-transponder systems for medical applications. *Ieee Transactions on Instrumentation and Measurement* 2001;50(6):1551–1555.
 - 90. Flick B.B., R. Orglmeister. A portable microsystem-based telemetric pressure and temperature measurement unit. *Ieee Transactions on Biomedical Engineering* 2000;47(1):12–16.
 - 91. Esashi M., S. Sugiyama, K. Ikeda, Y.L. Wang, H. Miyashita. Vacuum-sealed silicon micromachined pressure sensors. *Proceedings of the Ieee* 1998;86(8):1627–1639.
 - 92. Chen P.J., D.C. Rodger, R. Agrawal, S. Saati, E. Meng, R. Varma, M.S. Humayun, Y.C. Tai. Implantable micromechanical parylene-based pressure sensors for unpowered intraocular pressure sensing. *Journal of Micromechanics and Microengineering* 2007;17(10):1931–1938.
 - 93. Chen L., A. Manz, P.J.R. Day. Total nucleic acid analysis integrated on microfluidic devices. *Lab on a Chip* 2007;7(11):1413–1423.
 - 94. Heyries K.A., M.G. Loughran, D. Hoffmann, A. Homsy, L.J. Blum, C.A. Marquette. Microfluidic biochip for chemiluminescent detection of allergen-specific antibodies. *Biosensors & Bioelectronics* 2008;23(12):1812–1818.
 - 95. Isoda T., I. Urushibara, M. Sato, H. Uemura, H. Sato, N. Yamauchi. Development of a sensor-array chip with immobilized antibodies and the application of a wireless antigen-screening system. *Sensors and Actuators B-Chemical* 2008;129(2):958–970.
 - 96. Prakash R., K. Kaler. An integrated genetic analysis microfluidic platform with valves and a PCR chip reusability method to avoid contamination. *Microfluidics and Nanofluidics* 2007;3(2):177–187.
 - 97. Zhang C.S., J.L. Xu, W.L. Ma, W.L. Zheng. PCR microfluidic devices for DNA amplification. *Biotechnology Advances* 2006;24(3):243–284.
 - 98. Sethu P., A. Sin, M. Toner. Microfluidic diffusive filter for apheresis (leukapheresis). *Lab on a Chip* 2006;6(1):83–89.
 - 99. Battrell C. F M.S.B.H.W.J.M.H.C.A.L.W.B.; Micronics, Inc., assignee. Method and System for Microfluidic Manipulation, Amplification and Analysis of Fluids, for Example, Bacteria Assays and Antiglobulin Testing. US. 2004
 - 100. Chamot S.R., C. Depeursinge. MEMS for enhanced optical diagnostics in endoscopy. *Minimally Invasive Therapy & Allied Technologies* 2007;16(2):101–108.
 - 101. Hupert M.L., M.A. Witek, Y. Wang, M.W. Mitchell, Y. Liu, Y. Bejat, D.E. Nikitopoulos, J. Goettert, M.C. Murphy, S.A. Soper. Polymer-based microfluidic devices for biomedical applications. *Proceedings of SPIE* 2003;4982:52–64.
 - 102. Mitchell M.W., X. Liu, Y. Bejat, D.E. Nikitopoulos, S.A. Soper, M.C. Murphy. Modeling and validation of a molded polycarbonate continuous flow polymerase chain reaction device. *Proceedings of SPIE* 2003;4982:83–98.
 - 103. Lee D.S., S.H. Park, H.S. Yang, K.H. Chung, T.H. Yoon, S.J. Kim, K. Kim, Y.T. Kim. Bulk-micromachined submicroliter-volume PCR chip with very rapid thermal response and low power consumption. *Lab on a Chip* 2004;4(4):401–407.
 - 104. Koh C.G., W. Tan, M.Q. Zhao, A.J. Ricco, Z.H. Fan. Integrating polymerase chain reaction, valving, and electrophoresis in a plastic device for bacterial detection. *Analytical Chemistry* 2003;75(17):4591–4598.

105. Krishnan M., D.T. Burke, M.A. Burns. Polymerase chain reaction in high surface-to-volume ratio SiO₂ microstructures. *Analytical Chemistry* 2004;76(22):6588–6593.
106. Woolley A.T., D. Hadley, P. Landre, A.J. deMello, R.A. Mathies, M.A. Northrup. Functional integration of PCR amplification and capillary electrophoresis in a microfabricated DNA analysis device. *Analytical Chemistry* 1996;68(23):4081–4086.
107. West J., B. Karamata, B. Lillis, J.P. Gleeson, J. Alderman, J.K. Collins, W. Lane, A. Mathewson, H. Berney. Application of magnetohydrodynamic actuation to continuous flow chemistry. *Lab on a Chip* 2002;2(4):224–230.
108. Hong J.W., T. Fujii, M. Seki, T. Yamamoto, I. Endo. Integration of gene amplification and capillary gel electrophoresis on a polydimethylsiloxane-glass hybrid microchip. *Electrophoresis* 2001;22(2):328–333.
109. Shen K.Y., X.F. Chen, M. Guo, J. Cheng. A microchip-based PCR device using flexible printed circuit technology. *Sensors and Actuators B-Chemical* 2005;105(2):251–258.
110. Daniel J.H., S. Iqbal, R.B. Millington, D.F. Moore, C.R. Lowe, D.L. Leslie, M.A. Lee, M.J. Pearce. Silicon microchambers for DNA amplification. *Sensors and Actuators a-Physical* 1998;71(1–2):81–88.
111. Northrup M.A., B. Benett, D. Hadley, P. Landre, S. Lehew, J. Richards, P. Stratton. A miniature analytical instrument for nucleic acids based on micromachined silicon reaction chambers. *Analytical Chemistry* 1998;70(5):918–922.
112. Sun K., A. Yamaguchi, Y. Ishida, S. Matsuo, H. Misawa. A heater-integrated transparent microchannel chip for continuous-flow PCR. *Sensors and Actuators B-Chemical* 2002;84(2–3):283–289.
113. Khandurina J., T.E. McKnight, S.C. Jacobson, L.C. Waters, R.S. Foote, J.M. Ramsey. Integrated system for rapid PCR-based DNA analysis in microfluidic devices. *Analytical Chemistry* 2000;72(13):2995–3000.
114. Lin Y.C., M.Y. Huang, K.C. Young, T.T. Chang, C.Y. Wu. A rapid micro-polymerase chain reaction system for hepatitis C virus amplification. *Sensors and Actuators B-Chemical* 2000;71(1–2):2–8.
115. Lin Y.C., C.C. Yang, M.Y. Huang. Simulation and experimental validation of micro polymerase chain reaction chips. *Sensors and Actuators B-Chemical* 2000;71(1–2):127–133.
116. Zhou Z.M., D.Y. Liu, R.T. Zhong, Z.P. Dai, D.P. Wu, H. Wang, Y.G. Du, Z.N. Xia, L.P. Zhang, X.D. Mei, B.C. Lin. Determination of SARS-coronavirus by a microfluidic chip system. *Electrophoresis* 2004;25(17):3032–3039.
117. Gulliksen A., L. Solli, F. Karlsen, H. Rogne, E. Hovig, T. Nordstrom, R. Sirevag. Real-time nucleic acid sequence-based amplification in nanoliter volumes. *Analytical Chemistry* 2004;76(1):9–14.
118. Matsubara Y., K. Kerman, M. Kobayashi, S. Yamamura, Y. Morita, E. Tamiya. Microchamber array based DNA quantification and specific sequence detection from a single copy via PCR in nanoliter volumes. *Biosensors & Bioelectronics* 2005;20(8):1482–1490.
119. Curcio M., J. Roeraade. Continuous segmented-flow polymerase chain reaction for high-throughput miniaturized DNA amplification. *Analytical Chemistry* 2003;75(1):1–7.
120. Sethu P., C.H. Mastrangelo. Cast epoxy-based microfluidic systems and their application in biotechnology. *Sensors and Actuators B-Chemical* 2004;98(2–3):337–346.
121. Swerdlow H., B.J. Jones, C.T. Wittwer. Fully automated DNA reaction and analysis in a fluidic capillary instrument. *Analytical Chemistry* 1997;69(5):848–855.
122. Zhang N.Y., E.S. Yeung. On-line coupling of polymerase chain reaction and capillary electrophoresis for automatic DNA typing and HIV-1 diagnosis. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* 1998;714(1):3–11.
123. Ferrance J.P., Q.R. Wu, B. Giordano, C. Hernandez, Y. Kwok, K. Snow, S. Thibodeau, J.P. Landers. Developments toward a complete micro-total analysis system for Duchenne muscular dystrophy diagnosis. *Analytica Chimica Acta* 2003;500(1–2):223–236.
124. Huhmer A.F.R., J.P. Landers. Noncontact infrared-mediated thermocycling for effective polymerase chain reaction amplification of DNA in nanoliter volumes. *Analytical Chemistry* 2000;72(21):5507–5512.

125. Oda R.P., M.A. Strausbauch, A.F.R. Huhmer, N. Borson, S.R. Jurrens, J. Craighead, P.J. Wettstein, B. Eckloff, B. Kline, J.P. Landers. Infrared-mediated thermocycling for ultrafast polymerase chain reaction amplification of DNA. *Analytical Chemistry* 1998;70(20):4361–4368.
126. Tanaka Y., M.N. Slyadnev, A. Hibara, M. Tokeshi, T. Kitamori. Non-contact photothermal control of enzyme reactions on a microchip by using a compact diode laser. *Journal of Chromatography A* 2000;894(1–2):45–51.
127. Schneegass I., R. Brautigam, J.M. Kohler. Miniaturized flow-through PCR with different template types in a silicon chip thermocycler. *Lab on a Chip* 2001;1(1):42–49.
128. Chou C.F., R. Changrani, P. Roberts, D. Sadler, J. Burdon, F. Zenhausern, S. Lin, A. Mulholland, N. Swami, R. Terbrueggen. A miniaturized cyclic PCR device - modeling and experiments. *Microelectronic Engineering* 2002;61–62:921–925.
129. Liu J., M. Enzelberger, S. Quake. A nanoliter rotary device for polymerase chain reaction. *Electrophoresis* 2002;23(10):1531–1536.
130. Shi Y.N., P.C. Simpson, J.R. Scherer, D. Wexler, C. Skibola, M.T. Smith, R.A. Mathies. Radial capillary array electrophoresis microplate and scanner for high-performance nucleic acid analysis. *Analytical Chemistry* 1999;71(23):5354–5361.
131. Waters L.C., S.C. Jacobson, N. Kroutchinina, J. Khandurina, R.S. Foote, J.M. Ramsey. Multiple sample PCR amplification and electrophoretic analysis on a microchip. *Analytical Chemistry* 1998;70(24):5172–5176.
132. Waters L.C., S.C. Jacobson, N. Kroutchinina, J. Khandurina, R.S. Foote, J.M. Ramsey. Microchip device for cell lysis, multiplex PCR amplification, and electrophoretic sizing. *Analytical Chemistry* 1998;70(1):158–162.
133. Perch-Nielsen I.R., D.D. Bang, C.R. Poulsen, J. El-Ali, A. Wolff. Removal of PCR inhibitors using dielectrophoresis as a selective filter in a microsystem. *Lab on a Chip* 2003;3(3):212–216.
134. Gascoyne P., C. Mahidol, M. Ruchirawat, J. Satayavivad, P. Watcharasit, F.F. Becker. Microsample preparation by dielectrophoresis: isolation of malaria. *Lab on a Chip* 2002;2(2):70–75.
135. Namasivayam V., R.S. Lin, B. Johnson, S. Brahmasandra, Z. Razzacki, D.T. Burke, M.A. Burns. Advances in on-chip photodetection for applications in miniaturized genetic analysis systems. *Journal of Micromechanics and Microengineering* 2004;14(1):81–90.
136. Kumar A., G. Goel, E. Fehrenbach, A.K. Puniya, K. Singh. Microarrays: The technology, analysis and application. *Engineering in Life Sciences* 2005;5(3):215–222.
137. Bulyk M.L. DNA microarray technologies for measuring protein-DNA interactions. *Current Opinion in Biotechnology* 2006;17(4):422–430.
138. Cretich M., F. Damin, G. Pirri, M. Chiari. Protein and peptide arrays: Recent trends and new directions. *Biomolecular Engineering* 2006;23(2–3):77–88.
139. Hoheisel J.D. Microarray technology: beyond transcript profiling and genotype analysis. *Nature Reviews Genetics* 2006;7(3):200–210.
140. Hultschig C., J. Kreutzberger, H. Seitz, Z. Konthur, K. Bussow, H. Lehrach. Recent advances of protein microarrays. *Current Opinion in Chemical Biology* 2006;10(1):4–10.
141. Stoughton R.B. Applications of DNA microarrays in biology. *Annual Review of Biochemistry* 2005;74:53–82.
142. Barbulovic-Nad I., M. Luente, Y. Sun, M.J. Zhang, A.R. Wheeler, M. Bussmann. Bio-microarray fabrication techniques - A review. *Critical Reviews in Biotechnology* 2006;26(4):237–259.
143. <http://www.affymetrix.com>
144. Anderson R.C., X. Su, G.J. Bogdan, J. Fenton. A miniature integrated device for automated multistep genetic assays. *Nucleic Acids Research* 2000;28(12):e60i–vi.
145. Lenigk R., R.H. Liu, M. Athavale, Z.J. Chen, D. Ganser, J.N. Yang, C. Rauch, Y.J. Liu, B. Chan, H.N. Yu, M. Ray, R. Marrero, P. Grodzinski. Plastic biochannel hybridization devices: a new concept for microfluidic DNA arrays. *Analytical Biochemistry* 2002;311(1):40–49.

146. Liu J., C. Hansen, S.R. Quake. Solving the "world-to-chip" interface problem with a microfluidic matrix. *Analytical Chemistry* 2003;75(18):4718–4723.
147. Martynova L., L.E. Locascio, M. Gaitan, G.W. Kramer, R.G. Christensen, W.A. MacCrenan. Fabrication of plastic microfluid channels by imprinting methods. *Analytical Chemistry* 1997;69(23):4783–4789.
148. Qi S.Z., X.Z. Liu, S. Ford, J. Barrows, G. Thomas, K. Kelly, A. McCandless, K. Lian, J. Goettler, S.A. Soper. Microfluidic devices fabricated in poly(methyl methacrylate) using hot-embossing with integrated sampling capillary and fiber optics for fluorescence detection. *Lab on a Chip* 2002;2(2):88–95.
149. Soper S.A., S.M. Ford, S. Qi, R.L. McCarley, K. Kelly, M.C. Murphy. Polymeric microelectromechanical systems. *Analytical Chemistry* 2000;72(19):642A–651A.
150. Situma C., M. Hashimoto, S.A. Soper. Merging microfluidics with microarray-based bioassays. *Biomolecular Engineering* 2006;23(5):213–231.
151. Thorsen T., S.J. Maerkl, S.R. Quake. Microfluidic large-scale integration. *Science* 2002;298(5593):580–584.
152. Duffy D.C., J.C. McDonald, O.J.A. Schueller, G.M. Whitesides. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Analytical Chemistry* 1998;70(23):4974–4984.
153. Unger M.A., H.P. Chou, T. Thorsen, A. Scherer, S.R. Quake. Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science* 2000;288(5463):113–116.
154. Marcus J.S., W.F. Anderson, S.R. Quake. Microfluidic single-cell mRNA isolation and analysis. *Analytical Chemistry* 2006;78(9):3084–3089.
155. Lee C.C., G.D. Sui, A. Elizarov, C.Y.J. Shu, Y.S. Shin, A.N. Dooley, J. Huang, A. Daridon, P. Wyatt, D. Stout, H.C. Kolb, O.N. Witte, N. Satyamurthy, J.R. Heath, M.E. Phelps, S.R. Quake, H.R. Tseng. Multistep synthesis of a radiolabeled imaging probe using integrated microfluidics. *Science* 2005;310(5755):1793–1796.
156. Balagadde F.K., L.C. You, C.L. Hansen, F.H. Arnold, S.R. Quake. Long-term monitoring of bacteria undergoing programmed population control in a microchemostat. *Science* 2005;309(5731):137–140.
157. Anderson M.J., C.L. Hansen, S.R. Quake. Phase knowledge enables rational screens for protein crystallization. *Proceedings of the National Academy of Sciences of the United States of America* 2006;103(45):16746–16751.
158. Hansen C.L., S. Classen, J.M. Berger, S.R. Quake. A microfluidic device for kinetic optimization of protein crystallization and in situ structure determination. *Journal of the American Chemical Society* 2006;128(10):3142–3143.
159. Hansen C.L., M.O.A. Sommer, S.R. Quake. Systematic investigation of protein phase behavior with a microfluidic formulator. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101(40):14431–14436.
160. <http://www.fluidigm.com>
161. Schorzman D.A., J.M. Desimone, J.P. Rolland, S.R. Quake, R.M. Van Dam. Solvent-Resistant Photocurable "Liquid Teflon" for Microfluidic Device Fabrication. *Journal of the American Chemical Society* 2004;126(8):2322–2323.
162. Melin J., S.R. Quake. Microfluidic large-scale integration: The evolution of design rules for biological automation. *Annual Review of Biophysics and Biomolecular Structure* 2007;36:213–231.
163. Kamotani Y., T. Bersano-Begey, N. Kato, Y.C. Tung, D. Huh, J.W. Song, S. Takayama. Individually programmable cell stretching microwell arrays actuated by a Braille display. *Biomaterials* 2008;29(17):2646–2655.
164. Song J.W., W. Gu, N. Futai, K.A. Warner, J.E. Nor, S. Takayama. Computer-controlled microcirculatory support system for endothelial cell culture and shearing. *Analytical Chemistry* 2005;77(13):3993–3999.
165. Gu W., X.Y. Zhu, N. Futai, B.S. Cho, S. Takayama. Computerized microfluidic cell culture using elastomeric channels and Braille displays. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101(45):15861–15866.

166. Enzelberger M.M., C.L. Hansen, J. Liu, S.R. Quake, C. Ma. Nucleic acid amplification using microfluidic devices, WO/2002/081729. World Intellectual Property Organization; 2002.
167. Lee C., G. Sui, A. Elizarov, H.C. Kolb, J. Huang, J.R. Heath, M.E. Phelps, S.R. Quake, H. Tseng, P. Wyatt. Microfluidic Devices with Chemical Reaction Circuits. EP Patent 1,838,431; 2007.
168. El-Ali J., P.K. Sorger, K.F. Jensen. Cells on chips. *Nature* 2006;442(7101):403–411.
169. Park T.H., M.L. Shuler. Integration of cell culture and microfabrication technology. *Biotechnology Progress* 2003;19(2):243–253.
170. Sims C.E., N.L. Allbritton. Analysis of single mammalian cells on-chip. *Lab on a Chip* 2007;7(4):423–440.
171. Cheng J.Y., M.H. Yen, C.T. Kuo, T.H. Young. A transparent cell-culture microchamber with a variably controlled concentration gradient generator and flow field rectifier. *Biomicrofluidics* 2008;2(2):12.
172. Petronis S., M. Stangegaard, C.B.V. Christensen, M. Dufva. Transparent polymeric cell culture chip with integrated temperature control and uniform media perfusion. *Biotechniques* 2006;40(3):368–376.
173. Park J., T. Bansal, M. Pinelis, M.M. Maharbiz. A microsystem for sensing and patterning oxidative microgradients during cell culture. *Lab on a Chip* 2006;6(5):611–622.
174. Maharbiz M.M., W.J. Holtz, S. Sharifzadeh, J.D. Keasling, R.T. Howe. A microfabricated electrochemical oxygen generator for high-density cell culture arrays. *Journal of Microelectromechanical Systems* 2003;12(5):590–599.
175. Vollmer A.P., R.F. Probstein, R. Gilbert, T. Thorsen. Development of an integrated microfluidic platform for dynamic oxygen sensing and delivery in a flowing medium. *Lab on a Chip* 2005;5(10):1059–1066.
176. Ges I.A., B.L. Ivanov, D.K. Schaffer, E.A. Lima, A.A. Werdich, F.J. Baudenbacher. Thin-film IrOx pH microelectrode for microfluidic-based microsystems. *Biosensors & Bioelectronics* 2005;21(2):248–256.
177. Taff B.M., J. Voldman. A scalable addressable positive-dielectrophoretic cell-sorting array. *Analytical Chemistry* 2005;77(24):7976–7983.
178. Voldman J., M.L. Gray, M. Toner, M.A. Schmidt. A microfabrication-based dynamic array cytometer. *Analytical Chemistry* 2002;74(16):3984–3990.
179. Wang X.B., J. Yang, Y. Huang, J. Vykoukal, F.F. Becker, P.R.C. Gascoyne. Cell separation by dielectrophoretic field-flow-fractionation. *Analytical Chemistry* 2000;72(4):832–839.
180. Gomez-Sjoberg R., A.A. Leyrat, D.M. Pirone, C.S. Chen, S.R. Quake. Versatile, fully automated, microfluidic cell culture system. *Analytical Chemistry* 2007;79(22):8557–8563.
181. Hung P.J., P.J. Lee, P. Sabourchi, R. Lin, L.P. Lee. Continuous perfusion microfluidic cell culture array for high-throughput cell-based assays. *Biotechnology and Bioengineering* 2005;89(1):1–8.
182. Lii J., W.J. Hsu, H. Parsa, A. Das, R. Rouse, S.K. Sia. Real-time microfluidic system for studying mammalian cells in 3D microenvironments. *Analytical Chemistry* 2008;80(10):3640–3647.
183. Madou M., J. Zoval, G.Y. Jia, H. Kido, J. Kim, N. Kim. Lab on a CD. *Annual Review of Biomedical Engineering* 2006;8:601–628.
184. Balaban N.Q., U.S. Schwarz, D. Riveline, P. Goichberg, G. Tzur, I. Sabanay, D. Mahalu, S. Safran, A. Bershadsky, L. Addadi, B. Geiger. Force and focal adhesion assembly: a close relationship studied using elastic micropatterned substrates. *Nature Cell Biology* 2001;3(5):466–472.
185. Tan J.L., J. Tien, D.M. Pirone, D.S. Gray, K. Bhadriraju, C.S. Chen. Cells lying on a bed of microneedles: An approach to isolate mechanical force. *Proceedings of the National Academy of Sciences of the United States of America* 2003;100(4):1484–1489.
186. Hellmich W., C. Pelargus, K. Leffhalm, A. Ros, D. Anselmetti. Single cell manipulation, analytics, and label-free protein detection in microfluidic devices for systems nanobiology. *Electrophoresis* 2005;26(19):3689–3696.

187. Munce N.R., J.Z. Li, P.R. Herman, L. Lilge. Microfabricated system for parallel single-cell capillary electrophoresis. *Analytical Chemistry* 2004;76(17):4983–4989.
188. Klaus J.W., S.M. George. SiO₂ chemical vapor deposition at room temperature using SiCl₄ and H₂O with an NH₃ catalyst. *Journal of the Electrochemical Society* 2000;147(7):2658–2664.
189. Senturia S.D. *Microsystem Design*. Springer Science+Business Media, LLC: New York, 2005.
190. Lim K.S., W.J. Chang, Y.M. Koo, R. Bashir. Reliable fabrication method of transferable micron scale metal pattern for poly(dimethylsiloxane) metallization. *Lab on a Chip* 2006;6(4):578–580.
191. Niu X.Z., S.L. Peng, L.Y. Liu, W.J. Wen, P. Sheng. Characterizing and patterning of PDMS-based conducting composites. *Advanced Materials* 2007;19(18):2682–2686.
192. Bowden N., S. Brittain, A.G. Evans, J.W. Hutchinson, G.M. Whitesides. Spontaneous formation of ordered structures in thin films of metals supported on an elastomeric polymer. *Nature* 1998;393(6681):146–149.
193. Trau D., J. Jiang, N.J. Sucher. Preservation of the biofunctionality of DNA and protein during microfabrication. *Langmuir* 2006;22(3):877–881.
194. Kentsch J., S. Breisch, M. Stezle. Low temperature adhesion bonding for BioMEMS. *Journal of Micromechanics and Microengineering* 2006;16(4):802–807.
195. Ghafar-Zadeh E., M. Sawan, D. Therriault. Novel direct-write CMOS-based laboratory-on-chip: Design, assembly and experimental results. *Sensors and Actuators a-Physical* 2007;134(1):27–36.
196. Zimmermann S., D. Fienbork, A.W. Flounders, D. Liepmann. In-device enzyme immobilization: wafer-level fabrication of an integrated glucose sensor. *Sensors and Actuators B-Chemical* 2004;99(1):163–173.
197. Linder V., S. Koster, W. Franks, T. Kraus, E. Verpoorte, F. Heer, A. Hierlemann, N.F. de Rooij. Microfluidics/CMOS orthogonal capabilities for cell biology. *Biomedical Microdevices* 2006;8(2):159–166.
198. Pan J.Y. Reliability considerations for the BioMEMS designer. *Proceedings of the Ieee* 2004;92(1):174–184.
199. Bhagat A.A.S., P. Jothimuthu, A. Pais, I. Papautsky. Re-usable quick-release interconnect for characterization of microfluidic systems. *Journal of Micromechanics and Microengineering* 2007;17(1):42–49.
200. Christensen A.M., D.A. Chang-Yen, B.K. Gale. Characterization of interconnects used in PDMS microfluidic systems. *Journal of Micromechanics and Microengineering* 2005;15(5):928–934.
201. Han K.H., R.D. McConnell, C.J. Easley, J.M. Bienvenue, J.P. Ferrance, J.P. Landers, A.B. Frazier. An active microfluidic system packaging technology. *Sensors and Actuators B-Chemical* 2007;122(1):337–346.
202. Puntambekar A., C.H. Ahn. Self-aligning microfluidic interconnects for glass- and plastic-based microfluidic systems. *Journal of Micromechanics and Microengineering* 2002;12(1):35–40.
203. Fujii T., Y. Sando, K. Higashino, Y. Fujii. A plug and play microfluidic device. *Lab on a Chip* 2003;3(3):193–197.
204. Igata E., M. Arundell, H. Morgan, J.M. Cooper. Interconnected reversible lab-on-a-chip technology. *Lab on a Chip* 2002;2(2):65–69.
205. Yuen P.K. SmartBuild—A truly plug-n-play modular microfluidic system. *Lab on a Chip* 2008;8:1374–1378.
206. Shaikh K.A., K.S. Ryu, E.D. Goluch, J.M. Nam, J.W. Liu, S. Thaxton, T.N. Chiesl, A.E. Barron, Y. Lu, C.A. Mirkin, C. Liu. A modular microfluidic architecture for integrated biochemical analysis. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102(28):9745–9750.
207. Ko W.H. *Packaging of Microfabricated Devices and Systems*. *Materials Chemistry and Physics* 1995;42(3):169–175.

208. Murarka S.P. Multilevel interconnections for ULSI and GSI era. *Materials Science & Engineering R-Reports* 1997;19(3–4):87–151.
209. Tong H.M. Microelectronics Packaging - Present and Future. *Materials Chemistry and Physics* 1995;40(3):147–161.
210. James C.D., A.J.H. Spence, N.M. Dowell-Mesfin, R.J. Hussain, K.L. Smith, H.G. Craighead, M.S. Isaacson, W. Shain, J.N. Turner. Extracellular recordings from patterned neuronal networks using planar microelectrode arrays. *Ieee Transactions on Biomedical Engineering* 2004;51(9):1640–1648.
211. Fair R.B. Digital microfluidics: is a true lab-on-a-chip possible? *Microfluidics and Nanofluidics* 2007;3(3):245–281.
212. Hartley L., K. Kaler, O. Yadid-Pecht. Hybrid integration of an active pixel sensor and microfluidics for cytometry on a chip. *Ieee Transactions on Circuits and Systems I-Regular Papers* 2007;54(1):99–110.
213. Huang Y., J.M. Yang, P.J. Hopkins, S. Kassegne, M. Tirado, A.H. Forster, H. Reese. Separation of simulants of biological warfare agents from blood by a miniaturized dielectrophoresis device. *Biomedical Microdevices* 2003;5(3):217–225.
214. Petrou P.S., I. Moser, G. Jobst. BioMEMS device with integrated microdialysis probe and biosensor array. *Biosensors & Bioelectronics* 2002;17(10):859–865.
215. Piruska A., I. Nikcevic, S.H. Lee, C. Ahn, W.R. Heineman, P.A. Limbach, C.J. Seliskar. The autofluorescence of plastic materials and chips measured under laser irradiation. *Lab on a Chip* 2005;5(12):1348–1354.
216. Bliss C.L., J.N. McMullin, C.J. Backhouse. Rapid fabrication of a microfluidic device with integrated optical waveguides for DNA fragment analysis. *Lab on a Chip* 2007;7(10):1280–1287.
217. Lee K.S., H.L.T. Lee, R.J. Ram. Polymer waveguide backplanes for optical sensor interfaces in microfluidics. *Lab on a Chip* 2007;7(11):1539–1545.
218. Chung K., M.M. Crane, H. Lu. Automated on-chip rapid microscopy, phenotyping and sorting of *C.elegans*. *Nat Meth* 2008;5(7):637–643.
219. El-Ali J., S. Gaudet, A. Gunther, P.K. Sorger, K.F. Jensen. Cell stimulus and lysis in a microfluidic device with segmented gas-liquid flow. *Analytical Chemistry* 2005;77(11):3629–3636.
220. Voskerician G., M.S. Shive, R.S. Shawgo, H. von Recum, J.M. Anderson, M.J. Cima, R. Langer. Biocompatibility and biofouling of MEMS drug delivery devices. *Biomaterials* 2003;24(11):1959–1967.
221. Brischwein M., E.R. Motrescu, E. Cabala, A.M. Otto, H. Grothe, B. Wolf. Functional cellular assays with multiparametric silicon sensor chips. *Lab on a Chip* 2003;3(4):234–240.
222. Szarowski D.H., M.D. Andersen, S. Retterer, A.J. Spence, M. Isaacson, H.G. Craighead, J.N. Turner, W. Shain. Brain responses to micro-machined silicon devices. *Brain Research* 2003;983(1–2):23–35.
223. <http://www.mchips.com>
224. Shawgo R.S., A.C.R. Grayson, Y.W. Li, M.J. Cima. BioMEMS for drug delivery. *Current Opinion in Solid State & Materials Science* 2002;6(4):329–334.
225. Fallahi D., H. Mirzadeh, M.T. Khorasani. Physical, mechanical, and biocompatibility evaluation of three different types of silicone rubber. *Journal of Applied Polymer Science* 2003;88(10):2522–2529.
226. Mata A., A.J. Fleischman, S. Roy. Characterization of polydimethylsiloxane (PDMS) properties for biomedical micro/nanosystems. *Biomedical Microdevices* 2005;7(4):281–293.
227. Lee J.N., X. Jiang, D. Ryan, G.M. Whitesides. Compatibility of mammalian cells on surfaces of poly(dimethylsiloxane). *Langmuir* 2004;20(26):11684–11691.
228. Millet L.J., M.E. Stewart, J.V. Sweedler, R.G. Nuzzo, M.U. Gillette. Microfluidic devices for culturing primary mammalian neurons at low densities. *Lab on a Chip* 2007;7(8):987–994.
229. Kim L., Y.C. Toh, J. Voldman, H. Yu. A practical guide to microfluidic perfusion culture of adherent mammalian cells. *Lab on a Chip* 2007;7(6):681–694.

230. Kasemo B. Biological surface science. *Surface Science* 2002;500(1-3):656–677.
231. Makamba H., J.H. Kim, K. Lim, N. Park, J.H. Hahn. Surface modification of poly(dimethylsiloxane) microchannels. *Electrophoresis* 2003;24(21):3607–3619.
232. Fritz J.L., M.J. Owen. Hydrophobic Recovery of Plasma-Treated Polydimethylsiloxane. *The Journal of Adhesion* 1995;54(1):33–45.
233. Hu S., X. Ren, M. Bachman, C.E. Sims, G.P. Li, N. Allbritton. Surface Modification of Poly(dimethylsiloxane) Microfluidic Devices by Ultraviolet Polymer Grafting. *Analytical Chemistry* 2002;74(16):4117–4123.
234. Sletz B.E., N.A. Penner, F.E. Regnier. Capillary electrochromatography of peptides on microfabricated poly(dimethylsiloxane) chips modified by cerium(IV)-catalyzed polymerization. *Journal of Chromatography A* 2002;948(1-2):225–233.
235. Ocvirk G., M. Munroe, T. Tang, R. Oleschuk, K. Westra, D.J. Harrison. Electrokinetic control of fluid flow in native poly(dimethylsiloxane) capillary electrophoresis devices. *Electrophoresis* 2000;21(1):107–115.
236. Dou Y.H., N. Bao, J.J. Xu, H.Y. Chen. A dynamically modified microfluidic poly(dimethylsiloxane) chip with electrochemical detection for biological analysis. *Electrophoresis* 2002;23(20):3558–3566.
237. Decher G. Fuzzy nanoassemblies: Toward layered polymeric multicomposites. *Science* 1997;277(5330):1232–1237.
238. Sung W.C., C.C. Chang, H. Makamba, S.H. Chen. Long-term affinity modification on poly(dimethylsiloxane) substrate and its application for ELISA analysis. *Analytical Chemistry* 2008;80(5):1529–1535.
239. Hanein Y., Y.V. Pan, B.D. Ratner, D.D. Denton, K.F. Bohringer. Micromachining of non-fouling coatings for bio-MEMS applications. *Sensors and Actuators B-Chemical* 2001;81(1):49–54.
240. Lopez G.P., B.D. Ratner, C.D. Tidwell, C.L. Haycox, R.J. Rapoza, T.A. Horbett. Glow-Discharge Plasma Deposition of Tetraethylene Glycol Dimethyl Ether for Fouling-resistant Biomaterial Surfaces. *Journal of Biomedical Materials Research* 1992;26(4):415–439.
241. Dhayal M., J.S. Choi, C.H. So. Biological fluid interaction with controlled surface properties of organic micro-fluidic devices. *Vacuum* 2006;80(8):876–879.
242. Bajaj P., D. Akin, A. Gupta, D. Sherman, B. Shi, O. Auciello, R. Bashir. Ultrananocrystalline diamond film as an optimal cell interface for biomedical applications. *Biomedical Microdevices* 2007;9(6):787–794.
243. Hoivik N.D., J.W. Elam, R.J. Linderman, V.M. Bright, S.M. George, Y.C. Lee. Atomic layer deposited protective coatings for micro-electromechanical systems. *Sensors and Actuators a-Physical* 2003;103(1-2):100–108.
244. Wang Y.L., J.H. Pai, H.H. Lai, C.E. Sims, M. Bachman, G.P. Li, N.L. Allbritton. Surface graft polymerization of SU-8 for bio-MEMS applications. *Journal of Micromechanics and Microengineering* 2007;17(7):1371–1380.
245. Wang Y.L., M. Bachman, C.E. Sims, G.P. Li, N.L. Allbritton. Simple photografting method to chemically modify and micropattern the surface of SU-8 photoresist. *Langmuir* 2006;22(6):2719–2725.
246. Nordstrom M., R. Marie, M. Calleja, A. Boisen. Rendering SU-8 hydrophilic to facilitate use in micro channel fabrication. *Journal of Micromechanics and Microengineering* 2004;14(12):1614–1617.
247. Joshi M., N. Kale, R. Lal, V.R. Rao, S. Mukherji. A novel dry method for surface modification of SU-8 for immobilization of biomolecules in Bio-MEMS. *Biosensors & Bioelectronics* 2007;22(11):2429–2435.
248. Chen C.S., M. Mrksich, S. Huang, G.M. Whitesides, D.E. Ingber. Micropatterned surfaces for control of cell shape, position, and function. *Biotechnology Progress* 1998;14(3):356–363.
249. Duncan A.C., F. Weisbuch, F. Rouais, S. Lazare, C. Baquey. Laser microfabricated model surfaces for controlled cell growth. *Biosensors & Bioelectronics* 2002;17(5):413–426.

250. Duncan A.C., F. Rouais, S. Lazare, L. Bordenave, C. Baquey. Effect of laser modified surface microtopochemistry on endothelial cell growth. *Colloids and Surfaces B-Biointerfaces* 2007;54(2):150–159.
251. Edell D.J., V.V. Toi, V.M. McNeil, L.D. Clark. Factors influencing the biocompatibility of insertable silicon microshafts in cerebral-cortex. *Ieee Transactions on Biomedical Engineering* 1992;39(6):635–643.
252. Hoogerwerf A.C., K.D. Wise. A 3-dimensional microelectrode array for chronic neural recording. *Ieee Transactions on Biomedical Engineering* 1994;41(12):1136–1146.
253. Schmidt S., K. Horch, R. Normann. Biocompatibility of silicon-based electrode arrays implanted in feline cortical tissue. *Journal of Biomedical Materials Research* 1993;27(11):1393–1399.

Packaging of Biomolecular and Chemical Microsensors

Peter J. Hesketh and Xiaohui Lin

Abstract Packaging of microsensors is not standardized and highly customized depending upon the requirements for a specific application. This chapter reviews selected illustrative examples for chemical and biosensors for a range of applications. Approaches range from modified semiconductor device packages, fully customized packages to packages produced by rapid prototyping stereolithography. Packaging for room temperature operation of sensors, such as ion selective FETs, microelectrochemical sensors, microcantilever sensors, in addition to sensors which operate at higher temperatures including microhotplate sensors and electrochemical gas sensors is discussed. Also biomedical sensors and microfluidic devices including microneedles, soft lithography for microchannel fabrication, bonding, and packaging materials are discussed.

Keywords Chemical sensor · Gas sensor · Biosensor · Stereolithography · Bonding · Microcantilever · CHEMFET · ISFET · PDMS · Microfluidic · Microneedle · Soft lithography · Microelectrode · Microhotplate · MEMS · Tin oxide · Microchannel

Abbreviations

3D	Three-dimensional
AFM	Atomic force microscope
ALD	Atomic layer deposition
BCB	Benzocyclobutene
CAAT	Silver-coated nickel particles in foil
CAD	Computer-aided design
CE	Capillary electrophoresis
CHEMFET	Chemically sensitive field effect transistor
COC	Cyclic olefin copolymer
CVD	Chemical vapor deposition

P.J. Hesketh (✉)

School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: peter.hesketh@me.gatech.edu

DIL	Dual in line
DRIE	Deep reactive ion etching
DMFC	Direct methanol fuel cell
FEA	Finite element analysis
FIB	Focused ion beam
ISFET	Ion selective field effect transistor
LPC	Liquid phase chromatography
MEMS	Microelectro-mechanical systems
MSI	Microfluidic system interface
MTS	Mercapto propyl trimethoxy silane
NEMS	Nano electro-mechanical systems
OD	Outside diameter
PCB	Printed circuit board
PCR	Polymerase chain reaction
PDMS	Polydimethylsiloxane, or silicone rubber
PI	Polyimide
PMMA	Polymethylmethacrylate
PTFE	Polytetrafluoroethylene
RIE	Reactive ion etching
SAW	Surface acoustic wave
SEM	Scanning electron microscope
SLA	Stereolithographic assembly
SOI	Silicon on insulator
SPE	Solid phase extraction
STL	Stereolithography interface format
SU-8	SU 8 epoxy resin
USB	Universal system bus
UV	Ultraviolet
VLSI	Very large scale integration
WLP	Wafer-level packaging

List of Symbols

Ag	Silver
Au	Gold
Co	Cobalt
Cu	Copper
Fe	Iron
Ni	Nickel
T_g	Glass transition temperature
TiN	Titanium nitride
Pd	Palladium
Pt	Platinum

1 Introduction

The purpose of packaging a chemical sensor or biosensors is to expose a region to the analyte of interest, while providing mechanical support and both electrical and fluid connections for the sensor. The mechanical support is important, for example, if a membrane has to be held in place against the sensor. There are additional specific functions for a package. It can also include selective filter(s) to remove specific interfering species, or a valve/pump to control the sample flow over the sensor or sensor array. The package also defines the volume of analyte in which a measurement is made, when configured in a microfluidic format. Finally, the thermal environment of the sensor is affected by the package, and it is important because the rate of reaction for chemical sensors is a function of temperature. In addition, many biological assays must be carried out at specific temperatures and have temperature limits to avoid denaturation of biological selective molecules.

Packages for electronic devices are well developed [1] and for electronic microsystems [2]. Commercially available semiconductor packages can be, and often are adapted for sensor applications. For example, by mounting the die in the semiconductor package header any customization can be accomplished with modifications to the cover. The sensor or sensor array will be exposed to the analytes while electrical contacts are made through wire bonding. This approach has been very successful in many applications, particularly for gas sensors. There are few standard packages defined for sensors. Low-cost packaging for sensors are discussed by Tomescu [3] and in a recent book chapter by Brand [4]. Packaging for BioMEMS is discussed by Velten et al. [5] and some case studies for MEMS packaging are described by Lee et al. [6] and suggestions for future work along with some of the unsolved problems in packaging.

The functions of a package can be summarized as follows:

- Provide mechanical support and protection for sensor in the intended application
- Provide electrical connections to the sensor(s)
- Deliver the sample to the sensing interface
- Control thermal environment by creating regions with differing operating temperatures when required
- Protect any electronic sensor interface circuits from the environment
- Provide convenient integration at the system level
- Provide ease of calibration for sensor(s)
- Improve reliability
- Automate assembly
- Identify component part and date of manufacture

For nanotechnology the issue of packaging has not yet been developed. A future challenging issue for packaging is that specifically how to take advantage of scaling for nanoscale sensors. In fact, both the alignment and electrical contacts to the nanowires have to be carefully considered. When the nanowires are grown in place with a catalytic seed, the placement is predetermined by the catalyst, however,

otherwise alignment has to be carried out with methods such as fluid flow, electric field, or AFM probes to direct the position of the nanowires.

When we consider the intended duration for use of the sensor it has an important influence on the package type and material selection. When short-term use is required then stability and corrosion resistance are not as much a concern; however, in long-term use, these issues, in addition to hermeticity for protection of electronic interface circuit, also will be a requirement. For chemical and biomolecular sensors non-hermetic packaging is necessary because the sensor interface needs to be in contact with the environment to provide the sensing function. However, other MEMS devices such as microresonators require hermeticity in an inert environment or a vacuum to improve performance and stability.

Electrical connections are made through wire bonding or flip-chip bonding. Methods developed for semiconductor devices are routinely used for sensors. Packaging is a large topic and can hardly be covered in one chapter; therefore the scope here is limited to packaging for chemical and biomolecular sensors. Selected examples are used to illustrate the range of different approaches that have been developed, often for specific applications.

2 Chemical Sensors

A plethora of chemical sensing principles have been applied in a miniature format. Many excellent reviews of chemical sensors have been published, for example, the texts by Gardner [7], Peirce [8], Kovacs [9], and the seven volume series published by Wiley-VCH [10]. Since the first development of the ion selective field effect transistor (ISFETs) [11], chemically sensitive field effect transistor (CHEMFETs) [12], ion-controlled diodes [13], palladium gate gas sensor [14], and miniature Clark electrode [15], there has been a host of new development. Other miniature chemical sensors are based on miniaturization of larger scale sensors electrochemical sensor electrodes [16] or the pelistor or semiconductor metal oxides gas sensor for high temperatures and metal semiconductor (MOS) gas sensors.

2.1 ISFET Packaging

A recent review of ISFET devices is provided by Poghossian and Schoning [17]. Here we are concerned with the development of suitable packaging for the ISFET. Approaches include that of Smith et al. [18] which are based on precision hand assembly. The packaging of ISFETs has been challenging to avoid instability and drift that arise due to changes in impedance and leakage currents [19, 20] or aluminum electrodes neighboring the gate region to indicate degradation of the encapsulation [21]. The extended gate FET developed by van der Spiegel et al. [22, 23] located the chemically sensitive region a distance from the electronic device to reduce the requirements on encapsulation, as shown in Fig. 1. This method has

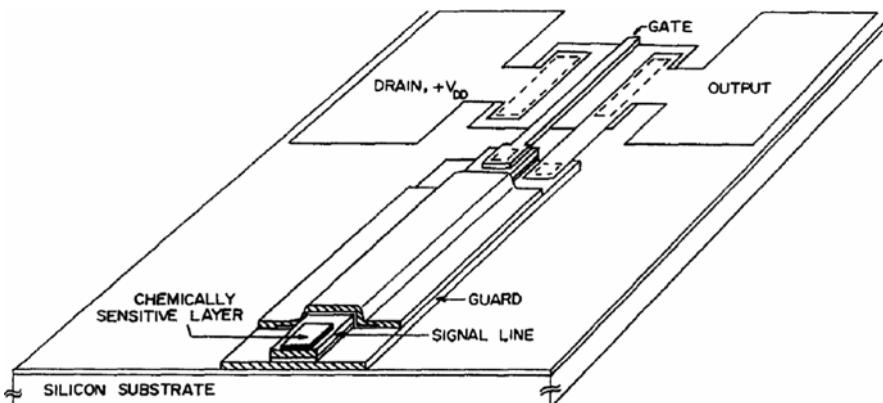


Fig. 1 Schematic diagram of extended gate field effect transistor. Reprinted from [22] with permission from Elsevier © (1983)

been applied to tin oxide sensing by Chi et al. [24]. One of the key issues of using ISFETs is the necessity for a reference electrode, unless the device is operated as a Kelvin Probe [25]. This implies the system should include packaging of the miniature reference electrode to provide a functional measurement system. The advantages of miniaturization of FETs will be important for arrays where the sensors can share the same reference electrode. Differential FETs where one device is passivated or a device is functionalized for different analyte sensitivity provide a method of addressing the reference without a reference electrode. Back contact ISFETs were developed by van den Vlekkt et al. [26] with a novel micromachining process to keep the electrical contact isolated from the working fluid by placing them on the other side of the wafer, see Fig. 2. In addition, needle-type ISFETs have been demonstrated by Matsuo and Esashi [27] and the needle-type ISFETs extend the device into a probe which might be particularly useful for biomedical applications.

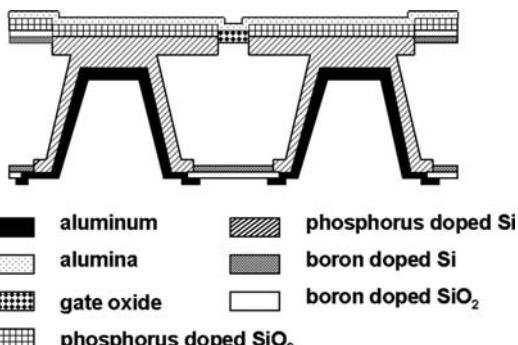


Fig. 2 Cross-sectional schematic diagram of a FET with back contact. Reprinted from [26] with permission from Elsevier © (1988)

Photocurable polymer encapsulation was investigated by several groups [28, 29, 30]. A novel three-dimensional ISFET built on a mesa structure has been developed to facilitate packaging of the devices [31] as shown in Fig. 3. A Kapton (a flexible polyimide film, trademark of DuPont) foil is defined over the device which has electrical contacts integrated. Assembly which involves a manual alignment and bonding with a controlled collapse structure and heating to 300°C with pressure applied for 2 min. Epoxy resin is poured over the back of the Kapton tape to support the structure.

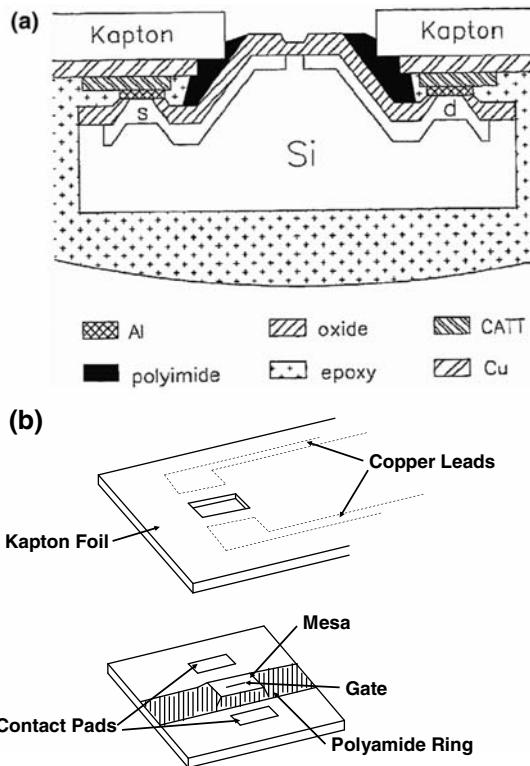


Fig. 3 (a) Cross-sectional view of a fully encapsulated mesa ISFET and (b) alignment of the Kapton foil over the mesa ISFET. Reprinted from [31] with permission MYU publishing (Tokyo)

The photopolymerization of encapsulation resin provides encapsulation that is effective and has low leakage currents. The process involves first coating with an adhesion promoter, mercapto propyl trimethoxy silane (MTS), to ensure attachment of the resin to the silicon oxide/nitride layer of the ISFET followed by several coatings, each of 30 μm thickness, exposure to UV radiation in the areas where the polymer should be removed from the gate and contact pads, and finally development

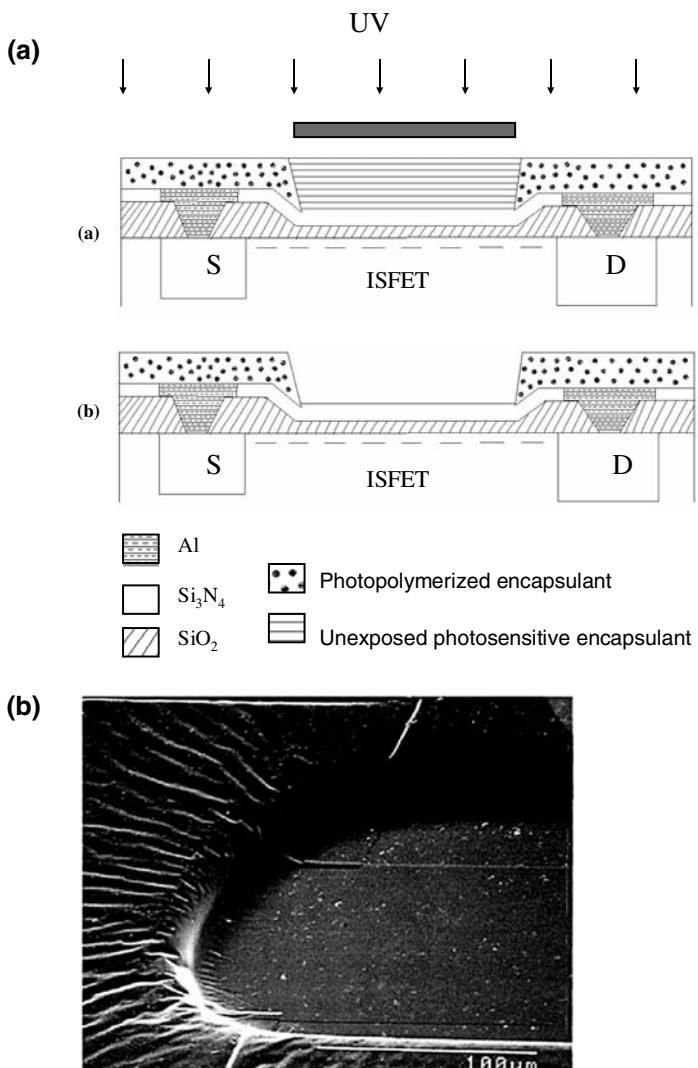


Fig. 4 (a) Schematic diagram of the direct photopolymerization of UV curable encapsulants on ISFET devices and (b) micrograph of gate-photocurable encapsulant for a specific formulation. Reprinted from [29] with permission from Elsevier © (2005)

in isopropyl alcohol and curing. Figure 4 shows the summary of the process and the encapsulated sensor package.

A large array of ISFETs has been studied for extra-cellular recording by Milgrew et al. [32]. Figure 5 shows the array of 16×16 ISFETS along with signal

conditioning and readout circuitry. The microfluidic chamber containing the cells is positioned over the array so that the cell culture can be grown over the array. A SU-8 photodefined channel is used to define the chamber, as indicated in the figure with a polydimethylsiloxane (PDMS)-molded structure defining the internal volume for the measurements. Electrical connections to the ISFETs are made with wire bonding and SU-8 (a photodefinable multifunction bis-A type epoxy) coating to provide electrical insulation on the PC board mounted die.

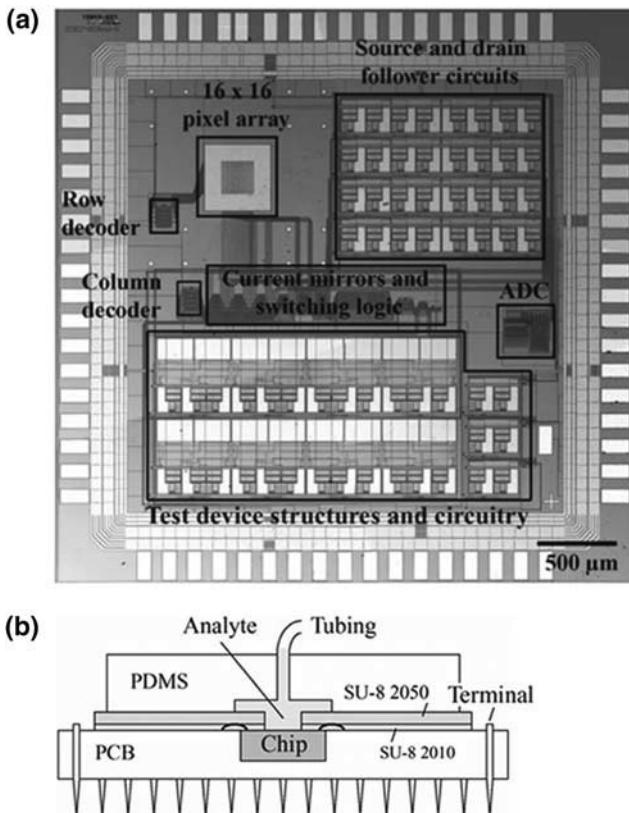
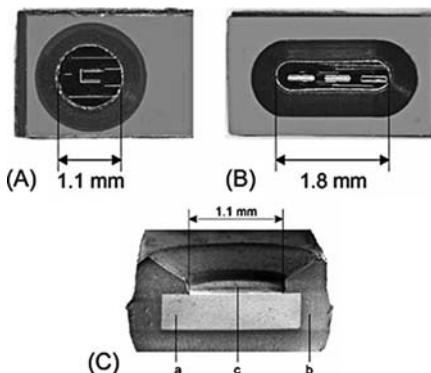


Fig. 5 (a) Micrograph of an integrated ISFET sensor array chip and (b) schematic diagram of the chip packaging process. Reprinted from [32] with permission of the ECS

A comprehensive review of ISFET packaging is provided by Oelßner et al. [33]. Technology for encapsulation of ISFETs includes the use of molds developed at KSI Meinsberg (Germany). The molds are a methylpentene copolymer produced with an injection molding. They provide the shape for casting of the epoxy resin over the sensor body and integration of electrical wires. Figure 6

Fig. 6 (a) Circular shape and 1.1 mm diameter measuring cavity; (b) elliptical-shaped cavity with three sensors; and (c) micrograph of an ISFET sensor across the cavity (a) ISFET chip, (b) encapsulation, (c) measuring cavity. Reprinted from [33] with permission from Elsevier © (2005)



shows the ISFET-encapsulated sensor package, the gate region is in a small circular region, or can be extended to include reference FET junctions. Probes can be placed in small fluid volumes and thereby allow study of corrosion, applications in food processes, or medical applications including ophthalmology, dermatology, gynecology, and dentistry. For example, Fig. 7 shows a needle-based encapsulation design with a pre-formed mold, as shown, for a single or multiple ISFET devices. Figure 8 shows the straight-forward encapsulation technology processing steps that use a thermoplastic elastomeric mold and heat-cured epoxy resin injected into the mold.

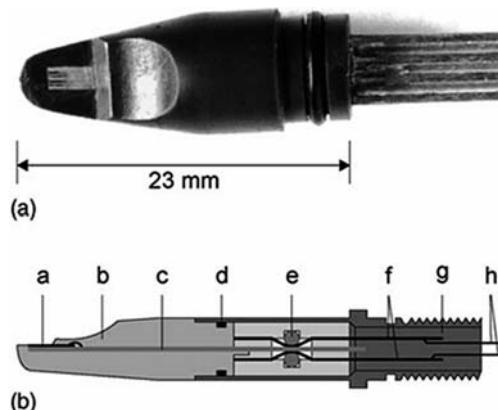


Fig. 7 (a) Photograph of an ISFET sensor element encapsulated according to molding process shown in Fig. 8 by KSI, Meinsberg, Germany. (b) Inserting and contacting the ISFET element of the sensor include (a) ISFET chip, (b) epoxy encapsulation, (c) printed circuit board, (d) O-ring sealing, (e) elastic contacting elements, (f) flexible printed circuit boards, (g) sensor stainless steel shaft, and (h) connecting wires. Reprinted from [33] with permission from Elsevier © (2005)

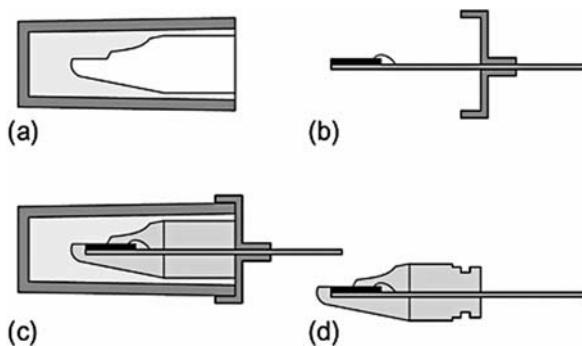


Fig. 8 Technology processing steps for hot-curing epoxy molding for the encapsulation of ISFET chips at the laboratory scale: (a) molds made from thermoplastic elastomer; (b) the chip is glued onto a circuit board and bonded and fixed to the cover of mold; (c) The chip carrier is positioned in the mold, cast, and hot-cured with epoxy under specific conditions; and (d) only minor finishing work is necessary after the mold release. Reprinted from [33] with permission from Elsevier © (2005)

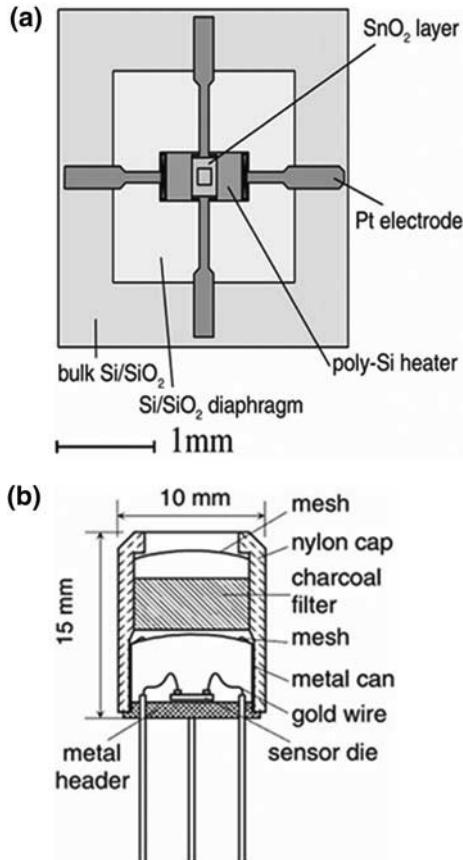
2.2 Microhotplate Packaging

Metal oxide chemical sensors require controlled elevated temperature for their correct operation. The package must provide thermal control and isolation in addition to fluidic and electrical connections to the device. Early work on microhotplates was pioneered by Suehle et al. [34] and with the improvements in tin oxide (SnO_2) film quality provided by nanostructured microspheres [35] and nanostructured SnO_2 grains [36] significantly improved sensing platform has been demonstrated. Excellent reviews are published by Simon et al. [37] and Semancik and Cavicchi [38].

Figure 9 shows a schematic diagram of a micromachined gas sensor developed by Motorola Inc. [39]. The tin oxide is formed by rheotaxial growth on a thin silicon nitride membrane to provide thermal isolation and hence reduce the power consumption in the range 30–150 mW. This allows the chip support to operate close to ambient temperature so that standard semiconductor package can be used as indicated in Fig. 9b. A charcoal filter is included to remove unwanted gases that might interfere with tin oxide response and a nylon cap with mesh opening for mechanical protection. The reliability of the thin silicon oxide–nitride membrane to thermal and mechanical shock was investigated with accelerated testing to establish reliability for this device in real world applications [39].

Figure 10 shows a microhotplate which utilizes a polysilicon heater and operates up to 500°C with just 22 mW of electrical power. Femtomolar isothermal desorption has been carried out by Shirke et al. [40] with heating rates of up to

Fig. 9 Schematic diagram of the MGS 1100 sensor from Motorola Inc. The micromachined element on the *left* and the sensor housing on the *right*. The sensitive film was obtained by rheotaxial growth and thermal oxidation of tin layers deposited on to the silicon oxide–nitride membrane. Reprinted from [37] with permission Elsevier © (2001)



10⁶°C/s and minimal power consumption due to the small thermal mass of the microhotplates. The technique has been demonstrated for the isothermal desorption of benzoic acid from a single SnO₂ covered microhotplate at 23–174°C. In addition, the microhotplate has been demonstrated as a sensitive device for the detection of chemical warfare agents at 5 nmol/mol (ppb) concentration levels and it shows stable operation for 14 h of use at elevated temperatures and with pulsed operation of 200 ms over the temperature range of room temperature to 480°C [41].

Sensitive and selective response using microhotplates with ultra-thin metal film coatings (Ag, Co, Cu, Fe, Ni) to volatile organics has been reported by Panchapakesan et al. [36]. The SnO₂ is deposited by selective CVD by heating the microhotplates to 500°C, see Fig. 11. The amount of material deposited was monitored by recording when the resistance reaches 2 k ohms. The resultant SnO₂ grain

Fig. 10 Photograph of a typical four-element microhotplate to sense film deposition. The thin silicon dioxide structure $100 \times 100 \mu\text{m}$ with a serpentine polysilicon heater is suspended with four silicon dioxide beams which have embedded polysilicon runners. The four TiN contact electrodes are defined on the top of the microhotplate. Reprinted with permission from [41] © 2005 IEEE

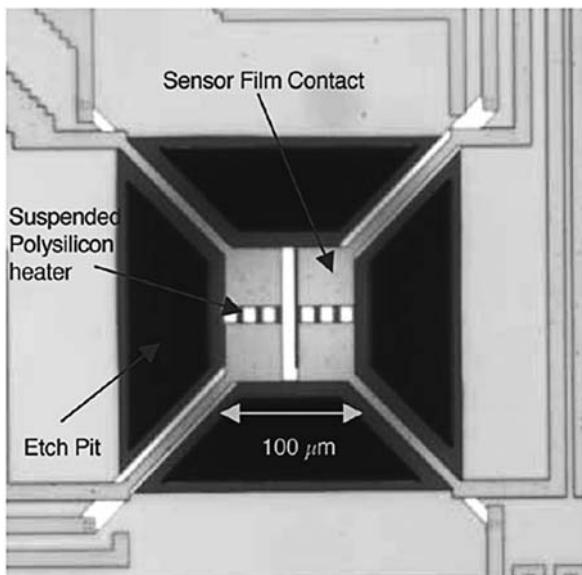
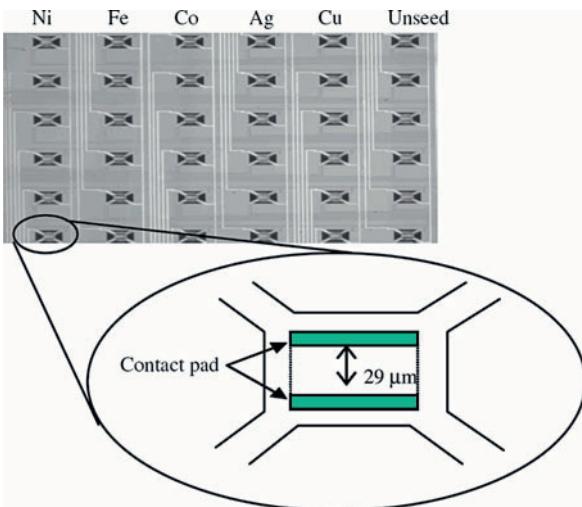
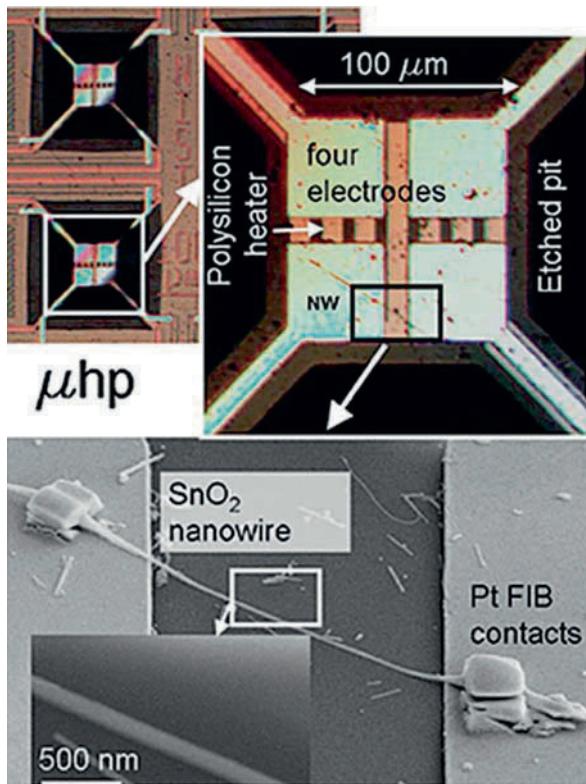


Fig. 11 Optical image and schematic illustration of the 36-element microhotplate array. Each hotplate has a different metal catalyst deposited onto it. The amount of material is monitored by measuring the conductivity between pair of contact pads. Reprinted from [36] with permission of IOP Publishing



size was in the range of 20–121 nm. Stable sensor operation for 200 h with methanol is reported at 80 ppm concentration levels. Finally, the combination of nanowires of SnO_2 with the microhotplates has been demonstrated to produce enhanced performance at lower operating temperature, reduced power consumption, and good

Fig. 12 (a) Optical image of a typical microhotplate and (b) SEM image of the individual tin oxide nanowire integrated as a chemiresistor using FIB technology. Reprinted from [42] with permission of AIP



stability [42]. Figure 12 shows the hotplate with a single SnO_2 nanowire integrated with FIB deposition necessary to make an electrical contact to the nanowire. The sensor response to reducing gases was observed and due to the low thermal mass a rapid examination of response as a function of temperature could be carried out to determine optimum operating conditions for highest gas sensitivity.

The KAMINA microarray tin oxide sensor developed by Goschnick [43] has 38 segments and operates at 1 W power consumption with a thin film of porous silica over the array with a gradient of thickness to produce differential response to different gases. The sensors operate at an elevated temperature, typically 300°C, maintained by thin film Pt heaters on the back of the substrate, as indicated in the insert of Fig. 13a. The sensor is wire bonded into a ceramic board, shown in Fig. 13b. The surface temperature is a function of the gas flow in this package, which can vary by up to 10°C; however, with the addition of a heat exchanger at the inlet, built-in stainless steel inlet, this variation is greatly reduced [44].

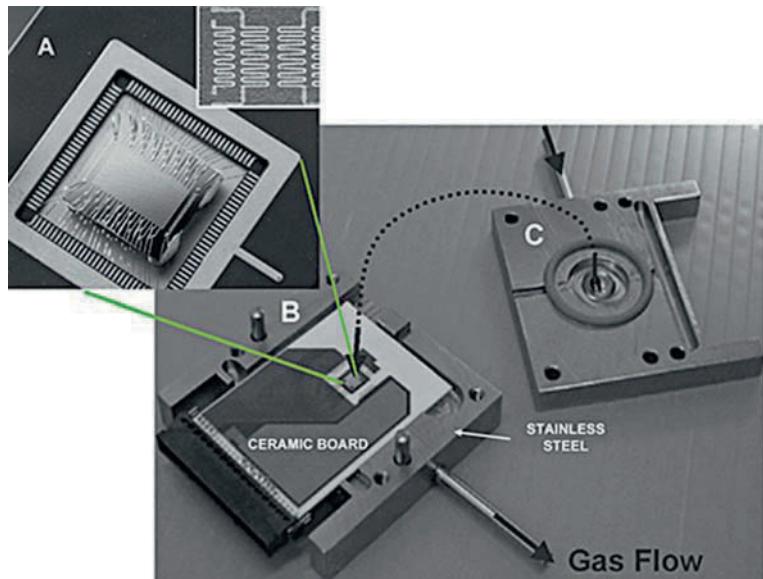


Fig. 13 The package of the KAMINA chip mounted with the ceramic board which is enclosed by two stainless steel parts providing a cavity above and below the chip through which the gas is delivered to the sensor. Reprinted from [44] with permission Elsevier © (2007)

Finally, in the lid of the housing there is also included a temperature and humidity sensor (Fig. 13c).

Polyimide-based microhotplates offer low cost and low thermal conductivity substrate for tin oxide gas sensors, thereby lowering the power consumption compared to silicon-based sensors. The operating and fabrication temperatures are less than 450°C [45]. There are several designs considered, as shown in Figs. 14 and 15. The hotplate is defined on polyimide sheets of Upliex-S ($T_g \sim 500^\circ\text{C}$, 50 μm thick, Hitachi Chemical Co.) with a spin on film of PI2731 ($T_g \sim 350^\circ\text{C}$ from DuPont) and the spin-coated polyimide on a silicon substrate. The heating element is a thin Pt film photolithographically defined and the tin oxide solution was deposited by dropping a microliter volume of a tin-containing solution onto the Pt conductivity sensor, followed by annealing at 450°C for 10 min. The power input to achieve a temperature of 325°C was in the order of 60 mW and the calculated temperature distribution over the sensing area for the silicon-based device as shown in Fig. 15a is plotted in Fig. 15b. The polyimide layer was spin coated on to the oxidized silicon substrate and the removal of the bulk silicon under the heated platform carried out with deep RIE etching. Packaging was achieved by assembly

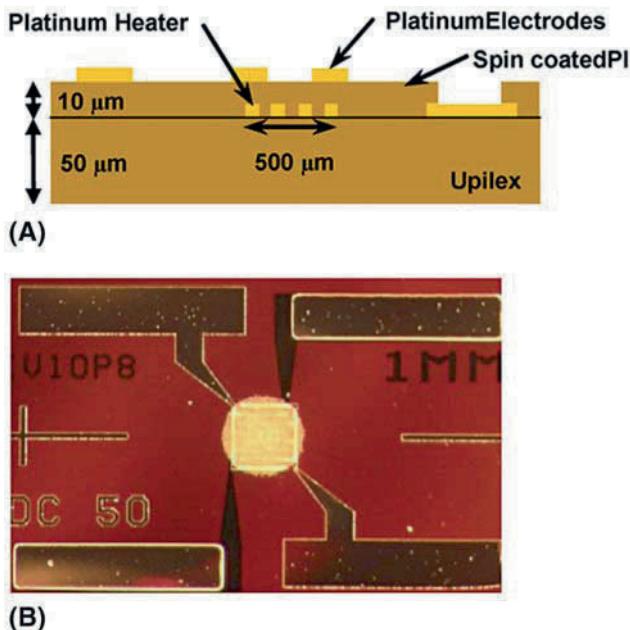


Fig. 14 Complete metal oxide microhotplate gas sensor made on a polyimide sheet, (a) cross-sectional diagram and (b) photograph of completed sensor, with 0.5 mm wide heating area. Reprinted from [45] with permission Elsevier © (2008)

of a gas permeable membrane to the sensors, placed into a TO-5 (transistor-outlined package-5) header with an epoxy, and wire bonding for electrical connections [46].

2.3 Microcantilever Packaging

Vacuum packaging of innovative hollow microcantilever resonators has been demonstrated by Burg et al. [47]. When the analyte flows through the inside of the cantilever resonator a frequency shift is observed. They are operated at a reduced pressure or vacuum environment, to ensure a high Q, and therefore sensitive measurement of the frequency shifts due to selective adsorption inside the hollow cantilever. SEM micrographs of the cantilever are shown in Fig. 16. The electrical contacts are located on the glass lid and fluid is injected into the hollow channel through pits defined through the thickness of the wafer. The sensors are attached to an adhesive backed gold-plated printed circuit board (PCB) and wire bonded. The assembly is clamped with Teflon (trademark from DuPont) against a PTFE manifold holding an array of 1/32" tubes through a polytetrafluoroethylene

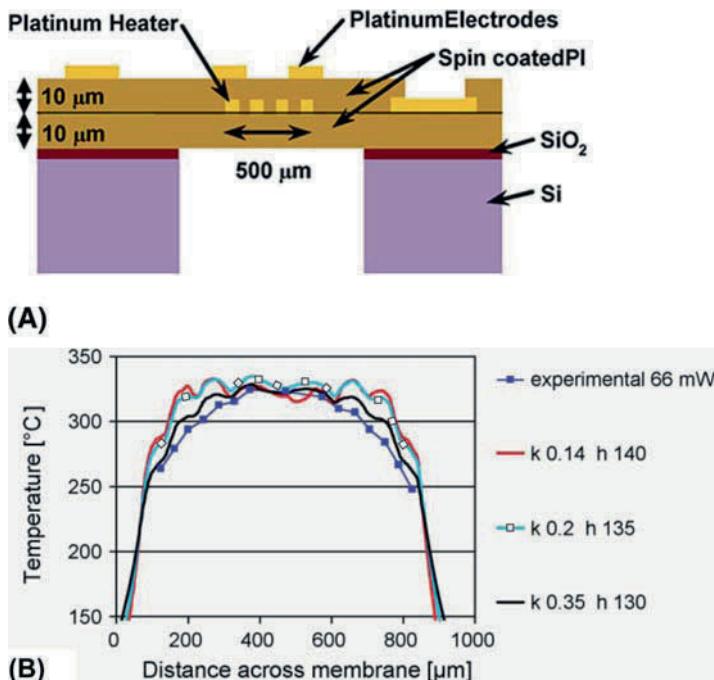


Fig. 15 (a) Schematic diagrams of the microhotplate gas sensor structure realized on a silicon substrate and (b) temperature distribution over the sensing area obtained from simulations, with thermal conductivity, k ($\text{W}/\text{m K}$) of Pt layer as a parameter and for several values of convection coefficient, h ($\text{W}/\text{m}^2 \text{ K}$). Reprinted from [46] with permission Elsevier © (2008)

(PTFE) block as shown in Fig. 17 and the fluid connections have an “O” ring-sealed interface.

Work has also been performed at ETH (Switzerland) on CMOS-integrated microcantilever sensor arrays includes their novel packaging [48]. Modified semiconductor packages such as the dual-in-line (DIL) packages wire bonded into customized packages with a ceramic support and a partial glob top to protect the electronic components while leaving the sensor area exposed. An array of piezoresistive microcantilevers with standard packaging is shown in Fig. 18. The piezoresistive detection method provides an attractive alternative to optical beam deflection methods normally used with atomic force microscope (AFM) probes. Here the probes have integrated silicon tips which are formed at the beginning of the cantilever micromachining process on a silicon on insulator (SOI) wafer [49]. A microheater is defined to allow actuation of the cantilever, with a piezoresistive sensor to detect deflection, thereby allowing feedback control of probe motion for an AFM imaging.

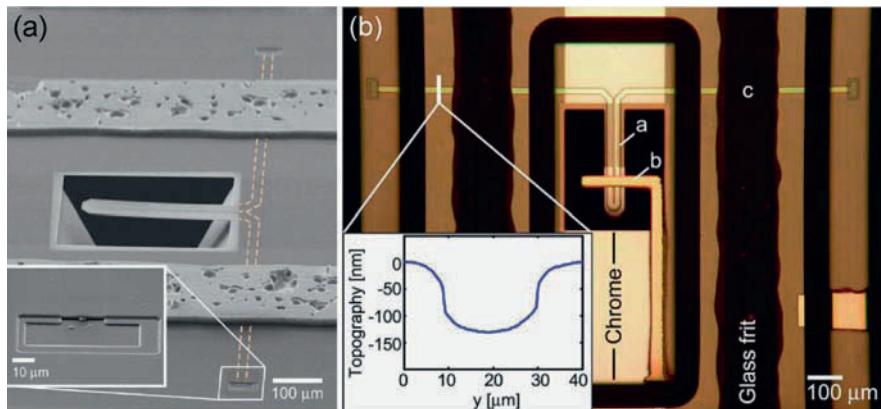


Fig. 16 (a) SEM micrograph of cantilever and glass frit seal after removal of the glass lid. The inset shows the channel inlet, and (b) SEM micrograph of packaged cantilever resonator. The 300 μm long beam contains a $1 \times 20 \mu\text{m}$ microfluidic channel; a – electrode on the glass surface above the cantilever enables electrostatic actuation; b – glass frit conforms to the surface topography (inset) and does not collapse the thin channel locations; and c – during bonding. Reprinted with permission from [47] © 2006 IEEE

Packaging for resonant beam sensor arrays has also been developed [50]. The array of four resonant beam sensors with an integrated CMOS electronics were designed for measurement of volatile organics in a liquid water sample. Packaging is accomplished with a custom polymethylmethacrylate (PMMA) fluid housing fitted with threaded connectors for miniature liquid phase chromatography (LPC) as shown in Fig. 19. The cantilever vibration is excited by an external magnetic field and based upon generation of a Lorenz force from the current flow in a conductor loop on the beam, I, as indicated in Fig. 19b. A miniature microcantilever sensor for electronic nose applications is packaged with SU-8 gasket and stainless steel housing by Pinnaduwage et al. [51], see Fig. 20.

Customized three-dimensional packages can be made by stereolithographic assembly (SLA) which is discussed in detail in a later section of this chapter. Figure 21 shows a typical gas sensor package made using SLA. A cantilever gas sensor is placed in the chip housing with its cantilever array exposed to gas which flows through the channel. A ten microcantilever array with piezoresistive readout is built on a SOI wafer with standard micromachining processes [52]. The sensors are typically 450 μm in length and positioned on one side of the $6 \times 13 \text{ mm}$ die. A recessed region is defined in the SLA base to provide alignment of the die so that the cantilevers are positioned near the middle of the flow channel. Contact pads were then wire bonded to metal pins for signal output. Details of the SLA fabrication process are given later in this chapter in Section 3.4.

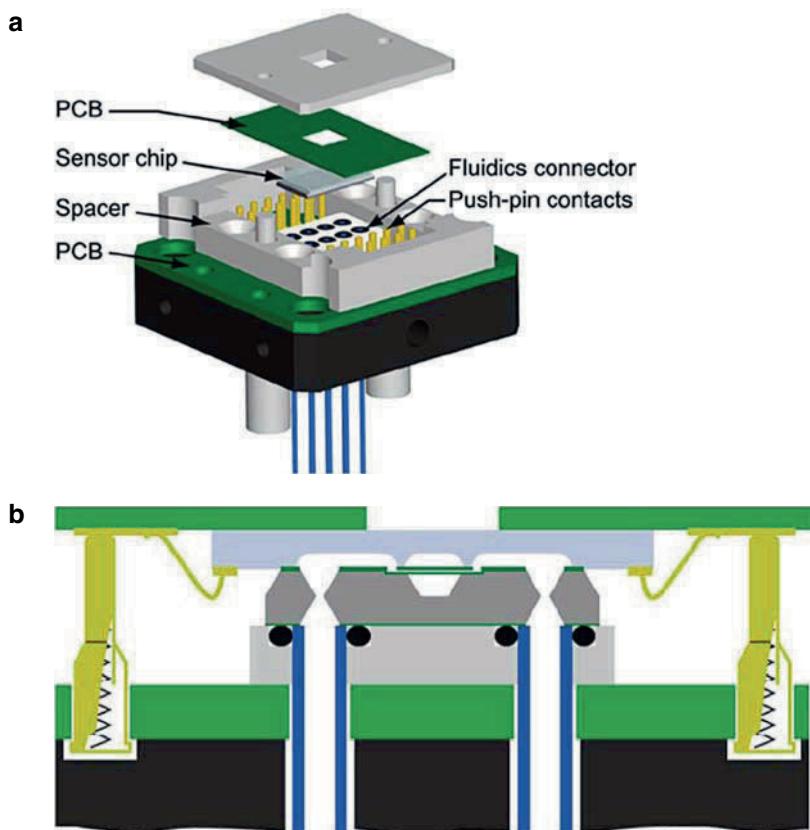


Fig. 17 (a) Bottom view of the sensor chip wire bonded to PCB. A colloidal silver paste is applied to the corner of the chip to connect the chromium layer on the resonator surface and a gold trace on the glass lid. (b) The sensor chip is clamped onto a Teflon manifold holding an array of standard 1/32" OD Teflon tubes. Perfluoroelastomer O-rings seal the manifold to the chip, and an array of spring loaded contacts provides the electrical connections to the PCB. Reprinted from with permission from [47] © 2006 IEEE

2.4 Metal Oxide High-Temperature Sensor Packaging

Higher temperature sensors have unique challenges and requirements for high-reliability applications. In particular for space applications a number of sensors have been developed by Hunter et al. at NASA [53]. A careful study at elevated temperatures was carried out to determine the effect of alloy composition on the sensor response and stability. To this end standard packages have been modified very successfully to provide reliable working systems, as shown in Fig. 22. A series of other analytes, including hydrazine, hydrogen, and oxygen has also been investigated. Figure 22a shows the hydrogen sensor which operates at 425°C [54]. The

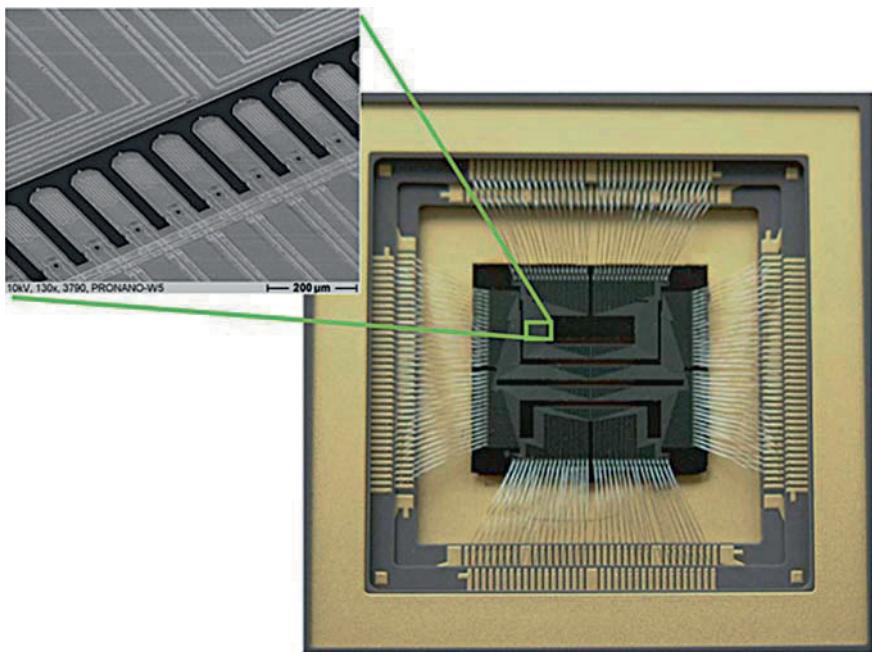


Fig. 18 A packaged fully addressable self-actuated piezoresistive cantilever array VLSI NEMS-chip incorporating 32 and 128 proximal probes and details of the array with the integrated silicon tips. Reprinted from [49] with permission Elsevier © (2001)

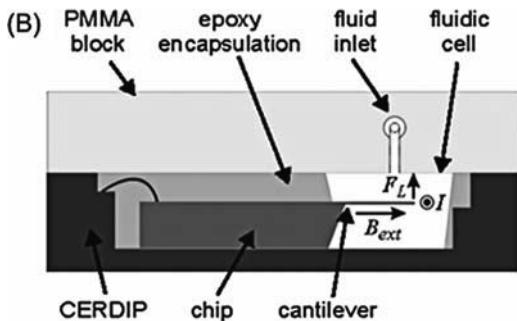
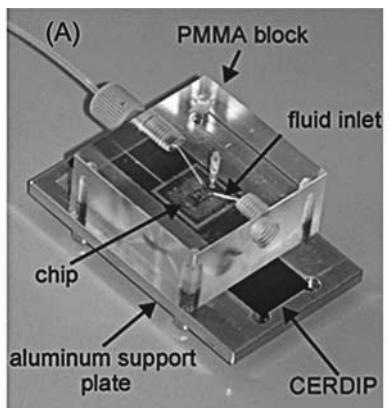
oxygen sensor, shown in Fig. 22b, utilizes conductivity measurements of a zirconia and operates at higher temperatures with a ceramic package.

2.5 Microfluidic Sensor Array Packaging

The function of a package is to direct fluid flow and provide mixing, flow splitting, filtering functions, and expose the sensor array to the analyte under controlled conditions. The design must accommodate flow of sample to the sensor array with necessary sampling volume, and there may be a distribution of concentration over the surface of the sensing interface depending upon specific flow conditions. Such deliberate modulation of the flow of analyte through the channel gives rise to a higher order chemical sensing [55] in which the transient response is recorded as a function of flow, rather than the steady-state response. Finally, the thermal design for constant temperature operation may also be critical for correct sensor operation.

Novel packaging with SU-8 for optical sensors that provide both fluid channels and optical to fiber connectors is described by Ruano-Lopez et al. [56]. In

Fig. 19 (a) Photograph of the packaged single chip system with mounted PMMA block for fluid handing; (b) schematic cross-section of the packaged cantilever chip. The direction of current, I , at the cantilever tip indicates out of the plane [50]. Reprinted with permission from the American Chemical Society © (2007)



addition optical sensors built with Mach-Zehnder interferometer on a CMOS chip are described by Blanco et al. [57]. Chang-Yen and Gale reported a microfluidic package for an optical biosensor array [58]. On this package the optical waveguide sensor was printed on a glass substrate and optical fibers are aligned on the die with tapered SU-8 channels. Fluidic package was constructed of PDMS, and is shown schematically in Fig. 23.

A variety of polymers, including SU8, had been applied to packaging of surface acoustic wave (SAW) sensors by Georgel et al. [59]. In this work the SAW device is flip-chip bonded to a PCB substrate providing electrical connections and a polymer sealing ring on the same substrate. The photosensitive SU-8 lithographically patterned was compared with a non-photosensitive benzocyclobutene (BCB, from Dow Chemical) for sealing the SAW sensors to the substrate. The SU-8 layer was 30 μm in thickness and after bonding shear strength test was performed and results are plotted in Fig. 24. The stud bump formed of NiAu provides electrical connections from the SAW sensor to the PCB to facilitate testing of the sensor.

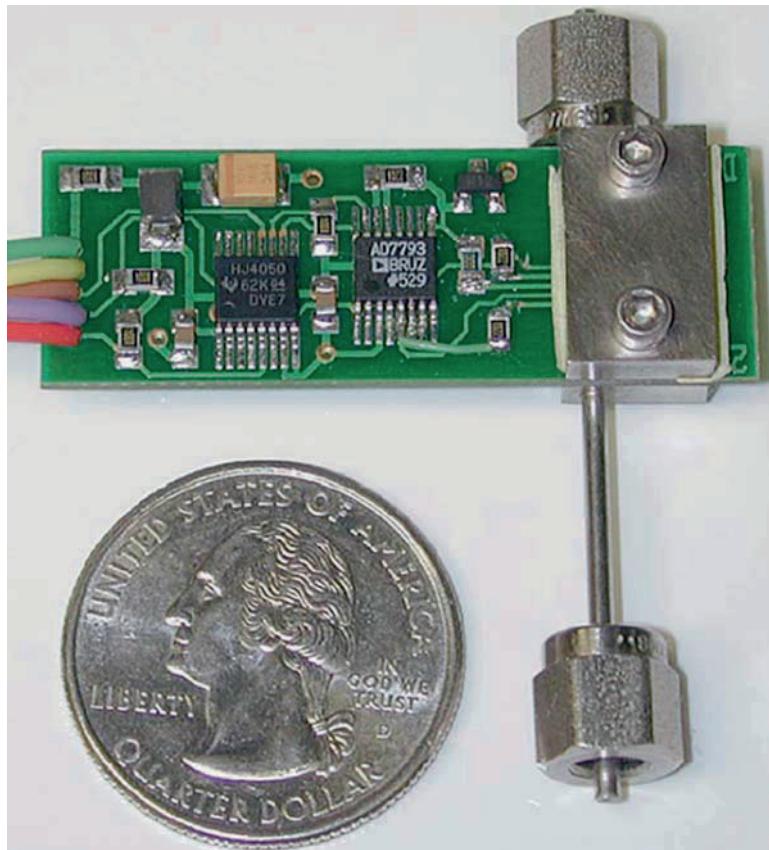


Fig. 20 Photograph of the sensor board with flow cell and interface electronics which allow the sensor to be connected to a computer USB port. Reprinted from [51] with permission of AIP

Miserendino and Tai [60] describe a fully integrated microgaskets for MEMS sensors which make use of an SU-8 layer and a photosensitive silicone layer. The microgaskets fabrication process is shown schematically in Fig. 25. The microgasket is integrated with 3 μm thick Parylene film to form microfluidic channels defined with a photoresist sacrificial layer. The microgaskets were capable of deforming approximately 25% and this behavior was characterized through a novel capacitance method. The microgaskets are re-usable and able to form leak-free seals operating at pressures up to 50 psi. The dead volume was estimated with finite element analysis (FEA) and found to be 9 nl per connection via, however, this can be reduced with a reduction in channel diameter.

Low-cost packaging has been manufactured by injection molding of cyclic olefin copolymer resin (COC) [61]. The COC polymer resin is heated to the liquid state and then injected into the mold. After the molding process the mold is cooled and the package is formed. Figure 26 shows an example of modular package for

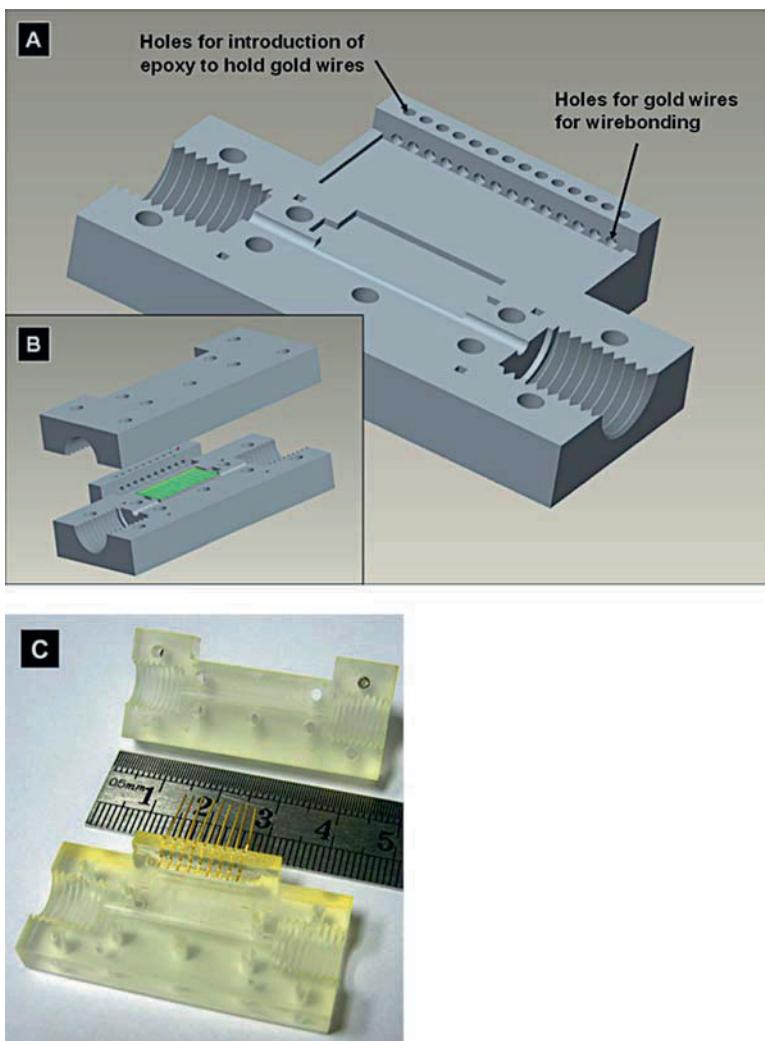


Fig. 21 Microcantilever packages built by stereolithography. **(a)** CAD image of the base of the flow cell threaded for connection to standard LPC connectors and indicating position for inserting gold wires for electrical connections. **(b)** CAD image of the two parts of flow cell where the sensor die is shown positioned next to the flow channel. **(c)** Photograph of the SLA package parts with the wire inserted [96]

fluidic components that can be snapped together to connect multiple sensors with standard diameter 1/8" OD tubing. The snap-in connector with standard tubing is combined with an electrode for an electrochemical-based chlorine sensor and a thin film resistor temperature sensor for an aqueous sample, as shown in Fig. 26c.

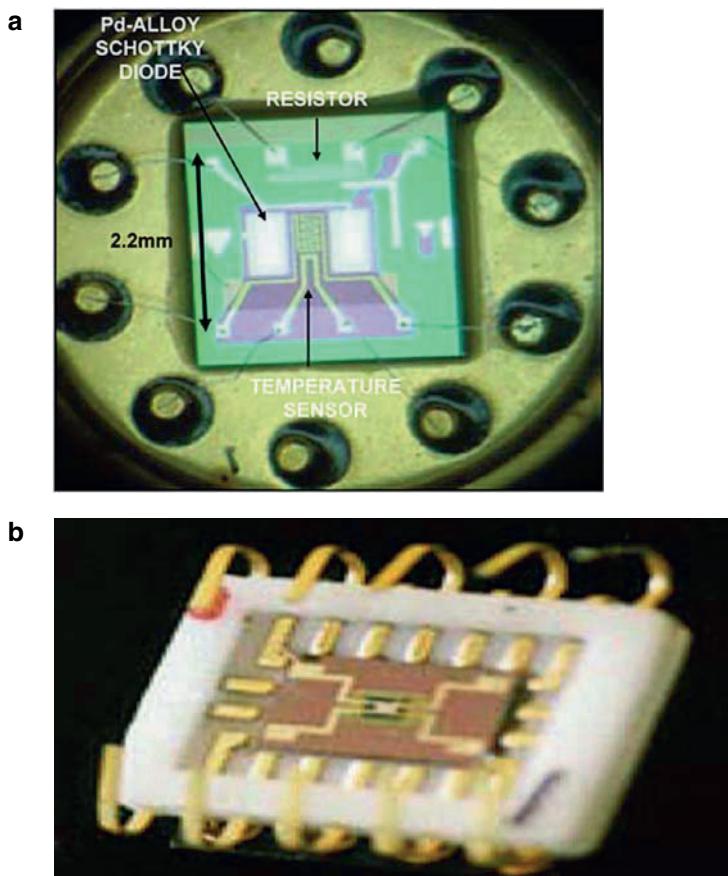


Fig. 22 (a) Picture of the packaged silicon-based hydrogen sensor. The Pd alloy Schottky diode (rectangular regions) resides symmetrically on either side of a heater and temperature sensor. The Pd alloy resistor is included for high concentration measurements. (b) Picture of zirconia-based oxygen sensor [54]. Reprinted with permission from the American Institute of Aeronautics and Astronautics © (2005)

2.6 Biomedical Devices

Requirements for biosensors are application specific and there are additional issues to consider including the biological fluid compatibility, the biocompatibility of materials which may be in contact with the body, and need for sterilization. There are a host of requirements for implantable devices including chemical stability, corrosion resistance, and the effects on the host immune system. The harsh environment in the body is a result of the high salt concentrations and the foreign

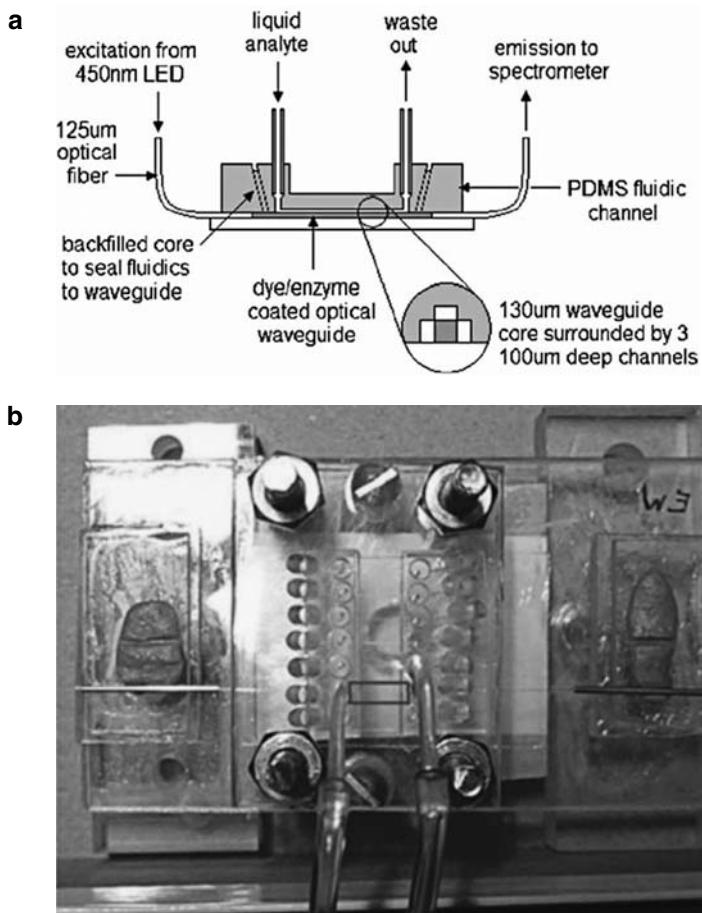


Fig. 23 (a) Diagram of the assembled waveguide sensor. The unusual cross-sectional area of the fluidic channel was created to limit the analyte exposure to the waveguide surfaces, allowing only one-dimensional analyte diffusion to the immobilized dye/enzyme layer [58]. (b) Photograph of assembled and packaged sensor. One of the seven channels is boxed in the center of the package [58]. Reprinted with permission from the SPIE

body reaction which takes place at the site of the implant. These issues are less stringent for semi-invasive devices, such as microneedles.

Microneedles and microdevices for drug delivery also require custom packages. A micropump integrated with microneedle array has been developed for painless drug delivery [62]. Figure 27a shows the needle array and Fig. 27b shows an SEM micrograph of the microneedles. The needles are approximately 200 μm in length to provide penetration of the stratum corneum to allow drugs to be delivered subcutaneously, however with very little pain as the nerves are located

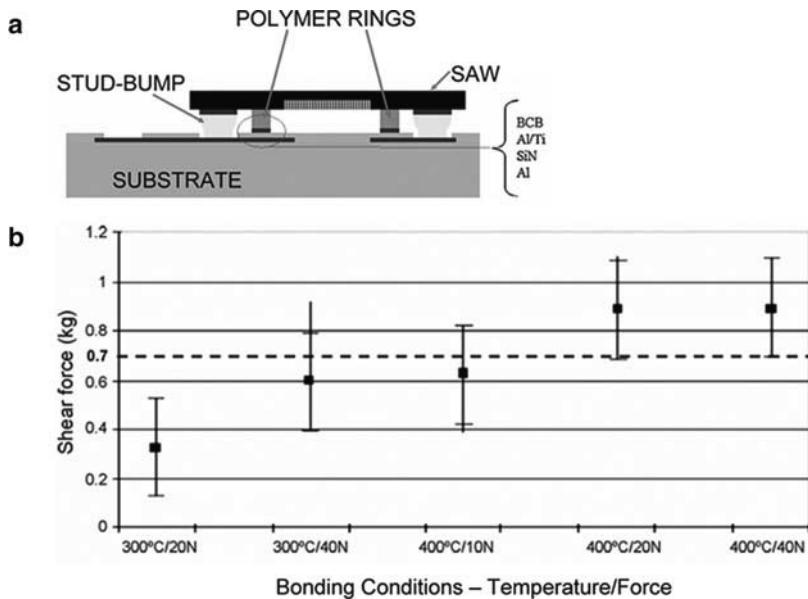


Fig. 24 (a) Cross-section of the different layers for the sealing of the SAW sensor to the substrate, (b) arrangement for the shear test experiment, and (c) shear force at different bonding forces and process temperatures with SU-8 material bond. Reprinted from [59] with permission Elsevier © (2008)

deeper in the skin. The housing provides a volume of drug and the piezoelectric pump, manufactured separately, provides the pressure to induce the flow of drug through the needles. Packaging of the pump includes gluing of the reservoir to the piezodisk and assembly to an outlet port. Packaging of the microneedle array is shown schematically in Fig. 27c. The silicon needle array is laminated to a polyimide support which has a hole cut out to allow drug flow into the needles. This assembly is then glued to the carrier base support piece with an epoxy.

Integration of a cannula, drug reservoir, and miniature pump on a single substrate has been examined by P.-Y. Li et al. [63] for the treatment of ocular diseases. The novel electrochemical pump generates pressure through electrolysis of water at low voltage and operates at low power. The gas pressure generated is sufficient to generate flow of drug at a controlled rate through the flexible Parylene cannula, see Fig. 28. The drug reservoir is made of PDMS and was molded with soft lithography mold [64]. The pumping rate was determined in the range of $2 \mu\text{L}/\text{min}$ suitable for the ocular drug delivery application. A range of flows could be generated from pL to μL making this a worthwhile approach, however, to further develop the full potential of this technology extensive animal testing would be required. Other MEMS devices are finding increasing use in medicine, in particular for surgery [65] and drug delivery [66].

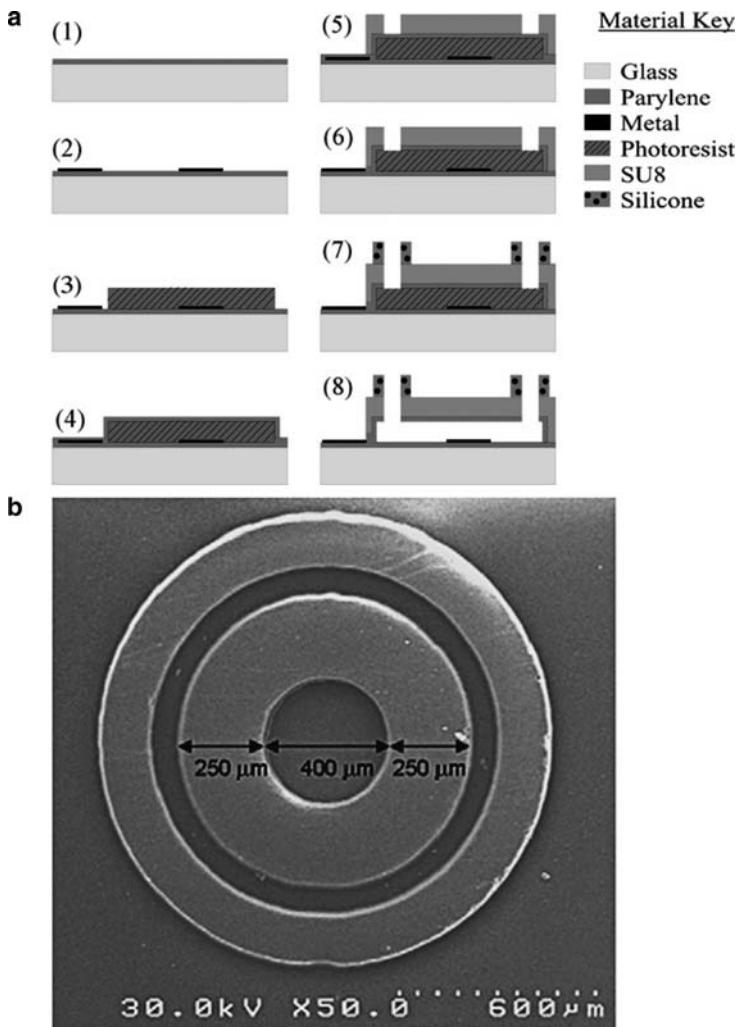
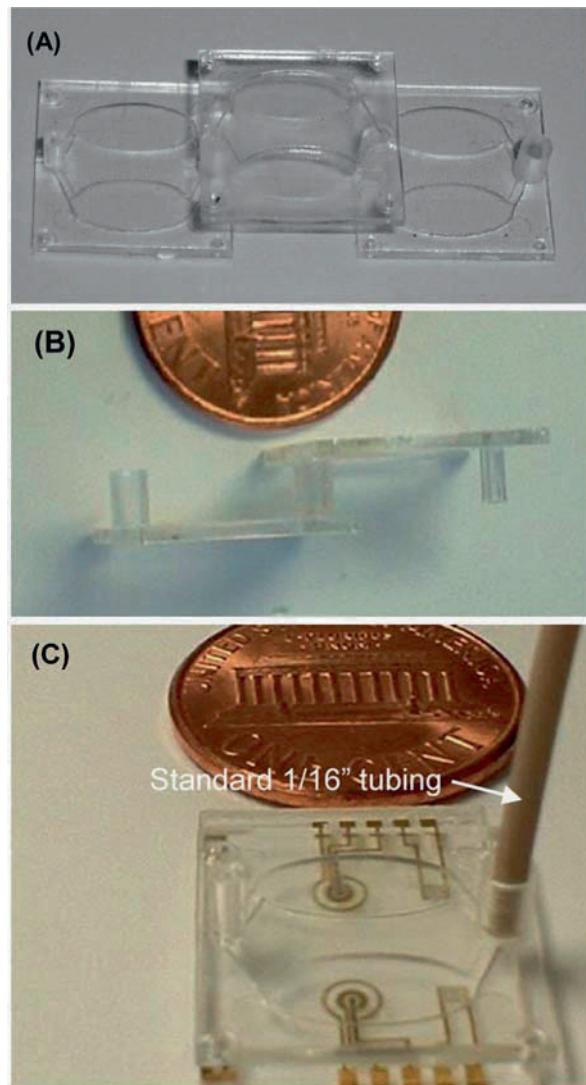


Fig. 25 (a) Cross-sectional diagram of device fabrication process: (1) Parylene-C deposited on substrate with adhesion promotor, (2) metal layer defined by lift-off, (3) thick photoresist layer defines microchannel, (4) 3- μm thickness layer of Parylene-C, (5) define fluidic vias 65- μm thick SU-8 layer, (6) oxygen plasma etching of Parylene, (7) photodefine silicone layer for gasket and descum with RIE, (8) dissolve photoresist in alcohol. (b) Scanning electron micrograph of a single photodefined silicone rubber MEMS O-ring, of thickness approximately 30 μm . Reprinted from [60] with permission Elsevier © (2008)

Fig. 26 (a) Fabricated interconnecting fluidic package showing package to package connection with mating inlet and outlet, (b) side view connectors, and (c) assembled sensor with microfluidic channels and snap-in connection. Reprinted from [61] with permission Elsevier © (2007)



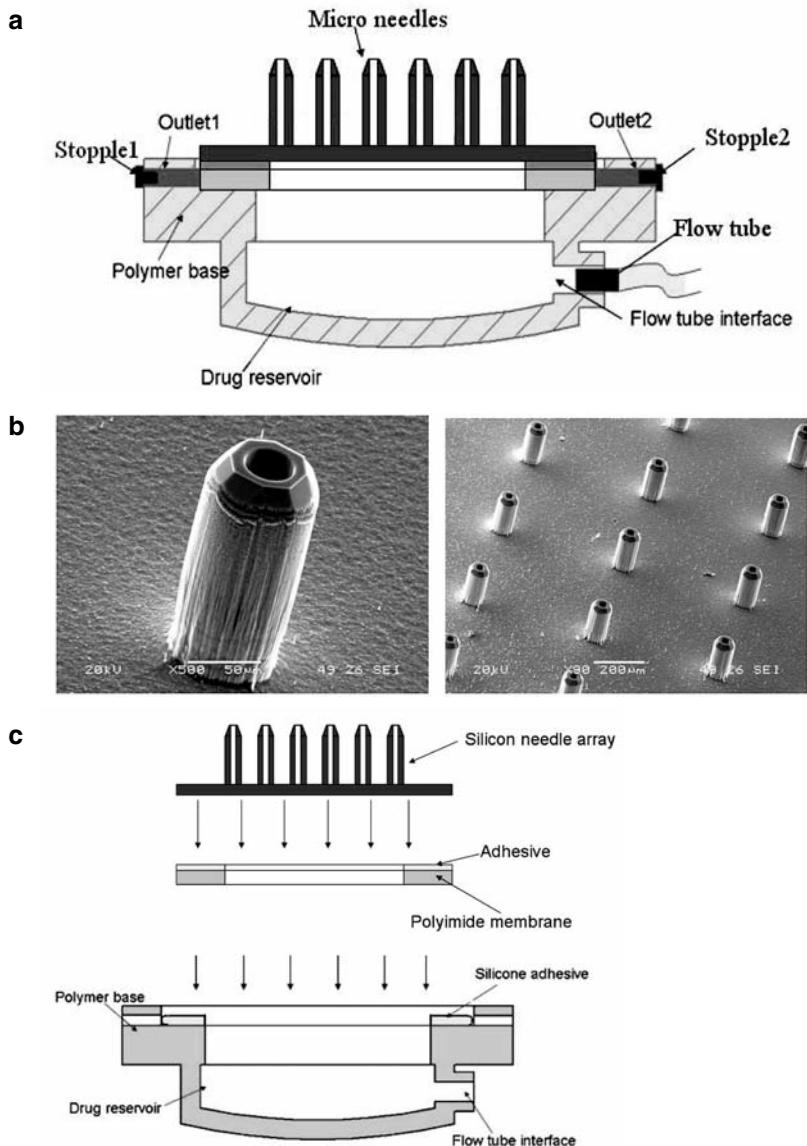


Fig. 27 (a) Packaged microneedle array assembled with drug reservoir; (b) SEM micrographs of silicon microneedles fabricated by micromachining, approximately 20 μ m in length, inner diameter of 30 μ m, and pitch 250 μ m. (c) Packaging process for microneedle array includes silicon microneedle die to polyimide bonding, dispensing epoxy on the microneedle base, removal of excess epoxy, curing, and insertion of the tube. Reprinted with permission from [62] © 2006 IEEE

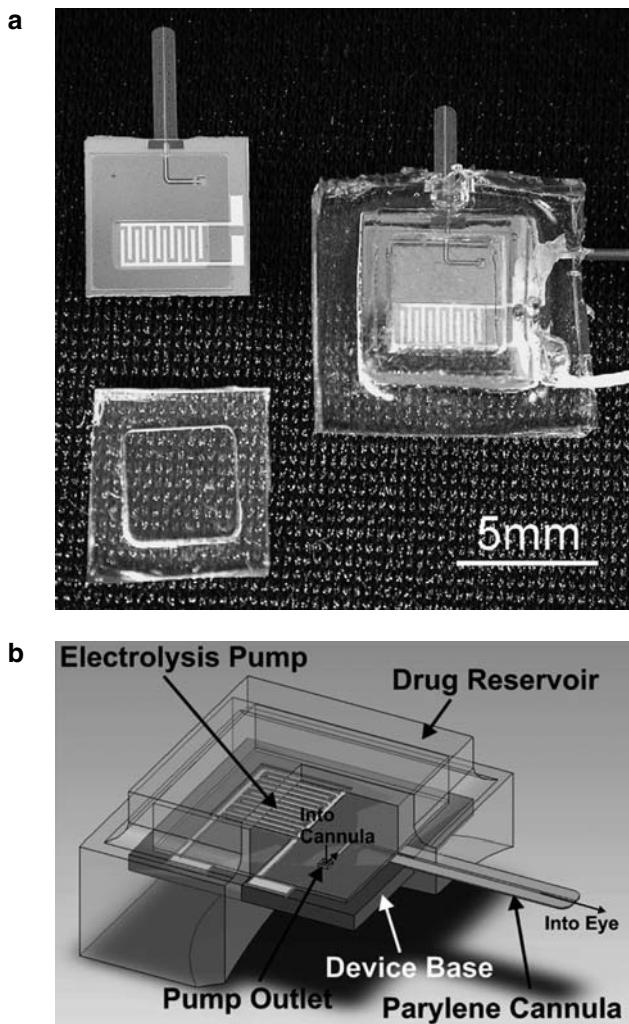


Fig. 28 (a) Pump and cannula chip (*upper left*), silicone drug reservoir (*lower left*), and packaged drug delivery device with attached wires for electrolysis actuation (*right*). (b) Schematic diagram illustrating the components of the drug delivery device with 1-mm wide cannula. The arrows follow the drug path from the reservoir through the channel of dimensions $100 \times 25 \times 5 \mu\text{m}$ in the Parylene cannula. Reprinted from [63] with permission Elsevier © (2008)

3 Packaging Technologies

An important part of packaging is the packaging technology including new materials and bonding of materials to form seals. A number of bonding methods have been developed, these include glass–silicon anodic bonding [67], silicon–silicon direct bonding [68], polymer bonding [69], solder bonding, and

adhesive bonding. The temperature at which bonding takes place and any stresses generated by the bonded interface can have an influence on sensor performance.

3.1 Wafer-Level Bonding

With the development of wafer bonding technology, wafer-level packaging (WLP) is showing its potential advantage in system integration to reduce packaging cost as well as increasing the yield of devices after dicing and packaging, since WLP would allow a wafer-level burn-in and wafer-level testing as such, it solves the known good die (KGD) issue. A WLP has the advantage that multiple dies are packaged in a single step prior to dicing. This technology has been widely used in MEMS area such as the packaging of microscanning mirrors [70], RF MEMS switch [71], and other applications. First, it can serve as a functional mechanical protection for the chip. This cap wafer is often featured by vias filled with sputter deposited metals that connect the pad on the chip to top solder bumpers. Yun et al. [72] demonstrated two approaches, so-called via-last and via-first method, to package micromachined accelerometers at the wafer level. The main character of this type of packaging is that the dicing procedure occurs after bonding which means a number of devices can be packaged at the same time. Second, the wafer-level package can be employed to realize the 3D structures. Takegawa et al. [73] presented a wafer-level packaging process using surface-activated bonding for silicon diaphragms. Diced chips were obtained with a high yield without damaging the interface and diaphragms. The structure also features sufficient bonding strength and low leak rate ($1.2 \times 10^{-12} \text{ Pa}\cdot\text{m}^3/\text{s}$) which indicate its ability to satisfy the requirements for MEMS hermetic packaging. They also pointed out that this technology is very suitable for applications requiring low stress, stability, and hermeticity, such as bioimplants and mechanical sensors.

When it comes to the packaging of chemical sensors, wafer-scale package has also been developed for gas sensors [74]. The sensor elements are fabricated by micromachining in the same manner as before, with the additional space for a Pyrex glass ring that is anodically bonded to the die. This supports a gas permeable membrane defined over the chemical sensor to allow analyte to diffuse to the sensor surface, while the bonding pads are not covered to facilitate electrical connections. Here the anodic bonding and membrane assembly are carried out at the wafer level, followed by dicing. The completed sensor is illustrated schematically in Fig. 29. This package has been further modified to be a smaller and a more cost efficient design as shown in Fig. 29d in which the gas-sensitive layer was placed in cavities fabricated by deep reactive ion etching (DRIE) from the back and the whole wafer with these cavities is sealed with a gas permeable membrane. The modified concept makes use of the silicon as the supporting rim thus takes away the need for the Pyrex glass frame to realize the WLP of the sensors and simplify therefore their processing [75].

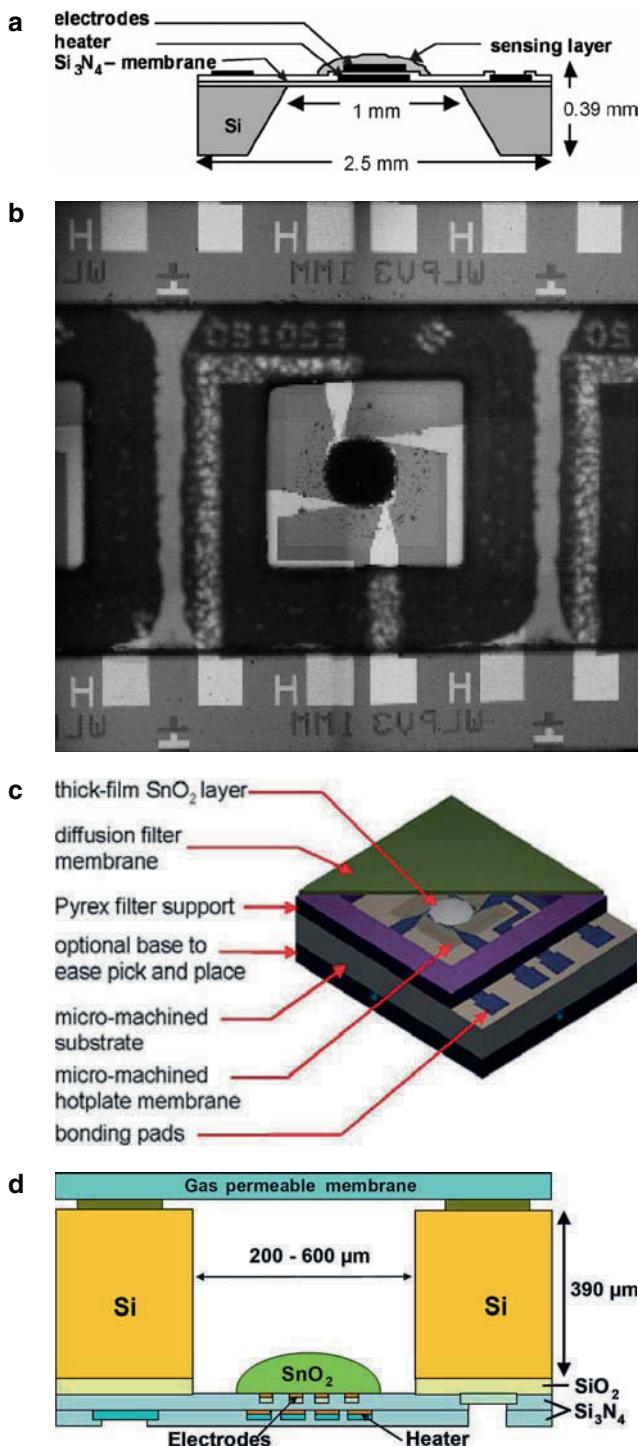


Fig. 29 (continued)

3.2 Localized Bonding

Localized anodic bonding of micromachined glass packaged is demonstrated by Zaire et al. [76] for encapsulation and protection of microelectronic circuits. Another example of localized bonding has been successfully demonstrated by Lin et al. [77, 78] in the form of bonding to form hermetic packages. Localization of heated regions can promote interdiffusion, the formation of intermetallics, and even localized chemical vapor deposition (CVD) to form new sealing materials at the interface. Polysilicon heaters were defined photolithographically to define local heated regions. The temperature in the microheater can increase to 1000°C while in the neighboring areas less than several micrometer away are only at 100°C. Building microheaters on the desired bonding area and applying pressure and electric current, the localized bonding process is activated, as shown in Fig. 30. Silicon-to-glass localized fusion bonding using polysilicon as microheater and silicon-to-gold eutectic bonding using gold as microheater are achieved with bonding strength higher than 10 MPa [77]. The local heating method has also been applied to bond solder and was able to overcome the surface roughness and create excellent step coverage by the reflow of the solder material. When applied as localized CVD bonding, this method is capable of activating the decomposition of silane locally by localized heating, thus connects the cap and device substrate firmly [78].

3.3 PDMS Fabrication

Polydimethylsiloxane (PDMS) or so-called silicon rubber is a material widely used in nowadays microdevices manufacturing and packaging. Here we are trying to reveal some of the functions of PDMS by describing its application in biosensors including gas sensor, MOSFET-type biosensor, optical biosensor array, and blood cell biochip.

3.3.1 Main Advantages and Disadvantages

PDMS has several attractive specific properties for the fabrication of microfluidic channels for use with aqueous systems of biochemicals and cells: it is transparent to visible and ultraviolet light, with good chemical and thermal stability, soft, flexible, electrically insulating, unreactive, permeable to gases. Also, its surface can be rendered hydrophilic that allows it to adhere to other materials like polymers and glass

Fig. 29 (continued) (a) Schematic cross-sectional diagram of the tin oxide hotplate gas sensor, (b) image of the sensor after bonding to the Pyrex cover and before sealing and dicing, and (c) schematic drawing of the completed packaged gas sensor. Reprinted with permission from [74] © 2007 IEEE. (d) Schematic diagram of the miniaturized wafer-level packaged metal oxide gas sensors combining DRIE of silicon and drop coating of the gas-sensitive film [75]

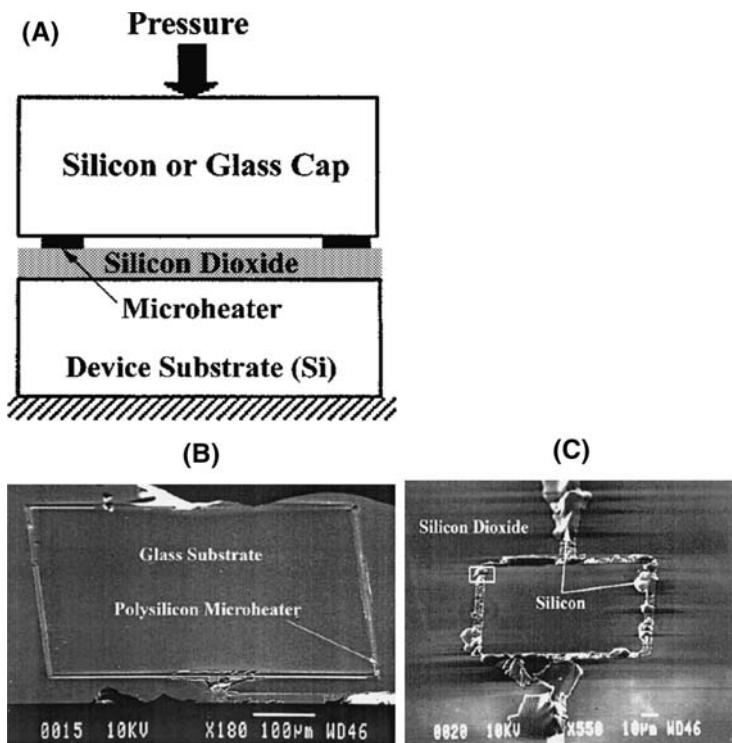


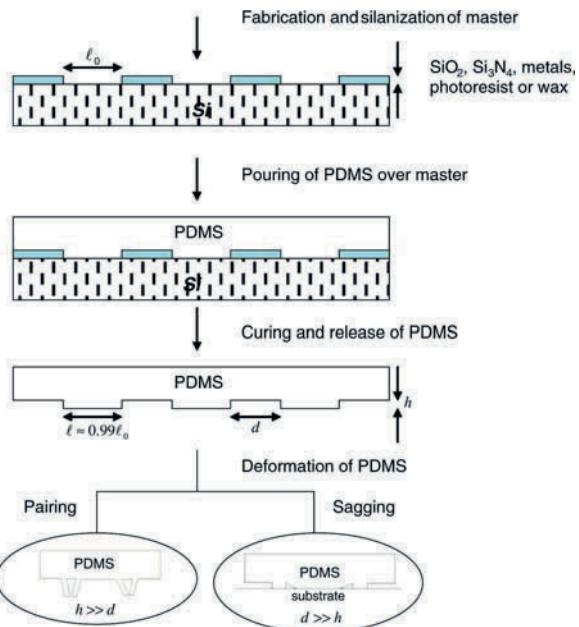
Fig. 30 (a) Schematic diagram showing the principle of localized bonding based upon microheater on insulating silicon dioxide film; (b) SEM micrograph showing the glass cap substrate is softened and has the shape of the polysilicon heater; (c) SEM micrograph shows localized silicon–gold eutectic formation. After the bond is forcefully broken, the silicon cap is torn apart and transferred to the device substrate. Reprinted with permission from [77] © 2000 IEEE

easily. The main disadvantages of PDMS are that it absorbs a range of organic solvents and compounds (particularly alkyl and aryl amines) [79] and that its surface properties can be difficult to control [80].

3.3.2 Soft Lithography and Multilayer Devices

Soft lithography represents a non-photolithographic strategy based on self-assembly and replica molding for carrying out micro- and nanofabrication structures with feature sizes ranging from 30 nm to 100 µm. In soft lithography, PDMS has always been chosen as the material for fabricating elastomeric stamps with structures on its surface. The elastomeric stamp is prepared by cast molding from a master which has been lithographically defined, as shown in Fig. 31[81]. The master is silanized by exposure to vapor for 30 min. The PDMS elastomer frequently used

Fig. 31 Schematic illustration of the procedure for fabricating PDMS stamps from a master having relief structures on its surface [81]. Reprinted with permission from Annual Reviews



is Sylgard 184 (Dow Corning Inc.). It is supplied in a two part kit: a liquid silicon rubber base; and a catalyst or curing agent. Once mixed and pored over the master and heated to elevated temperature the liquid mixture becomes a cross-linked elastomer in a few hours. The cured PDMS is then peeled off the master so that each master can be used multiple times. In soft lithography, the featured PDMS stamp is often used as a master to further transport specific structures attached to it to another substrate. However, in some of the bio or chemical application, the PDMS layers, flat or structured, serve directly as the body of the devices. In C.W. Cho's work, single wall carbon nanotube network gas sensors were fabricated on PDMS [82].

Zhang et al. [83] present a PDMS microdevice with built-in optical biosensor array. Here the merging of different technologies including fluorescence-based optical sensing, layer-by-layer encapsulation, and PDMS-based microfluidics to produce a hybrid system which can provide on-site monitoring of the cellular microenvironment within the microchannels. The layout of their device is shown in Fig. 32. The top cover layer could be a thin PDMS slab or a microscope glass slide; the middle layer is the functional PDMS layer with fabricated microchannels and nanowell where both cell culture and in situ sensing take place. The bottom layer is fiber optics alignment layer which can also be made from PDMS. As mentioned above, PDMS material can be rendered hydrophilic thus here is applied as a

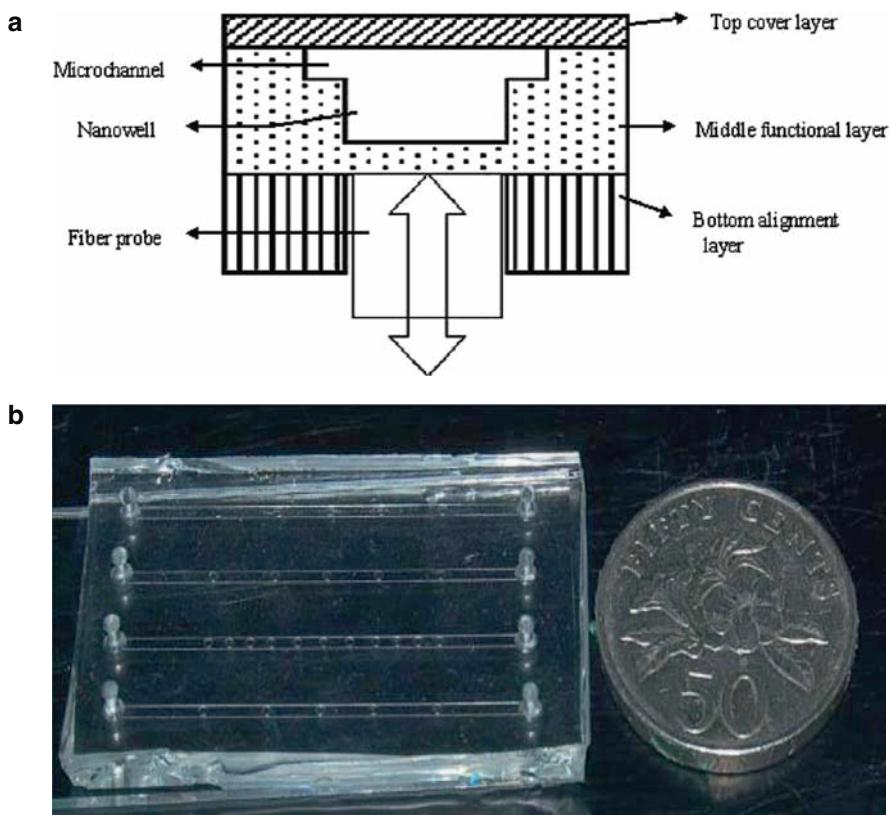


Fig. 32 (a) Cross-section of the tri-layer microdevice, where the *top layer* is glass, *middle layer* is PDMS, and *bottom layer* alignment layer for the fiber optics. (b) Molded PDMS device (without the *bottom layer*) [83]. Reprinted with permission from the SPIE

pretreatment to achieve water proof sealing between the top and middle layer with the aid of proper pressure. Figure 32 shows a molded PDMS device without the bottom layer. What should be mentioned here is that the PDMS is non-fluorescent and transparent down to a wavelength of 230 nm that satisfy the need of fluorescence measurement.

Another example comes from J.K. Shin's [84] work which reported a MOSFET-type biosensor assembled in a PDMS microfluidic channel package for the electrical detection of nanoscale biomolecules. As can be seen in Fig. 33, the upper PDMS layer with the microchannel was aligned and bonded along with the sensing area of the dual-gated MOSFET device implanted in lower layer, which is also PDMS, except for bonding pad.

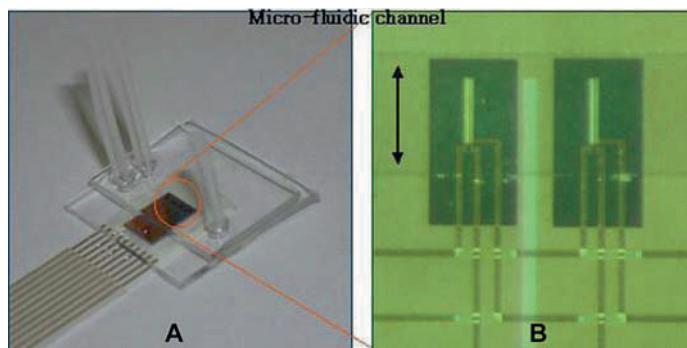


Fig. 33 (a) Photograph of the sensor device packaged with a metalized alumina header and (b) image of the MOSFET-type sensor and PDMS microfluidic channel formed on the device [84]. Reprinted with permission from the SPIE

3.3.3 PDMS as a Sealing Material

Another important function for PDMS is serving as the sealing layer. Thanks to its properties of transparency and elasticity, PDMS layers are capable of performing both sealing and mechanical support function to the chip and device. Y. H. Cho [85] demonstrated the sealing function of PDMS in a MEMS-based biochip for the characterization of a single red blood cell. As shown in Fig. 34, PDMS is served as the sealing layer that can be flat or with a 3D structure. The 3D sealing layer required 3D master mold and higher requirement for alignment while the flat one is easy to align. Another packaging sample showing the functions of PDMS can be found in Zhong's [86] work which presents a microdirect methanol fuel cell (μ DMFC)

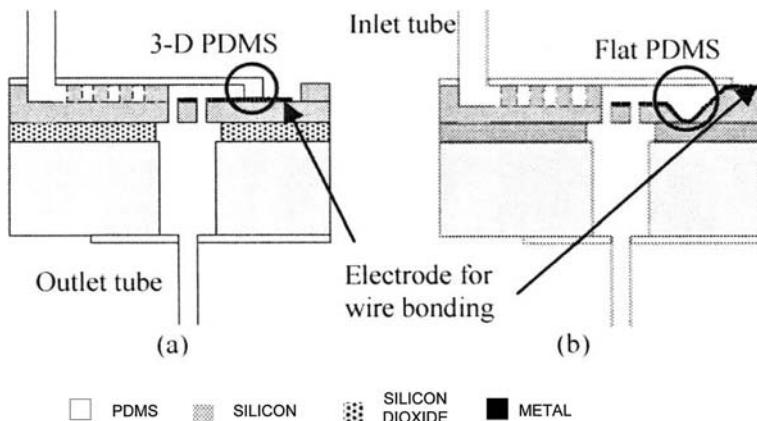


Fig. 34 Schematic view of the electrode sensor packaging, with use of (a) 3D-molded PDMS structure and (b) 2D flat PDMS structure. Reprinted with permission from [85] © 2005 IEEE

stack implemented by microfabrication process and PDMS assembly method. The prefabricated PDMS pieces, approximately 200–300 μm in thickness were used in the packaging for three purposes: (1) keeping the MEAs, leading channels and feeding holes from leakage by the PDMS sealing layers, (2) protecting the fragile silicon plates against mechanical stress with the PDMS buffering layers, (3) fixing the feeding tubes with PDMS blocks. The assembly process of each parts of the μ DMFC can be found in Fig. 35. The components are held together with an aluminum plate for testing the operation of the DMFC.

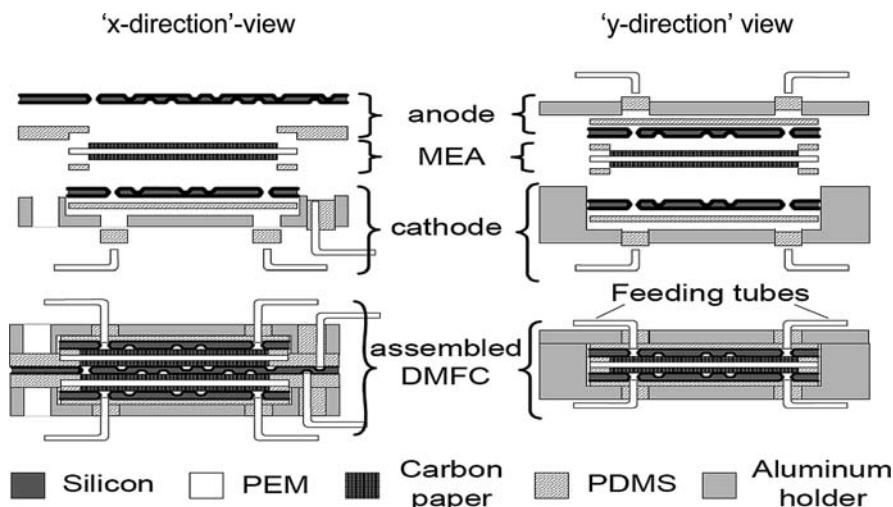


Fig. 35 Schematic cross-sections in orthogonal planes, showing the components of the miniature direct methanol fuel cell stack. The prefabricated MEA and silicon anode with integrated electrodes are assembled with the molded PDMS membranes, approximate thickness 0.2–0.3 mm. The feed tubes and electrodes are connected to complete the assembly process. Reprinted with permission from [86] © 2007 IEEE

3.4 Stereolithography-Based Packaging

There are few reviews of earlier work on SLA for 3D MEMS [87] and custom packages [88]. A typical process for SLA manufacturing includes the following steps. First a 3D model should be built using any 3D modeling software which is able to convert its own format into STL file (3D Systems specific file format), such as Solidworks, Pro/Engineer, IDEAS, or AutoCAD. Then, with the help of “3D Lightyear SLA Platform Preparation Software” you can prepare STL three-dimensional model representations into build files that are to be built on an SLA solid imaging system. There are several things that need to be done in this preparation step, including verified your 3D model, determining the build direction, editing the support layer, and slicing the part. Finally, at the SLA station, the liquid resin

is solidified using a computer-directed, ultraviolet (UV) laser beam. By exposing successive layers of photosensitive polymer resin the CAD design is turned into a solid plastic object, which is rinsed in solvent and cured in a UV oven.

An alternative to SU-8 is the use of stereolithographic assembly (SLA) to define customized microfluidic channels and packages for sensor arrays. Han et al. [89, 90] describe microfluidic channels with integrated microvalves and plug-in connections with O-ring seals, as shown in Fig. 36. The valves are pneumatically actuated and have a dead volume of 24 nL. One-step plug-in microfluidic connections are made with an O-ring seal on each port. The lower O-ring is to prevent clogging when the microchannels are bonded with adhesive to the glass chip. One advantage of using this approach is that the glass microfluidic component can be fabricated independently from the microfluidic system interface (MSI) layer. SLA system enables the manufacture of complex microfluidic structures with very few design constraints and sufficient resolution (25 μm vertical, $\sim 1 \mu\text{m}$ horizontal). Some of the important features should be considered and designed in the SLA parts to prevent leakage (space designed for O-ring or silicon gasket)

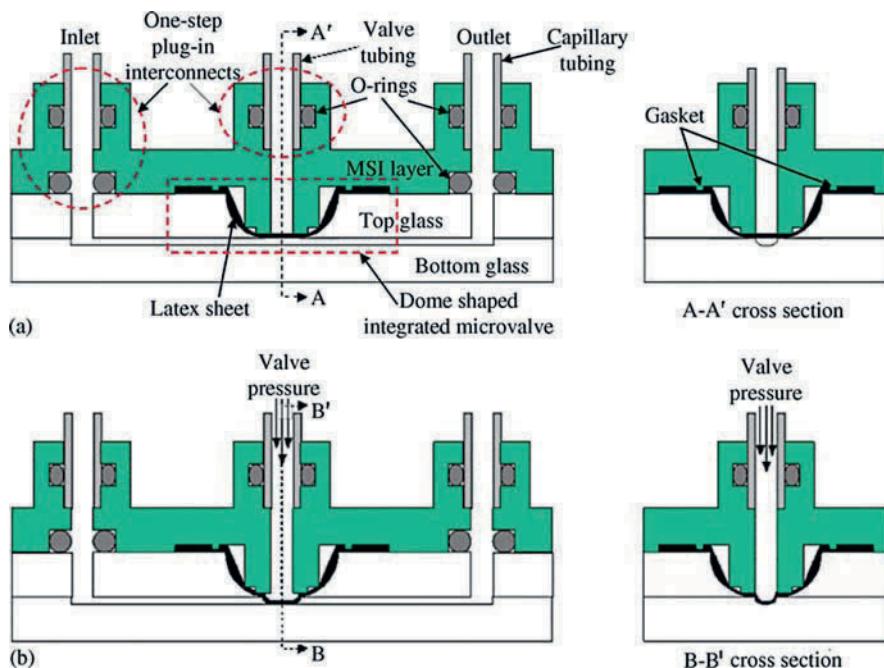


Fig. 36 Examples of the one-step plug-in microfluidic interconnects fabricated with SLA combined with two layers of glass that form microfluidic channels. The plug-in interconnects and dome-shaped monolithic microvalve are shown in cross-section: (a) when the microvalve is open and (b) when the microvalve is closed. Reprinted from [90] with permission Elsevier © (2007)

as well as accomplishing their functions (microvalve). The MSI can be configured with integrated microvalves to facilitate fluid manipulation for chemical and bioassays.

Figure 37 shows an example of a customized package for a AFM microprobe that provides a fluid chamber attached to the probe [91]. The surface of SLA resin can be coated with Parylene to improve chemical resistance and potential for biocompatibility. Another example is a package that fits the chip containing a sensor array and then is inserted into a standard 1/8" diameter tube connector as shown in Fig. 38.

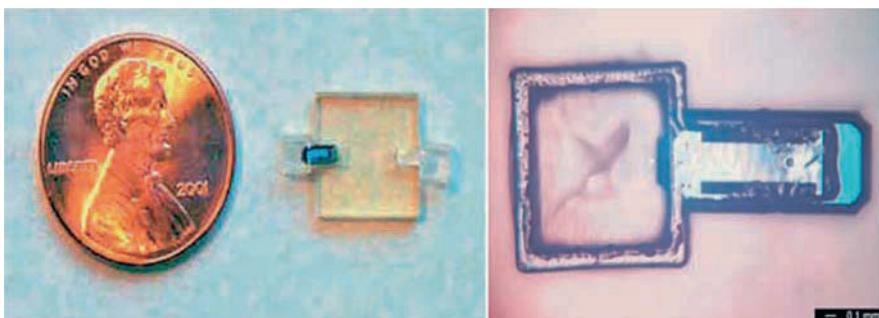


Fig. 37 SLA package for a AFM cantilever sensor that allows fluid introduction for the probe [91]

The advantage of SLA is that customized packaging geometries can be configured and the integration of microvalves can be easily accomplished to facilitate fluid manipulation, sample routing, and mixing, for example, for PCR analysis, as shown in Fig. 39. The MSI component was aligned and connected to other parts of the fluidic system made of glass with epoxy adhesive. Thermal management for PCR is important. Singh and Ekaputri [92] describe an ANSYS simulation of temperature profiles in a multilayer packaged system.

The flexibility of SLA is to build custom fluidic channel designs as indicated by the example of mimicking the sense of olfaction [93]. Building a series of flow channels with a stationary phase of Parylene-C approximately 10 μm thick, providing an analog to the mucus coating of the nasal passages, as shown in the CAD drawing in Fig. 40a. The prototype olfactory mucosa is assembled onto a microsensor array, such that each sensor element is beneath one of the channels. Thus the sample is exposed to each sensor in series depending upon the flow and interaction with the stationary phase. Hence the temporal response of the sensors indicates some attributes of the gas species being detected. The SLA flow channels are attached to the die with additional uncured epoxy resin, as shown in Fig. 40b.

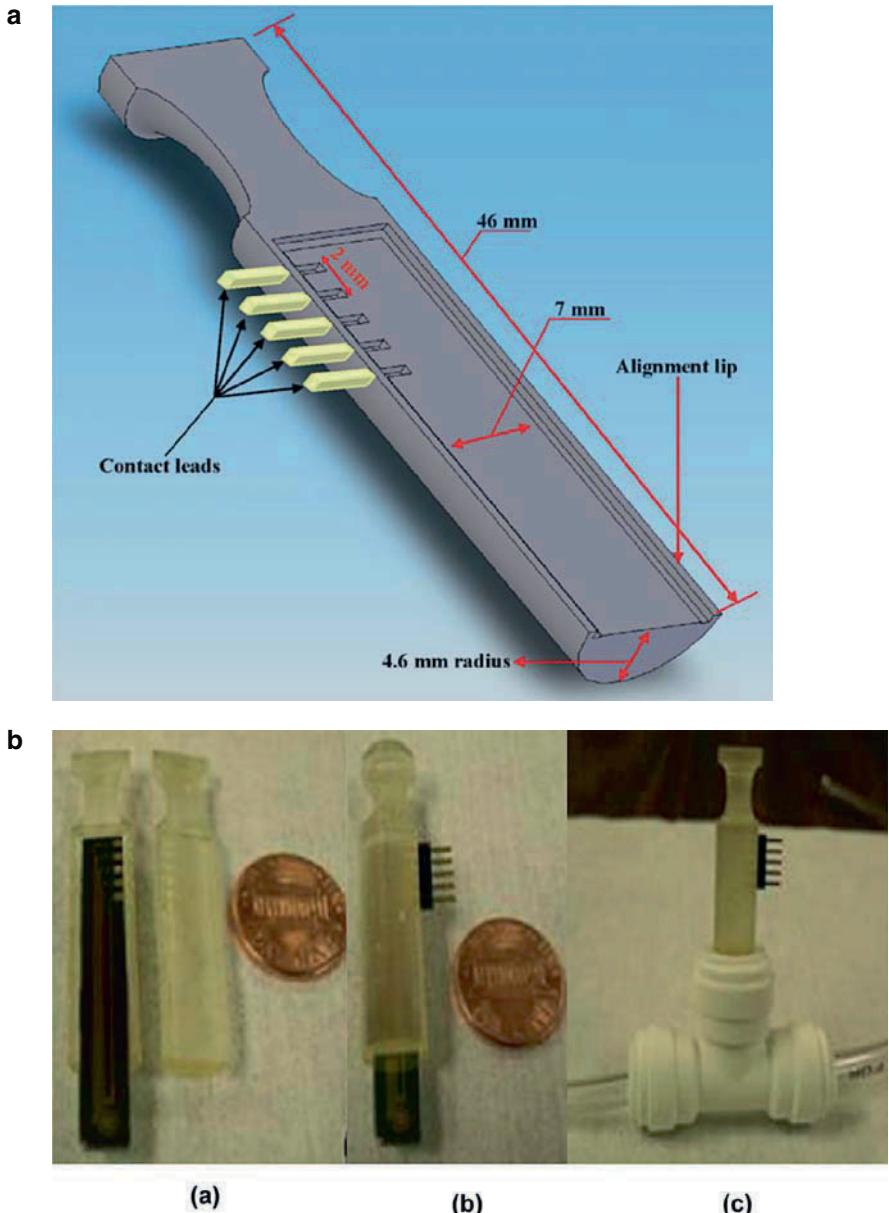
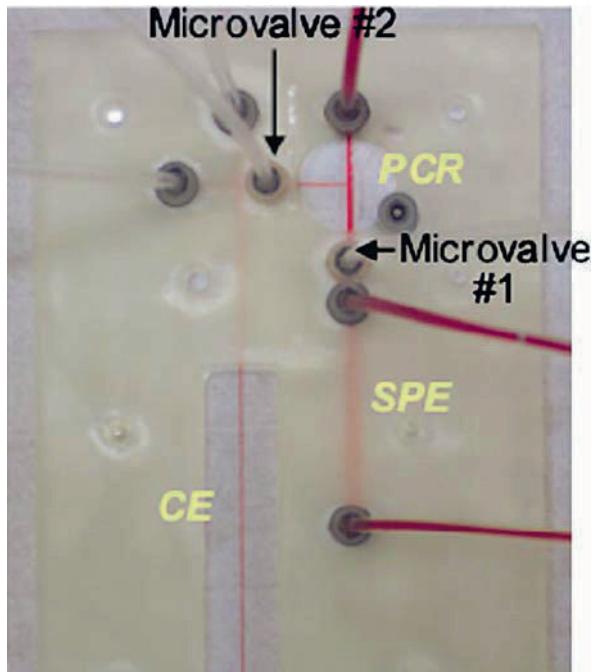


Fig. 38 (a) Solidworks 3D model of package for the sensor with clamping action and electrical contacts and (b) a snap-in connection of a standard tubing to the package for a sensor: (a) before assembly, (b) after assembly, and (c) in standard $3/8''$ diameter tube connector. Reprinted from [61] with permission Elsevier © (2007)

Fig. 39 Red dye flowing through the microfluidic compartment in the DNA separation system with microvalve #1 and #2 open. This simulates the sequence of events necessary for a PCR analysis. The integrated valves are operated at pressures of 400 kPa and when valve #1 and #2 were closed dye fills the SPE chamber. When valve #1 is open sample is directed from SPE to PCR thermal cycling chamber, and finally when valve #2 is opened sample is injected into CE separation column. Reprinted from [90] with permission Elsevier © (2007)



4 Discussion of Future Prospects and Conclusions

The selection of an appropriate method for packaging of a sensor or sensor array is an important, and a complex issue that is often underestimated when designing systems. One must consider the mechanical design, thermal issues, chemical and biocompatibility, ease of manufacture, and cost. Based upon the intended application, material selection can have an influence on performance. For example, the influence of the stresses generated by package may act on the sensor to change performance or reliability, and should be carefully considered. Important material properties include hermeticity, corrosion resistance, inertness of package, and possibility of absorption of chemicals and subsequent leaching from the package during the operational lifetime of sensor. Issue of reliability of packaging has been investigated by de Reus et al. [94] and more work is needed.

Many chemical sensor packages are hand assembled and require precision alignment under the microscope. Lower cost packaging with injection molding of polymers has been successful for ISFET sensors in applications that require short-term use. The cost of materials and level of automation have an impact on the cost of assembly and manufacture.

The packaging requirement for nanosensors may not be that much different from microsensors. The size of the nanowires suggests that a smaller die could result and hence smaller scale packages. One of the current issues for the further successful

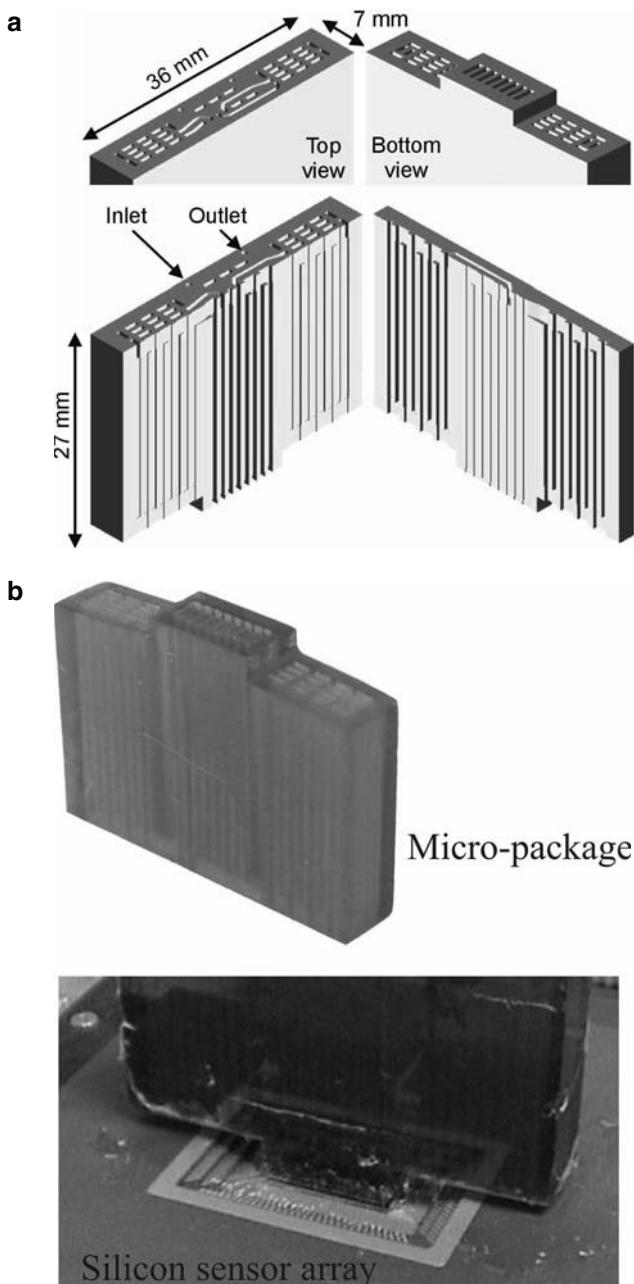


Fig. 40 (a) CAD drawing of MSL package, dimensions $36 \times 27 \times 7$ mm, which was built by SLA and coated with Parylene-C and (b) photograph of the micropackage and final assembly to the sensor array providing channels for the gas separation [93]. Reprinted with permission of IET Publishers

application of nanowire sensors is the ability to make reproducible and reliable electrical connections to the nanowires. However, packages suitable for MEMS chemical sensors [6] may also be useful because the ultimate function is identical.

Given the diversity of sensor platforms and the range of different applications it appears that currently a package must be specifically designed for each individual application. This level of customization leads inevitably to increased cost. Until a generic packaging approach is developed for chemical and biomolecular sensors the cost of the packaging will invariably exceed the cost of the sensor. New methods for packaging need to be developed that are amenable to automated assembly and mass production. An area where progress in wafer-level packaging has occurred includes new materials, such as polycrystalline diamond for integrated packaging [95]. The potential for further development in sensors with integrated packages is a fruitful one for study where, perhaps progress in novel materials processes such as atomic layer deposition (ALD)/CVD and use of flexible substrates would be useful. Nevertheless, packaging remains a multidisciplinary area of study and needs development of generic methods in order to achieve reduced cost for chemical and biomolecular sensors.

References

1. Harper C. A. (ed.). Electronic Packaging and Interconnection Handbook. McGraw-Hill: New York, 2004.
2. Tummala R. (ed.). Fundamentals of Microsystems Packaging. McGraw-Hill: New York, 2001.
3. Tomescu C. C. Low-cost Sensor Packaging. Delft University Press: Delft, The Netherlands, 2001.
4. Brand O., Packaging. In: Gianchandani Y. B, O. Tabata, H. Zappe (eds.). Comprehensive Microsystems. Elsevier: Amsterdam, The Netherlands, 2008:431–460.
5. Velten T, H. H. Ruf, D. Barrow, N. Aspragathos, P. Lazarou, E. Jung, C. K. Malek, M. Richter, J. Kruckow, M. Wackerle. Packaging of bio-MEMS: Strategies, technologies, and applications. IEEE Transaction on Advanced Packaging 2005;28:533–546.
6. Lee Y. C., B. A. Parviz, J. A. Chiou. Packaging for microelectromechanical and nanomechanical systems. IEEE Transactions on Advanced Packaging 2003;26:217–226.
7. Gardner J. W., F. Udrea. Microsensors: Principles and Applications. Wiley: London, 2009.
8. Pearce T.C., S. S. Schiffman, H. T. Nagle, J. W. Gardner Handbook of Machine Olfaction: Electronic Nose Technology. Wiley-VCH: Weinheim, Germany, 2003.
9. Kovacs G. T. A. Micromachined Transducer Sourcebook. McGraw Hill: New York, 1998.
10. Göpel W, T. A. Jones, M. Kleitz, I. Lundström, T. Seiyama. Sensors: A Comprehensive Survey, Vol. 2, Pt. I, Chemical and Biochemical Sensors. Wiley-VCH: Weinheim, Germany, 1997.
11. Bergveld P. Development of an ion-sensitive solid-state device for neurophysiological measurements. IEEE Transactions on Biomedical Engineering 1970;17:70–71.
12. Moss S. D., C. C. Johnson, J. Janata. Hydrogen, Calcium, and Potassium Ion-Sensitive FET Transducers: A Preliminary Report. IEEE Transactions on Biomedical Engineering 1978;25:49–54.
13. Chern G-C, J. N. Zemel. Temperature dependence of the gate-controlled portion of ion-controlled diodes. IEEE Transactions on Electron Devices 1982;29:115–123.
14. Yamamoto N., S. Tonomura, T. Matsuoka, H. Tsubomura. A study on a palladium-titanium oxide Schottky diode as a detector for gaseous components. Surface Science 1980;92: 400–406.

15. Koudelka M. Performance Characteristics of a planar 'Clark-type' oxygen sensor. *Sensors and Actuators* 1986;9:249–258.
16. Maclay G.J., W. J. Buttner, J. R. Stetter. Microfabricated amperometric gas sensors. *IEEE Transactions on Electron Devices* 1988;35:793–799.
17. Poghossian A, M. J. Schonung. Chemical and biological field-effect sensors for liquids – a status report. In: Marks R. S., C. R. Lowe, D. C. Cullen, H. H. Weetall, I. Karube, (eds.). *Handbook of Biosensors and Biochips*. John Wiley: Chichester, UK, 2007.
18. Smith R.L., J. Janata, R. J. Huber. Transient phenomena in ion sensitive field effect transistors. *Journal of Electrochemical Society* 1980;127:1599–1603.
19. Grisel A., C. Francis, E. Verney, G. Mondin. Packaging technologies for integrated electrochemical sensors. *Sensors and Actuators* 1989;17:285–295.
20. Chovelon J.M., N. Jaffrezic-Renault, Y. Cros, J. Fombon,D. Pedone. Monitoring of ISFET encapsulation aging by impedance measurements. *Sensors and Actuators B* 1991;3:43–50.
21. Gracia I, C. Cane, M. Lozano, J. Esteve. On-line determination of the degradation of ISFET chemical sensors. *Sensors and Actuators B* 1993;8:218–222.
22. van der Spiegel J, I. Lauks, P. Chan, D. Babic. The extended gate chemically sensitive field effect transistor as multi-species microprobe. *Sensors and Actuators* 1983;4:291–298.
23. Gracia I, C. Cane, E. Lora-Tamayo. Electrical characterization of the aging of sealing materials for ISFET chemical sensors. *Sensors and Actuators B* 1995;24–25:206–210.
24. Chi L-L, Jung-Chuan Chou, Wen-Yaw Chung, Tai-Ping Sun, Shen-Kan Hsiung. Study on extended gate field effect transistor with tin oxide sensing membrane. *Materials Chemistry and Physics* 2000;63:19–23.
25. Janata J. Thirty years of CHEMFETs – A Personal View. *Electroanalysis* 2004;16:1831–1835.
26. van den Vlekkert H. H, B. Kloek, D. Prongue, J. Berthoud, B. Hu, N. F. de Rooij, E. Gilli, P. H. de. Crousaz. A pH-ISFET and an integrated pH-pressure sensor with back-side contacts. *Sensors and Actuators* 1988;14:165–176.
27. Matsuo T, M. Esashi. Methods of ISFET fabrication. *Sensors and Actuators* 1981;1:77–96.
28. Bratov A, J. Mufioz, C. Dominguez, J. Bartroli. Photocurable polymers applied as encapsulating materials for ISFET production. *Sensors and Actuators B* 1995;24–25: 823–825.
29. Munoz J, A. Bratov, R. Mas, N. Abramova, C. Dominguez, J.. Barfroli. Planar compatible polymer technology for packaging of chemical microsensors. *Journal of Electrochemical Society* 1996;143:2020–2025.
30. Munoz J, C. Jimenez, A. Bratov, J. Bartroli, S. Alegret, C. Dominguez. Photosensitive polyurethanes applied to the development of CHEMFET and ENFET devices for biomedical sensing. *Biosensors & Bioelectronics* 1997;12:577–585.
31. van Hal R.E.G., P. Bergveld, J. F. J. Engbersen. Fabrication and packaging of mesa ISFETs. *Sensors and Materials* 1996;8:455–468.
32. Milgrewa M.J., M. O. Riehle, D. R. S. Cumming. A large transistor-based sensor array chip for direct extracellular imaging. *Sensors and Actuators B* 2005;111–112:347–353.
33. Oelßner W., J. Zosel, U. Guth, T. Pechstein, W. Babel, J.G. Connery, C. Demuth, M. Grote Gansey, J.B. Verburg. Encapsulation of ISFET sensor chips. *Sensors and Actuators B* 2005;105:104–117.
34. Suehle J.S., Richard E. Cavicchi, Michael Gaitan, Steve Semancik. Tin oxide gas sensor fabricated using CMOS micro-hotplates and in -situ processing. *IEEE Electron Device Letters* 1993;14:118–120.
35. Martinez C.J., B. Hockey, C. B. Montgomery, S. Semancik. Porous tin oxide nanostructured microspheres for sensor applications. *Langmuir* 2005;21:7937–7944.
36. Panchapakesan B., R. Cavicchi, S. Semancik, D. L. DeVoe. Sensitivity, selectivity and stability of tin oxide nanostructures on large area arrays of microhotplates. *Nanotechnology* 2006;17:415–425.
37. Simon I., N. Barsan, M. Bauer, U. Weimar. Micromachined metal oxide gas sensors: opportunities to improve sensor performance. *Sensors and Actuators B* 2001;73:1–26.

38. Semancik S., R. Cavicchi. Kinetically controlled chemical sensing using micromachined structures. *Accounts of Chemical Research* 1998;31:279–287.
39. Bosc J.M., J. P. Odile. Micromachined structure reliability testing specificity the Motorola MGSLI00 gas sensor example. *Microelectronics Reliability* 1997;37:1791–1794.
40. Shirke A.G., R. E. Cavicchi, S. Semancik, R. H. Jackson, B. G. Frederick, M. Clayton Wheeler. Femtomolar isothermal desorption using microhotplate sensors. *Journal of Vacuum Science Technology A* 2007;25:514–526.
41. Meier D.C., C. J. Taylor, R. E. Cavicchi, E. White, M. W. Ellzy, K. B. Sumpter, S. Semancik. Chemical warfare agent detection using MEMS-compatible microsensor arrays. *IEEE Sensors Journal* 2005;5:712–725.
42. Meier D.C., S. Semancik. Coupling nanowire chemiresistors with MEMS microhotplate gas sensing platforms. *Applied Physics Letters* 2007;91:063118
43. Goschnick J., An electronic nose for intelligent consumer products based on a gas analytical gradient microarray. *Microelectronic Engineering* 2001;57–58:693–704.
44. Gmur R., J. Goschnick, T. Hocker, H. Schwarzenbach, M. Sommer. Impact of sensor packaging on analytical performance and power consumption of metal oxide based gas sensor microarrays. *Sensors and Actuators B* 2007;127:107–111.
45. Briand D., S. Colin, A. Gangadharaiah, E. Velab, P. Dubois, L. Thiery, N. F. de Rooij. Micro-hotplates on polyimide for sensors and actuators. *Sensors and Actuators A* 2006;132: 317–324.
46. Briand D., S. Colin, J. Courbat, S. Raible, J. Kappler, N.F. de Rooij. Integration of MOX gas sensors on polyimide hotplates. *Sensors and Actuators B* 2008;130:430–435.
47. Burg T.P., Amir R. Mirza, Nebojsa Milovic, Christine H. Tsau, George A. Popescu, John S. Foster, Scott R. Manalis. Vacuum-packaged suspended microchannel resonant mass sensor for biomolecular detection. *Journal of Microelectromechanical Systems* 2006;15: 1466–1476.
48. Song W.H. Packaging techniques for CMOS-based chemical and biochemical microsensors. Ph. D. Dissertation in Physical Electronics Laboratory. Swiss Federal Institute of Technology: Zurich, Switzerland, 2005:104.
49. Rangelow I.W., Tzv. Ivanov, K. Ivanova, B.E. Volland, P. Grabiec, Y. Sarov, A. Persaud, T. Gotszalk, P. Zawierucha, M. Zielony, D. Dontzov, B. Schmidt, M. Zier, N. Nikolov, I. Kostic, W. Engl, T. Sulzbach, J. Mielczarski, S. Kolb, Du P. Latimer, R. Pedreau, V. Djakov, S.E. Huq, K. Edinger, O. Fortagne, A. Almansa, H.O. Blom. Piezoresistive and self-actuated 128-cantilever arrays for nanotechnology applications. *Microelectronic Engineering* 2007;84:1260–1264.
50. Vancura C., Y. Li, J. Lichtenberg, K.-U. Kirstein, A. Hierlemann. Liquid-phase chemical and biochemical detection using fully integrated magnetically actuated complementary metal oxide semiconductor resonant cantilever sensor systems. *Analytical Chemistry* 2007;79:1646–1654.
51. Pinnaduwage L.A., A. C. Gehl, S. L. Allman, A. Johansson, A. Boisen. Miniature sensor suitable for electronic nose applications. *Review of Scientific Instruments* 2007;78: 55101–55104.
52. Choudhury A., P. J. Hesketh, Z. Hu, T. Thundat. Low noise chemical detection with a Piezoresistive Microcantilever. *Transactions of the ECS* 2006;3:473–482.
53. Hunter G.W., C.C. Liu, D. Makel. Microfabricated Chemical Sensors for Aerospace Applications. In: Gad-el-Hak M, (ed.). *MEMS Handbook: Design and Fabrication*. CRC Press: Boca Raton, USA, 2006.
54. Hunter G.W., J. C. Xu, P. G. Neudeck, D. B. Makel, B. Ward, C. C. Liu. Intelligent chemical sensor systems for in-space safety applications. First International Forum on Integrated System Health Engineering and Management in Aerospace. American Institute of Aeronautics and Astronautics: Napa, CA, USA, 2005.
55. Hierlemann A., R. Gutierrez-Osuna. Higher-order chemical sensing. *Chemical Reviews* 2008;108:563–613.

56. Ruano-Lopez J.M., M. Aguirregabiria, M. Tijero, M. T. Arroyo, J. Elizalde, J. Berganzo, I. Aranburu, F. J. Blanco, K. Mayora. A new SU-8 process to integrate buried waveguides and sealed microchannels for a Lab-on-a-Chip. *Sensors and Actuators B* 2006;114:542–551.
57. Blanco F.J., M Agirregabiria, J. Berganzo, K. Mayora, J. Elizalde, A. Calle, C. Dominguez, L. M. Lechuga. Microfluidic-optical integrated CMOS compatible devices for label-free biochemical sensing. *Journal of Micromechanics and Microengineering* 2006;16:1006–1016.
58. Chang-Yen D.A., B. K. Gale. Design, fabrication, and packaging of a practical multianalyte-capable optical biosensor. *Journal of Microlithography, Microfabrication, Microsystems* 2006;5:21101–21107.
59. Georgel V, F.Verjus, E.C.E. van Grunsven, P. Pouliche, G. Lissorgues C Pellet S. Chamaly, T. Bourouina. A SAW filter integrated on a silicon passive substrate used for system in package. *Sensors and Actuators A* 2008;142:185–191.
60. Miserendino S, Y.-C. Tai. Modular microfluidic interconnects using photodefinition silicone microgaskets and MEMS O-rings. *Sensors and Actuators A* 2008;143:7–13.
61. Pepper M, N. S. Palsandram, P. Zhang, M. Lee, H.-J. Cho. Interconnecting fluidic packages and interfaces for micromachined sensors. *Sensors and Actuators A* 2007;134:278–285.
62. Ma B, S. Liu, Z. Gan, G. Liu, X. Cai, H. Zhang, Z. Yang. A PZT insulin pump integrated with a silicon micro needle array for transdermal drug delivery. *IEEE Electronic Components and Technology Conference*, 2006:677–681.
63. Li P-Y, J. Shih, R. Lo, S. Saati, R. Agrawal, M. S. Humayun, Y.-C. Tai, E. Meng. An electrochemical intraocular drug delivery device. *Sensors and Actuators A* 2008;143:41–48.
64. Unger M.A., H.-P. Chou, T. Thorsen, A. Scherer, S. R. Quake. Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science* 2000;288:113.
65. Rebello K.J. Applications of MEMS in surgery. *Proceedings of the IEEE* 2004;92:43–55.
66. Gilla H.S., M. R. Prausnitz. Pocketed microneedles for drug delivery to the skin. *Journal of Physics and Chemistry of Solids* 2008;69:1537–1541.
67. Knowles K.M., A. T. J. van Helvoort. Anodic bonding. *International Journal of Materials Review* 2006;51:273–311.
68. Alexe M, U. Gosele, U. Gosele. *Wafer Bonding*. Springer: New York, 2004.
69. Kim H, K. Najafi. Characterization of low temperature wafer bonding using thin film parylene. *Journal of Microelectromechanical Systems* 2005;14:1347–1355.
70. Oldsen M, Hofmann, U., Quenzer, H.J.; Janes, J.; Stolte, C.; Gruber, K.; Ites, M.; Sorensen, F.; Wagner, B.. A novel fabrication technology for anti-reflex wafer-level vacuum packaged microscanning mirrors. *Proceedings of the SPIE – The International Society for Optical Engineering*. SPIE, 2008:688201–688211.
71. Kim J-M, Lee, S., Baek, C.-W., Kwon, Y., Kim, Y.-K.. BCB-based wafer-level packaged single-crystal silicon multi-port RF MEMS switch. *Electronics Letters* 2008;44:118–119.
72. Yun C.H., Martin, J. R., Chen, T., Davis, D. MEMS wafer-level packaging with conductive vias and wafer bonding. *TRANSDUCERS '07 & Eurosensors XXI* 2007 14th International Conference on Solid-State Sensors, Actuators and Microsystems, 2007:2091–2094.
73. Takegawa Y., Baba, T., Okudo, T., Suzuki, Y. Wafer-level packaging for micro-electro-mechanical systems using surface activated bonding. Source: *Japanese Journal of Applied Physics*, Part 1: Regular Papers and Short Notes and Review Papers, *Solid State Devices and Materials* 2007;46:2768–2770.
74. Raible S, D. Briand, J. Kappler, N. F. de Rooij. Wafer level packaging of micromachined gas sensors. *IEEE Sensors Journal* 2006;6:1232–1235.
75. Briand D., Guillot, L., Raible, S., Kappler, J., de Rooij, N.F. Highly integrated wafer level packaged MOX gas sensors. *TRANSDUCERS '07 & Eurosensors XXI*, 14th International Conference on Solid-State Sensors, Actuators and Microsystems, 2007: 2401–2404.
76. Ziae B., Jeffrey A. Von Arx, J. A., Dokmeci, M. R., K. Najafi. A hermetic glass-silicon micropackage with high-density on-chip feedthroughs for sensors and actuators. *Journal of Microelectromechanical Systems* 1996;5:166–179.
77. Lin L. MEMS post-packaging by localized heating and bonding. *IEEE Transactions on Advanced Packaging* 2000;23:608–616.

78. Cheng Y.T., L. Lin, K. Najafi. Localized silicon fusion and eutectic bonding for MEMS fabrication and packaging. *Journal of Microelectromechanical Systems* 2000; 9:3–5.
79. Weibel D.B., G. M. Whitesides. Applications of microfluidics in chemical biology. *Current Opinion in Chemical Biology* 2006;10:584–591.
80. Lee J.N., C. Park, G. M. Whitesides. Solvent compatibility of poly(dimethylsiloxane)-based microfluidic devices. *Analytical Chemistry* 2003;75:6544–6554.
81. Xia Y, G. M. Whitesides. Soft lithography. *Annual Review Materials Science* 1998;28: 153–184.
82. Cho C-W, Lim, C.-H., Woo, C.-S., Jeon, H.-., S. Park, B., Ju, H., Lee, C.-J., Maeng, S., Kim, K.-C., Kim, S. H., Lee, S.-B. Highly flexible and transparent single wall carbon nanotube network gas sensors fabricated on PDMS substrates. *Materials Research Society Symposium Proceedings*, v 922, Organic and Inorganic Nanotubes: From Molecular to Submicron Structures. MRS, 2006:95–101.
83. Zhang L., Cheung, K. Y., Chua, Y. R., Trau, D. PDMS microdevice with built-in optical biosensor array for oil-site monitoring of the microenvironment within microchannels. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, v 6445, Optical Diagnostics and Sensing VII, 2007:64450 W
84. Shin J-K, Kim, D.-S., Lim, G., Shoji, S.. Electrical detection of biomolecules in a PDMS micro-fluidic channel using a MOSFET-type biosensor. *The International Society for Optical Engineering*, v 6416, Biomedical Applications of Micro- and Nanoengineering III, 2007:641612.
85. Cho Y.H., N. Takama, T. Yamamoto, T. Fujii, B. J. Kim. MEMS based biochip for the characterization of single red blood cell. *Proceedings of the 3rd Annual International IEEE EMBS Special Topic Conference on Microtechnologies in Medicine and Biology*: Kahuku, Oahu, Hawaii, 2005:60–63.
86. Zhong L., X. Wang; Y. Jiang; X. Qiu, Y. Zhou, L. Liu. A silicon-based micro direct methanol fuel cell stack with compact structure and PDMS packaging. *Micro Electro Mechanical Systems IEEE*, 2007:891–894.
87. Varadan V., X. Jiang, V. V. Varadan. *Microstereolithography and other Fabrication Techniques for 3D MEMS*. Wiley, New York, 2001.
88. Rosen D.W.. *Stereolithography and Rapid Prototyping*. In: Hesketh P.J., (ed.). *BioNanoFluidic MEMS*. Springer: New York, 2008:175–196.
89. Han A., O. Wang, M. Graff, S. K. Mohanty, T. L. Edwards, K.-H. Han, A. B. Frazier. Multi-layer plastic/glass microfluidic systems containing electrical and mechanical functionality. *Lab on a Chip* 2003;3:150–157.
90. Han K.H., R. D. McConnell, C. J. Easley, J. M. Bienvenue, J. P. Ferrance, J. P. Landers, A. B. Frazier. An active microfluidic system packaging technology. *Sensors and Actuators B* 2007;122:337–346.
91. Tse A.L., P. J. Hesketh, J. L. Gole, D. W. Rosen. Rapid prototyping of chemical sensor packages with stereolithography. *Microsystems Technology* 2003;9:319–323.
92. Singh J, M. Ekaputri. PCR thermal management in an integrated Lab on Chip. *Journal of Physics: Conference Series* 2006;34:222–227.
93. Covington J.A., J.W. Gardner, A. Hamilton, T.C. Pearce, S.L. Tan. Towards a truly biomimetic olfactory microsystem: an artificial olfactory mucosa. *IET Nanobiotechnology* 2007;1:15–21.
94. de Reus R, C. Christensen, S. Weichel, S. Bouwstra, J. Janting, G. Friis Eriksen, K. Dyrbye, T. Romedahl Brown, J.P. Krog, O. Sondergaard Jensen, P. Gravesen. Reliability of industrial packaging for microsystems. *Microelectronics Reliability* 1998;38:1251–1260.
95. Zhu X, D. M. Aslam, Y. Tang, B. H. Stark, K. Najafi. The fabrication of all-diamond packaging panels with built-in interconnects for wireless integrated microsystems. *Journal of Microelectromechanical Systems* 2004;13:396–405.
96. Choudhury A. A piezoresistive microcantilever array for chemical sensing applications. Ph. D. Dissertation in Mechanical Engineering. Georgia Institute of Technology: Atlanta, USA, 2007, 178.

Nanobiosensing Electronics and Nanochemistry for Biosensor Packaging

Dasharatham G. Janagama and Rao R. Tummala

Abstract The nanobiosensor technology, an integral part of bioelectronics, is revolutionizing the health care industry, forensic medicine, home-land security, food and drink industries, environmental protection, genome analysis of organisms and communications. In this chapter, the biosensing electronics building blocks such as biosensing, signal conversion and signal processing, and biosensing mechanisms have been described. The use of smart nanomaterials such as Zinc oxide (ZnO) nanowire and carbon nanotubes (CNTs) in fabrication of biosensors has been described. A critical account on biological probe design, probe preparation and interfacing with signal transducer element of the biosensor is given. Various transducer elements and biosensing methods have been enumerated. The application of nanochemistry for biosensing packaging, with special reference to surface modification and biofunctionalization of sensor devices, and integration of biosensor system with system-on packaging (SOP) are detailed. Finally, the chapter concludes with future trends in biosensor technologies and a brief summary.

Keywords: Bioelectronics · nanobiosensor building blocks · bioprosbes · nanochemistry · surface modification · biofunctionalization · nanobiosensor packaging

Chapter Objectives

- Define bioelectronics, nanobioelectronic sensors, and their applications.
- Describe fundamentals of biology related to interfacing biological material with electronic systems.

D.G. Janagama (✉)

Package Research Center (PRC), Georgia Institute of Technology, Atlanta, GA, USA
e-mail: jdgoud@hotmail.com

- Present the elements or building blocks of nanobioelectronic sensors, and fabrication methodologies.
- Describe functionality of nanobioelectronic sensors and nanopackaging
- Identify the challenges in interface mechanisms between biosensor element and transducer; signal transduction, processing, and monitoring.

1 Introduction

1.1 Bioelectronics and Biosensors

Bioelectronics, including biosensors, is an emerging interdisciplinary field encompassing bioscience, chemistry, physics, material science, electronics, and engineering. A key factor of bioelectronics is in understanding the mechanisms of interfaces between biological materials and electronics. The miniaturization of electronic circuits laid the foundation for building microelectronic devices and modules. Typical bioelectronics devices are biosensors, biochips, electronic pills, electronic eye, electronic nose, genetic toggle switches, biocomputing devices, and bioactuators (Fig.1). Bioelectronics is revolutionizing the health care industry, forensic medicine, food and drink industries, environmental protection, genome analysis of organisms, and communications.

Biosensor is an analytical device, which converts a biological response into an electrical signal. It facilitates selective identification of biological materials such as peptide (two or more amino acids linked by the carboxyl group of one amino acid to the amino group of another), protein, deoxyribonucleic acid (DNA), and ribonucleic acid (RNA). The biochip is a collection of miniaturized test sites (microarrays), which perform thousands of biological reactions involving protein or DNA in a few

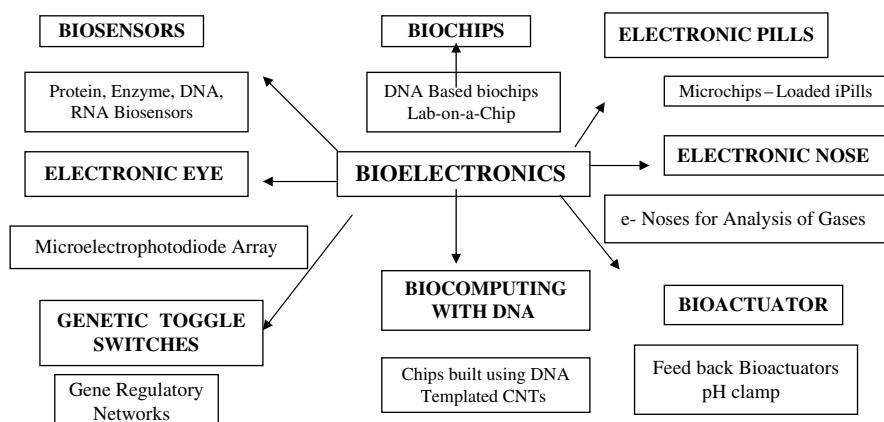


Fig. 1 Schematic of bioelectronics

seconds. The concept of electronic pills is the development of noninvasive microelectronics for monitoring physiological processes *in situ*. Electronic eye consists of a smart camera-computer device, which facilitates blind people to perform visual tasks. Electronic nose device typically works as an animal nose. It recognizes any compound or combination of compounds. This device is very useful in the detection of spoilage of food stuff and pathogenic contamination, oil spills, and environmental pollutants. Genetic toggle switches are gene-regulatory circuits with addressable cellular memory unit. They are used in biotechnology, biocomputing, and gene therapy applications. The strands of DNA (or RNA), the computer code of life, are used to create molecular-size circuits of self-assembly, which revolutionized the computing capabilities. It is known to surpass the capability of a super computer. Bioactuators are microsensor-based feedback devices, which are used, for example, in the detection of tumor microenvironment and biophysical cancer treatment. Biofuel cell is a device that converts the biochemical energy into electrical energy. They are used as small power sources for implanted medical sensors, actuators, and telemetry devices. Biofuel cell could replace rechargeable batteries in the near future.

1.2 Fundamentals of Biosensors and Micro/Nanosystems Packaging

Among the bioelectronic devices, biosensors came into prominence because of vast application potentials. A sensor is a device that senses or measures a real-world condition such as physical contact, motion, heat or light, and converts the impulses into analog or digital representation. The biosensors are used for detection, identification, quantification, and analysis of target molecules. A multifunctional sensory system, the nervous system, is evolved in the animals for sensing physical touch or contact, taste, smell, vision, and hearing. For example, the sense of smell is initiated by sensory receptors in the nasal epithelium (sensing element) upon inhaling certain molecules, carried through the neuronal pathway (transducer), and passed to the olfactory portion of the brain for processing the information as to the nature of gaseous molecules that come into contact with nasal epithelium in the nose (Fig. 2). These sensory organs in animals help in sensing the nature of surrounding environment, ultimately aimed at self-preservation and survival.

A typical biosensor is made up of three basic parts, namely biosensing, transducer or signal conversion, and signal processing elements (Fig. 3). In a biosensor, the sensing or biorecognition element (biosensing) is a protein, peptide, deoxyribonucleic acid (DNA), ribonucleic acid (RNA), whole cell or tissue, which interacts or hybridizes specifically with its complementary target or analyte molecules. It is the primary part of the sensor, and determines the selectivity and sensitivity of the sensor. The transducer or signal conversion is a device such as an electrode, piezoelectric crystal, photosensitive element, nanobelts/wire, carbon nanotubes (CNTs) or thermistor, which interfaces with the recognition element, and converts the magnitude of the biological relevant information or response signal into an electronic

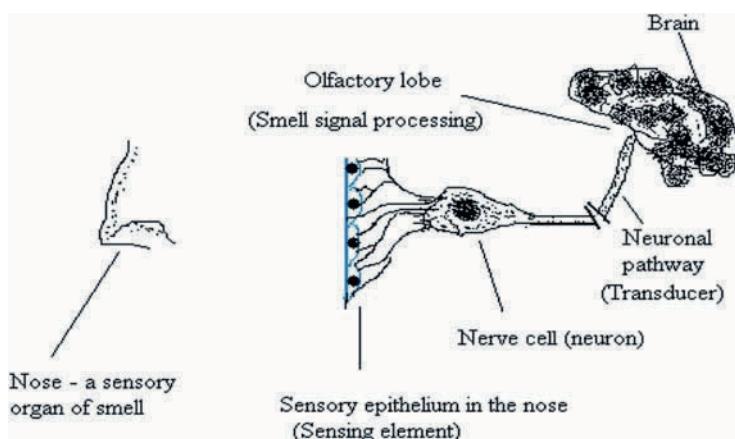


Fig. 2 Schematic of mechanism of sense of smell – a sensory system

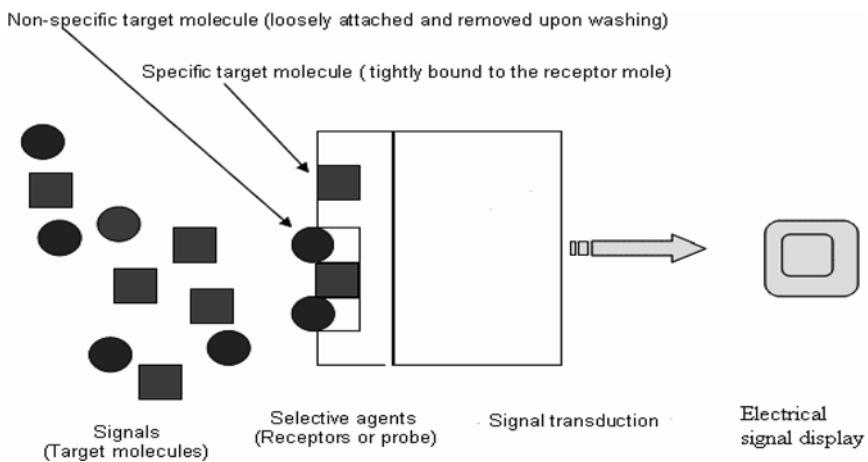


Fig. 3 Principle of a biosensor

signal. The vast gene diversity among animals and plants is manifested in the major biomolecules such as proteins, peptides, DNA, and RNA. Biosensors are used for detection, identification, quantification, and analysis of biological samples.

Biosensors are categorized into two main types based on the location of lodging of the device: *In vivo* or implantable biosensor, if inserted into the cell, tissue, or body for sensing the analyte; and *in vitro* biosensors, if placed in an artificial environment external to the cell, tissue, or body. The biosensor devices have been further categorized based on the types of molecules detected into nucleic acid sensors and DNA chips, immunosensors, enzyme-based biosensors, organism and whole cell-based biosensors, natural and synthetic receptors-based biosensors

(Fig. 4). Innovative signal transduction technology, systems integration, proteomics, bioelectronics, and nanoanalytical systems are essential for the development of smart biosensors.

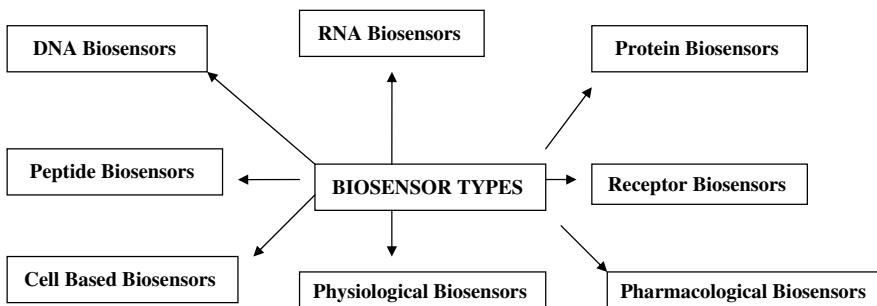


Fig. 4 Schematic of different types of biosensors

The conventional detection system involves labeling of the probe molecules, requirement of large amount of target molecules to enable detection, and elaborate signal detection methods. Further, the conventional technologies based on the “top-down” approaches are difficult to continue to scale down due to limitations in lithography and interconnect schemes for design and fabrication of biosensors.

The novel properties or unusual combinations of properties possessed by nanostructured materials are utilized for fabrication of nanobiosensors. A “bottom-up” technology incorporating the integration of nanotechnology with biology and bioelectronics constitute the third generation of smart biosensors (Fig. 5).

Nanoscience and nanotechnology involves in manipulating the materials at the scale of 100 nm or less. The advancement in micro/nanofabrication technologies enables fabrication and packaging of varieties of functional devices using nanomaterials for biological and nonbiological applications. The applications range across material science, information technologies, environmental protection, and medicine. As many biological functions occur at micro and nanoscale in the cell and subcellular systems, the nanodevices help in potential life science applications including improved drug and gene delivery, biocompatible materials for implants, and advanced sensors for early disease detection of diseases and therapies. These nanodevices made with smart nanomaterials are efficient, cost effective, and easily implantable.

The biosensor packaging deals with integrating a number of functional active and passive components, such as biosensing elements, biosignal conversion elements, signal processing elements, microfluidic systems, reagent amplification system into a single sensor chip and package. Lab-on-chip (LOC) concept refers to translating the entire laboratory testing processes and functions onto a tiny microelectronic device. The microfluidics incorporated into the package facilitates the flow of the samples and reagents in the chip. Biosensor detection systems

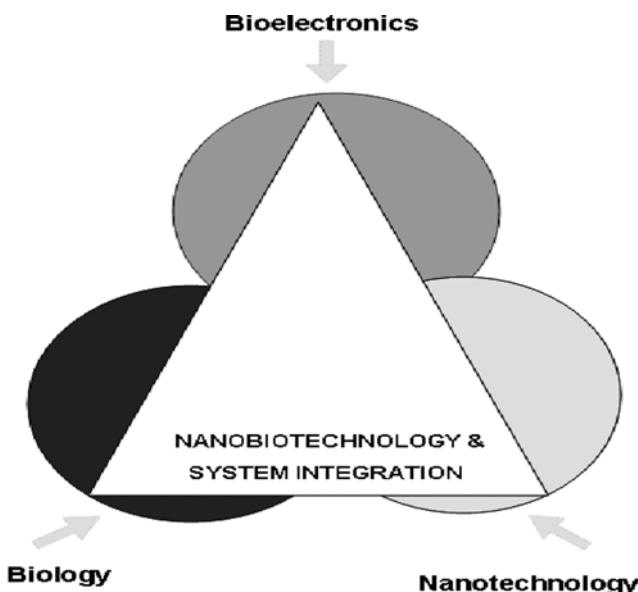


Fig. 5 Integration of nanotechnology with biology and bioelectronics to produce nanotechnology-based biosensors

applied to analyze DNA–DNA (deoxy ribonucleic acid), antibody–antigen, protein–protein, protein–nucleic-acid, protein–lipid and protein–small-molecule interactions, and enzyme–substrate interactions. The biosensor packing includes surface modification chemistry, capture molecule attachment (biofunctionalization) with detection methods and high throughput. However, for complete biosystem integration, the signal and power distribution, I/Os, and suitable interfaces are to be incorporated.

In this chapter of biosensors, we describe the application of biosensors in various walks of life, fundamentals of biology as applied to biosensors, biomechanisms, building blocks of biosensors, biological probe design and interfacing, various transducer mechanisms, challenges in integration of biomolecules with interface mechanisms and signal transduction, integration of nanobiosensors with electronics and fluidics using system-on-package (SOP) platform technologies, and future trends in biosensors. The following nine sections will briefly describe the above in a systematic way.

In Section 1, a brief introduction to bioelectronics and biosensors with emphasis on the interdisciplinary nature of the subject, and historical account of the development of biosensors systems is presented.

In Section 2, the applications and importance of bioelectronics, and biosensors in health care, food hygiene, environmental safety, industrial bioprocess control and monitoring, genome analysis, drug discovery, and drug abuse are described.

In Section 3, fundamentals of living cells as applied to different types of sensors are described citing the examples of prokaryotic cells and eukaryotic cells, DNA, RNA, proteins, and peptides.

In Section 4, building blocks of biosensors such as target molecules, probe molecules and their biomechanisms, interface mechanisms, and signal transduction are mentioned. The molecular aspects of biorecognition elements such as antibodies, nucleic acids (DNA/RNA), natural and synthetic receptors, enzymes and their biomechanisms are discussed.

Detailed account of biosensing mechanisms is dealt in Section 5.

Section 6 deals with various transducer types and their mechanism of interfacing and signal transduction, and challenges in biofunctionalization of the devices and signal transduction are mentioned.

Nanostructures, biomolecules, and cells as components of biosensors are described in Section 7

In Section 8, design and preparation of biological probes such protein and DNA for biofunctionalization of nanostructures, fabrication methods, and nanopackaging are described.

Section 9 deals with materials characterization of biosensor structures.

Section 10 deals with the future trends in nanobiosensors. As anticipated, discovery of novel materials and innovative approaches, it may progress by leaps and bounds, as in the case of the computer industry after the development of integrated circuits. A brief summary of nanobiosensors is given.

2 Applications of Bioelectronics

The recent advances in miniaturization of electronic circuits provided the technological capability to utilize biological materials for fabricating sensors and circuitry. The microelectronics are the basis for biosensors, biochips, implantable medical devices, prosthetic devices, artificial organs, including electronic nose and electronic eye, electronic pills, surgical and medical devices, biofuel cells, molecular motors, molecular electronics, biofabrication templates, which endow enormous power to the detection system.

2.1 Applications of Biosensors

The biosensors with electrical detection system are superior over the conventional system of immunohistochemical method of detection of biological samples. The importance of critical detection of hazardous pathogen and toxic molecules in the environment was realized by the events of September 11, 2001 attack on US soil, and the subsequent anthrax attacks, and ricin (a potent poison from castor beans) attacks in the United Kingdom, which highlighted the value of using sensitive biosensors and bioelectronics all over the world. The important applications of biosensors are as follows:

(a) Clinical Diagnostics

The clinical biosensors offer sensitive and precision detection of physiological, prenatal, and postnatal diagnosis of genetic diseases and abnormalities. These biosensors are capable of detecting single-nucleotide polymorphisms, which are caused by gene mutations, including single-base substitution and small insertions or deletion bases in the genome. Rapid and high-throughput detection methods using biosensors are in dire need for early detection of breast cancer, prostate cancer, AIDS, genetic diseases, bacterial and viral infections, and bacterial and viral drug resistance. The other infectious and communicable disease detection include amebiasis, chickenpox, cholera, cyclosporiasis, diphtheria, encephalitis, hand, foot-and-mouth disease, hepatitis A, malaria, measles, meningitis, rubella, and West Nile virus. The clinically relevant point mutations are detected by a novel piezoelectrical biosensor [1].

(b) Drug Discovery

The cell-based biosensors (CBB) are engineered to optically or electrically report specific biological activity. The cellular biosensors are comprised of living cells and can be used in various applications, including screening chemical libraries for drug discovery. Panels of biosensors may also be useful for elucidating the function of novel genes.

(c) Drug Abuse

Nanoscale detection and quantification systems for various drug abuses are essential for law enforcement authorities to control the menace of drug abuse. Cocaine sensor has been developed for rapid detection and quantification of coca alkaloids [2]. DNA biosensors and DNA chips are the potential detection devices for investigation and evaluation of DNA–drug interaction mechanisms.

(d) Genome Analysis

The use of biosensors in basic research to measure biomolecular interactions in real time is increasing dramatically. Applications include protein–protein, protein–peptide, DNA–protein, DNA–DNA, and lipid–protein interactions. Such techniques have been applied to, antibody–antigen, receptor–ligand, signal transduction, and nuclear receptor studies. Biosensors are also useful for the detection of genetically engineered products, genome analysis [3], and detection of mutations in plants and animals by identifying DNA sequences. There is a great demand for cheap, portable, and easy-to-use devices for point-of-care analysis. These systems integrate on the same substrate the sites of the reaction, the sensors providing electric signals and the circuits for signal conditioning and amplification.

(e) Food Hygiene

Food hygiene is of priority for any nation aspiring to be healthy and sound. The protection of dairy products from Botulinum toxins; foods from poisonous bacterial and fungal contamination; pathogens devastating animal livestock; protection of meat from mad cow disease causing pathogens; castor bean ricin that are poisonous to people, animals, and insects are some of the well-known hazards, which should be monitored and detected in time to prevent a national calamity. These biomolecules can be offered as the powerful weapons and “road map for terrorists” to destroy humans and livestock indiscriminately. *Salmonella* bacteria biosensor was developed, which detects the pathogen in food and water as low as 10^4 cell colony forming unit/ml [4]. The timely detection of bird flu caused by the virus H5N1 will save the poultry industry and public health all over the world. Biosensors are also useful in monitoring the freshness of aquatic foods including fish, and fermentation processes.

(f) Industrial Bioprocess Control and Monitoring

Bioprocess control and monitoring biosensors are important technological inputs in biotechnology industry. DNA chip array biosensors based on microelectrode arrays, which are fabricated with silicon technology and activated with oligonucleotide probes that bind to the target sequence, monitors the expression of selected genes. These devices are critical for production processes using microorganisms in bioreactors, particularly the protein products generated by recombinant DNA technology during microbial cell or mammalian cell fermentation. Further, they are used for controlling/monitoring extracellular parameters in a bioreactor. Thermal biosensors, for instance, are used for specific temperature monitoring and controlling of bioprocesses. Addition of a flow injection analysis system (FIC) to the biosensor facilitates the control delivery of all reagents to the bioreactor. We described electrical detection of enzymatic hydrolysis of starch using ZnO nanobelts/wire. The detection of the degree of hydrolysis of starch converting into glucose, cyclodextrins, and other monomer products is very critical in food and pharmaceutical industries. Further, real-time and on-line detection diction of starch hydrolysis events could be monitored through electronic detection system [5].

(g) Environmental Safety

Biosensors are widely used in protecting the environment by monitoring of pollutants in water, air, oils, soil, toxic vapor, acid rain, and odor in the atmosphere. Analysis of explosive trinitrotoluene (TNT) in ground water with detection limit of TNT of 20 $\mu\text{g}/\text{ml}$ using portable fiber-optic biosensor was reported [6]. Pesticides in the water are detected using in situ optical biosensor with a detection limit of 2 ppm for atrazine and alachlor [7]. Ozone detection system was developed using fiber-optic near-infrared evanescent wave, which detects ozone over the range of 0.02–0.35 vol%, in the response time of about 1 min

[8]. Sensitive and rapid biosensors are required for identification of pathogenic organisms in dealing with bioterrorism. Metro subways and other sensitive places need biosensors, which should detect toxins and explosive materials in trace amounts. The detection of industrial toxic effluents and human and animal waste into water bodies is an important aspect of environmental safety and protection.

2.1.1 Advantages of Biosensors over Conventional Detection Systems

The use of nanotechnology-based biosensors confers several advantages over the conventional detection systems. Magnetic materials show super magnetic behavior at nanoscale [9]. The semiconducting nanoparticles show resonance tunneling and Coulomb blockade effects [10]. Nanoparticles undergo widening their band gap energy as the material dimensions reduce, resulting in blueshifts in optical properties caused by a quantum confined effect. These unique properties do not exist in their bulk counterparts. Therefore, the functional nanostructures composed of metals, semiconductors, magnetic materials, quantum dots, molecular dyes, and other polymers used in fabricating nanobiosensors, which show the potential advantages over the conventional types:

- In a biosensor, no involvement of exogenous molecules or labels such as conjugation with enzyme, radioactive fluorescence, or chemiluminescence molecules.
- The probe and target-binding events or enzyme and substrate reactions are recorded by the transducer in the absence of any label requirement.
- A polymerase chain reaction (PCR) free detection systems for DNA identification.
- It is a nanoscale detection of the target molecules, a key factor in early detection of diseases such as breast cancer and AIDS.
- Rapid and high-throughput detection.
- Detection processes are simple, user friendly, fast, and cost effective.
- Reduced material requirement to fabricate the devices.
- Reduced power requirement, and easier recycling.
- Novel properties and new capabilities.
- Repetitive, portability, and stability.

2.2 Applications of Other Bioelectronic Sensing Devices

(a) Multielectrode Array (MEA)

Microelectrode array is a conventional passive MEA tool to record electrical activity in an intact cellular environment of cell cultures, brain slices, and other tissues. It provides a simple interface for monitoring the electrical activity and impedance characteristics of populations of cultured cells over extended periods. The cell contact with electrode is created as cells attach directly to planar electrode structures.

It is used in drug screening, drug delivery, and monitoring of drug side effects. MEA technology is used in the study of neurite outgrowth, synaptogenesis and regeneration, and modulation of these processes by drug, neuronal stimulation, and recording. Figure 6 shows the layout of the electrodes fabricated on the cell culture dish (a), and cell culture set up (b), may be neuronal or any cell-type experimental cultures.

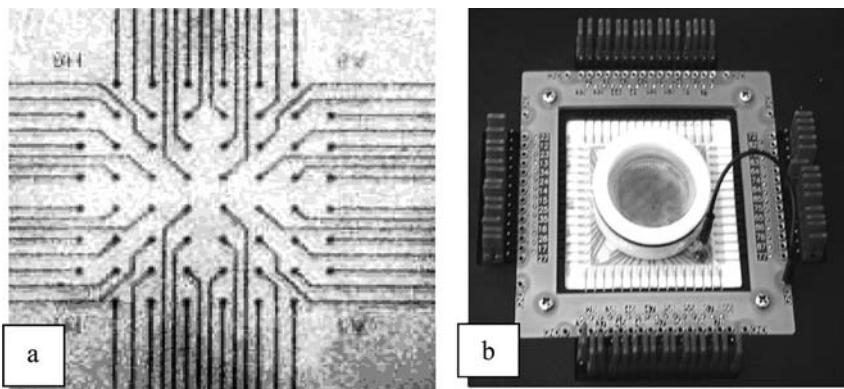
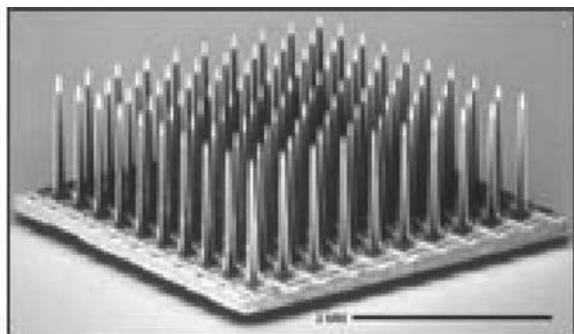


Fig. 6 Shows 2D multielectrode array (MEA): (a) layout of electrodes, (b) MEA assembly with cell culture system (Source: Courtesy of Dr. Daniel A. Wagenaar, 2003)

Microelectrodes array (MEA) of microscale dimensions of Si-fabricated structures is advanced microdevices. They are used as retinal prosthetic devices to stimulate retinal nerve cells as a viable treatment for degenerative retinal diseases, and several other applications. Figure 7 shows miniature array of microelectrodes. The array consists of a 4 mm × 4 mm base that contains 100 silicon spikes of 1–1.5 mm long.

Fig. 7 Shows 3D MEA
(Source: Courtesy of Cyber Kinetics)



(b) Microphtodiode Array (MPDA)

Microphtodiode array (MPDA) is a micromachined photovoltaic device, used as retinal prosthetics to replace degenerated photoreceptors of the retina of the eye.

MPDA in the subretinal space transforms light stimuli into electrical current for the stimulation of still healthy retinal neurons. They are fabricated as silicon-based multilayered microstructures.

It is made of silicon measuring 2 mm^2 , implanted into the retina of the eye. As the silicon material is not biocompatible, other suitable material for implantation is highly sought after. The photodiodes are connected in series and can be independently stimulated. Figure 8 shows the schematic of MPDA implant in the retina of the eye.

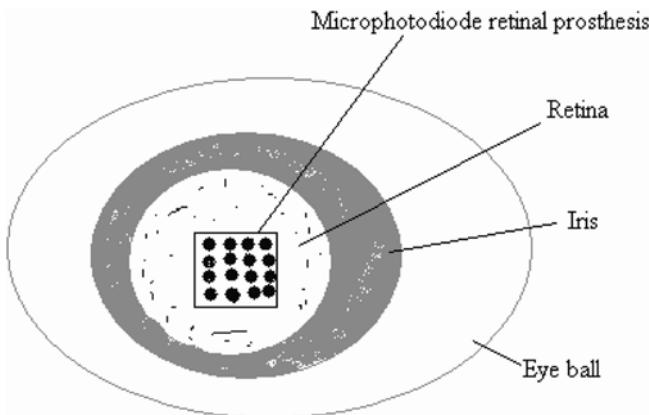
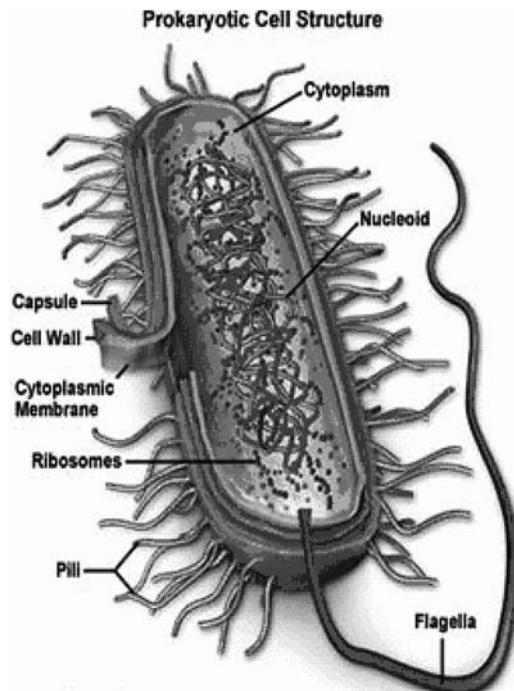


Fig. 8 Schematic of MPDA implant in the eye

3 Fundamentals of Biological Materials as Sensing Elements

The knowledge of fundamentals of living cells and biomolecules is a prerequisite for understanding the concepts of biosensors. All biological functions in the cell or organism such as digestion, excretion, respiration, reproduction, and sensing are carried out in the structures at nanoscale. The living cells in organisms have been categorized into prokaryotic cells with out-defined nucleus, and eukaryotic cells with well-defined nucleus. The prokaryotes are less evolved unicellular organisms such as bacteria and virus with relatively simple subcellular structure and organization. Most cells are between $10 \mu\text{m}$ (micrometers) and $100 \mu\text{m}$ in diameter. The human cell is about $10 \mu\text{m}$ in diameter, and the bacterium is generally rod shaped with dimensions $1 \times 2 \mu\text{m}$. The bacterial cell consists of a cell wall enclosing the cytoplasm, in which the genetic material (DNA/RNA), mesosomes, ribosomes, and other organelles are distributed (Fig. 9). The functional aspects of prokaryotic cells are related directly to the structure and organization of the macromolecules in their cell make-up, i.e., DNA, RNA, phospholipids, proteins, and polysaccharides unlike in eukaryotes in which specific organelles are involved. The molecular diversity in the primary structure of these functional molecules is responsible for the diversity that exists among prokaryotes. The bacteria play a vital role in maintaining the

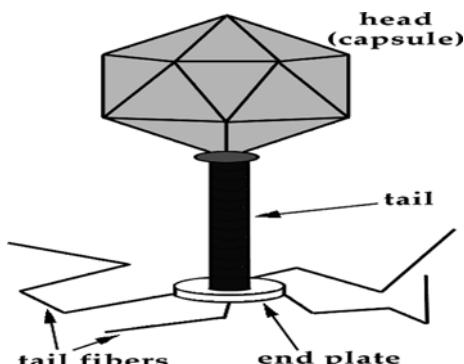
Fig. 9 Schematic diagram of a typical bacterial cell (note the absence of nucleus)
 (Source:<http://www.williamsclass.com/images/JPGImages/BacteriaCell.jpg>)



earth as a suitable place for inhabitation by other forms of life as they involve in putrefaction or decay, and cleaning up of dead organisms. Many biotechnology products are obtained from useful bacteria, such as antibiotics, recombinant gene products, therapeutics, dairy products, and breweries.

A virus particle, also known as a virion, is essentially a nucleic acid (DNA or RNA) enclosed in a protein shell or coat (Fig. 10). Viruses are extremely small, approximately 15–25 nm in diameter.

Fig. 10 Schematic diagram of a typical virus (Source: Emiliani, C. 1993. *Extinction and viruses*)



The eukaryotic cells are represented in protozoa (single celled), mesozoa (few celled), and metazoan (multicelled) – the major branches of the animal kingdom. The eukaryotic cell, as in higher animals, has well-defined substructures or organelles with specific cellular functions. Unlike in bacteria and virus, the cell is covered with a double membrane, which is made up of mainly protein, lipid, and carbohydrate molecules.

DNA is a double helix of complementary strands, which contain the code for operation of all living organisms. As shown in Fig. 11, DNA is made up of four standard bases in long chains, known as adenine (A), guanine (G), thymine (T), and

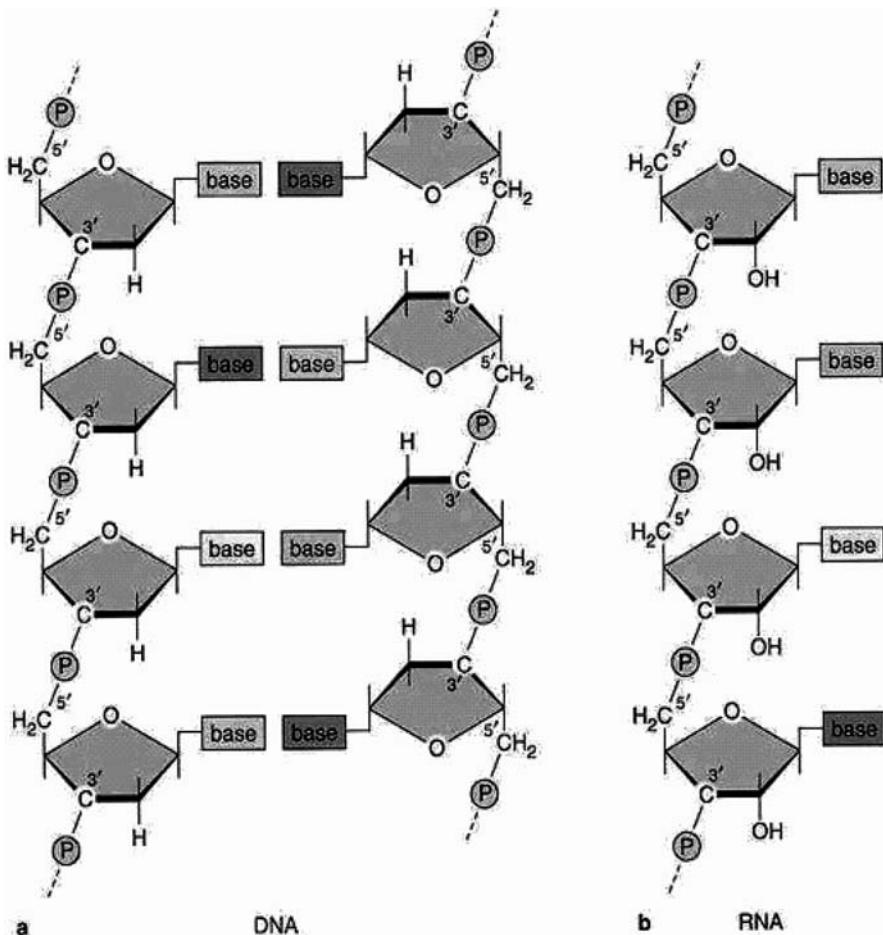


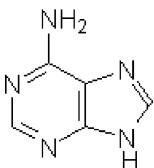
Fig. 11 Schematic of DNA and RNA structures. DNA double helix interconnected by bases (A-T, G-C). RNA is a single-stranded structure with uracil (U) base in place of thymine (T) (Source: <http://faculty.uca.edu/johnc/DNA%20structure.jpg>)

cytosine (C). A–T and G–C are the complementary base pairs. Adenosine and guanine are purine bases, whereas cytosine, thymine, and uracil are pyrimidine bases. The sequences of nucleotides define “genes.” The genes define RNA and proteins. Of the 22 amino acids commonly found in nature, a series of 3 DNA bases (triplet code) is enough to correspond one amino acid, which is called universal code, means the code applied to all organisms to follow ($4 \times 4 \times 4 = 64$ possible combination of four DNA bases).

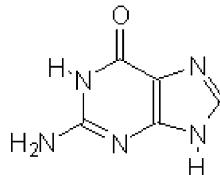
DNA carries the genetic information of a cell and consists of thousands of genes. Each gene serves as a recipe on how to build a protein molecule. RNA is a single-strand molecule. As in DNA, RNA contains four nucleotides, except the base uracil instead of thymine (Fig. 12).

The purines

Adenosine

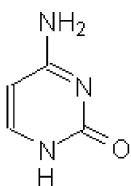


Guanine

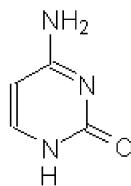


The Pyrimidines

Cytosine



Thymine



Uracil

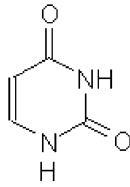


Fig. 12 Purine and pyrimidine bases in nucleic acids

A crucial property of the purines and pyrimidines is their ability to form hydrogen-bonded pairs composed of one purine and one pyrimidine. The constant distance of 1.085 nm between the two helical strands, as rungs in a ladder, is due to similar base-pairing pattern (G–C and A–T).

Amino acids are the building blocks (monomers) of proteins. Proteins and enzymes are essential structural and functional building blocks of all organisms. Generally, 22 different amino acids are used to synthesize proteins. The shape and other properties of each protein are dictated by the precise sequence of amino acids in it. By possible combinations of these amino acids, thousands of proteins are generated as thousands of words are formed using 26 alphabets in the English language.

Thus, variability in shape and size among different organisms is due to the vast diversity in their protein and enzyme structures.

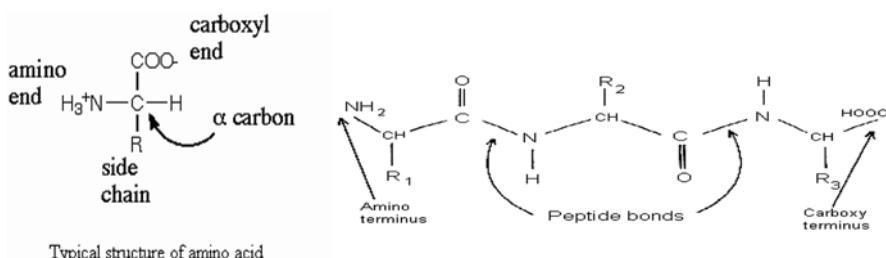


Fig. 13 Structure of a typical amino acid with amide and carboxylic group at the end. Two amino acids linked by peptide bond

Several amino acids linked through peptide bonds forming a functional gene product, namely protein (Fig. 13). Proteins are synthesized in the cytoplasm of the cell under the direction of DNA. The chemical messages in DNA are transcribed into RNA (m-RNA) in the nucleus. The m-RNA move out of the nucleus through nuclear pores into the cytoplasm, where with the help of ribosomes and transfer RNA (t-RNA), and other factors translate the m-RNA information into a specific long chain of amino acids called polypeptide. The polypeptide undergoes post-translational modifications and forms 3D functional protein and enzyme molecules.

Proteins perform important tasks for the cell functions or serve as building blocks in the form of structural proteins and enzymes. Antigens are functional proteins on the cell surface as well as within the cell. Antibodies are particular proteins made by cells in response to the antigens.

4 Building Blocks of Biosensors

As shown in Fig. 14, the main building blocks of biosensors are biosensing element, signal conversion or transducer, and signal processing.

In a cell based biosensors (CBB), the cell with receptor molecules acts as primary transducer, and the secondary transducer [11] is typically an electrical device (Fig. 15).

The interface and transduction works by electrical, optical, acoustic, mechanical, thermal, chemical, and magnetic methods. As shown in Fig. 16, the bifunctional membrane coated with probe molecules such as enzymes, antibodies, DNA, RNA, tissue, cells, organelles, receptors (natural or synthetic). Upon hybridization or affinity binding with the target molecules, a signal is generated in the form of production of chemical substance, weight change, light, sound, heat, or electrical impulse, which ultimately converts into electrical signal.

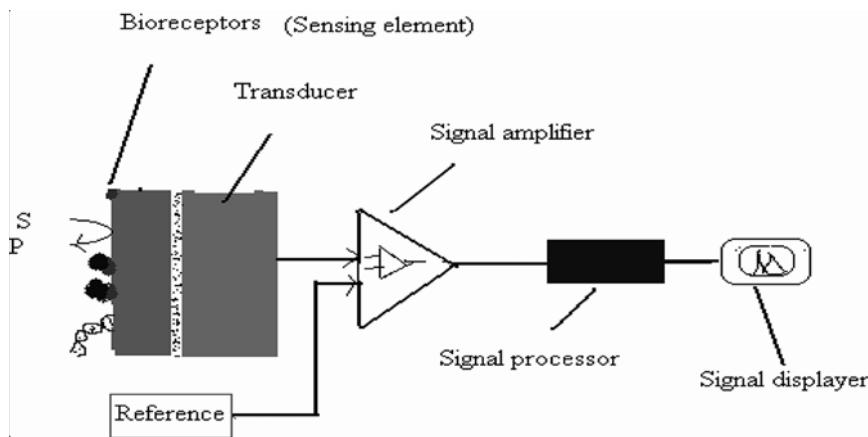


Fig. 14 Schematic of a typical biosensor: bioreceptors (note DNA, protein receptor molecules attached; s – substrate, p – products in an enzyme-based sensor), transducer, signal amplifier, signal processor, and signal displayer

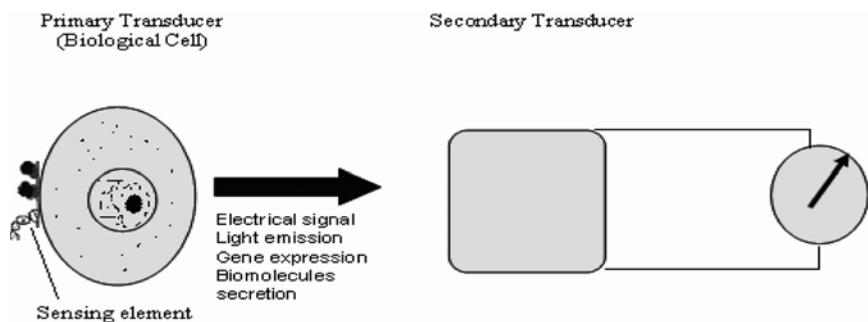


Fig. 15 Schematic of a typical cell-based biosensor showing primary and secondary components of the biosensor

The important biological component of a biosensor is a receptor, which recognizes the analyte or target molecules. These biorecognition elements are specific and sensitive for a particular substance but not others. There are four main groups of biological elements:

- 1 Antibodies (proteins)
- 2 Nucleic acids (DNA/RNA)
- 3 Natural and synthetic receptors
- 4 Enzymes (biocatalysts)

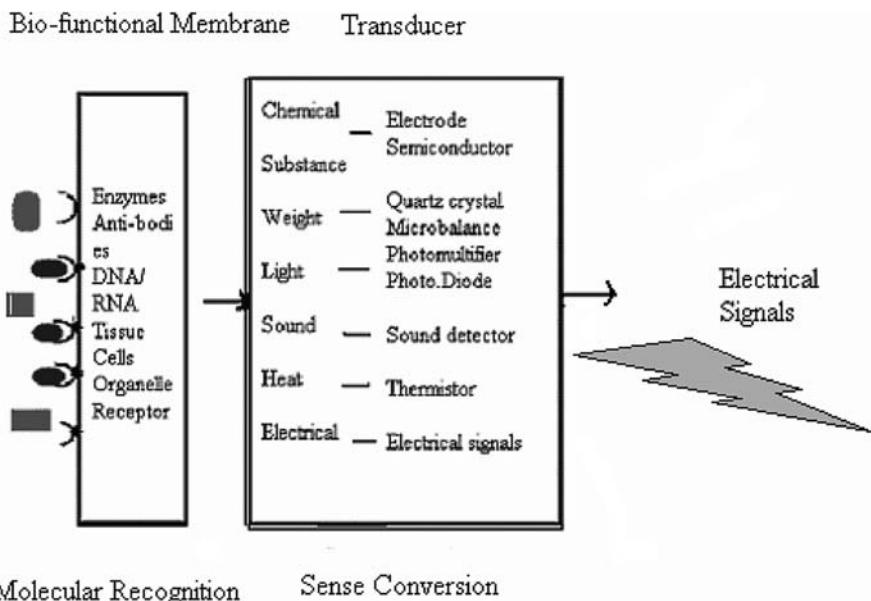


Fig. 16 Illustration of biomechanisms of biorecognition and signal transduction

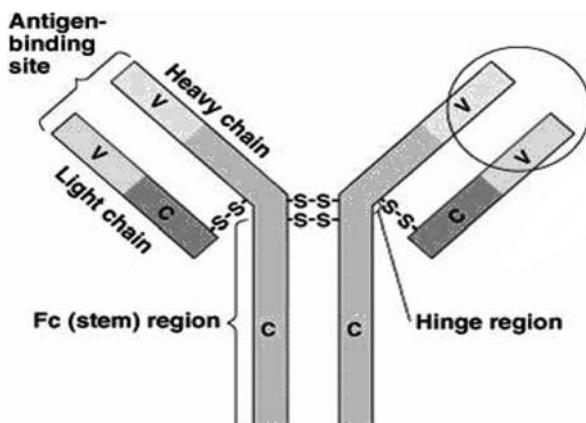
5 Biosensing Mechanisms

The biomechanisms for recognition of the important biomolecules such as proteins, DNA, and RNA lie in their molecular structure. They have a naturally evolved selectivity to biologically active analytes. The complementary structure of DNA and specific binding sites on antigen and antibody are utilized for the detection of these molecules for analytical purposes.

5.1 Immune Complex Formation

Antigen–antibody reactions are the binding processes of an antigen molecule with its specific antibody. This reaction involves site or epitope recognition on the antigen by the antibodies and then binding (Fig. 17). Epitope is a portion of a molecule to which an antibody binds. It can be composed of sugars, lipids, or amino acids. The epitopes can be visualized by using monoclonal antibodies specific for that epitope. The affinity reactions are recognized by labeling the antibodies usually with radioisotopes, enzymes, fluorescent probes, chemiluminescent probes, or metal probes. Generally, it is visualized in gel, western blot, or labeling via immunofluorescence. It is a reversible chemical reaction with high specificity. The main forces that play a role in these reactions are hydrogen bonds, electrostatic forces, Van der Waals bonds, and hydrophobic bonds. The stability of the antigen–antibody

Fig. 17 Schematic diagram of immunoglobulin molecular structure. Binding sites (epitopes) of antigen and antibody reaction lead to the formation of covalently bonded immune complexes, which are neutralized proteins (Courtesy of Indian River Community College, Health Science Center, US)



binding depends on the valency of antibodies, which are termed as avidity and affinity. Antibody upon cleavage with papain enzyme, separates into two proteolytic fragments in the hinge region, namely Fab and Fb. The Fab fragment retains the antigen-binding activity, binding to a monovalent antigen with an affinity nearly as high as that of the entire antibody molecule. The Fc fragment mediates class-specific functions, such as complement fixation, exhibited by IgG and IgM antibodies, leading to lysis of the intruding cell. The *in vitro* antigen–antibody reactions (serological tests) are widely used in blood grouping, diagnosis of infectious diseases, and detection of specific protein of interest. These tests are mainly based on the mechanisms of agglutination, precipitation, complement fixation, and solid-phase immunoassays such as enzyme-linked immunosorbent assay (ELISA) and immunosensors.

In a conventional method of immunodetection of protein, the primary antibodies (probe molecules) are coated onto the microtiter plate, and allowed the hybridization to take place with labeled specific target molecules (antigen). The molecular labeling could be radioisotopes, enzyme, digoxigenin, fluorescent, or chemiluminescent. After hybridization with the target molecules, the plate is washed to remove the unbound nonspecific target molecules. Figure 18 shows only the specific target molecules hybridized with probe molecules, and the nonspecific molecules cannot hybridize with the receptor molecules.

5.2 DNA Hybridization

DNA and DNA molecules are hybridized with complementary DNA, RNA probes (Fig. 19).

The sequence-specific DNA, RNA hybridized molecules are detected in Southern and Northern blots, respectively. The nonspecific DNA sequences are removed by washing. In conventional detection system, nucleic acids require radioisotopes,

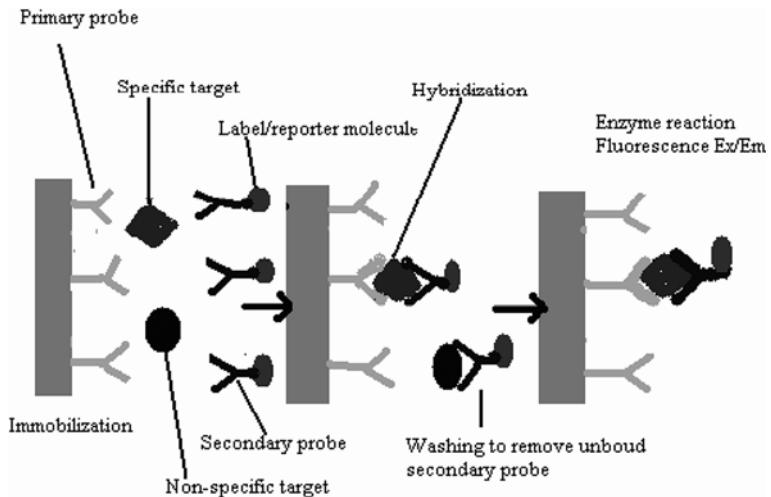


Fig. 18 Schematic of biosensing mechanism of antigen and antibody binding and immune complex formation

Watson-Crick base pairing: **A T T T C A T C G G (Target)**

T A A A G T A G C C (Probe) Complementary structure

Fig. 19 DNA and RNA molecules are hybridized with complementary DNA, RNA probes (Fig. 19b).

enzyme, fluorescent, chemiluminescent, or metal labeling for the detection of hybridization signals.

5.3 Natural and Synthetic Receptors

Templating techniques, such as molecular imprinting (MIP) on polymers is used to form natural or synthetic receptor sites on polymers [12]. The concept of cell imprinting is based on stamping of cells with minimum force into a polymerizing sensor layer. The patterned surface remained after the polymerization recognizes cells depending on geometrical fit and chemical interaction [13].

Yeast-imprinted polyurethane sensor layers in combination with microbalances were shown to detect yeast cells without cross-selectivity to bacteria [14]. Mass-sensitive detection of cells, viruses, and enzymes with artificial receptors was performed using lysozymes, tobacco mosaic virus (TMV), and erythrocytes-imprinted polymers [15].

5.4 Biocatalysts

The catalytic biosensor is based on the recognition and binding of an analyte followed by catalytic conversion of the analyte. The reaction is detected by the transducer. Enzymes are the natural protein biocatalysts that specifically bind to a unique substrate or group of substrates. Enzyme stability is the key factor in determining the stability of enzyme biosensors. Biocatalysts include enzymes, organelles, whole cells, microorganisms, and tissues.

6 Functionality of Various Transducers and Challenges

The biosignals are converted into electrical, optical, and electromagnetic signals by a variety of signal conversion methods as described below:

- Electrochemical (conductimetric, amperometric, and potentiometric)
- Optical
- Acoustic wave
- Microelectromechanical (resonant)
- Magnetic
- Thermal

6.1 Electrochemical Method

Electrochemical fluid cell is used for the field-effect transistor (FET) detection of DNA or protein. It involves electrical impedance measurements of the probe and target molecules using a three-electrode electrochemical cell containing DNA or protein-modified Si sample as working electrode, Ag/AgCl reference electrode, and Pt counterelectrode. A poly dimethylsiloxane (PDMS) fluid cell ($5\text{ }\mu\text{l}$ volume) is pressure sealed between DNA-modified Si sample ($8 \times 4\text{ mm}$) and the Pt counterelectrode (Fig. 39). Impedance is measured using a three-electrode potentiostat coupled to an impedance analyzer. All electrochemical potentials considered with respect to the saturated Ag/AgCl reference electrode [16]. The multiwalled carbon nanotubes dispersed on the working electrode forms the efficient working electrode (Fig. 20). The impedance is proportional to the target–probe molecular hybridization.

6.2 Optical Method

In optical signal conversion method, the bioprobe and target hybridization signals are converted into optical signals. For example, surface plasmon resonance (SPR) is an advanced optical signal conversion technique. The basic SPR configuration

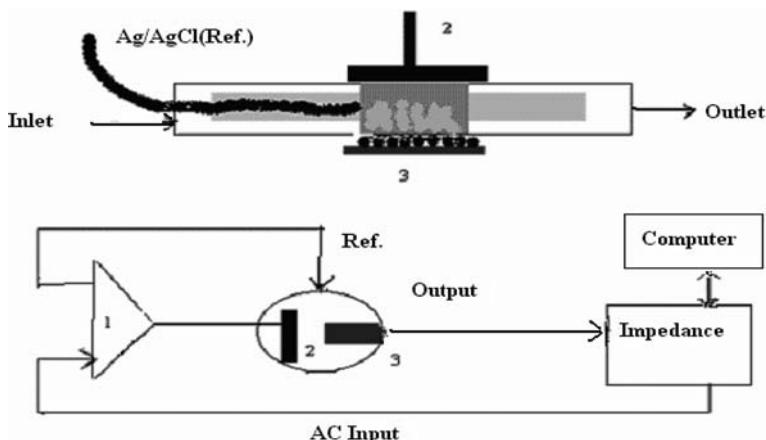


Fig. 20 Three-electrode chemical fluid cell biosensor. 1 – Potentiostat; 2 – platinum foil counter-electrode; 3 – working electrode as DNA-modified silicon sample (Source: Courtesy of Cai et al., 2003)

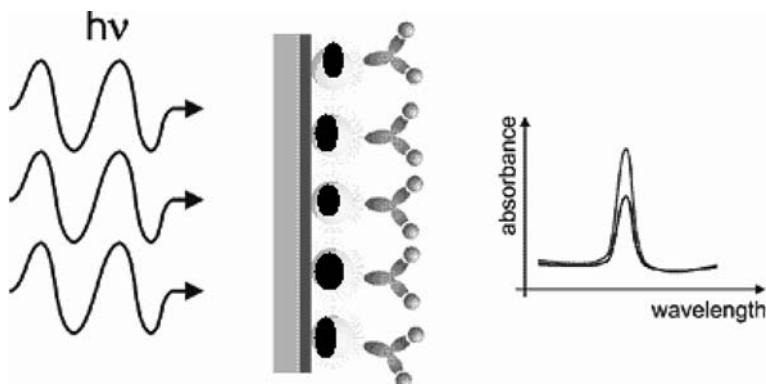


Fig. 21 Schematic of optical biosensor. Gold and silver nanoparticles are deposited on a quartz substrate. The substrate is coated with mercaptosilane adhesion layer and functionalized using gold or silver nanoparticles, self-assembled monolayers of functional thiols, antibodies, and antigen. The resulting absorption spectrum increase upon binding of analytes to the functional nanoparticles (Courtesy of A. Campitelli, C. Van Hoof, IMEC)

consists of a prism or glass slide coated with a thin metal film, usually silver or gold (Fig. 21). Bioprobe such as DNA or protein is immobilized on the thin gold film. When the light is passed through the prism and onto to the gold surface at angles and wavelengths near the surface plasmon resonance condition (the intensity of the reflected light is reduced at a specific incident angle producing a

sharp shadow), the optical reflectivity of the gold changes very sensitively in the presence of the gold surface. Similar effect is found using quartz substrate coated with gold and silver nanoparticles. The high sensitivity of the optical response is due to very efficient collective excitation of conduction electrons near the gold surface.

The optical components in an optical sensor could be fiber optics, wave guides, photodiode, spectroscopy, charge-coupled device (CCD), and interferometers. The changes in intensity, frequency, phase shift, and polarization are measured. The type of measurements involved are absorbance (oligo – 260 nm, peptide – 280 nm), fluorescence, refractive index, and light scattering. Surface plasmon resonance, particularly useful for the detection of biological molecules: (1) they can be extremely sensitive (nanomoles or less) and (2) they are nondestructive to the sample. Optical conversion technique offers an inexpensive system for the detection of heavy metals, particularly the high toxic arsenic in the environment [17]. In optical method, no reference electrode is needed, and there is no electrical interference.

6.3 Acoustic Wave Method

Acoustic wave is a kind of wave used in piezoelectric component, such as crystal, in circuits. The frequency characteristic in the crystal varies in response to the surrounding environment. In acoustic wave signal conversion method, an applied radio frequency produces mechanical stress in the crystal. As a result surface acoustic wave (SAW) is induced. The SAW is received by the electrodes and is translated to voltage (Fig. 22). The probe and target molecular hybridization on the surface of quartz platform causes signal conversion change into the voltage. These are typically signal conversion components that detect the presence of low mass such as single molecule or cell.

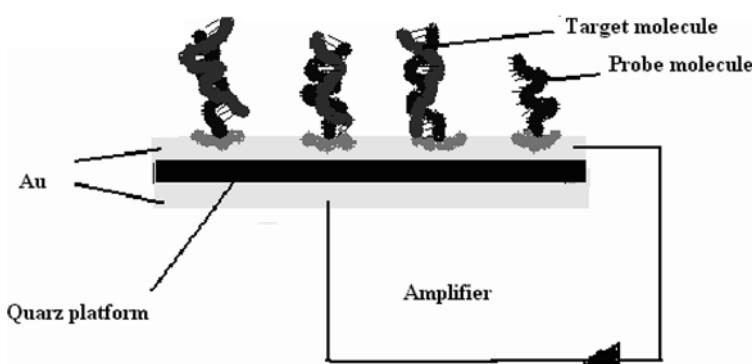


Fig. 22 Schematic of acoustic wave biosensor detects probe and target molecular binding on the surface of quartz platform caused difference in the voltage

6.4 Electromechanical Method

Nanocantilevers and quartz crystal microbalance sensors are extremely sensitive mass sensors, which are capable of measuring attogram (10^{-18} g) mass sensitivity. When target binding occurs between protein, DNA or RNA target molecules, and probe molecules on the cantilever beam, change of adsorption is directly related to the cantilever deflection. The molecular binding surface tension causes bending of cantilever beam. The deflection of the beam is measured optically (Fig. 23). The resonance frequency varies as a function of mass loading. Nanobelts, nanorings, and nanosprings with piezoelectric properties have been synthesized, for use as nanoscale transducers, actuators, and sensors.

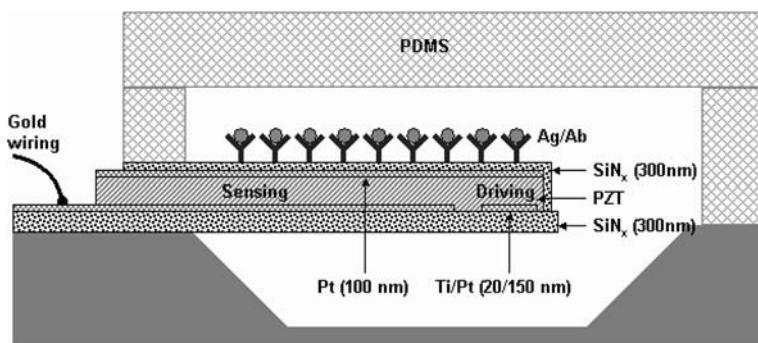
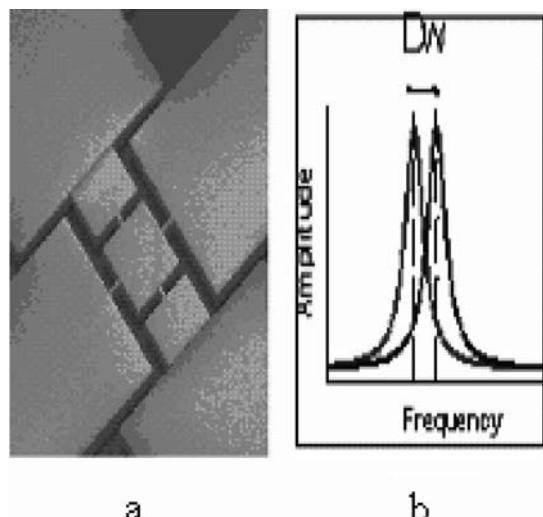


Fig. 23 Piezoelectric cantilever (Source: Courtesy of Drs. Il-Han Hwang and Jong-Hyun Lee, 2006)

The cantilever biosensor can also be operated in a dynamic mode, where its frequency shift is measured before and after the binding of biomolecules on the cantilever surface [18].

Single-crystal zinc oxide (ZnO) and SnO₂ nanobelts that spontaneously rolled them up for piezoelectricity into helical structures with widths in the range of 10–60 nm wide and 5–20 nm thick. These nanobelts/wires are quasi-one-dimensional semi-conducting and piezoelectric nanostructures used to fabricate field effect transistor (FET). These nanobelts induce a field effect in the presence of charged biological molecules such as proteins, peptides, DNA, and RNA. The ZnO and Si nanowires [19], and Si microresonators (Fig. 24) are in recent use for nanobiosensor devices fabrication. They are synthesized by simply evaporating oxides of zinc, tin, indium, cadmium, and gallium, whose electrical performance is extremely sensitive to the surface-adsorbed molecules. The FET-based device offers as a potential technique for *in situ*, real-time, wireless, and implantable biosensors [20]. These devices facilitate detection of subnanomolar quantities of proteins, DNA, and RNA from anthrax, salmonella, smallpox, AIDS virus, and other pathogens and toxins, including ricin.

Fig. 24 Si microresonator, capacitive, or piezoelectrical (Source: Courtesy of Dr. Farrokh Ayazi, Georgia Tech.)



6.5 Magnetic Method

Magnetic nanoparticles are also used in force-amplified biological sensor (FABS). FABS use magnetic particles to label antibodies, sequence-specific DNA or RNA. Atomic force microscope (AFM) is employed to pull on and measure the strength of single DNA–DNA or antibody–antigen bonds [21]. Thus, the detection of binding and hybridization events of probe and target molecules is done using magnetic nanoparticles (Fig. 25).

The magnetic particles used for FABS must be smooth and spherical, so that the entire surface of each particle is available for binding to the flat cantilever surface. Irregular particles effectively have a reduced active area. The particles should exert

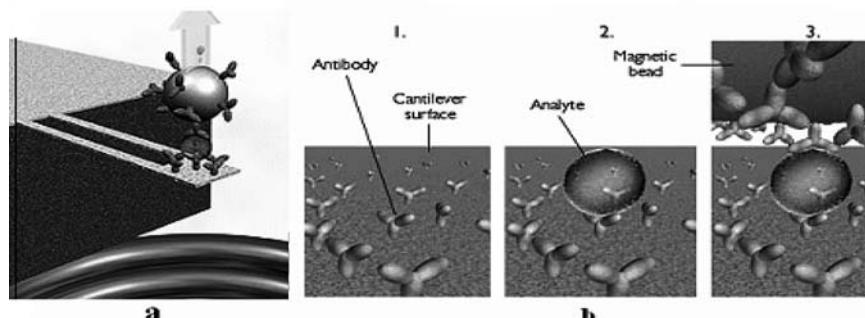


Fig. 25 Magnetic biosensor, in which the antibodies labeled magnetic particles are attracted in the magnetic field (Source: Courtesy of Drs. Baselt et al. 1998)

a uniform force in a magnetic field, and a high magnetic moment for maximum signal. They should not corrode when immersed in physiological salt solutions; low density and initially nonmagnetic to avoid coagulation of the solution.

6.6 Chemical and Enzymatic Method

Variety of glucose oxido-reductases and lactic dehydrogenases are used as a component for various enzyme sensors. Electron mediator-dependent glucose dehydrogenases (GDH) recognized as ideal constituents of mediator-type enzyme sensors. The dehydrogenases utilize the NAD⁺/NADH couple. The NADH generated (Fig. 26) is detected by the amperometric method.

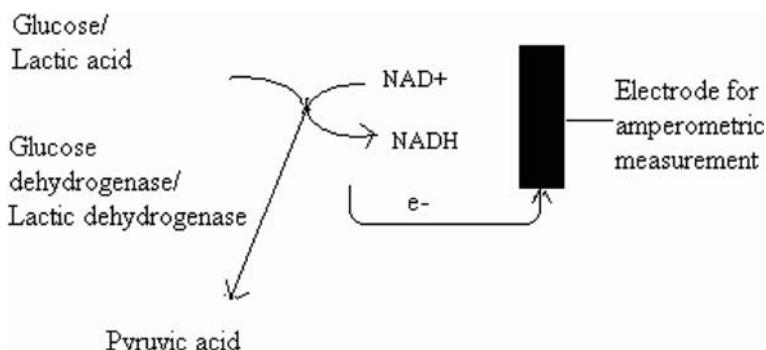


Fig. 26 Enzymes catalyzing oxido-reductions. The electrons produced in the reaction are a measure of glucose

In simple glucose biosensors, the enzymatic oxidation of glucose produces hydrogen peroxide, which in turn generates electrons by electrode reaction. The current density is used as a measure of glucose in the sample (Fig. 27).

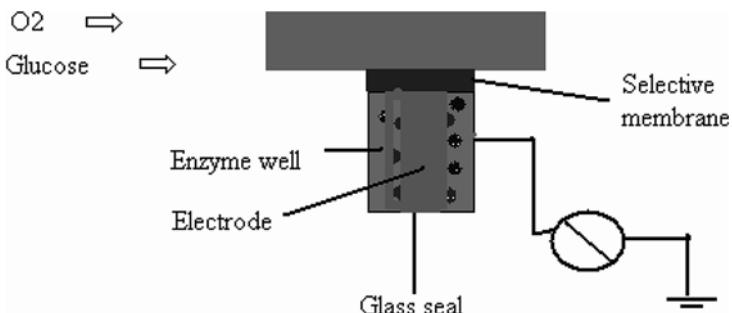


Fig. 27 The selective membrane in contact with glucose and the enzyme acts as a transducer

6.7 Thermal Method

The thermal biosensor is an enzyme thermistor modified for split-flow analysis. A miniaturized thermal biosensor involves in a flow-injection analysis system for the determination of glucose in whole blood. The blood glucose is determined by measuring the heat evolved when samples containing glucose passed through a small column with immobilized glucose oxidase and catalase. A sample of whole blood as small as 1 μl can be measured directly using this thermal biosensor [22]. The detailed account of nanomaterials, fabrication methodologies of nanodevices are available [23,24].

6.8 Challenges in Integration of Biomolecules with Interface Mechanisms and Signal Transduction

To make any biosensor device commercially viable, the following criteria should be met:

- Relevance of output signal to measurement environment
- Speed of response, accuracy, sensitivity, dynamic range, and repeatability
- Amenable to testing and calibration
- Reliability and self-checking capability
- Insensitivity to temperature (or temperature compensation)
- Insensitive to electrical and other environmental interference
- Biocompatible in vivo application
- Development of a new generation of biomimetic sensors for single-molecule detection.

The technical issues that need to be addressed in integrating biomolecules with interface mechanisms and signal transduction include

- Integration of different technologies from the interdisciplinary fields (electronics, chemistry, physics, and biology) for proper functioning of a biosensor. The interface between the transducer and the biological receptor may be made more exclusive thus reducing interference, and development of new receptors with improved or new affinities.
- The scaling issues involved in miniaturizing process without detriment to overall functionality of the biosensor.
- The control of size, shape, and magnetic properties of magnetic nanoparticles in fabricating nanoparticles-based biosensors.
- Proper retention of specificity and sensitivity of complex biological molecules that have their inherent instability. Necessary care should be taken to ascertain their viability on the attachment surfaces.

- Toxicity of nanoparticles for in vivo applications such as targeted drug delivery, MRI contrasts agents, and pathogen cleansing.
- Magnetic “traps” or read/write architectures that is biocompatible and capable of manipulating single magnetic particles with nanoscale precision.
- Preventing the biosensor devices from biofouling, as protein build up on the biological active interfaces.
- Finding efficient interfaces between biological materials and electronics.
- Evaluation of new manufacturing procedures for large-scale production.
- Optimized performance of the sensor is supported by associated electronics. Fluidics, and integrating with system-on-package (SOP) platform.

7 Nanostructures, Biomolecules, and Cells as Components of Biosensors

The novel properties possessed by nanostructured materials open several possibilities toward fabricating advanced nanobiosensors with superior sensitivity and specificity. For instance, single-walled carbon nanotubes (SWNTs) and multiwalled carbon nanotubes (MWNTs) exhibit unique properties such as chemical stability, good mechanical properties, outstanding charge-transport characteristics, and high surface area that makes them ideal for the design of nanobiosensors for detecting sub-femtogram quantities of target protein, DNA, and RNA. Among the approaches for exploiting nanotechnology in medicine, nanoparticles offer some unique advantages as sensing, image enhancement, and delivery agents [25].

The remarkable biorecognition capabilities of the biomolecules such as DNA, RNA, and proteins can be utilized for the fabrication of ultra miniature biological electronic sensing devices, by conjugating them with nanoscale solids (particles, dots, or wires). Varieties of nanoparticles are available commercially, including polymeric nanoparticles, metal nanoparticles, liposomes, micelles, quantum dots, dendrimers, and nanoassemblies. The gold nanoparticles of size 1.4 nm are used in the fabrication biosensors. In the fabrication of immunosensors, the nanoparticles are conjugated with the protein probe, which involves in antigen–antibody immune complex formation. Nanoparticles are also used for conjugating DNA or RNA probe involving sequence-specific DNA or RNA hybridization detection, in the construction of nucleic acid probes. Nanoparticles-based biosensors are sensitive, inexpensive, noninvasive detection systems for measuring biological molecules, and advanced over the existing technologies due to their nanosize, smart performance, and constituting the third-generation biosensors. Aptamers are oligonucleotides (DNA and RNA) that can bind with high affinity and specificity to wide range of target molecules (proteins, peptides, drugs, vitamins, and other organic or inorganic compounds). The biosensors fabricated with resonant oscillating quartz crystal can detect minute changes in resonance frequency when probe ss-DNA hybridizes with target DNA. DNA-based supramolecular chemistry allowed the replacement of natural bases in the DNA molecule with artificial nucleotides or nucleotide mimics

[26]. Further, new generation of such nucleotide mimics was reported in which the hydrogen bonding interactions were replaced by metal-mediating base pairs [27]. The metal incorporation into the interior of the DNA confers DNA metal-based functions such as thermal stability and magnetic properties (Fig. 28).

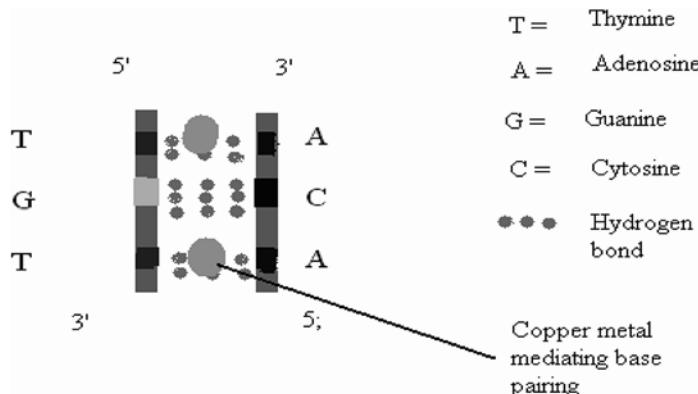


Fig. 28 Schematic of copper metal-mediated DNA base pairing

The cell-based biosensors are composed of two transducers, where the primary transducer is cellular and the secondary transducer is typically electrical. It is known that olfactory neurons respond to odorant molecules, and the retinal neurons in the eye are triggered by photons. These cells in the body act as primary transducers capturing the signals and in turn passed on to the secondary transducer.

7.1 The Gold Nanoparticles

Nanoparticles are used for fabrication of DNA-based biosensors [28,29]. The hybridization or binding events of sequence-specific probe single-stranded DNA (ss-DNA) functionalized with gold particles and target molecules demonstrated changes in electrical conductivity. The target and probe-binding events localize gold nanoparticles in an electrode gap. The silver deposition facilitated by this gap leads to measurable conductivity changes. These nanoparticles-based devices could detect target DNA or RNA at concentration as low as 500 fmol. As shown in Fig. 29, sequence-specific DNA or RNA probe is attached to the gold nanoparticles and subjected to hybridization with target oligonucleotides.

The gold nanoparticles coated with probe protein or antibody is used for the detection of specific antigen in immunosensor system.

The enzyme electrodes are constructed by introducing nanoparticles into the critical sites of the enzyme, such as in redox center of glucose-oxidizing enzymes. When the enzyme oxidizes glucose, electrons flow through the nanoparticle and into the electrode. The amount of current indicates the level of glucose present. These

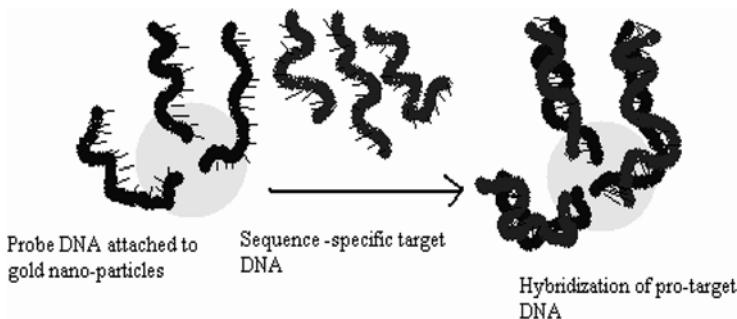


Fig. 29 Schematic diagram of hybridization of probe DNA on gold nanoparticles with sequence-specific target DNA

enzyme-conjugated electrodes are ultimately used in fabrication of implantable enzyme sensors inside the body for monitoring blood glucose.

Precipitation, physical vapor deposition (PVD), chemical vapor deposition (CVD), and organic encapsulations are the common methods of synthesizing nanoparticulate gold film on various supports.

7.2 Magnetic Nanoparticles

Magnetism-based biosensors use magnetic nanoparticles functionalized with biological molecules such as DNA, RNA, and protein for fabrication of variety of biosensors. Metals such as nickel (Ni), cobalt (Co), Fe, Fe_3O_4 and $\gamma\text{-Fe}_2\text{O}_3$, and their oxides are used to synthesize magnetic beads or particles of length scales ranging from 1 to 100 nm. As shown in Fig. 30, the target DNA is biotinylated and hybridized with sequence-specific ss-DNA probe [30]. The streptavidin functionalized magnetic beads interact with biotinylated hybridized DNA,

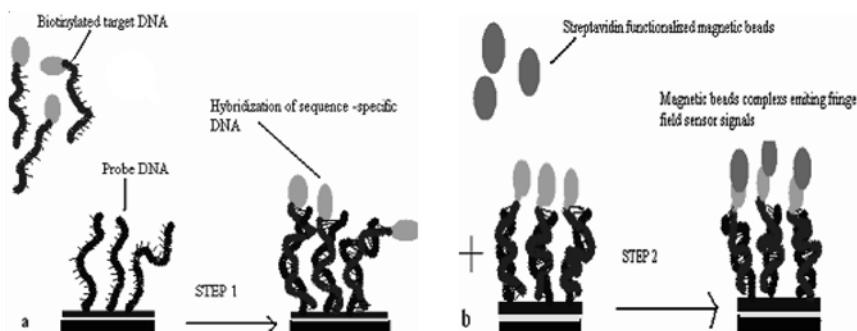


Fig. 30 Schematic diagram of hybridization of target DNA with probe DNA, followed by hybridization with streptavidin functionalized magnetic beads

and the hybridization event is detected by a change in resistance from a giant magnetoresistive (GMR).

Nanoparticles conjugated with biomolecules are used in the detection of antigens. Europium (III)-labeled nanoparticles coated with antibodies or streptavidin are used to detect prostate-specific antigen (PSA) in serum [31].

The important brain molecules, gamma-amino butyric acid (GABA), are involved in Huntington's disease and Parkinson's, seizures, myoclonic discharges, and alcohol addiction. In real-time measurements of the molecules, ultra-violet (UV)-visible spectroscopic determination methods are insensitive. The nanoparticles-based neurotransmitter sensors fabricated using nanoparticles will facilitate accurate measurements of neurotransmitting molecules.

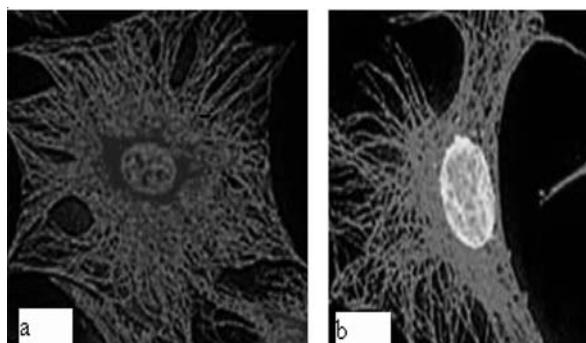
The magnetic nanoparticles with heterogeneous size from 200 to 400 nm are shown to be advantageous over magnetic microspheres. They exhibit high magnetization per unit weight, remain in suspension for longer periods of time without aggregating, and are faster – an important requirement for the more sensitive sensors and measurements.

7.3 Semiconductor Quantum Dots

The quantum dots are inorganic semiconductor nanocrystals. They are created by confining free electrons in a 3D semiconducting matrix. Those tiny islands or droplets of confined free electrons show many interesting electronic properties.

A key feature of these materials is that they can be modified with large number of small molecules and linker groups to optimize the functionality for particular applications. The quantum dots can be covalently linked with biorecognition molecules such as DNA, RNA, antibodies, peptides, and small-molecule inhibitors for use as fluorescent probes, which exhibit improved brightness, photo stability, and broader excitation spectra (Fig. 31a, b) Thus, quantum dots are superior over the existing imaging techniques, which use natural molecules such as organic dyes and proteins (Fig. 31).

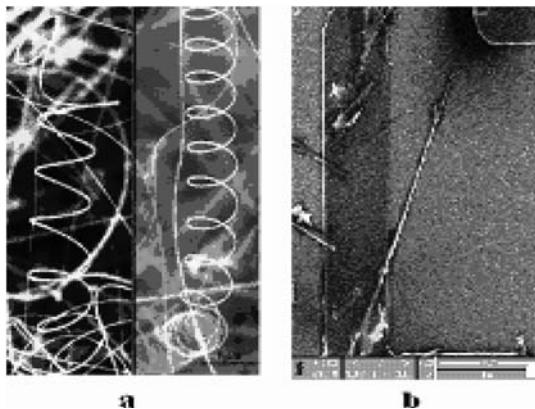
Fig. 31 (a) The microtubules of the cell were stained with 605-nm fluorescent quantum-dot conjugate, and the nuclei were counterstained with Hoechst dye, (b) nucleus and microtubules were labeled with red and green quantum-dot conjugates (Source: Quantum Dot Corp)



7.4 Nanowire

Semiconducting and piezoelectric nanostructures such as single-crystal zinc oxide (ZnO) nanobelts (Fig. 32) are used as sensitive signal conversion components, capable of measuring signal conversion from a single molecule or cell. ZnO nanobelts/wires (10–60 nm wide and 5–20 nm thick) induce a field effect in the presence of charged biological molecules such as proteins, peptides, DNA, and RNA [32]. ZnO and Si nanowires are used as efficient signal conversion components [33].

Fig. 32 (a) Single-crystal zinc oxide (ZnO) nanobelts; (b) ZnO nanowire mounted between the electrodes (Source: Courtesy of Dr. ZL Wang, Georgia Institute of Technology)



7.5 Carbon Nanotubes (CNTs)

The carbon nanotubes and silicon nanowires and actinyl peroxide self-assembled structures are the products of novel innovative technologies. Self-assembled carbon nanotubes produced in the arc discharge process are widely used in catalytic processes [34]. Based on the wall structure of the nanotubes, two types of tubes are categorized: single-walled and multiwalled. A single-walled carbon nanotube (SWNTs) consists of a single graphene cylinder, whereas a multi-walled carbon nanotube (MWNTs) comprises several concentric graphene cylinders. SWNTs can possess efficient electrical conductivity. The carbon atoms are strongly (covalently) bound to each other (carbon–carbon distance 0.14 nm). As shown in Fig. 33, rolling up the sheet along one of the symmetry axis gives either a zig-zag ($m = 0$) tube or an armchair ($n = m$) tube.

The carbon nanotubes (CNTs) possess unique properties such as chemical stability, good mechanical properties, high surface area, and outstanding charge-transport characteristics that make them ideal for fabrication of nanobiosensors.

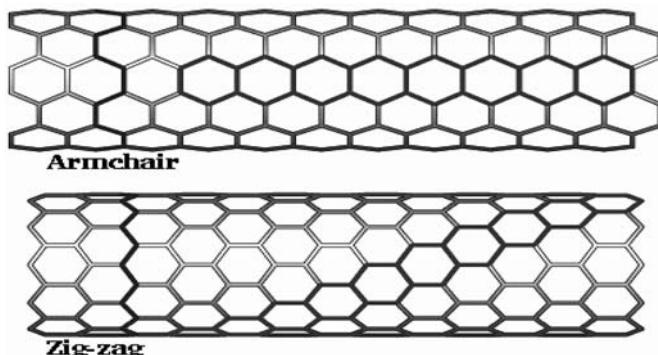


Fig. 33 Structure of carbon nanotubes (Courtesy of Technion, Department of Physics, Israel University (<http://phycomp.technion.ac.il/~talimu/structure.html>)

7.6 Actinyl Peroxide Compounds

Actinyl peroxide compounds, are known to self-assemble into nanosized, hollow cages contain actinides of other elements such as potassium. Nanowire containing a step every five silicon atom rows and a row of gold atoms in the middle of the terrace is reported [35].

The fabrication of nanostructures from deoxyribonucleic acid (DNA) – DNA dendrimers, DNA nanotubes, and DNA diamond adducts, and assemble these nanostructures into large-scale functional constructs used in biosensor devices.

7.7 Conductive Polymer Membrane

Polymer conductive property has been utilized for fabricating conductive polymer-based biosensor array (electronic nose). When conductive polymers come into contact with analyte molecules, their conductivity changes, creating a current within the sensor that is proportional to the concentration of the analyte. These currents are picked up by signal processing circuits in the instrument, and compiled into a response profile that characterizes the atmosphere around the sensor. The Langmuir–Blodgett films method is used for nanofilm fabrication in which monolayers of desired material efficiently built one at a time. The electronic nose consists of two components: (1) An array of chemical sensors (usually gas sensors) and (2) A pattern-recognition algorithm. This sensor can be useful for detecting odor or smell emitting from biological organisms and chemical sources. The following are some of the examples of electron nose applications:

- Smelling foods for their freshness, including freshness of fish and other sea foods
- Bacterial growth on foods such as meat and fresh vegetables
- Process control of cheese, sausage, beer, and bread manufacture
- Identification of quality of coffee. Monitoring of roasting process

- Rancidity measurements of oils
- Identification of spilled chemicals in the environment, particularly in gas transportation, and leakage of utility gases
- Diagnosis of pulmonary infections (e.g., TB or pneumonia)
- Diagnosis of ulcers in the lungs by breath tests

7.8 DNA, Proteins, and Cells

DNA, proteins, and cells are the basis for fabrication of biosensors and biodevices for various applications.

8 Biochemical Probe Design and Interfacing with Biosensors

The production, proper attachment, and optimization of probe biomolecules such as DNA, RNA, and antibodies on electrode surfaces are challenging tasks in fabricating a biosensor. Several critical steps are involved in immobilization and assessment of proper immobilization and viability.

8.1 Preparation of Protein Probe

The protein of interest to be used as protein probe is produced for biofunctionalization of biosensor devices. Proteins are produced mainly by three methods:

8.1.1 Proteins Produced from Cell and Tissue Cultures

The cells with the protein of interest are cultured using appropriate culture conditions. The cells are lysed in a cell lysis buffer with 0.1 mg/ml lysozyme, and proteins are extracted. A protein of interest is isolated and purified using standard protocols.

8.1.2 Proteins Produced by Gene Activation

As shown in Fig. 34 when a regulatory gene for the protein of interest is in “off” condition, the corresponding gene does not produce any protein in the bacteria. Addition of a new regulatory gene makes the gene in “on” position, and facilitates the protein production.

8.1.3 Recombinant Protein Production

Recombinant gene technology is used for the production of a specific single protein, unlike in other two methods where total proteins are in the pool. To overexpress the gene, the plasmid is transformed into the *E. coli* strain or any other expression vector. As shown in Fig. 35, the plasmid is cut open at an appropriate position using a

Production of protein by gene activation technology

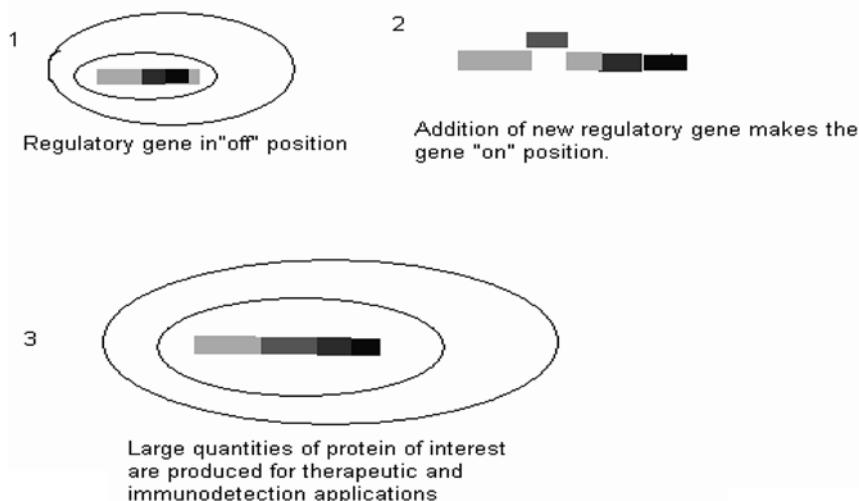


Fig. 34 Schematic of production of protein by gene activation

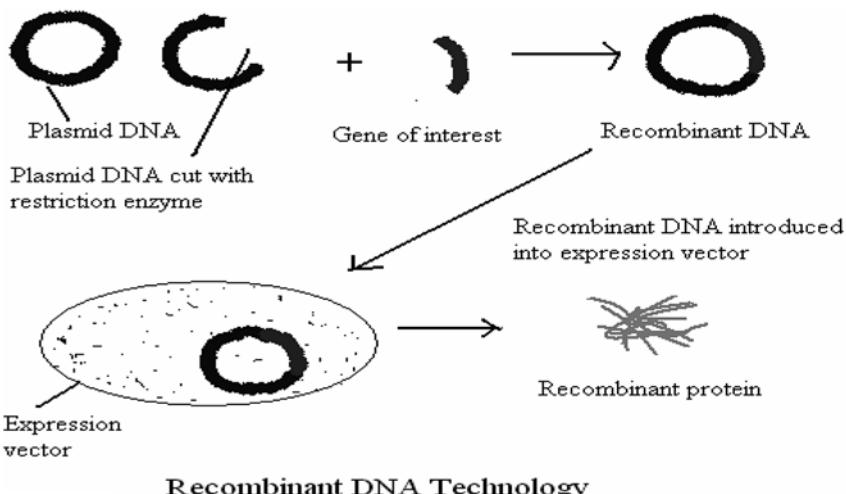


Fig. 35 Schematic of production of protein using recombinant gene technology

restriction enzyme. The gene responsible for the protein of interest is inserted and attached to the cut ends of the plasmid using ligase enzyme. The resulting recombinant DNA is used for expression of the protein in an appropriate expression vector.

The proteins obtained by various methods are purified by standard protein purification methods.

8.1.4 Production of Polyclonal and Monoclonal Antibodies

The polyclonal antibodies are produced *in vivo* in response to immunization with different epitopes on the protein injected. Antiserum can be raised in mouse, rat, rabbit, sheep, goat, and multiple injections of the antigen along with the adjuvant (a nonspecific enhancer of the immune response). After repeated immunizations, the antibodies produced will be predominantly IgG high affinity.

The monoclonal antibodies provide single epitope specificity. They are produced using hybridoma technology. The antibody-producing cells from the mouse's spleen and cancerous immune cells fuse to form hybridoma cells. Individual hybridoma cells are cloned and tested to find out the clones that produce the desired antibody.

8.2 Synthesis of DNA/RNA Oligonucleotide Probes

DNA can be synthesized using polymerase chain reaction (PCR). The appropriate primer, Taq polymerase, d-NTPs, and other ingredients in the reaction condition can generate desired oligos synthesis. The complementary oligonucleotides for a particular target gene are synthesized in commercially available DNA/RNA synthesizer. Figure 36 shows the rapid multiplication of the DNA molecules in PCR. Starting from single molecule, an astronomical figure of 68 billion molecules are obtained in 35 cycles of DNA synthesis.

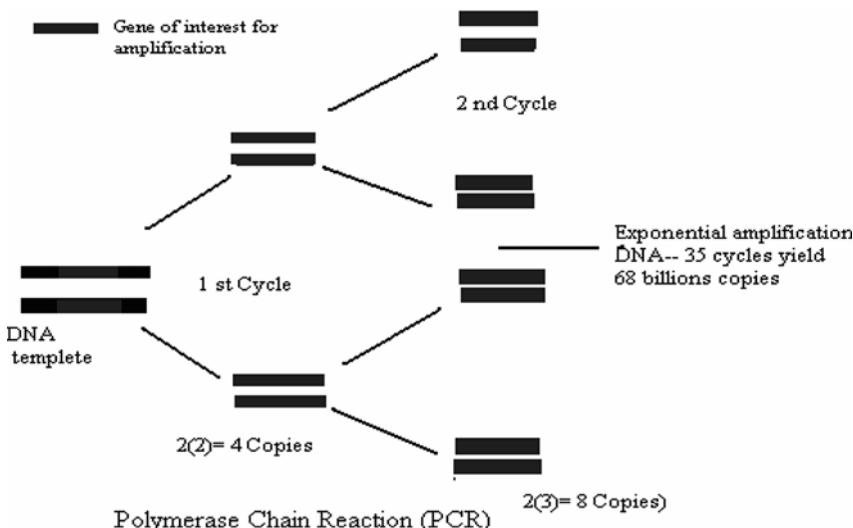


Fig. 36 Schematic diagrams of polymerase chain reaction (PCR). The reaction set 30 – 40 amplification cycles of 3 min each. Denaturation for 1 min heating at 94°C; annealing for 45 s at 54°C, and extension for 2 min at 72°C; first produced 2 DNA molecules

All nucleotides are purified by gel filtration and reverse-phase chromatography using an Invitrogen RP-C18 column.

8.3 Electrically Conductive Contact Surface Materials

Nobel metals such as gold, platinum, and palladium are known to be biocompatible. Titanium is widely used as implants and surgical staples. These metals are also widely used as surface materials. Use of conductive polymers for surface coating with the ligands offer great advantage over metals and inorganic semiconductors, including higher electrical conductivity than other conductive materials, plasticity, elasticity, low mass density, low coefficient of expansion, resistance to chemicals and corrosion, tunable optical properties, and electrochromism. The functionalization of their nanoparticles with FITC-streptavidin is achieved by carbodiimide chemistry, similar to that of enzyme binding to nanoparticles.

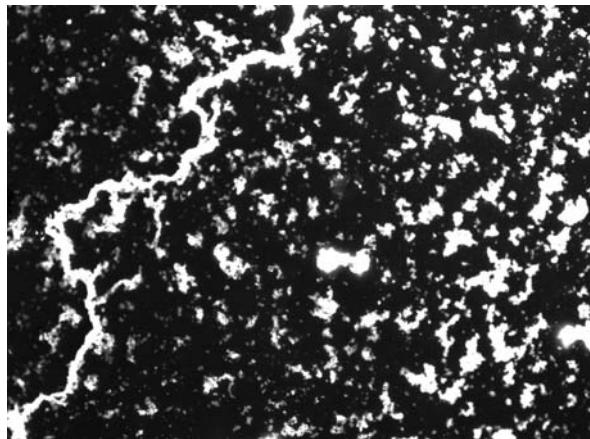
8.4 The Process of Attachment or Immobilization of Biomolecules to the Surfaces

The attachment of probe biomolecules to the surface of electrodes is a critical step in the construction of a nucleic acid and protein-based biosensors. DNA, RNA, and proteins may be passively immobilized through hydrophobic and ionic interactions. However, covalent immobilization is often necessary for proper adsorption, orientation, and conformation to the noncovalent surfaces. The immobilization process should occur rapidly and selectively in the presence of common functional groups, including amines (NH_2), carboxyl (COOH), hydroxyl (OH), thiol (SH), and aldehyde (CHO). The surface density of the molecules should be optimized. Low-density surface coverage yields a correspondingly low frequency of binding sites. High density may result in inaccessible binding sites to the target molecules. The correct orientation of ligand molecules on the surface is required to facilitate the ligand available for the target. For efficient immobilization, the maximum biochemical activity and minimum nonspecific interactions are required. A brief description of attachment methods is as follows.

8.4.1 Surface Modification and Biofunctionalization of Carbon Nanotubes

The interface between biological molecules and novel nanomaterials is important to developing new types of miniature devices for biological applications. The carbon nano-tubes are surface functionalized and biofunctionalized as shown in Fig. 37. Carbon nanotubes are surface modified for incorporating carboxylic group (-COOH) by refluxing in concentrated nitric acid for 4 h. The carboxylic groups formed on the surface of the nanotube giving a chemical handle for attaching carbon nanotubes to protein or DNA molecules. The modified carbon nanotubes are attached on glassy graphite electrode using *N,N*, dimethylformamide (DMF). The CNTs are biofunctionalized with antibodies or ss-DNA (probe) and hybridized with antigen or ss-DNA (analyte or target), respectively. Hybridization is detected with

Fig. 37 Carbon nanotube (O.D. \times wall thickness \times length 20–50 nm \times 1–2 nm \times 0.5–2 μm). Carbon nanotubes are biofunctionalized with protein, and visualized using a confocal microscope



fluorescein isothiocyanate (FITC)-labeled IgG or cyanine (Cy3 or Cy5)-labeled target DNA using confocal microscope. In case of unlabeled probe molecules, the change in electrical measurement indicates probe–target molecular hybridization.

The streptavidin/biotin system is used to investigate the adsorption behavior of proteins on the sides of single-walled carbon nanotubes (SWNTs) [36]. Functionalization of SWNTs by coadsorption of a surfactant and poly (ethylene glycol) is found to be effective in resisting nonspecific adsorption of streptavidin. Specific binding of streptavidin onto SWNTs is achieved by cofunctionalization of nanotubes with biotin and protein-resistant polymers (Fig. 38).

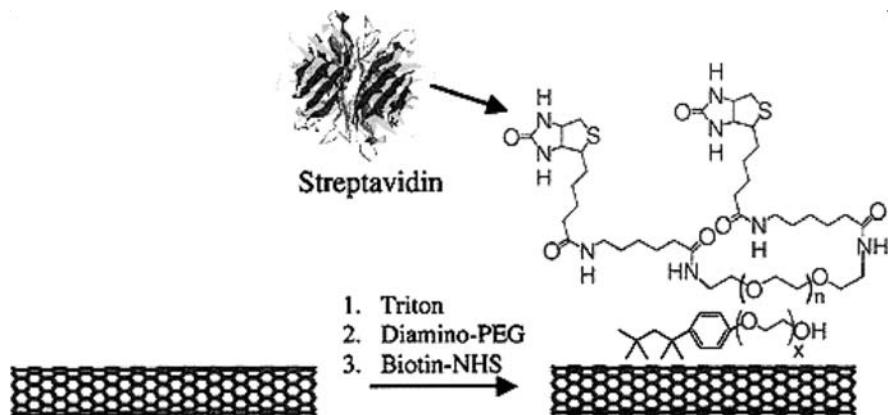


Fig. 38 Surface modification, biofunctionalization and detection of biotin–streptavidin interactions (Source: Courtesy of Moonsub Shim et al. 2002)

Similarly, carbon nanotubes functionalized with a carboxylic acid group (MWNTs-COOH) for covalent DNA immobilization and enhanced hybridization

[37]. The MWNTs–COOH-modified glassy carbon electrode (GCE) was fabricated and oligonucleotides with the 5'-amino group were covalently bonded to the carboxyl group of carbon nanotubes. The hybridization reaction on the electrode was monitored by differential pulse voltammetry (DPV) analysis using an electroactive intercalator daunomycin as an indicator (Fig. 39).

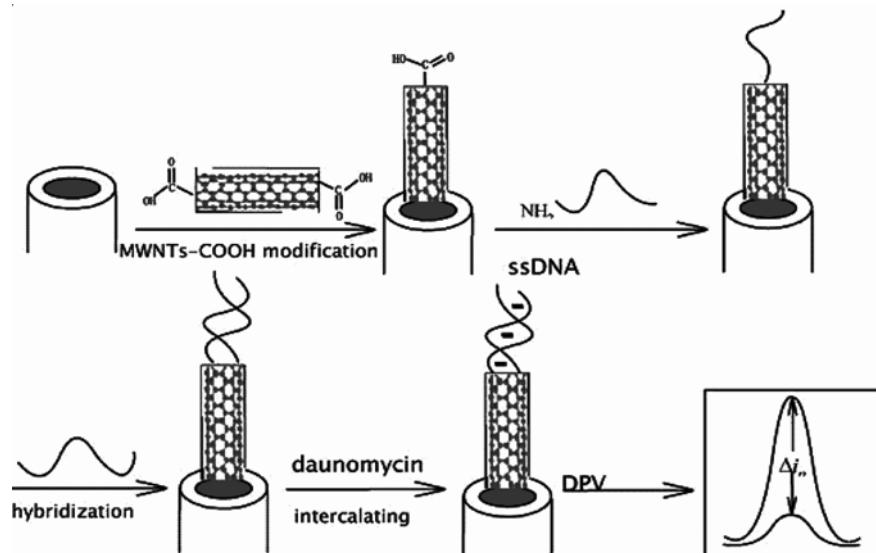


Fig. 39 Surface modification, biofunctionalization and detection of DNA hybridization. (Source: Courtesy of Hong Cai et al. 2003)

8.4.2 Thiolated DNA for Self-Assembly on to Gold Transducers

Functionalized oligonucleotides are widely used in diagnostics, therapeutics, and genetic analysis. Thiolation of DNA involves condensation of DNA with thiolated polyethylenimine (PEI-SH) and the resulting nanoparticles surface-coated using thiol-reactive poly [*N*-(2-hydroxypropyl) methacrylamide] (PHPMA) with 2-pyridyldisulfanyl or maleimide groups, forming reducible disulfide-linked or stable thioether-linked coatings, respectively.

8.4.3 Covalent Linkage to the Gold Surface

Self-assembling process of alkylthiols on gold is initiated by the strong chemical interactions between the sulfur and the gold surface. It involves chemisorption that forced a thiorate molecule to adsorb commensurate with a gold lattice. Then, the tail-to-tail interaction of the molecules created by lateral interchain nonbonded interactions, such as by van der Waals, steric, repulsive, and electrostatic forces, is strong enough to align the molecules parallel on the gold surface and create a crystalline

film. In most cases, organic disulfides, thiols, and sulfides have been utilized for the preparation of stable SAMs on the gold surfaces. Methyl-terminated alkanethiol, biotin-terminated alkanethiol, COOH-terminated alkanethiol, CONHS-terminated alkanethiol, NH₂-terminated alkanethiol are main functional groups that establish co-valent linkage. The photonic activation of disulfide bridges was shown to orient protein immobilization on biosensor surfaces [38].

8.4.4 Covalent (Carbodiimide) Coupling to Carbon Electrodes

The carbodiimides are hetero bifunctional coupling reagents that are utilized for coupling reactions through carboxyl and amino groups. The procedures generally used for carbodiimide coupling of peptides and proteins. Most peptides and proteins contain exposed carboxyl groups, so carbodiimides such as EDC are useful reagents for linking carboxyl groups to primary amine-containing matrices. The cross-linker, 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDAC), is widely used in immobilization of peptide, proteins, and sugars to various supports such as polystyrene plates, beads, gels, and biosensors. Cross-link is also used for structural studies, with intramolecular, inter-subunit, or between proteins and DNA. The following is the typical RGD peptide (arginine–glycine–aspartic acid) cross-linking to anilin in the presence of EDAC (Fig. 40).

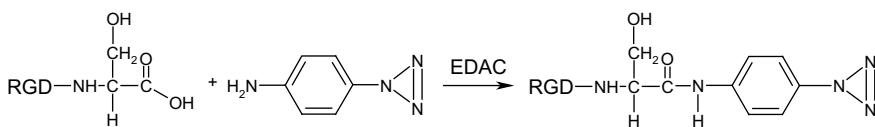


Fig. 40 Peptide cross-linkage to aniline in the presence of linker molecules EDAC

The contact surface of the electrode is washed thoroughly and coated with poly-terthiophene-bearing carboxylic acid groups. This coated surface is used to attach DNA/RNA nucleotides. The conducting polymer (poly TTCA)-coated electrode is immersed in a 30 mM acetic acid/acetate buffer (pH 5.2) containing 4.0 mM EDAC for 1 h. The EDAC-attached conducting polymer-coated electrode is rinsed with a buffer solution and subsequently incubated in 11 μM probe DNA/RNA/acetate buffer solution for 6 h at 25°C. In this process, NH₂-linked C6-modified DNA/RNA is immobilized on the poly TTCA film by the formation of covalent bonds with carboxyl groups on the polymer.

Different cross-linkers such as cysteamin, glutaraldehyde, and bovine serum are used for attaching probe biomolecules on to gold layer on the silicon nitride electrodes. Figure 41 shows attachment of the herbicide protein, 2, 4-dichlorophenoxy acetic acid using glutaraldehyde cross-linker. The gold layer is about 40 nm in thickness, whereas the probe biomolecules along with linker molecules form 4 nm thickness.

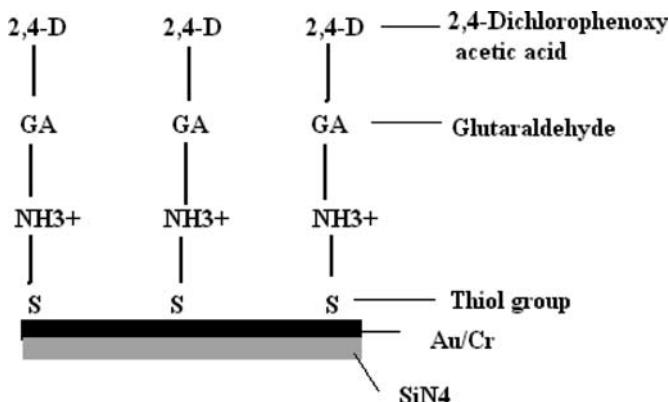


Fig. 41 Schematic of attachment of the probe molecules on gold layer coated onto silicon nitride electrode

8.4.5 Biotinylated DNA for Complex Formation with Avidin or Streptavidin

Single stranded DNA (ss-DNA) is biotinylated on nylon membrane or electrode surfaces using UV light. The membrane or surface is subjected to a single-wash step. The synthesized probe is completely free of unincorporated precursors. If biotinylation is performed under gentle conditions, the biological activity of the protein can be preserved.

Biotinylation of the target oligonucleotides can be achieved either enzymatically or chemically. Avidin is an egg-white-derived glycoprotein with an extraordinarily high affinity (affinity constant $> 10^{15} \text{ M}^{-1}$) for biotin. Streptavidin is similar in properties to avidin but has a lower affinity for biotin. Many biotin molecules can be coupled to a protein, enabling the biotinylated protein to bind more than one molecule of avidin. By covalently linking avidin with different ligands such as fluorochromes, enzymes or electron microscopy (EM) markers can be utilized to study a wide variety of biological structures and processes. The extraordinarily high affinity between avidin or streptavidin and biotin results in rapidly formed and stable complex between the streptavidin conjugate and the biotin-labeled protein or DNA.

8.4.6 Polymers for Attachments of Probe Molecules

Hydrogels, filter membranes, dendrimeric polymers, poly-L-lysine, and polyethylene glycol (PEG) have been used as solid support for microarray assays (Fig. 42). In one-dimensional coating, poly-L-lysine is used for the development of antibody/antigen microarray assays [39]. Polyethylene glycol is used in two-dimensional coating. The PEG-treated surface provides large spacer molecule, which results in higher analyte capture capacity [40]. Dendrimeric polymers, filter membranes, and hydrogels facilitate the formation of monolayer of probe molecules [41]. Three-dimensional matrix structure reduces denaturation of protein because of

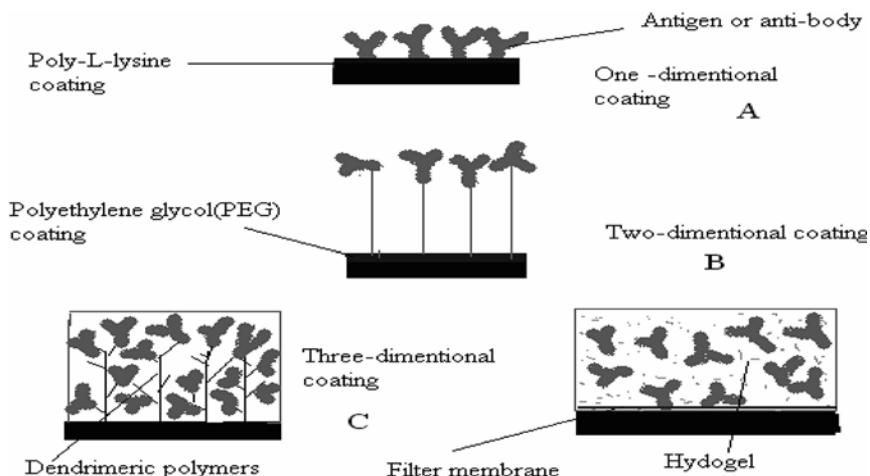


Fig. 42 Schematic of attachment of probe molecules. (a) One-dimensional coating; (b) two-dimensional coating; and (c) three-dimensional coating

the aqueous environment in the coating. The disadvantage with this coating is long incubation time required to probe target molecular interactions.

8.4.7 Simple Adsorption to Carbon Surfaces

Self-assembled monolayer (SAM) coatings can be prepared at room temperature in the laboratory by simply dipping the desired substrate in the required millimolar solution for a specific period of time followed by rinsing with the same solution, water, ethanol, and drying using a jet of dry nitrogen. The high affinity of gold for sulfur-containing molecules generates self-assembled monolayers. The quality of monolayers depends on several factors such as nature and roughness of the substrate, solvent used, temperature, concentration of the adsorbate [42]. The polyamidoamine (PAMAM) dendrimer-modified magnetite nanoparticles are used to improve the efficiency of protein immobilization. First, aminopropyltrimethoxysilane (APTS) is immobilized on nanoparticles. Then, methylacrylate and ethylenediamine are added to form a dendritic structure on nanoparticles surface. It has been demonstrated that PAMAM dendrimer-coated magnetite nanoparticles are 3.9–7.7 times more efficient than aminosilane-modified magnetite nanoparticles [43].

8.5 Biosensor Device Fabrication Methods

The nanodevices are fabricated using one or more of the following common methods:

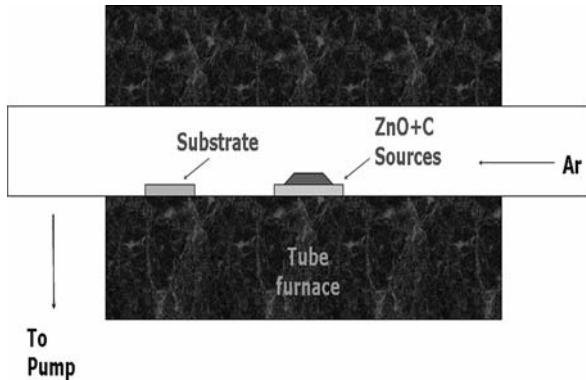
- Soft lithography, which is a new high-resolution patterning technique
- Thermal diffusion bonding of structural ceramics to superalloys

- UV laser micromachining for 2D and 3D medical devices.
- Deep reactive ion etching in the substrate.
- Electronic pick and place for heterogeneous integration
- Microfabricated fluidic nanotiterplate
- Micromechanical milling
- Microplastic injection molding
- Polymer embossing, and polymer injection molding
- Micromanufacturing of cellular systems

8.5.1 Fabrication Nanowire-Based Biosensors

The ZnO nanowire is synthesized through a high-temperature vapor–solid (VS) process that has been developed in Georgia Tech [44] and used for fabrication of ZnO nanowire-based biosensor system for detection of biological molecules and cancer cells [45]. The system setup is shown in Fig. 43.

Fig. 43 The setup of the vapor–solid (VS) synthesis process (Courtesy of PRC, Georgia Tech)

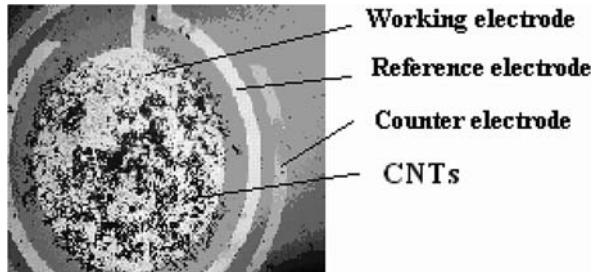


The setup of the vapor-solid (VS) synthesis process consists of a horizontal high-temperature tube furnace with the length of ~50 cm, an alumina tube of ~75 cm, a rotary pump, and a gas-controlling system. About 2 g of the mixture of commercial ZnO powder (Alfa Aesar) and carbon powder is loaded into a polycrystalline Al₂O₃ boat, which is placed in the center of the alumina tube. The furnace is heated up to ~1400°C in a ramp rate of 20°C/min from room temperature. Gas Ar at a flow rate of 50 standard cubic centimeters per minute (sccm) is introduced into the tube as carrying gas. After holding at ~1400°C for about 2 h, the furnace is slowly cooled down to room temperature. Polycrystalline Al₂O₃ substrates which are positioned ~11 cm from the ZnO source at the downstream side are used to collect the as-synthesized ZnO nanowires. The whole system is maintained at a pressure of ~200 mbar throughout. The as-synthesized sample is then inspected under a LEO 1530 field-emission scanning electron microscope (SEM) operated at 5 kV.

8.5.2 Fabrication of Microelectrochemical Biosensors

As shown in Fig. 44, microelectrochemical sensor device is fabricated patterning the Pt electrodes using lift-off process. The patterning of the silver electrode is done by a semiadditive process (Evaporate Ti/Ag, photoresist patterning, electroplating Ag, photoresist, and seed layer removal). Chloridized in FeCl_3 to form Ag/AgCl [46].

Fig. 44 Three-electrode structure of electrochemical biosensor (Courtesy of PRC, Georgia Tech)



Sol–Gel System for Encapsulation of the Enzyme

Sol–gel chemistry is gaining importance in acting as an interface between biological materials and nonbiological components in biosensors. Zirconia/Nafion sol–gel encapsulation of enzyme is used in fabrication of enzyme-based sensors. The zirconia sol–gel solution is prepared by mixing 8.0 ml of deionized water with 1.0 ml of 0.12 M ZrO₂ dissolved in isopropanol and 0.25 ml of 11 M HCl solution. The zirconia/nafion composite is prepared by mixing zirconia solution with nafion. Glucose oxidase enzyme at 100 mg/ml is prepared in 0.05 M phosphate buffer at pH 7.0. The enzyme encapsulating matrix is prepared by mixing 40 μl of zirconia/nafion stock solution and 40 μl of the enzyme solution, casted on the working electrode (Fig. 45). The advantage of the sol–gel encapsulation are as follows:

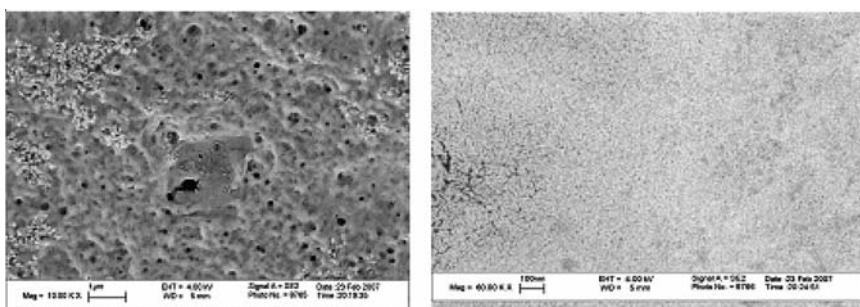


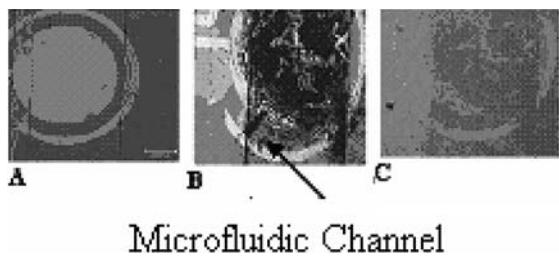
Fig. 45 FESEM micrographs of zirconia-nafion gels before (*top*) and after (*bottom*) enzyme entrapment (Courtesy of PRC, Georgia Tech)

- It resists chemical attack and withstands high temperature.
- Nafion is highly conductive to cations, making it ideal for many membrane applications.
- Pure nafion films revealed large calcification and instability of the cast films.
- Incorporate zirconia into anionic nafion assembly in order to prevent calcification.
- Encapsulation of the enzyme with the sol–gel protects from heat and

8.6 Micro-fluidic Channels

Microfluidics system is the manipulation of liquids in channels having cross-sectional dimensions on the order of 10–100 μm . Microfluidic channels are essential for guiding and helping biofluids carrying the target or the sample to be tested onto sensing element of the biosensor. As all biological molecular reaction takes place in liquid environment, microfluidic channels facilitate controlled flow of micro-quantities of fluids containing desired test molecules (Fig. 46, 47). The common fluids used as target in microfluidic devices include whole blood samples, bacterial cell suspensions, protein or antibody solutions, and various buffers. Fluidics carries the biological molecules without involving mechanical moving parts that will wear out.

Fig. 46 Planar microelectrodes: (a) without CNTs and sol–gel enzyme; (b) CNTs and sol–gel enzyme, open microfluidic channel; (c) CNTs and sol–gel enzyme, sealed microfluidic channel



Materials for implantable devices and microfluidic channels need to be biocompatible, which is not producing a toxic, injurious, or host response in living tissue. Biocompatible materials such as polydimethylsiloxane (PDMS), LCP materials are commonly used for fabrication of microfluidics. Further, laminated plastic microfluidic components made of polyimide, poly methylmethacrylate (PMMA), and polycarbonate materials with thickness between 25 and 125 μm have also been developed and used. These devices can be potentially manufactured in high volume with low unit cost because of advancements in the materials and chemical processes such as high-speed full-field excimer laser ablation process (Anvik Incorporated) to generate the fluidic channels. Low-cost PDMS adhesive-based bonding is used to seal the channels. Fabrication of both the sub-micron features of the photonic crystal sensor structure and the >10 μm features of a flow channel network in one step at room temperature on a plastic substrate was demonstrated [47,48]. Polymer-based

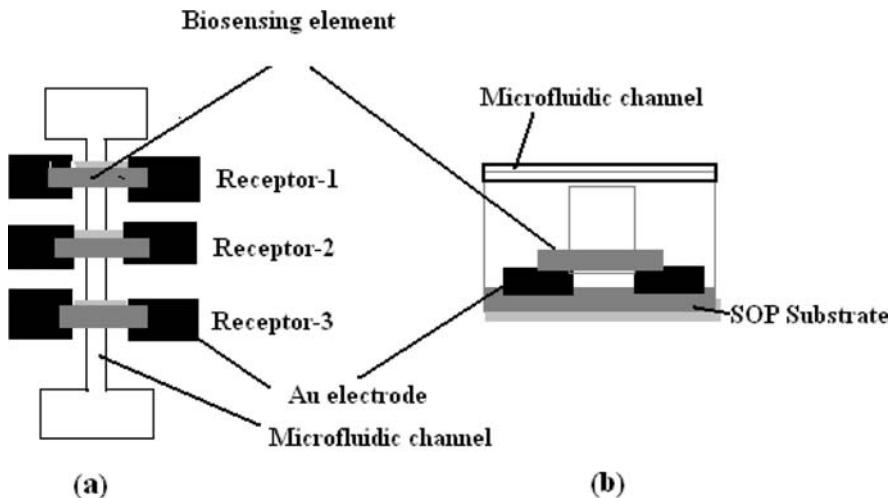


Fig. 47 (a) Schematic of microfluidics integrated, multiplexed biosensor system; (b) cross section of the device

microfluidic devices can be created by combining lithographic and conventional manufacturing techniques (e.g., injection molding or embossing). Polymeric channels can also be made transparent in the UV, visible, and IR ranges. This property enables the use of generic off-chip detection systems, thereby reducing the cost and complexity of a microanalysis chip. As an example, laminated plastic microfluidic components of polyimide, poly (methylmethacrylate) (PMMA), and polycarbonate materials with thickness between 25 and 125 μm have been developed [49]. These devices can be potentially manufactured in high volume with low unit cost because of advances in the materials and processing tools made by the electronic industry.

8.7 Biosensor Packaging

The highly miniaturized electronic system technology mostly relies on integrated circuits (IC) integration (Moore's law) for performance improvement and cost reduction. System-on-package (SOP) technology paradigm (the second law of electronics) pioneered by Georgia Tech Packaging Research Center, since the early 1990s provides system-level miniaturization in a package size that makes today's hand-held devices into megafunctional systems, with applications ranging from computing, wireless communications, health care, and personal security. The SOP is a system miniaturization technology that ultimately integrates nanoscale thin film components for batteries, thermal structures, active and passive components in low-cost organic packaging substrates, leading to micro-scale to nanoscale modules and systems. True miniaturization of products should take place not only at IC but also at system level, the latter made possible by thin film batteries, thermal structures,

and embedded actives and passives in package-size boards. This is the fundamental basis for the SOP concept. The traditional packaging by which components are integrated into systems today, presents several sets of barriers in cost, size, performance, and reliability. The SOP concept overcomes these barriers by the best of both the IC and the package integration at system level, the IC for transistor density and the system package for component density of RF, optical, digital, and biofunctions. In addition to miniaturization, such a concept leads to lower cost, higher performance, and better reliability of all electronic systems including biosensor systems described in this chapter. The attributes of traditional electronic system are bulky, poor reliability, moderate performance, medium cost compared to SOP-based systems, which show miniaturized, $3\times$ improvement in reliability, high performance, and low cost. SOP-based biosensor systems can be greatly enhanced in terms of the functionalities in both clinical and nonclinical domains. As shown in Fig. 48, the energy source for many biomedical microsystems is through an external RF link. The system communicates with the outside world over an inductively coupled bidirectional wireless link [50].

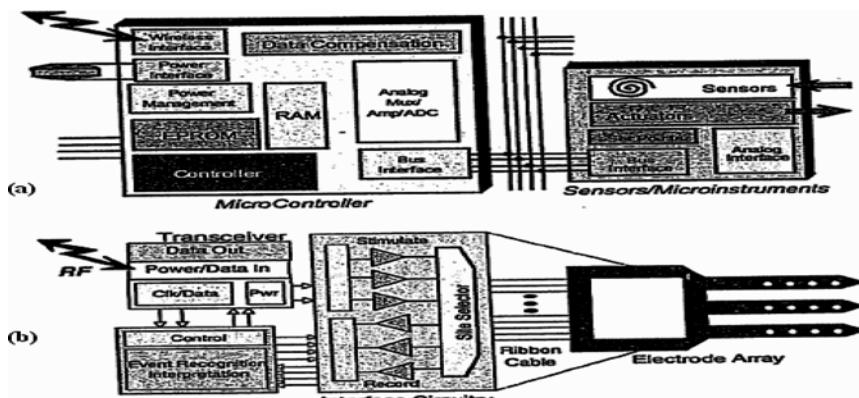


Fig. 48 (a) Schematic of a wireless integrated microsystem; (b) implantable version of the same microsystem (Courtesy of Dr. Kensall D. Wise, University of Michigan)

9 Characterization of Functionalized Biosensor Structures

The functionalized electrode surface is characterized by transmission electron microscopy (TEM), Fourier transform infrared spectroscopy (FT-IR), fluorescent microscopy, X-ray photoelectron spectroscopy (XPS), scanning electron microscopy (SEM), and atomic-force microscopy (AFM). Fourier transform infrared spectroscopy (FT-IR) is used to identify specific types of chemical bonds or functional groups based on their unique absorption signatures.

Transmission electron microscopy (TEM) is used to study the morphology and size of the particles.

Energy-dispersive X-ray spectrometer is used to confirm the composition of particles immobilized on gold surface. The ellipsometry is a real-time optical measurement technique, which is used to study several material characteristics such as layer thickness, optical constants (refractive index and extinction coefficient), and optical anisotropy. It is used for quantification and visualization of the lateral thickness distribution of thin (0–30 nm) transparent layers on solid substrates. The thinness of the biofunctionalized molecules on a nanodevice surface is measured at the same time with a high lateral resolution.

Atomic-force microscope (AFM) is used to take three-dimensional images of resolution on the order of 1–2 nm. It consists of cantilever tip assembly, piezo tube scanner, and laser deflection system (Fig. 49). The attachment of the biomolecules such as proteins and DNA, and the molecular hybridization events on the biosensor devices can be characterized using AFM [51].

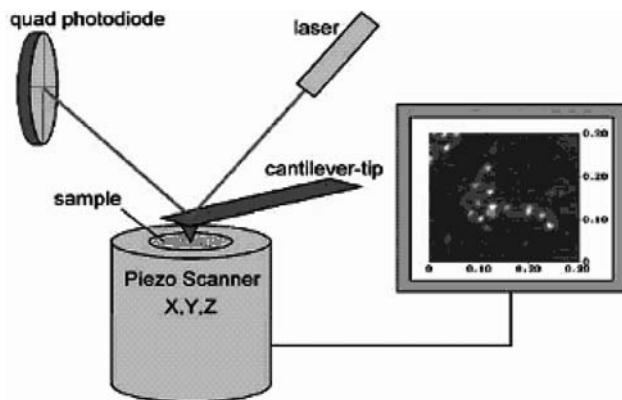


Fig. 49 Schematic of the atomic force microscope (Source: Courtesy of Hafner et al., 2001)

10 Future Trends and Summary

The tremendous advancements in the sensor technologies are due to the great technological demand for rapid, sensitive, and cost-effective biosensor systems in vital areas of human activity such as health care, genome analysis, food and drink, the process industries, environmental monitoring, defense, and security. In fact, nanobioconjugates that consists of various functional nanoparticles or nanostructures linked to biological molecules revolutionized the biosensor systems. Biosensors have enormous potentials and prospects in wide range of applications. These nanobioelectronic devices will revolutionize the detection system and signal processing, and enrich the quality of life.

At present, the nanotechnology-based biosensors are at the early stage of development. The vast applications of nanotechnology in such diverse fields such as semiconductors, biological and medical devices, polymer composites, optical

devices, dispersions, and coatings are amazing. With the discovery of novel materials and innovated approaches, it may progress by leaps and bounds, as in the case of the computer industry after the development of integrated circuits. As biosensors are essentially involved in revealing molecular information, they form a part of important information technology. The potentials and prospects of biosensors are effectively attracting myriad start-ups to the Fortune 500 companies to manufacture these devices.

In summary, the building blocks of biosensors, biomechanisms, biofunctionalization of nanodevices, search for suitable interfaces between biological material and electronics, intelligent signal processing of information transmitted by biosensors and other related issues are discussed. The methods for protein and DNA and RNA molecular probe design and syntheses, and microfabrication methods of the biosensor devices are enumerated. A detailed account on application of nucleic acid sensors and DNA chips, immunosensors, enzyme-based biosensors, organism and whole cell-based, ion sensors, natural and synthetic receptors for biosensors, new signal transduction technology, systems integration, proteomics and single-cell analysis, bioelectronics, nanoanalytical systems, and importance of nanoscale target molecular detection in cancer, AIDS and other diseases are discussed. The issues of technical challenges in designing and fabricating the smart biosensors, and integration of nanobiosensors with system-on-package (SOP) are discussed. Emphasis is made on nanochemistry applications for biosensor packaging. The potentials and prospects of biosensors, and their commercial value are mentioned.

References

1. Dell'Attì D, Tombelli S, Minunni M, and M Mascini: Detection of clinically relevant point mutations by a novel piezoelectric biosensor. *Biosens Bioelectron*, **2006**, 27(10): 1876–1879.
2. Toppozada A R, Wright J, Eldefrawi A T, Eldefrawi M E, Johnson E L, Emche S D, and C S Helling: Evaluation of a fiber optic immunosensor for quantitating cocaine in coca leaf extracts. *Biosens Bioelectron*, **1997**, 12(2):113–124.
3. Nuwaysir F, Huang W, Albert T J, Jaz S, Nuwaysir K, Pitas A, Richmond T, Gorski T, Berg J P, Ballin J, McCormick M, Norton J, Pollock T, Sumwalt T, Butcher L, Porter D, Molla M, Hall C, Blattner F, Sussman M R, Wallace R L, Cerrina F, and R D Green: Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. *Genome Res*, **2002**, 12(11):1749–1755.
4. Zhou C, Pivarnik P, Rand A G, and S V Letcher: Acoustic standing-wave enhancement of a fiber-optic *Salmonella* biosensor. *Biosens Bioelectron*, **1998**, 13(5):495–500.
5. Liu J, Janagama G D, Iyer M K, Tummala R R, and Z L Wang: ZnO Nanobelts/wire for Electronic Detection of Enzymatic Hydrolysis of Starch International Symposium and Exhibition on Advanced Packaging Materials Processes, Properties and Interfaces, Georgia Tech., Atlanta, USA (**2006**).
6. Shriner-Lake L C, Donner B L, and F S Ligler: On-site detection of TNT with a portable fiber optic biosensor. *Environ Sci Technol*, **1997**, 31:837–841.
7. Ragan F, Meaney M, Vos J G, MacCraith B D, and J E Walsh: Determination of pesticides in water using ATR-FTIR spectroscopy on PVC/chloroparaffin coatings. *Anal Chim Acta*, **1996**, 334:85–92.

8. Potyrailo R A, Hobbs S E, and G M Hieftje: Near-ultraviolet evanescent-wave absorption sensor based on a multimode optical fiber. *Anal Chem*, **1998**, 70:1639–1645.
9. Woods S I, Kirtley J R, Sun S, and R H Koch: Direct investigation of superparamagnetism in Co nanoparticle films. *Phys Rev Lett*, **2001**, 87(13), 137205, 4 Pages
10. Kaplan D K, Sverdlov V A, and K K Likharev: Coulomb gap, Coulomb blockade, and dynamic activation energy in frustrated single-electron arrays. *Phys Rev B*, **2003**, 68, 045321, 4 pages.
11. Pancrazio J J, Whelan J P, Borkholder D A, Ma W, and D A Stenger: Development and application of cell-based biosensors. *Ann Biomed Eng*, **1999**, 27:697–711.
12. Dickert F L and S Thierer: Molecularly imprinted polymers for optochemical sensors. *Adv Mater*, **2004**, 8(12):987–990.
13. Liebertzeit P A, Glanzning G, Jenik M, Gazda-Miarecka S, Dickert F L, and A Leidl: Softlithography in chemical sensing—analytes from molecules to cells. *Sensors*, **2005**, 5:509–518.
14. Hayden O and F L Dickert: Selective microorganism detection with cell surface imprinted polymers. *Adv Mater*, **2001**, 12:311–313.
15. Hayden O, Bindeus R, Haderspock C, Mann K, Wirl B, and F L Dickert: Mass-sensitive detection of cells, viruses and enzymes with artificial receptors. *Sens Actuators B*, **2003**, 91:316–319.
16. Cai W, Peck J R, van der Weide D W, and R J Hamers: Direct electrical detection of hybridization at DNA-modified silicon surfaces. *Biosens Bioelectron*, Article in Press, **2004**, 19(91):1013–1019.
17. Sahoo S K and V Labhasetwar: Nanotech approaches to drug delivery and imaging. *Drug Disc Today*, **2003**, 8:1112–1120.
18. Il-Han Hwang and Jong-Hyun Lee: Self-actuating biosensor using a piezoelectric cantilever and its optimization. *J Phys: Conference Series*, **2006**, 34:362–367.
19. Cui Yi, Qingqiao W, Hongkun P, and C M Lieber: Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. *Science*, **2001**, 293:1289–1292.
20. Wang Z L: Second International Workshop on nano-bio-electronics packaging held in Atlanta, (**2005**).
21. Baselt D R, Lee G u, Natesan M, Metzger S W, Sheehan P E, and R J Colten: A biosensor based on magnetoresistance technology. *Biosens Bioelectron*, **1998**, 13:731–739.
22. Harborn U, Xie B, Venkatesh R, and B Danielsson: Evaluation of a miniaturized thermal biosensor for the determination of glucose in whole blood. *Clin Chim Acta*, **1997**, 267(2):225–237.
23. Wilson M, Kannangara K, Smith G, Simmons M, and B Raguse: Nanotechnology, Basic Science and Emerging Technology. Chapman & Hall/CRC, NY (**2002**).
24. Goddard W A, Brenner DW, Lyshevski S E, and G J Lafrate Hand Book of Nanoscience, Engineering and Technology. CRC Press, USA (**2003**).
25. Jones E F, He Jiang, VanBrocklin H F, Franc B L, and Y Seo: Nanoprobes for medical diagnosis: current status of nanotechnology in molecular imaging. *Current Nanoscience*, **2008**, 4(1):17–29(13).
26. Rist M J and J P Marino: Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions. *Curr Org Chem*, **2002**, 6(9):775–793.
27. Kool E T: Artificial DNA through metal-mediated base pairing: structural control and discrete metal assembly. *Acc Chem Res*, **2002**, 35:936.
28. LaVan D A, Lynn D M, and R Langer: Moving smaller in drug discovery and delivery. *Nat Rev*, **2002**, 1:77–84.
29. Niemeyer C M: Self assembled nanostructures based on DNA: towards the development of nanobiotechnology. *Curr Opin Chem Biol*, **2000**, 4:609–618.
30. Langer P R, Waldrop A A, and D C Ward: Enzymatic synthesis of biotin-labeled polynucleotides: novel nucleic acid affinity probes. *Proc Natl Acad Sci USA*, **1981**, 78:6633–6637.

31. Huhtinen P, Soukka T, Lovgren T, and H Harma: Immunoassay of total prostate-specific antigen using europium (III) nanoparticle labels and streptavidin-biotin technology. *J Immunol Methods*, **2004**, 294(1-2):111-122.
32. Wang Z L: Second International Workshop on nano-bio-electronics packaging held in Atlanta, **2005**.
33. Cui Yi, Qingqiao W, Hongkun P, and C M Lieber: Nanowire nanosensors for highly sensitive and selective detection of biological and chemical species. *Science*, **2001**, 293:1289-1292.
34. Ebessen T W(ed.): Carbon Nanotubes: Preparation and Properties. CRC Press, Boca Raton, FL, USA (**1997**).
35. Riikonen S and D Sanchez-Portal: Structural model for Si (553)- Au atomic chain reconstruction. *Nanotechnology*, **2005**, 16:5218-5223.
36. Shim Moonsub, Kam Nadine Wong Shi, Chen J Robert, Li Yiming, and Hongjie Dai: Functionalization of carbon nanotubes for biocompatibility and biomolecular recognition. *Nano Letters*, **2002**, 2(4), 285-288.
37. Cai Hong, Cao Xuni, Jiang Ying, He Pingang, Fang Yuzhi, Cai Hong, Cao Xuni, Jiang Ying, He Pingang, and Yuzhi Fang: Carbon nanotube-enhanced electrochemical DNA biosensor for DNA hybridization detection. *Anal Bioanal Chem*, **2003**, 375:287-293.
38. Neves-Petersen MT, Snabe T, Klitgaard S, Duroux M and S B Petersen: Photonic activation of disulfide bridges achieves oriented protein immobilization on biosensor surfaces. *Protein Sci*, **2003**, 15(2):343-351.
39. Langer R and D A Tirrell: Designing materials for biology and medicine. *Nature*, **2004**, 428:487-492.
40. Kusnezow W and J D Hoheisel: Solid support for microarray immunoassays. *J Mol Recognit*, **2003**, 16:165-176.
41. Beier M and J D Hoheisel: Versatile derivatisation of solid support media for covalent bonding on DNA-microchips. *Nucl Acids Res*, **1999**, 27:1970-1977.
42. Mourougou-Candoni N, Naud C, and F Thibaudau: Adsorption of thiolated oligonucleotides on gold surfaces: an atomic force microscopy study. *Langmuir*, **2003**, 19(3):682-686.
43. Pan B F, Gao F, and H C Gu: Dendromer modified magnetic nanoparticles for protein immobilization. *J Colloid Interface Sci*, **2005**, 284(1):1-6.
44. Pan Z W, Dai Z R, and Z L Wang: "Nanobelts of semiconducting oxides". *Science*, **2001**, 291:1947.
45. Liu J, Janagama D G, Iyer M K, Tummala R R, and Z L Wang: ZnO Nanobelts/Wire for Electronic Detection of Enzymatic Hydrolysis of Starch International Symposium and Exhibition on Advanced Packaging Materials Processes, Properties and Interfaces, Georgia Tech., Atlanta, USA, (**2006**).
46. Janagama D G, Markondeya Raj J P, Iyer M K, and R R Tummala: Bio SOP with embedded Electrochemical Biosensors using MWCNT electrodes and Zirconia/Nafion based enzyme encapsulants. ECTC 57th Electronic components and Technology conference, **2007**.
47. Choi C J and B T Cunningham: Single-step fabrication and characterization of photonic crystal biosensors with polymer microfluidic channel. *Lab Chip*, **2006**, 6:1373-1380.
48. Zaytseva N V, Goral V N, Montagna R A, and J Baeumner: Development of a microfluidic biosensor module for pathogen detection. *Lab Chip*, **2005**, 5:805-811.
49. Martin P M, Laminated plastic microcomponents for biological and chemical systems. *J Vac Sci Technol*, **1999**, 17(4):2262-2269.
50. Wise K D: Wireless integrated microsystems: coming breakthroughs in health care. Proceedings of International Electron Devices meeting, **2006**.
51. Hafner J H, Cheung C L, Woolley A T, and C M Lieber: Structural and functional imaging with carbon nanotube AFM probes. *Prog Biophys Mol Biol*, **2001**, 77(1):73-110.

Molecular Dynamics Applications in Packaging

Yao Li, Jeffrey A. Hinkley, and Karl I. Jacob

Abstract Development of integrated circuits (IC) and semiconductor chips has narrowed dimensions of electronic components to length scales of nanometer size. However, design and manufacture of nano-scale electronic components need extensive understanding of material properties at small length scales, preferably at the atomistic level. Behavior of materials at nanometer scale is significantly influenced by size effects; behaviors of materials significantly change at nanometer level compared in macroscopic world due to size effects. As a result, electronic packaging has to adopt significant design requirements associated with the tremendous reduction in size.

Molecular dynamics (MD) simulation [1–5] was invented as a tool to account for the interactions between basic particles, generally atoms, in small systems of interest. Many attributes of materials can be obtained from MD simulation, at least on a qualitative level and sometimes in more quantitative manner. MD has been utilized as a powerful tool to narrow the number of possible candidate materials for many components in electronic packaging, based on the critical requirements on the material, such as resistance of materials to moisture, stress and thermal cycling, and strength of interfaces. This is achieved by understanding the behavior of the materials with hypothetical structure and composition, reducing or eliminating the need for synthesizing these materials at least during the initial material selection process. Thus, in addition to understanding material behavior, MD can be used to address the inverse problem of designing materials with specific properties for a chosen end use.

In the past decade, MD procedure has experienced explosive growth in its usage in a variety of areas. Here, we give a brief review of MD simulation procedure with particular emphasis to its applications in electronic packaging.

Among all the micro-scale or nano-scale simulation methods (MD, Monte Carlo methods and quantum mechanics), MD serves as a major simulation tool in electronic packaging area. Selection of an appropriate simulation method depends on

K.I. Jacob (✉)

Department of Polymer, Textile and Fiber Engineering, Georgia Institute of Technology, Atlanta, GA, USA

e-mail: jacob@gatech.edu

the length scale of the systems under consideration and the time scale associated with the process of interest that the system is subjected to. Many problems of interest in the packaging area fall within the length scale limits of MD as it is comparable to the relevant size of many electronic packaging devices. Compared to other methods, MD can model most properties and processes at the equilibrium state as well as many nonequilibrium phenomena (such as water diffusion, heat transfer, mechanical deformation) in the packaging area. In addition, MD can be coupled with (1) Monte Carlo methods to incorporate complementary effects within the MD time/length scale domain, (2) quantum mechanics to capture events even at smaller scale and the effects of changes in electronic structures from the changes in atomic positions or structure, and (3) with finite element methods to embed quantum and molecular effects with continuum structural behavior for a material or a device. Depending on the particular property, in most cases molecular simulations can give excellent qualitative results and in many cases very good quantitative results as well. Since there are limitations on time and domain size in MD simulations, one may find difficulties in accurately simulating some of the properties that are realized in a larger domain and longer duration in real materials.

Keywords Molecular dynamics · Molecular simulations · Electronic packaging

1 Molecular Dynamics Procedure

Molecular dynamics (MD) is a simulation methodology based on the well-familiar Newton's law, relating force with the rate of change of momentum, namely:

$$\vec{F} = m\vec{a} \quad (1)$$

If all the atoms and/or molecules in the simulation box are treated as classical particles obeying Newton's equations, then the positions and velocities of each atom can be calculated knowing the positions and velocities at a time immediately before the current time, once the forces acting on each atom are computed. Forces are derived from a conservative potential function, also known as force fields, arising from the interactions between the chemical units present in the simulation box.

Assuming that we have N particles in the system, we have $6N$ independent variables: positions $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N$ and velocities $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_N$. Derived from the classical dynamics, positions, and velocities of particles are related as

$$\frac{d\vec{r}_i}{dt} = \vec{v}_i \quad (2)$$

and

$$\frac{d\vec{v}_i}{dt} = \frac{\vec{f}_i}{m_i} = \frac{d^2\vec{r}_i}{dt^2} \quad (3)$$

where \vec{f}_i is the force vector acting on particle i . In order to implement Equations (2) and (3) in a computer simulation, numerical integration is employed instead of analytical integration. If the position of particle i at time t is known as $\vec{r}_i(t)$, using the Taylor's expansion, the position of the particle at time $t+\Delta t$ is given by

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta t \cdot \vec{v}_i(t) + \frac{(\Delta t)^2}{2} \vec{r}_i(t) + \frac{(\Delta t)^3}{6} \vec{a}_i(t) + O((\Delta t)^4) \quad (4)$$

Substituting Equations (2) and (3) into Equation (4), we can get

$$\vec{r}_i(t + \Delta t) = \vec{r}_i(t) + \Delta t \cdot \vec{v}_i(t) + \frac{(\Delta t)^2}{2} \cdot \frac{\vec{f}_i(t)}{m_i} + \frac{(\Delta t)^3}{6} \vec{a}_i(t) + O((\Delta t)^4) \quad (5)$$

Thus, the position of atoms at time $t+\Delta t$ can be calculated by knowing the position at t and the forces at t , provided the time interval Δt is short. A similar relation can be derived for calculating the velocity at a given time, knowing the velocity at a short interval prior to that time.

Another assumption made by MD is the reversibility of time, i.e., if we trace particles' trajectories in the reverse time direction, all particles will follow their original trajectories. Under this assumption, we can obtain the particle's position at time $t-\Delta t$ as

$$\vec{r}_i(t - \Delta t) = \vec{r}_i(t) - \Delta t \cdot \vec{v}_i(t) + \frac{(\Delta t)^2}{2} \cdot \frac{\vec{f}_i(t)}{m_i} - \frac{(\Delta t)^3}{6} \vec{a}_i(t) + O((\Delta t)^4) \quad (6)$$

Adding the Equations (5) and (6) and omitting the higher-order derivatives will give us

$$\vec{r}_i(t + \Delta t) + \vec{r}_i(t - \Delta t) = 2\vec{r}_i(t) + (\Delta t)^2 \cdot \frac{\vec{f}_i(t)}{m_i} \quad (7)$$

Thus, as long as we know the positions at least at time t and preferably at $t-\Delta t$ as well as the net force acting on the particles at time t , we can compute the positions at time $t+\Delta t$. Equation (7) is the widely used Verlet algorithm in MD simulation, which is similar to the finite difference numerical procedure.

Since higher-order terms are neglected, a requirement for Equation (7) to be useful is to have very small time interval Δt , comparable to the characteristic time scale of actual atoms' movement. A typical time step of 1–10 fs is adopted in most MD simulations.

The remaining unknown in Equation (7) is $\vec{f}_i(t)$, which is the gradient of the potential energy U field, given by

$$\vec{f}_i(t) = -\nabla_i U(t) \quad (8)$$

The interaction potential energy U in an MD simulation for polymers usually consists of a number of bonded and nonbonded terms, as given by

$$U = \sum_{i=1}^{N_{\text{bond}}} U_{\text{bond}} + \sum_{i=1}^{N_{\text{angle}}} U_{\text{angle}} + \sum_{i=1}^{N_{\text{torsion}}} U_{\text{torsion}} + \sum_{i=1}^{N_{\text{inversion}}} U_{\text{inversion}} + \sum_{i=1}^{N-1} \sum_{j>i}^N U_{\text{vdw}} \\ + \sum_{i=1}^{N-1} \sum_{j>i}^N U_{\text{electrostatic}} \quad (9)$$

Since these energy terms may involve two or more particles, to compute $\vec{f}_i(t)$, we need to go through all the energy terms, calculate corresponding net force acting on particle i by other particles. The potential energy function or the coefficients used in this expression are generally called force field in MD. Besides the functional form of the potentials, force field defines a set of parameters for the interactions involved in each type of particle that contribute to the potential energy. Both the functional form and parameters can be obtained from empirical, semi-empirical, or ab initio methods. In Equation (9), the first four terms stand for interactions between bonded particles, i.e., bond stretching U_{bond} , bond angle bending U_{angle} , dihedral angle torsion U_{torsion} , and inversion interaction $U_{\text{inversion}}$; while the last two terms are interactions between unbonded particles, i.e., van der Waals energy U_{vdw} and electrostatic energy $U_{\text{electrostatic}}$; N_{bond} , N_{angle} , N_{torsion} , and $N_{\text{inversion}}$ are the total numbers of these specific interactions in the simulated system; i and j in the van der Waals and electrostatic terms indicate the pair-wise particles involved in the interaction. A number of force fields have been developed for different types of systems. Popular force fields include: AMBER [6] – widely used for proteins and DNA; CHARMM [7] – widely used for both small molecules and macromolecules; EAM [8] – broadly used for fcc metals; COMPASS [9] – developed specially for polymers; CFF [10] – adapted to a broad variety of organic compounds, to name a few. Thus, a conservative force field is assumed, arising from the interactions of a given atom with other atoms.

Subtracting Equation (6) from Equation (5) will give the following equation:

$$\vec{r}_i(t + \Delta t) - \vec{r}_i(t - \Delta t) = 2\Delta t \cdot \vec{v}_i(t) + O((\Delta t)^3) \quad (10)$$

Neglecting the last term and dividing the Equation (10) by $2\Delta t$ leads to

$$\vec{v}_i(t) = \frac{\vec{r}_i(t + \Delta t) - \vec{r}_i(t - \Delta t)}{2\Delta t} \quad (11)$$

The total kinetic energy then can be calculated as

$$\vec{E}_{\text{kinetic}}(t) = \sum_{i=1}^N \frac{1}{2} m_i \vec{v}_i^2(t) \quad (12)$$

Many thermodynamics properties (temperature, pressure, etc.) of the simulated system can also be derived from the time-dependent position and velocity data of

atoms based on statistical mechanics. For an isolated system, the total energy and linear momentums are conserved.

MD simulations are carried out within the framework of an appropriate ensemble, such as microcanonical ensemble (NVE), canonical ensemble (NVT), and isothermal–isobaric ensemble (NPT), depending upon which thermodynamic variables are fixed. For example in the isothermal–isobaric ensemble, the number of moles of the substance (N), pressure (P), and temperature (T) are kept constant. There are many subtle procedures involved in keeping the average value of these variables fixed during the simulation, which is achieved through various “thermostats.” Since the number of interactions between atoms increase with the number of atoms present in the system, for a system containing N particles the number of times forces are calculated at a given time step vary at least with N^2 [2]. This will limit the number of particles one could simulate and the simulation time one can afford, even with the availability of larger and faster massively parallel clusters. Although there are many exceptions, the highest number of particles or atoms generally studied in MD is in the order of millions of atoms and the simulation time frame has been within nanosecond duration. Even with a very large number of atoms in a system, many of those atoms will reside on the surface of the simulation box, giving rise to the influence by surface energy in the simulation of such nanosystems. Thus, if the interest is to study the bulk properties, the surface effects must be removed, which is achieved by using the periodic boundary condition. For periodic boundary conditions, the same simulation box or unit cell is repeated in every direction from the primary simulation box to fill the entire domain. When periodic boundary conditions are applied, an atom that moves from the primary simulation cell to the next cell will reappear like an atom entering from the opposite side of the primary simulation cell as all cells are images of the same cell.

MD is a convenient tool to study the static and dynamic attributes of materials. Researchers have been applying MD to most materials science areas, including metals, small molecules, polymers, biomaterials, etc.

2 Structures and Properties of Epoxy

The earliest materials used in IC packaging were ceramic flat packs. During the development of the IC industry, the need for encapsulation shifted packaging materials from ceramic flat packs to plastic flat packs. Encapsulation plastics primarily used in the last step of IC packaging are epoxy resins. In addition, epoxys also serve as strong adhesives applicable to most interfaces. Flexible epoxy resins can be excellent electrical insulators, good conductors of heat, offer reliable protection to moisture, and strong adhesion to metals. Detailed studies about the thermal and mechanical properties of epoxy resins are crucial to choosing suitable epoxys for packaging applications.

With MD simulations, when done with proper forcefields, one can predict how material properties will change by changing functional groups, backbone structures, crosslink densities, and other structural parameters. There are a variety of epoxy

compounds in the market, and the resulting variability of epoxy structures by using various epoxy compounds is highly useful for practical applications, but a challenging problem for simulations. Epoxy resins are crosslinked polymer networks without a universal long-range repeat unit. Construction of a realistic epoxy initial structure is the first step in the simulation process. If the initial structure is not a realistic one with energies close to the minimum energy for that simulation cell, then MD results could be affected, or simulation time will be significantly higher for realistic results. In order to avoid the lengthy simulation time requirement, an initial configuration should be generated as close to the final lowest energy state as possible. One of the major problems of generating initial structures for epoxy resins is the complexity of crosslinking reactions between epoxy groups and curing agents. The reaction process is even more complex considering the presence of various types of epoxy groups and curing agents and possible side reactions.

Recently, new algorithms were developed to build realistic initial configurations of epoxy polymer networks at different length scales accessible to MD. The first category of these methods adopt “step growth” algorithm to create networks from basic repeat units. The assumption made here is that those repeating units are identical and therefore can be replicated and extended in three dimensions. Starting from one central unit, surrounding units are grown from the reactive points of the central unit such that each head atom should be linked to a tail atom. This method is straightforward and easy to implement. However, this procedure only works for certain networks and the network produced is too regular and flawless to be realistic. Spatial hindrance encountered in the growth process may be another problem if the repeating unit is short or spatially compact [11].

In the second approach, segments between fixed crosslinking junctions are generated first according to the junction-to-junction distance of segments between them. This technique is appropriate for networks with uncertain number of repeating units. If atoms at crosslinking points are the same and branches growing from the crosslinking points are chemically identical, efficiency of this approach can be very good. Linkages at each crosslinking point should be determined while the structures of each branch are different. The computational cost is expected to be very high when the network is complex and large and as in the previous case, presence of defects and unreacted functional groups, are not easy to introduce using this approach [12].

Another way to generate initial epoxy structures is to mimic the actual polymerization process. Generally, as the first step monomers are created in a simulation box, and an equilibrium simulation is performed to blend and relax monomers. Then, using a crosslinking criteria (e.g., the distance between nearby atoms is within a chosen limit) adopted to generate bonds between possible reactive groups, crosslinking is implemented. This type of method can build either all atom models or coarse grain models. They are able to simulate networks with multifunctional groups and capture the complexity of these networks. Usually, an imperfect network will be generated [13]. All these methodologies to generate initial polymer network configurations rely on using a valid force field for polymers, although reasonable

configurations can also be created using other methods such as Monte Carlo methods using less rigorous force fields. COMPASS and Dreiding [14] are the two major popular force fields commonly used for polymers, among others.

Once the initial structure is generated, then a series of MD simulations are carried out under various conditions of temperature and pressure (or stresses). The data generated from these simulations, namely position and velocity of each atom in the simulation cell at various time steps, can be used to calculate a number of material properties. Many reports have introduced different methodology to simulate polymer networks, for understanding the structure and confirmation of polymers as well as to extract physical and mechanical properties of a polymer. Generation of initial structures under various conditions is generally enough to study the conformational aspects of polymers, for example, Yarovsky et al. [15] developed a way to construct low-molecular weight water-soluble epoxy resins. Recently, Xu et al. [16] conducted atomic simulations of epoxy resins using an alternate MM (molecular mechanics)/MD method. A four-step procedure (monomer construction, monomer equilibrium, crosslinking, and the final network equilibrium) was employed in these papers to understand the conformational aspects of polymer chains. Reasonable agreements between computational results and experimental results have shown the validity of these methods. Although many researchers use custom codes for MD simulations, industrial research and even undergraduate instructions are usually done with commercial simulations codes [17], such as the molecular simulation package Cerius2 (now Material Studio) [18] from Accelrys to study conformational aspects and to predict material properties. However, the size of the simulation cell is generally limited when such commercial software is used mainly with PCs, and control of the detailed crosslinking process is also difficult to implement. In the last few years, researchers tried to combine and modify previous methodologies to build polymer networks, taking efficiency and robustness as criteria. Komarov et al. [19] and Varshney et al. [20] both successfully deployed different mechanisms to construct highly crosslinked, relatively large-scale polymer networks with fairly consistent properties with experiments. Figure 1 is a typical example of the initial structure of a simulation box filled with crosslinked epoxy resins.

There are a number of material properties that are relevant in the IC packaging area. Thermal stability of epoxy resins is the one of the most critical properties. Glass transition temperature (T_g) is one physical parameter that we can obtain easily from MD simulation, for epoxy with various compositions. To obtain the glass transition, a simulated annealing process can be performed with molecular dynamics to find the volume of the simulation box for various temperatures. The temperature at which a sudden change in the slope of the volume and temperature relationship occurs indicates the T_g . Figure 2 is an example showing the change of volume as a function of temperature, where the intersection point of the two fitting lines gives the T_g . Another attribute that is important in packaging systems is the volume thermal expansion coefficient (VTEC), which can be obtained from the volume–temperature relationship given in Fig. 2. VTEC is different above and below the T_g for the same sample. Linear thermal expansion coefficient is related to VTEC by a 1/3 factor for isotropic materials.

Fig. 1 Epoxy resins model system: stick-ball model, balls with different grades of black represent oxygen, carbon, hydrogen and nitrogen. Reprinted from [16], copyright 2006, with permission from Elsevier.

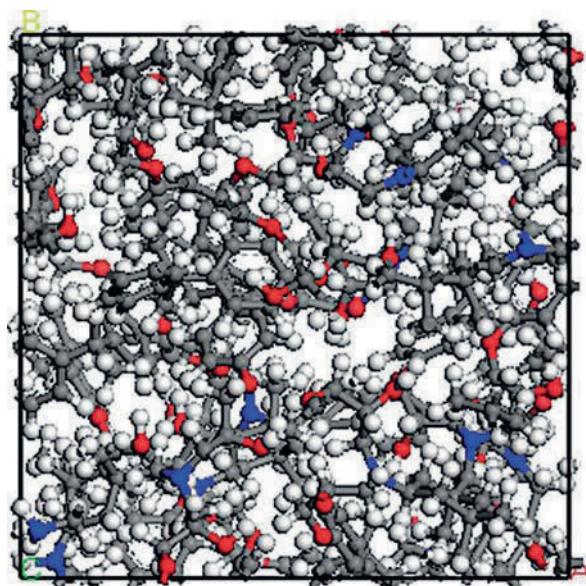
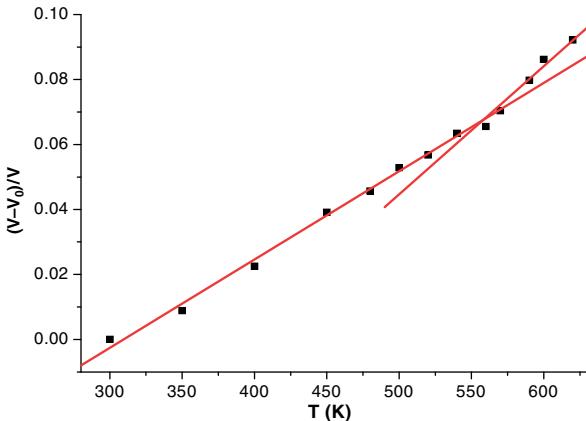


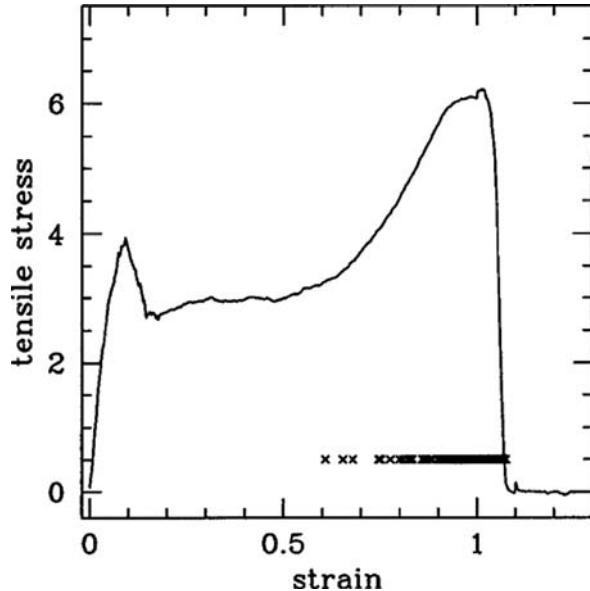
Fig. 2 Volume change vs. temperature of diaminodiphenyl methane/4,4'-diaminodiphenylsulfone (TGDDM/DDS) system



Important mechanical properties, such as bulk modulus, Young's modulus, and Poisson's ratio, also can be computed with molecular simulations. This can be done in two different ways, using molecular mechanics (MM) or with MD simulations. In the first approach, boundary atoms in the opposite sides are moved by a very small amount and the force that must be applied on those planes to keep the entire system at the minimum energy configuration can be calculated. Then by uniformly increasing the axial strain in the material (by moving the boundary layers progressively and calculating the corresponding force), the force-displacement relationship can be obtained. There are various modifications to this approach, such as instead of moving only the boundary atoms, all atoms are moved, in the direction the nearest

boundary was moved, proportional to their position from the center of the simulation box such that a uniform displacement is induced in the material from the displacement of the boundary atoms. Such a modification could allow one to apply larger displacements on the boundary layer, consequently larger strains, and thus fewer simulations. In the MD approach, stresses are applied on the boundary and MD simulations can be carried out with constant pressure constraint until internal equilibrium is reached, thus the corresponding strain can be obtained. Then the force is increased and the simulation process is continued. There are a number of coupling procedures, to connect applied stress with the simulation procedure, such as loose coupling technique [21] and Parrinello–Rahman technique [22]. As an example, a stress–strain curve drawn from a simulated uniaxial tensile test of a model sample is shown in Fig. 3. From this curve, the Young's modulus or the tangent or secant modulus at any strain can be calculated easily. Mechanical strength can also be calculated; but to obtain strength, these simulations must be repeated many times since strength is more a statistical property sensitive to polymer confirmation in the simulation cell.

Fig. 3 Tensile stress–strain curve for the fully bonded network. The points mark the strains at which bond breaking occurs. Stress is in LJ units. Reprinted with permission from [23]. Copyright 2001 American Chemical society



As implied earlier, crosslink densities that relate directly to the polymer morphology will have a significant influence on the thermal and mechanical properties. Usually, higher crosslink densities lead to higher T_g and better mechanical strength. M. Stevens et al. [23–26] manipulated connectivity of epoxy networks to investigate its influence on the strength of interfaces. It was also found that mechanical properties were dependent on the degree of curing. The factors affecting properties of epoxys as adhesives will be discussed again later.

3 Moisture Diffusion

Considering the environment ICs usually encounter, it's necessary to investigate the durability aspects of the system, which relates to the resistance of insulating materials to water intake and diffusion. Presence of water could create swelling, affect thermal expansion, and could cause hydro-thermal degradation of the polymeric material. Experimentally, it is relatively easy to measure the diffusion coefficient of small molecules through bulk material once the material is synthesized, although for interfaces it may pose some difficulties. However, compared to experiments, MD is an effective approach to get an insight into the diffusion behavior at an atomistic and fundamental level, if such materials are unavailable for diffusion experiments or we want to design and synthesize an epoxy composition with certain moisture transport properties. In this case, the power of MD is more on narrowing down the candidate materials with specific properties that needs to be synthesized for further experimental verification. The basic MD approach for moisture diffusion in polymers is fairly straightforward [27–29]:

According to Fick's law,

$$\vec{J} = -D\nabla c \quad (13)$$

where \vec{J} is the flux, D is the diffusion coefficient, and c is the concentration of the diffuser.

Conservation of mass gives:

$$\frac{\partial c}{\partial t} + \nabla \cdot \vec{J} = 0 \quad (14)$$

Combining Equations (13) and (14), we can get

$$\frac{\partial c}{\partial t} = D\nabla^2 c \quad (15)$$

Solution diffusion equation in three dimensions is given by

$$c(\vec{r},t) = (2\pi Dt)^{3/2} \exp\left[-\frac{r^2}{2Dt}\right] \quad (16)$$

The approach first used by Einstein will lead to the following equation:

$$\langle r^2(t) \rangle = 6Dt \quad (17)$$

If we define $\Delta\vec{r}_i(t) = \vec{r}_i(t) - \vec{r}_i(0)$, then

$$\langle \Delta r^2(t) \rangle = \frac{1}{N} \sum_{i=1}^N \langle \Delta r_i^2(t) \rangle = \frac{1}{N} \sum_{i=1}^N \langle [\vec{r}_i(t) - \vec{r}_i(0)]^2 \rangle \quad (18)$$

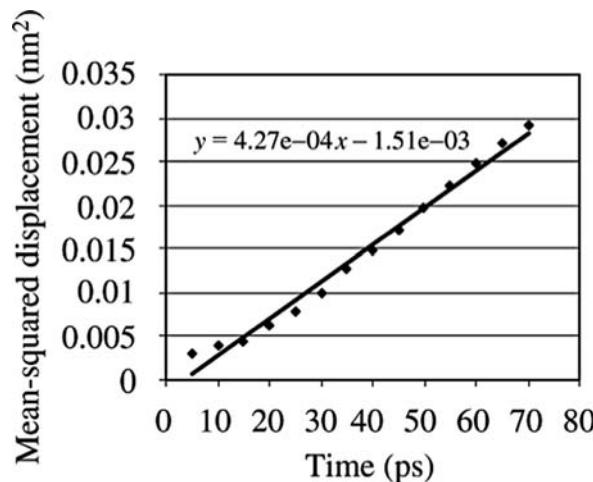
Substituting Equation (18) into Equation (17) gives

$$D = \lim_{t \rightarrow \infty} \frac{1}{6N} \sum_{i=1}^N \langle [\vec{r}_i(t) - \vec{r}_i(0)]^2 \rangle \quad (19)$$

This equation can be applied to get the diffusion coefficient D , using the positions of diffusing molecules in a medium. For the problem of interest, in Equations (18) and (19), N is the number of penetrating water molecules in the epoxy; $\vec{r}_i(t)$ is the position vector of water molecule i at time t ; $\langle \rangle$ means an ensemble average.

Diffusion of water molecules in polymers was studied by many researchers [30–37]. Diffusion coefficient is mainly related to temperature, relative humidity, and the structure of materials. In MD, one can compute mean square displacement (MSD) and plot it as a function of time. From this plot, diffusion coefficient can be derived from the slope of a linear fitting to data points according to Equation (17) or (19). Figure 4 shows an example for this method. Although the principle is fairly straightforward, the implementation and analysis of the result can be involved when using this method. Since there are various regions in the mean-squared displacement–time plot, one has to be careful in picking the appropriate region for calculating the diffusion coefficient. For example, diffusion coefficient calculated from the ballistic region may not be of any significance. In addition, many diffusion processes are non-Fickian, thus care must be taken in the interpretation of data to obtain meaningful results.

Fig. 4 Mean-squared displacements of water molecules against time in the epoxy molding compound (EMC) bulk and the fitted lines. Reprinted from [30] with permission from Koninklijke Brill NV.



For polymers, the total amount and the distribution of free volume cavities will directly influence how fast small molecules can diffuse. Diffusion coefficients of same penetrant molecules in polydimethylsiloxane (PDMS) is three times larger than in amorphous polyethylene (PE) due to larger free volume and broader free volume distribution in PDMS [35]. Since most polymers stay in a glassy state at

room temperature, the diffusion of small molecules always take a much longer time than in a simple liquid. Interactions with water molecules, such as hydrogen bonding, can significantly alter the diffusion process. This procedure is very useful in obtaining various parameters associated with the diffusion of water molecules, as shown by E. Dermitzaki et al. [38], who by systematically changing the structures of epoxy resins found out that the polarity is the dominant factor for water uptake.

Size effect of materials also plays a role in water transport. Penetrating capabilities of small molecules through bulky materials are believed to differ from that in thin films or membranes. A comparison between interface and bulk models of simulated rubbery and glassy polymers addressed by D. Hofmann et al. [32] revealed an experimentally important pervaporation feature of membranes, which means a combination of membrane permeation and evaporation. Comparison between epoxy molding compound (EMC) bulky samples and EMC/Cu interfaces [30] shows that water molecules are more prone to penetrate in thin interfaces. The diffusion coefficient of moisture through interfaces increased by one magnitude compared to that in bulk. Such behavior demands researchers to pay more attention to study and prevent moisture diffusion particularly at thin interfaces, and this is becoming an important issue as electronic devices are miniaturized.

Some small penetrants (such as water) in a polymer have aggregative attributes and intend to form large clusters. In one of the earliest papers in this field, Yoshinori Tamai et al. [35] studied the mechanism of small molecules diffusing through an amorphous polymer membrane. Diffusion of different penetrants (water, methane, and ethanol) in different polymers (PE and PDMS) was conducted. Diffusion coefficient of nonaggregated penetrant species coincided with experimental value; however, aggregation of water and ethanol in PDMS reduced diffusion coefficient by more than 1 magnitude. It suggested that water clusters forming in polymers played an important role in diffusion. Results by Fukuda et al. [31] also validated this observation. This study was also performed with MD simulation with a simulation time of about 10 ns. The diffusion coefficient of a water cluster containing 10 molecules in polyethylene was detected to be $7.1 \times 10^{-7} \text{ cm}^2/\text{s}$. Relevant to the topic on aggregation, the relative humidity affects the diffusion behavior through the formation of water aggregates. At a low moisture concentration, the number of water clusters in polymer was believed to be few. Higher relative humidity leads to a low moisture concentration gradient and more water aggregates. These two results will contribute to a low diffusion coefficient [30].

The result from water uptake can be divided into two categories: change in physical properties and loss of mechanical strength. It was discussed earlier that water update by epoxy will decrease glass transition temperatures [38]. Larger amount of moisture in epoxies means a larger reduction of glass transition temperature, which can change the state of some materials from glassy to rubbery. It was observed experimentally that the joint strength of epoxy adhesives was substantially weakened when exposed to moisture above a critical relative humidity for a long time [39]. However, the influence of moisture exposure on the mechanical properties of epoxies has not been adequately captured through MD simulation.

Since moisture diffusion will directly determine the reliability of ICs in long periods of time, it is worthwhile to clearly understand the mechanism of moisture penetration at a deeper level. Besides the diffusion coefficient, further studies on the mechanical behavior of materials, coupling of water molecular diffusion and polymer swelling, combination of diffusion, and chemical degradation of polymers are possible areas for the application of MD techniques.

4 Interconnection Alloys

Several materials are used as electrical interconnects between electronics devices in electronics packaging. A major part of these interconnect materials are metals with low melting point, because melting metals at high temperature will cause serious damage to the IC devices. β -Tin (β -Sn) and lead (Pb) are the two most commonly used materials among the metals with low melting temperature. While used as interconnects in the packaging area, both Sn and Pb have poor electrical conductivity. Another problem with Pb is its harmful impact to the environment. More and more lead-free alloys of tin and other metals with high electrical conductivities are used as interconnects. Copper and silver are the best two electrical conductors; while the drawback with them is the high melting temperature.

One solution to the problem addressed in the previous paragraph is to mix tin and silver/copper at a proper ratio to obtain a reasonable melting temperature and electrical conductivity. However, a commonly used Sn/Ag solder with a ratio of 96.5/3.5 has a melting point of 221°C [40], higher than temperature threshold of some electronic components. Considering nanosize effects on material properties, more and more researchers [40–45] tried to reduce the melting temperature by decreasing the size of metals. This is an area of considerable interest in nanoscale simulations. H. Dong et al. [40] employed MD simulation in conjunction with modified embedded atom method (MEAM) to investigate the intermix ability of nano-Ag/Sn alloy. MEAM is a modified approach based on embedded atom method (EAM) describing the energies between atoms, particularly for bcc metallic atoms. In these simulations, two 4-nm droplets of Ag and Sn were generated, equilibrated, and placed together at 500 K, a temperature that is higher than the melting temperature of Sn and lower than that of Ag. The initial configurations of Ag and Sn are shown in Fig. 5, where Ag maintained its ordered lattice structure and Sn had a liquid-like structure.

A 3 ns simulation time for interdiffusion suggested that no substantial intermix took place except collapse of both particles (Fig. 6). If the temperature was elevated to 1000 K, a total interdiffusion was observed. Using analytical tools and the calculated activation energy for interdiffusion, time need for intermix at 500 K was estimated to be 118 ns. This time scale is too large for MD simulation but easily feasible for practical solder application, but with increasing temperature this problem also falls within the capabilities of MD. Such MD simulations can be used to understand temperature dependence of the degree of mixing, speed of interdiffusion, and even activation energy.

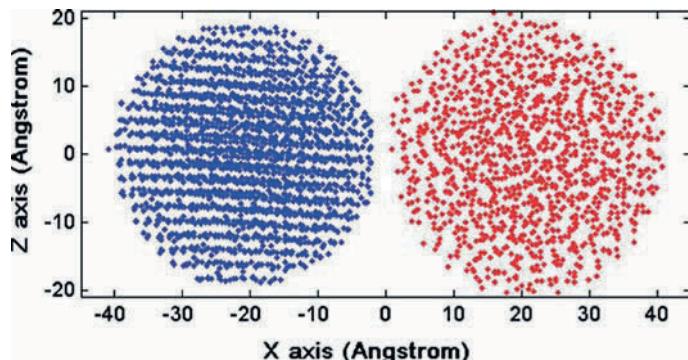


Fig. 5 Initial configurations of Ag and Sn. Reprinted from [40] with permission from Institute of Physics.

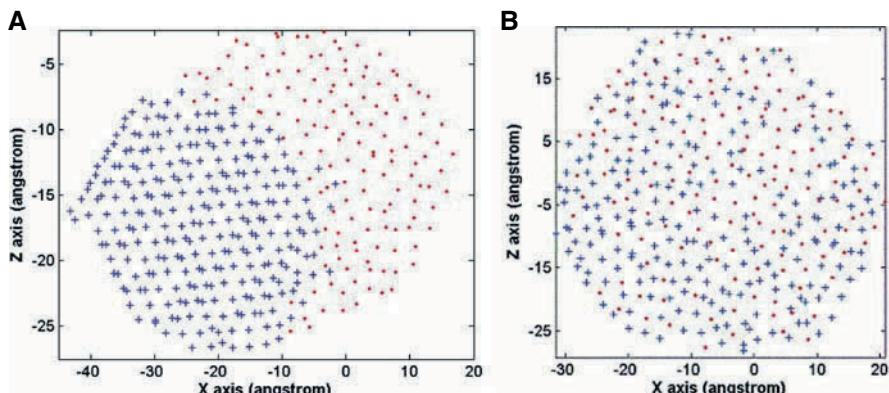


Fig. 6 Configurations of Ag/Sn after 3 ns at (a) 500 K and (b) 1000 K. Circles stand for silver particle; plus signs stand for tin particles. Reprinted from [40] with permission from Institute of Physics.

Another approach to replace the current Sn/Pb solder is by using conductive polymer/metal composites, which are often called electrically conductive adhesive (ECA). Two examples of ECA systems are shown in Fig. 7. One concern is the electrical conductivity of these composites, which is not only related to the conductive

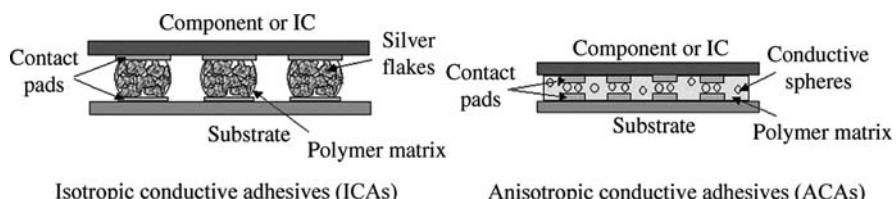


Fig. 7 Schematic diagram of electrically conductive adhesives (ECAs). Reprinted from [43] with permission from Koninklijke Brill NV.

polymer matrix but also with the dispersion state of metal fills. Interfaces between metal fills will significantly degrade the electrical conductivity. To interconnect isolated metal fills, one has to heat the polymer/metal composite assembly to fuse metal fills. The maximum temperature one can heat such a part is limited by the degradation temperature of polymers. Again, decreasing the size of metal particles is a viable means to suppress the melting temperature of metals. MEAM/EAM-coupled MD simulations were conducted by H. Dong et al. [43] to investigate the time evolution of Cu nanoparticles between two substrates in the same temperature range as discussed before. Results showed that a more ordered structure than the original one was formed under the applied strain, which might improve conductivity of ECA, although experimental verification is needed to validate this result.

Besides thermal properties of interconnect alloys, another problem to solve is the failure mechanism of Sn-based alloys. Many experimental investigations were reported recently, but a fundamental understanding of the universal failure mechanism for these systems is still lacking. MD simulations provide a convenient and powerful guiding tool to investigate the failure mechanism at the atomic level. W. Wang et al. [46] investigated uniaxial tension of nanocracked β -Sn single crystal using the MD method. The sample was prepared by first generating ideal β -Sn single crystal and then manually removing some atoms to create defects. Morse potential was used to compute the pairwise interatomic forces. Uniaxial tension applied along different directions was shown to induce change in dislocation mechanisms. In Fig. 8, simulation plots are given for two different tension cases. In the first case (force along the (010) direction), only a one-slip system was observed, which means a few dislocations. In the second case, two-slip system was observed

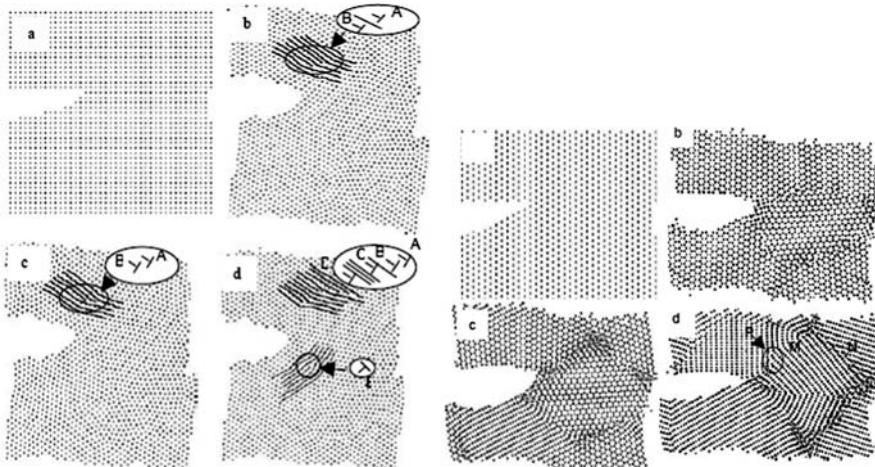


Fig. 8 *Left:* Simulation plots of the first case: applied tensile force along (010) direction (**a**) initial, (**b**) after relaxation, (**c**) after 4000 fs, (**d**) after 5,000 fs. *Right:* Simulation plots of the second case: applied tensile force along (001) direction (**a**) initial, (**b**) after relaxation, (**c**) after 5000 fs, (**d**) after another 5000 fs without loading. Reprinted with permission from [46], Copyright 2005 IEEE.

almost simultaneously. But the disorder induced by atomic positions from the two slips would initiate sub-boundaries to minimize energy. These results were similar to the experiments conducted in TEM.

5 Modification of Interface

The adhesives at the interface between multilayers are expected to provide sufficiently strong adhesion to undergo thermal cycling, moisture transport, and chemical and radiation degradation. However, interactions between plastics and metals are observed in general to be weak, which is a significant challenge for electronic packaging. Efforts were made to improve joint adhesion by introducing strong connections between polymers and metals. Among all the physical and chemical interactions, metallic bonds, ionic bonds as well as covalent bonds are three candidates for adhesive purpose. When making modifications for the adhesive layers, it is convenient to examine the effectiveness of newly introduced interfacial connections using calculations with MD [47–50]. Using MD simulations, with appropriate modifications, one can also observe bond formation and bond breakage at the interfaces (under appropriate conditions the system is subjected to), compare changes in chemical bonds and to calculate the strength of these bonded joints. Although the procedure appears straightforward, there are a number of difficulties in implementing MD to study interface bonding, but still the results are expected to provide qualitative and sometimes quantitative patterns. Normally, the interactions between plastics and metals were believed to be the van der Waals and the weaker type of electrostatic forces. These pair forces are dependent on distances between neighboring atoms for a specific system. However, some researchers[51] argued that the bonding between epoxy adhesive and metal was strongly related to the content of metal oxides, although the mechanism was unclear. Development of the force field involved in simulating the interface between dissimilar materials is a major challenge. Unless appropriate force fields are used, the results will not be representative for the system we are investigating.

Fan et al. [52–55] (cf. Figs. 9, 10) tracked the interfacial energies of epoxy/copper system in a thermal cycling test via MD method. Interfacial energies were calculated from the difference between total energy of combined copper/epoxy system and the sum of energies of copper and epoxy considered separately. The bonding energy was found to decrease with a decreasing percentage of cuprous oxide and additional complications were found from voids occurring at the epoxy/copper interface. So, the key factor governing the interfacial adhesion strength is still under considerable debate.

In general, addition of a third layer between the two layers in an interface (like between epoxy and copper) can create stronger interactions between the neighboring layers. If suitable force fields are available, this problem provides a fertile area for MD simulations, in order to understand the enhancement of strength with the nature of the newly added third layer. An alternative way to enhance interactions is to make an epoxy/copper composite instead of pure epoxy to provide metallic interactions.

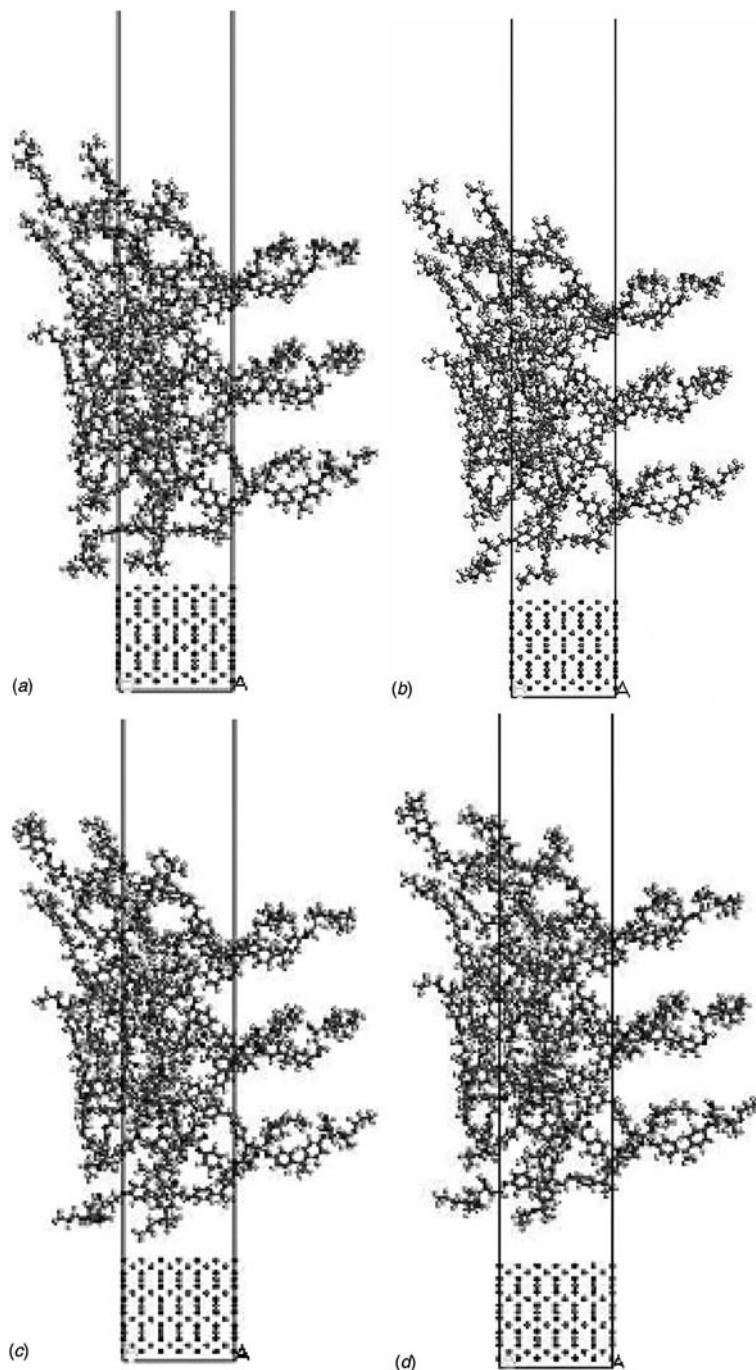


Fig. 9 Snapshots of the system after (a) 0 thermal cycle, (b) 600 thermal cycles, (c) 1200 thermal cycles, and (d) 1800 thermal cycles from the MD simulations. Reprinted from [53] with permission from ASME.

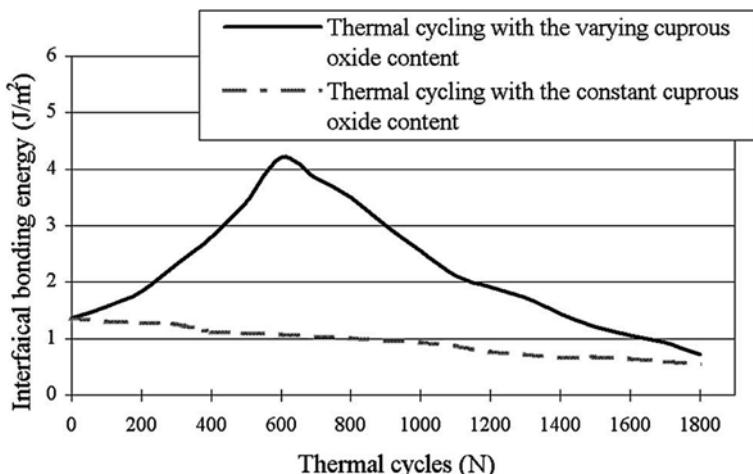


Fig. 10 Interfacial bonding energy against thermal cycles. Reprinted from [53] with permission from ASME.

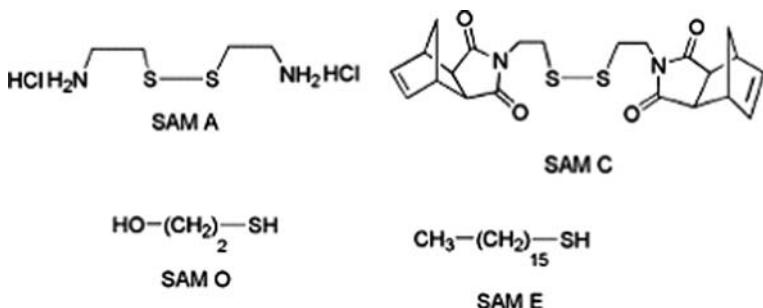


Fig. 11 Chemical structure for four types of SAMs. Reprinted with permission from [54], Copyright 2008 IEEE.

Wong et al. [54] used a self-assembly monolayer (SAM) to enhance joint adhesion in the epoxy/copper system. Four different types of thiols/disulfides, copper surface modifiers, were chosen as the SAM candidates. Their structures are shown in Fig. 11.

Change of potential energies before and after a covalent bond forming between SAM and epoxy was used as a criterion to determine how many interfacial bonds were formed. SAM density on the copper substrate was estimated by the same method. Interfacial bonding energies between SAM A and SAM C modified copper and epoxy showed a significant enhancement of adhesion, which was also consistent with shear tests (Fig. 12). Problems like this are suitable for MD simulations, to understand how certain functional groups could enhance the adhesion and the resulting structure of the adhesive layer.

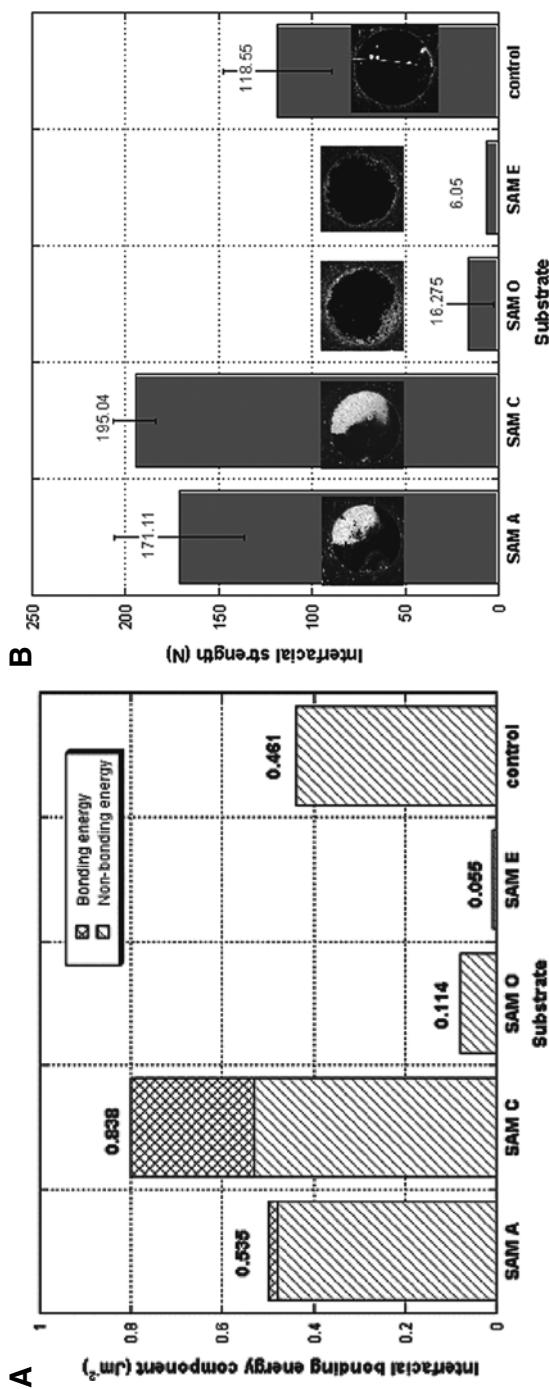


Fig. 12 (a) Energy component of the interfacial bonding energy predicted by MD, (b) interfacial bonding strength data evaluated by the button shear test.
Reprinted with permission from [54], Copyright 2008 IEEE.

6 Thermal Conductivity

Fast technological development in the semiconductor industry and tougher new requirements governed by the market place have reduced the size of chips to nanometer scale, that allows to pack more dense chips in a given area. Since many of the devices containing these chips are continuously used, these chips and associated connective circuits will produce a significant amount of heat. Performance of IC chips will be weakened if the heat produced during the process is not dissipated quickly. As described below, heat conduction can also be studied using the MD approach. Theoretically, the parameter that describes the speed of thermal dissipation is called thermal conductivity λ , which is defined by the following equation:

$$\lambda = -\frac{J}{\Delta T} \quad (20)$$

where J is the heat flux, which is defined as the amount of energy transferred in a given time through a given surface area and ΔT is the temperature gradient. But this Fourier law is valid only for macroscopic materials. When the size of devices is reduced to nanoscale, thermal conductivity is believed to depend not only on the properties of materials but also the size of the materials. Thus, considering the complexities involved on the dependence of properties on the size, this is an active area of research. It is relatively easy to measure the thermal conductivity of bulk materials, but finding size-dependent thermal conductivity coefficients for nanosized materials pose some experimental difficulties. So, instead of using numerous experiments, it may be preferable to use simulation methods to study thermal transport in IC chips. MD deals with the smallest scale simulation of heat dissipation compared to other continuum models as well as Boltzmann transport equation (BTE) models. Dependence of various parameters in MD and their relationship between heat-transfer phenomenon is summarized in the sketch given in Fig. 13.

In MD simulations, the thermal conductivity can be computed either using nonequilibrium MD (NEMD) or equilibrium MD (EMD) approach. In NEMD, the thermal conductivity is computed by adding a thermal gradient or a heat flux to the heat carrier; while in EMD one calculates the thermal conductivity from current thermal fluctuation in the system using the so-called Green–Kubo method [57].

Schelling et al. [57] measured thermal conductivities of silicon films using the methods mentioned above. In NEMD, a temperature gradient was added in a single well-defined direction. The potentials (force fields) in silicon were described with Stillinger and Weber model. Thermal conductivities were found to depend on the system size (finite system effect) using either NEMD or EMD method. But reliable results could be derived by extrapolation to infinite system size. It is well known that the thermal conductivity of a thin film is lower than that of a bulk (see Fig. 14), which was further validated by the work of Schelling et al.

The other NEMD method to compute thermal conductivities is by imposing heat flux on the system reversely. MüllerPlathe [58] used this kind of NEMD to measure thermal conductivities. The advantage is relatively easy to implement. Also, the

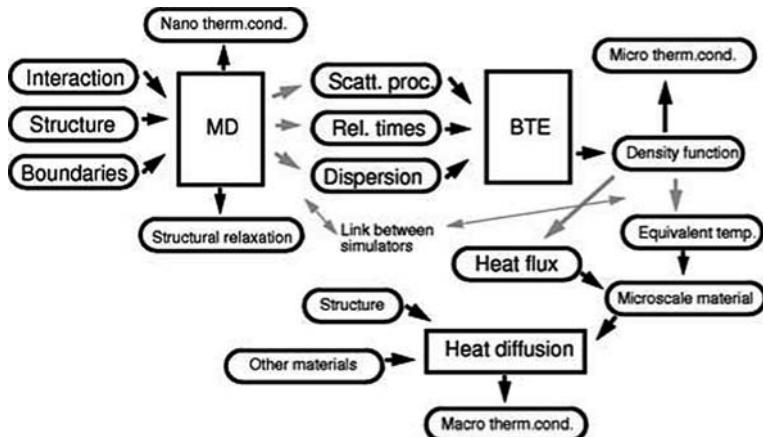
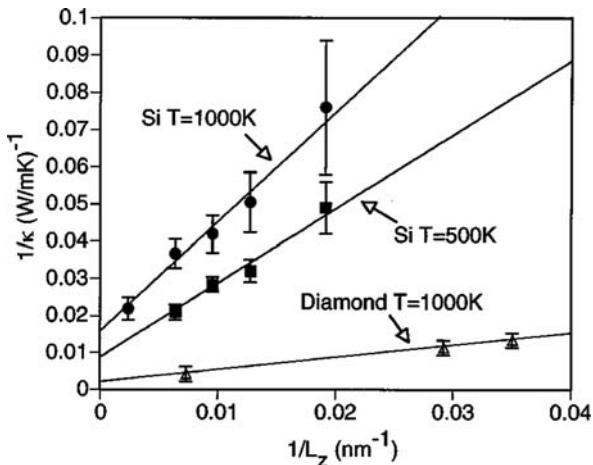


Fig. 13 Input and output quantities required in different models. Reprinted from [56] with permission from Springer Science+Business Media.

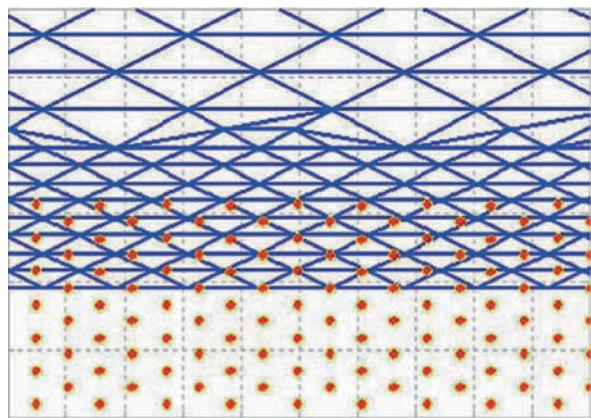
Fig. 14 System size dependence of $1/k$ on $1/L_z$. Reprinted with permission from [57], Copyright 2002 by the American Physical Society.



system was compatible with simulation boxes with or without periodic boundary conditions.

Yang et al. [59] proposed a MD/FEM (finite element method) coupled methodology to investigate heat conductivities through a copper/aluminum interface. The advantage of coupled method is making implication of boundary conditions easier, and this method can extract nanoscale aspects of certain regions in the bulk properties of the whole system, since the domain covered by MD is embedded in the finite element domain that applies to larger elements and larger domains. This approach is particularly effective when a nanoscale second phase (or different material) is embedded in the larger domain. Although this is a powerful technique in linking domains with different length scales, linking the MD and finite element domains is

Fig. 15 The configuration of the coupled region. Reprinted from [59] with permission from Springer Science+Business Media.



one of the challenges in this process. Various approaches have been proposed. As an example, Fig. 15 shows a method where the intermediate region between MD and finite element is covered by both FEM meshes and MD particles, to keep the continuity from the smaller MD regions to the finite element region.

Using such an approach, the thermal conductivity was calculated for a Cu/Al interface and the results are summarized in Fig. 16. An approximately linear relationship between thermal conductivities and thermal flux was found during the primary stage. When the thermal flux achieved certain critical point, thermal

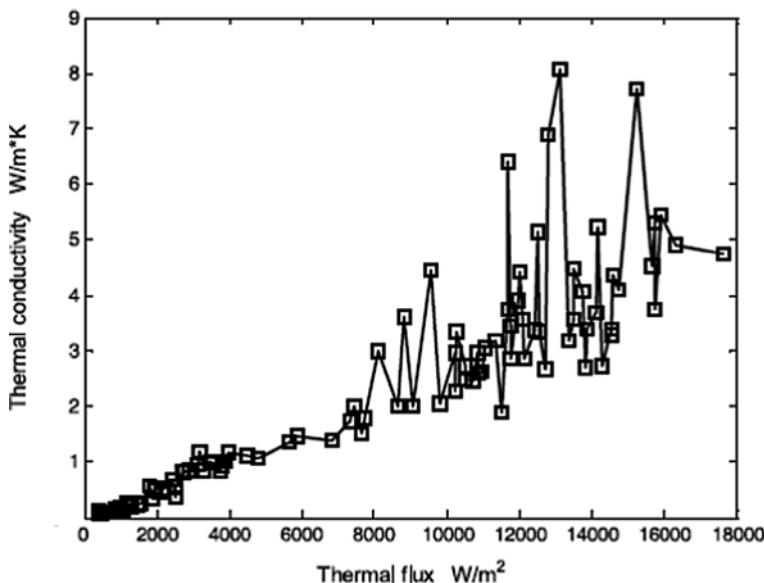


Fig. 16 The relation of thermal conductivity and thermal flux. Reprinted from [59] with permission from Springer Science+Business Media.

conductivities would fluctuate intensely, showing the instability and dependence of thermal conductivity on thermal flux.

7 Applications of Carbon Nanotubes

Carbon nanotubes (CNTs) have been extensively studied [60–71] as a potential alternative/additive of refills in composites used in the electronic industry due to its excellent mechanical strength, heat dissipation and electrical transport, and low thermal expansion. Since theoretically calculated thermal conductivity of multiwall carbon nanotubes (MWNT) at room temperature can be as high as 6600 W/mK [72], researchers are paying more and more attention to promising applications of CNTs as components of thermal interface materials (TIMs). However, the actual thermal conductivity depends on a number of parameters and varies from tens to thousands W/mK. For isolated CNTs, those factors governing thermal conductivity include length of CNTs [66, 67], temperature [60], defects [61], and other factors. For a CNT-based composite, a possible challenge to deal with is the high thermal resistance at the interfaces between fillers and matrix. An experimentally measured thermal conductivity coefficient is often smaller than the value predicted by mixing rules. MD is an effective way to help to understand the enhanced heat dissipation behaviors of TIMs. Algorithms to compute thermal conductivity are described in the previous subtopic. Some results of CNT-based composites obtained from MD simulation are presented here and compared with experimental measurement.

It has been shown that defects in isolated CNTs will scatter phonons (the quantized modes of crystal vibration associated with heat transfer) and decrease intrinsic tube conductivity [60]. As a major type of TIMs, CNT/polymer composites have higher thermal conductivity than polymer bulks, lower elastic modulus, and thermal expansion coefficient than metals. However, the interface between the tubes and soft matrix plays an important role in the overall performance of TIM. Increasing the number of covalent bonding between CNTs and polymer will reduce interface resistance and provide durable adhesion but also induce defects into CNTs, since this is generally done through incorporating functionalized groups to the nanotubes that are capable of forming covalent bonds. MD studies of CNTs with varied degree of chemical functionalization from different research groups validated each other – namely, functionalization of CNTs reduces thermal conductivity and interfacial resistance. For example, Shenogin et al. [71] grafted octane molecules to single-wall carbon nanotubes (SWNT) to investigate dependency of thermal conductivity and the MD simulation results are shown in Fig. 17. It can be seen that CNT with 10% functionalized atoms has a thermal conductivity of about half of that of perfect tubes, while interface resistance is also reduced by a factor of about 2. Similar results are also reported by other researchers [62, 69]. Although the percentages of conductivity and resistance degraded by functionalized atoms vary case by case, the trends drawn from these researches are identical. Hence, the problem is how to find an optimal percentage of functionalized atoms to balance conductivity and resistance.

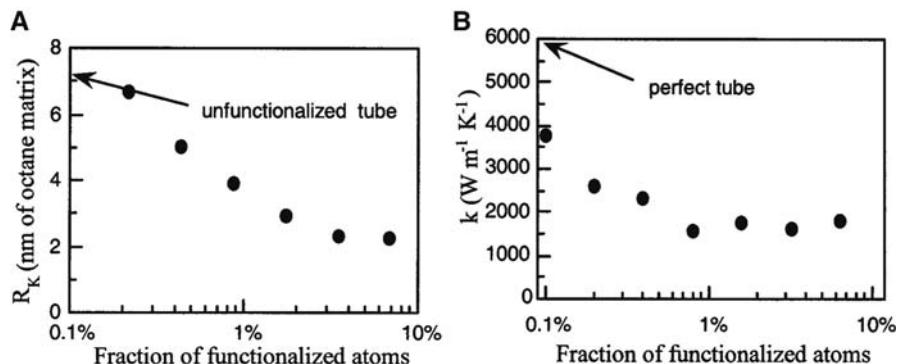


Fig. 17 (a) Interfacial resistance in units of equivalent matrix thickness vs. fraction of tube carbon atoms with covalently attached octane molecules, (b) carbon nanotube conductivity vs. fraction of functionalized tube carbon atoms, obtained from MD simulations. Reprinted with permission from [71], Copyright 2004, American Institute of Physics.

Precise experimental manipulation of arraying or growing CNTs makes it possible to align vertically to the surface. Figure 18 gives an example of a pure CNT TIM with controllable thickness. When a pure CNT TIM was placed between two aluminum surfaces, the CNT TIM had less thermal contact resistance than in the case of direct contact or with graphite coating. Observed from experiments, the thermal conductivity was also enhanced by more than three times [70] with this construction. Under this situation, CNTs are linked to planar surfaces with point contacts. The size effects of the contact points, and many other aspects associated with CNT growth and behavior can be studied by MD.

Saha et al. [73] built two planar silicon substrates with constriction to mimic the point contact situation using MD, where the focus was how the radius of constriction influences thermal resistance with other conditions fixed. Theoretically, the

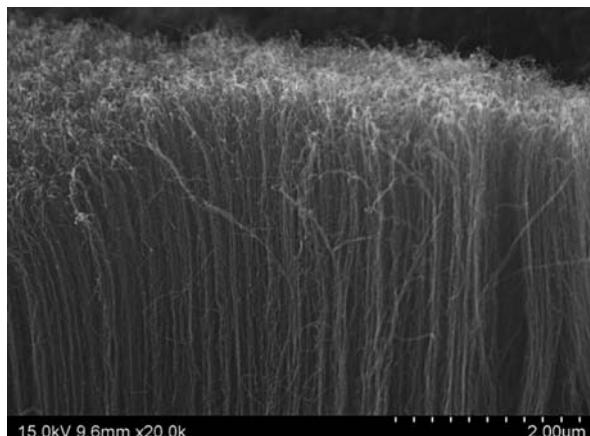


Fig. 18 Scanning electron microscopy image of low-density aligned CNT. Reprinted from [70], Copyright 2007, with permission from Elsevier.

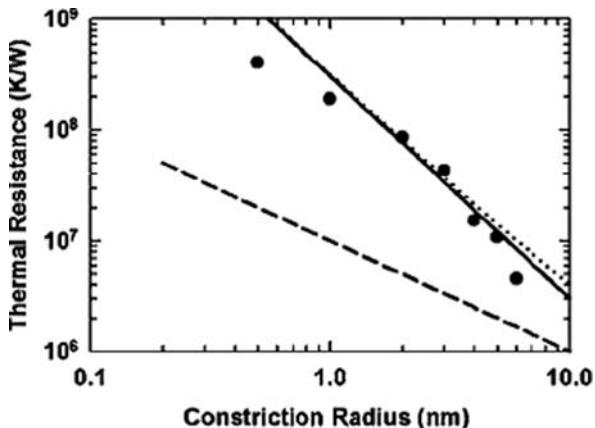
total thermal resistance R is a summation of two parts: the diffusive resistance R_d and the ballistic resistance R_b , which can be calculated based on the following two equations:

$$R_d = 1/(2ka) \quad (21)$$

$$R_b = 4l/(3\pi ka^2) \quad (22)$$

where k is thermal conductivity, a is radius of constriction, l is phonon mean free path (phonon group velocity and phonon mean free path are assumed independent of frequency). The general trend is that a larger radius will lead to a lower resistance (Fig. 19). The pattern for the variation of thermal resistance from the theoretical approach (Equations 21 and 22) is also shown in the figure. MD with appropriate idealizations can be used to study properties such as thermal resistance.

Fig. 19 The MD thermal resistance (filled circles) measured as a function of constriction radius. Also plotted are theoretically calculated curves (solid line: R_b , dashed line: R_d , dotted line: R). Reprinted with permission from [73], Copyright 2007, American Institute of Physics.



8 Summary

In this chapter, we reviewed applications of molecular dynamics techniques in the electronic packaging area. In general, MD has been used to help to understand events at the atomistic length scales, such as atomistic mechanisms of diffusion, adhesion, heat conduction and failure, predict material properties, and design new structures. MD can be applied to almost all types of materials in all states, although simulating certain phenomenon in solid state is more difficult than in the liquid state. It can generally compute both statistical and dynamic properties, and explain physical and mechanical behaviors. As long as initial configurations are correctly generated and proper force fields are selected, MD will give fairly good results compared to experiments. However, the interpretation of MD output to obtain relevant parameters has to be done properly; otherwise, the results may not be useful. In general, at least useful qualitative patterns can be obtained and in many problems quantitative information can be extracted by the judicious application of MD.

The advantage of MD is that it is less expensive and much more convenient than experiments, if the objective is materials design. When the scale of interest comes to nanoscale, MD can explore interactions at atomistic level, which is not easily achieved through experiments. MD can also be employed to measure parameters associated with a phenomenon, which cannot easily be characterized experimentally. However, MD does have some shortcomings. First of all, it is a molecular modeling method that cannot deal with mesoscopic and macroscopic simulations. Second, generally chemical reactions cannot be modeled by classical MD, although some aspects of it can be done through specific modifications. Third, results of MD rely on the accuracy of the force fields and initial configurations. Currently, researchers try to combine MD with other simulation methods (such as finite element methods, quantum mechanics simulation) to overcome some of the drawbacks of the classical MD, although that could remove some of the difficulties of classical MD, but it can create its own difficulties. Also, significant effort has been devoted to increase the efficiency and effectiveness of MD algorithms.

The main application of MD methods in “nano packaging,” where there are nanoscale components, is to give another tool to understand conformational aspects, properties, and in aiding the design of new materials such as insulators, conductors, adhesives, and solders. Currently, MD simulations are also routinely applied to study nanoparticle/polymer composites [74], lead-free solder [75], nanometals [76, 77], that are particularly relevant to the packaging area. In summary, applications of MD techniques in the packaging area are many and growing. We have provided only an outline of the applications of MD in this area due to space limitations, and the literature contains vast amount of information about the procedure as well as applications of MD in the packaging area.

Acknowledgement We gratefully acknowledge the financial support from NASA Langley Research Center.

References

1. Alder B.J. and Wainwright T.E., Studies in molecular dynamics.1. General method. *Journal of Chemical Physics* 1959, 31 (2), 459–466.
2. Frenkel D. and Smit B., *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. Academic Press: San Diego, 2002; p xxii, 638 p.
3. Haile J.M., *Molecular Dynamics Simulation: Elementary Methods*. Wiley: New York, 1992; p xvii, 489 p.
4. Rapaport D.C., *The Art of Molecular Dynamics Simulation*, 2nd ed. Cambridge University Press: Cambridge; New York, 2004; p xiii, 549 p.
5. Sadus R.J., *Molecular Simulation of Fluids: Theory, Algorithms, and Object-Orientation*, 1st ed. Elsevier: Amsterdam; New York, 1999; p xxvii, 523 p.
6. Ponder J.W. and Case D.A., Force fields for protein simulations. *Protein Simulations* 2003, 66, 27–86.
7. MacKerell A.D., Bashford D., Bellott M., Dunbrack R.L., Evanseck J.D., Field M.J., Fischer S., Gao J., Guo H., Ha S., Joseph-McCarthy D., Kuchnir L., Kuczera K., Lau F.T.K., Mattos C., Michnick S., Ngo T., Nguyen D.T., Prodhom B., Reiher W.E., Roux B., Schlenkrich M., Smith J.C., Stote R., Straub J., Watanabe M., Wiorkiewicz-Kuczera J., Yin D., and Karplus M.,

- All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B* 1998, **102** (18), 3586–3616.
- 8. Daw M.S., Foiles S.M., and Baskes M.I., The embedded-atom method – a review of theory and applications. *Materials Science Reports* 1993, **9** (7–8), 251–310.
 - 9. Sun H., COMPASS: an ab initio force-field optimized for condensed-phase applications – overview with details on alkane and benzene compounds. *Journal of Physical Chemistry B* 1998, **102** (38), 7338–7364.
 - 10. Niketic S.R. and Rasmussen K., *The consistent force field: a documentation*. Springer-Verlag: Berlin; New York, 1977; p ix, 212 p.
 - 11. Hamerton I., Heald C.R., and Howlin B.J., Molecular modelling of the physical and mechanical properties of two polycyanurate network polymers. *Journal of Materials Chemistry* 1996, **6** (3), 311–314.
 - 12. Leung Y.K. and Eichinger B.E., Computer-simulation of end-linked elastomers. 1. Trifunctional networks cured in the bulk. *Journal of Chemical Physics* 1984, **80** (8), 3877–3884.
 - 13. Cheng K.C. and Chiu W.Y., Monte-Carlo simulation of polymer network formation with complex chemical-reaction mechanism – kinetic approach on curing of epoxides with amines. *Macromolecules* 1994, **27** (12), 3406–3414.
 - 14. Mayo S.L., Olafson B.D., and Goddard W.A., Dreiding – a generic force-field for molecular simulations. *Journal of Physical Chemistry* 1990, **94** (26), 8897–8909.
 - 15. Yarovsky I. and Evans E., Atomistic simulation of the sol formation during synthesis of organic/inorganic hybrid materials. *Molecular Simulation* 2002, **28** (10–11), 993–1004.
 - 16. Wu C.F. and Xu W.J., Atomistic molecular modelling of crosslinked epoxy resin. *Polymer* 2006, **47** (16), 6004–6009.
 - 17. Fan H.B. and Yuen M.M.F., Material properties of the cross-linked epoxy resin compound predicted by molecular dynamics simulation. *Polymer* 2007, **48** (7), 2174–2178.
 - 18. <http://accelrys.com/products/materials-studio/>.
 - 19. Komarov P.V., Chiu Y.T., Chen S.M., Khalatur P.G., and Reineker P., Highly cross-linked epoxy resins: an atomistic molecular dynamics simulation combined with a mapping/reverse mapping procedure. *Macromolecules* 2007, **40** (22), 8104–8113.
 - 20. Varshney V., Patnaik S.S., Roy A.K., and Farmer B.L., A molecular dynamics study of epoxy-based networks: cross-linking procedure and prediction of molecular and material properties. *Macromolecules* 2008, **41** (18), 6837–6842.
 - 21. Brown D., and Clarke J.H.R., A loose-Coupling, Constant-pressure, molecular-dynamics algorithm for use in the modeling of polymer materials. *Computer Physics Communications* 1991, **62** (2–3), 360–369.
 - 22. Parrinello M. and Rahman A., Polymorphic transitions in single-crystals – a new molecular-dynamics method. *Journal of Applied Physics* 1981, **52** (12), 7182–7190.
 - 23. Stevens M.J., Interfacial fracture between highly cross-linked polymer networks and a solid surface: effect of interfacial bond density. *Macromolecules* 2001, **34** (8), 2710–2718.
 - 24. Stevens M.J., Manipulating connectivity to control fracture in network polymer adhesives. *Macromolecules* 2001, **34** (5), 1411–1415.
 - 25. Tsige M., Lorenz C.D., and Stevens M.J., Role of network connectivity on the mechanical properties of highly cross-linked polymers. *Macromolecules* 2004, **37** (22), 8466–8472.
 - 26. Tsige M. and Stevens M.J., Effect of cross-linker functionality on the adhesion of highly cross-linked polymer networks: a molecular dynamics study of epoxies. *Macromolecules* 2004, **37** (2), 630–637.
 - 27. Drozdov A.D., Christiansen J.D., Gupta R.K., and Shah A.P., Model for anomalous moisture diffusion through a polymer-clay nanocomposite. *Journal of Polymer Science Part B, Polymer Physics* 2003, **41** (5), 476–492.
 - 28. Marsh L.L., Lasky R., Seraphim D.P., and Springer G.S., Moisture solubility and diffusion in epoxy and epoxy-glass composites. *IBM Journal of Research and Development* 1984, **28** (6), 655–661.

29. Yu Y.T. and Pochiraju K., Three-dimensional simulation of moisture diffusion in polymer composite materials. *Polymer-Plastics Technology and Engineering* 2003, 42 (5), 737–756.
30. Fan H.B., Chan E.K.L., Wong C.K.Y., and Yuen M.M.F., Investigation of moisture diffusion in electronic packages by molecular dynamics simulation. *Journal of Adhesion Science and Technology* 2006, 20 (16), 1937–1947.
31. Fukuda M. and Kuwajima S., Molecular-dynamics simulation of moisture diffusion in polyethylene beyond 10 ns duration. *Journal of Chemical Physics* 1997, 107 (6), 2149–2159.
32. Hofmann D., Fritz L., Ulbrich J., Schepers C., and Bohning M., Detailed-atomistic molecular modeling of small molecule diffusion and solution processes in polymeric membrane materials. *Macromolecular Theory and Simulations* 2000, 9 (6), 293–327.
33. Lin Y.C. and Chen X., Investigation of moisture diffusion in epoxy system: experiments and molecular dynamics simulations. *Chemical Physics Letters* 2005, 412 (4–6), 322–326.
34. Soles C.L. and Yee A.F., A discussion of the molecular mechanisms of moisture transport in epoxy resins. *Journal of Polymer Science Part B Polymer Physics* 2000, 38 (5), 792–802.
35. Tamai Y., Tanaka H., and Nakanishi K., Molecular simulation of permeation of small penetrants through membranes. 1. Diffusion-coefficients. *Macromolecules* 1994, 27 (16), 4498–4508.
36. Vanlandingham M.R., Eduljee R.F., and Gillespie J.W., Moisture diffusion in epoxy systems. *Journal of Applied Polymer Science* 1999, 71 (5), 787–798.
37. Zannideffarge M.P. and Shanahan M.E.R., Diffusion of water into an epoxy adhesive – comparison between bulk behavior and adhesive joints. *International Journal of Adhesion and Adhesives* 1995, 15 (3), 137–142.
38. Dermitzaki E., Wunderle B., Bauer J., Walter H., and Michel B., Structure property correlation of epoxy resins under the influence of moisture and comparison of diffusion coefficient with MD-simulations. In *9th Int. Conf. on Thermal, Mechanical and Multiphysics Simulation and Experiments in Micro-Electronics and Micro-Systems EuroSimE*, 2008.
39. Comyn J., Groves C.L., and Saville R.W., Durability in high humidity of glass-to-lead alloy joints bonded with an epoxide adhesive. *International Journal of Adhesion and Adhesives* 1994, 14 (1), 15–20.
40. Dong H., Fan L.H., Moon K.S., Wong C.P., and Baskes M.I., MEAM molecular dynamics study of lead free solder for electronic packaging applications. *Modelling and Simulation in Materials Science and Engineering* 2005, 13 (8), 1279–1290.
41. Dong H., Moon K.S., and Wong C.P., Molecular dynamics study on the coalescence of Cu nanoparticles and their deposition on the Cu substrate. *Journal of Electronic Materials* 2004, 33 (11), 1326–1330.
42. Dong H., Moon K.S., and Wong C.P., Molecular dynamics study of nanosilver particles for low-temperature lead-free interconnect applications. *Journal of Electronic Materials* 2005, 34 (1), 40–45.
43. Dong H., Zhang Z.Q., and Wong C.P., Molecular dynamics study of a nano-particle joint for potential lead-free anisotropic conductive adhesives applications. *Journal of Adhesion Science and Technology* 2005, 19 (2), 87–94.
44. Dong H., Fan L., Moon K., and Wong C.P., Molecular dynamics simulation of lead free solder for low temperature reflow applications. In *55th Electronic Components and Technology Conference*, 2005; pp 983–987.
45. Dong H., Moon K., and Wong C.P., Molecular dynamics study on coalescence of silver (Ag) nanoparticles and their deposition on gold (Au) substrates. In *9th International Symposium on Advanced Packaging Materials: Processes, Properties and Interfaces*, 2004; pp 152–157.
46. Wang W., Ding Y., and Wang C., Molecular dynamics (MD) simulation of uniaxial tension of β -Sn single crystals with nanocracks. In *6th International Conference on Electronic Packaging Technology*, 2005.
47. Iwamoto N. and Pedigo J., Property trend analysis and simulations of adhesive formulation effects in the microelectronics packaging industry using molecular modeling. In *48th IEEE Electronic Components and Technology Conference*, USA, 1998; pp 1241–1246.

48. Iwamoto N. Applying polymer process studies using molecular modeling. In *4th International Conference on Adhesive Joining and Coating Technology in Electronics Manufacturing*, 2000; pp 182–187.
49. Iwamoto N. Advancing materials using interfacial process and reliability simulations on the molecular level. *International Symposium on Advanced Packaging Materials: Processes, Properties and Interfaces*, 2000; pp 14–17.
50. Mustoe G.G.W., Nakagawa M., Lin X., and Iwamoto N. Simulation of particle compaction for conductive adhesives using discrete element modeling. In *49th Electronic Components and Technology Conference*, 1999; pp 353–359.
51. Su Y.Y. and Shemenski R.M., The role of oxide structure on copper wire to the rubber adhesion. *Applied Surface Science* 2000, *161* (3–4), 355–364.
52. Chan E.K.L., Fan H., and Yuen M.M.F., Effect of interfacial adhesion of copper/epoxy under different moisture level. In *7th International Conference on Thermal, Mechanical and Multiphysics Simulation and Experiments in Micro-Electronics and Micro-Systems, EuroSimE*, 2006.
53. Fan H.B., Chan E.K.L., Wong C.K.Y., and Yuen M.M.F., Molecular dynamics simulation of thermal cycling test in electronic packaging. *Journal of Electronic Packaging* 2007, *129* (1), 35–40.
54. Wong C.K.Y., Fan H.B., and Yuen M.M.F., Interfacial adhesion study for SAM induced covalent bonded copper-EMC interface by molecular dynamics simulation. *IEEE Transactions on Components and Packaging Technologies* 2008, *31* (2), 297–308.
55. Wong C.K., Fan H., and Yuen M.M.F., Investigation of adhesion properties of Cu-EMC interface by molecular dynamic simulation. In *6th International Conference on Thermo, Mechanical and Multiphysics Simulation and Experiments in Micro-Electronics and Micro-Systems, EuroSimE*, 2005.
56. Heino P., Simulations of nanoscale thermal conduction. *Microsystem Technologies Micro- and Nanosystems-Information Storage and Processing Systems* 2009, *15* (1), 75–81.
57. Schelling P.K., Philpot S.R., and Kebinski P., Comparison of atomic-level simulation methods for computing thermal conductivity. *Physical Review B* 2002, *65* (14), 144306.
58. MullerPlathe F., A simple nonequilibrium molecular dynamics method for calculating the thermal conductivity. *Journal of Chemical Physics* 1997, *106* (14), 6082–6085.
59. Yang P. and Liao N.B., Research on characteristics of interfacial heat transport between two kinds of materials using a mixed MD-FE model. *Applied Physics A, Materials Science and Processing* 2008, *92* (2), 329–335.
60. Bi K.D., Chen Y.F., Yang J.K., Wang Y.J., and Chen M.H., Molecular dynamics simulation of thermal conductivity of single-wall carbon nanotubes. *Physics Letters A* 2006, *350* (1–2), 150–153.
61. Che J.W., Cagin T., and Goddard W.A., Thermal conductivity of carbon nanotubes. *Nanotechnology* 2000, *11* (2), 65–69.
62. Fan H.B., Zhang K., and Yuen M.M.F., Thermal performance of carbon nanotube-based composites investigated by molecular dynamics simulation. In *57th Electronic Components and Technology Conference*, Reno, NV, 2007; pp 269–272.
63. Grujicic M., Cao G., and Gersten B., Atomic-scale computations of the lattice contribution to thermal conductivity of single-walled carbon nanotubes. *Materials Science and Engineering B, Solid State Materials for Advanced Technology* 2004, *107* (2), 204–216.
64. Lopez M.J., Rubio A., and Alonso J.A., Deformations and thermal stability of carbon nanotube ropes. *IEEE Transactions on Nanotechnology* 2004, *3* (2), 230–236.
65. Ma P.C., Tang B.Z., and Kim J.K., Effect of CNT decoration with silver nanoparticles on electrical conductivity of CNT-polymer composites. *Carbon* 2008, *46* (11), 1497–1505.
66. Maruyama S., A molecular dynamics simulation of heat conduction of a finite length single-walled carbon nanotube. *Microscale Thermophysical Engineering* 2003, *7* (1), 41–50.
67. Mingo N. and Broido D.A., Length dependence of carbon nanotube thermal conductivity and the “problem of long waves”. *Nano Letters* 2005, *5* (7), 1221–1225.

68. Ngo Q., Cruden B.A., Cassell A.M., Sims G., Meyyappan M., Li J., and Yang C.Y., Thermal interface properties of Cu-filled vertically aligned carbon nanofiber arrays. *Nano Letters* 2004, 4 (12), 2403–2407.
69. Padgett C.W. and Brenner D.W., Influence of chemisorption on the thermal conductivity of single-wall carbon nanotubes. *Nano Letters* 2004, 4 (6), 1051–1053.
70. Shaikh S., Lafdi K., and Silverman E., The effect of a CNT interface on the thermal resistance of contacting surfaces. *Carbon* 2007, 45 (4), 695–703.
71. Shenogin S., Bodapati A., Xue L., Ozisik R., and Kebelinski P., Effect of chemical functionalization on thermal transport of carbon nanotube composites. *Applied Physics Letters* 2004, 85 (12), 2229–2231.
72. Berber S., Kwon Y.K., and Tomanek D., Unusually high thermal conductivity of carbon nanotubes. *Physical Review Letters* 2000, 84 (20), 4613–4616.
73. Saha S.K. and Shi L., Molecular dynamics simulation of thermal transport at a nanometer scale constriction in silicon. *Journal of Applied Physics* 2007, 101 (7), 074304-1–074304-7.
74. Starr F.W., Schroder T.B., and Glotzer S.C., Molecular dynamics simulation of a polymer melt with a nanoscopic particle. *Macromolecules* 2002, 35 (11), 4481–4492.
75. Sellers M.S., Schultz A.J., Kofke D.A., and Basaran C., Molecular dynamics modeling of grain boundary diffusion in Sn-Ag-Cu solder. In *AIChe Annual Meeting*, Salt Lake City, Utah, 2007.
76. Heino P. and Ristolainen E., Molecular dynamics study of thermally induced shear strain in nanoscale copper. *IEEE Transactions on Advanced Packaging* 1999, 22 (3), 510–514.
77. Liu H.H., Jiang E.Y., Bai H.L., Wu P., Li Z.Q., and Sun C.Q., The kinetics and modes of gold nanowire breaking. *Journal of Computational and Theoretical Nanoscience* 2008, 5 (7), 1450–1453.

Nanoscale Deformation and Strain Analysis by AFM/DIC Technique

Y.F. Sun and John H.L. Pang

Abstract Digital image correlation (DIC) technique with the aid of scanning probe microscopes has become a very promising tool for deformation analysis of micro- and nanoscale components. The scanner drift of the atomic force microscope (AFM) is a great disadvantage to the application of digital image correlation to micro/nanoscale deformation measurements. This chapter has addressed the image distortion induced by the scanner drifts and developed a method to reconstruct AFM images for the successful use of AFM image correlation. The proposed AFM/DIC method is to generate a corrected image from two correlated AFM images scanned at the angle of 0° and 90°, respectively. The method has been validated by two simple AFM/DIC experiments. The application of this AFM/DIC technique was demonstrated by the deformation measurement on the micro-interconnection in a micro thermoelectric cooler. AFM images of the scan region of interest were obtained separately when the microelectronic device was before and after operating at both its cooling and heating stages. The AFM images were then used to obtain the in-plane deformation fields in the observed region of the micro-assembly. AFM image correlation is performed for nanoscale deformation analysis using the authors' AFM-DIC program. The results show that the observed region was subjected to cyclic strains when the device worked between its cooling and heating stages, and cyclic strain in the vertical direction was found to be significant deformation mode. The thermally induced deformation behavior of the micro-assembly device was modeled by finite element analysis (FEA). Both thermal-electric analysis and thermal stress analysis were conducted on a 3D finite element model of the device. It is shown that the experimental results were able to validate the finite element analysis results.

Keywords Atomic force microscope · AFM · Digital image correlation · DIC · Nanoscale

Y.F. Sun (✉)

School of Mechanical & Aerospace Engineering, Nanyang Technological University, Singapore, Singapore

e-mail: sunyaofeng@pmail.ntu.edu.sg

1 Introduction

Digital image correlation (DIC) deformation measurements employ optical image correlation and deformation analysis and provide a noncontact, full-field displacement measurement technique. The digital image correlation (DIC) technique can be employed with different digital imaging instruments for macro/micro/nanodeformation measurements. The deformation measurement at the submicron scale is a challenging need for micro-to-nanoscale deformation analysis in characterizing packaging and interconnection materials.

The DIC method was first developed by Peters et al. [1–3]. The DIC technique has great capabilities in adapting to digital images obtained from different imaging tools. For example, combined with optical microscope, digital image correlation can be used for deformation measurements at micrometer scale. With the aid of scanning electron microscope (SEM)/scanning probe microscope (SPM), the submicron- and nanoscale resolution of deformation measurements by DIC can be applied if distortions in SEM [4–6] or scan drifts in SPM [7, 8] could be calibrated and corrected.

With the shrinking size of the packages and increasing number of components integrated into the packages, more challenges are attributed to experimental deformation/strain analysis for the design and reliability study of such packages. There is a potential to employ digital image correlation technique to handle AFM (a typical scanning probe microscope) images in expectation of the nanoscale deformation measurement with the aid of high spatial resolution of AFM. There are few successful nanoscale applications [9, 10] of digital image correlation with the aid of atomic force microscope (AFM). The failures in the use of AFM/DIC technique result from the AFM scanner drift effect. For example, a drift of 1 nm when scanning a length of 100 nm causes a false strain of 1%, which can flood the actual deformation to be measured. In order to apply AFM/DIC technique successfully for deformation measurements, a method to compensate for the AFM scanner drifts is needed [11, 12].

The spurious distortion of AFM images results not only from the nonlinear motions of the AFM piezoelectric scanner when responding to the applied voltage, but also from its hysteresis and creep effects, even from the complex cross coupling between x , y , and z motions. The nonlinearity corrections of AFMs have been investigated in some studies [4–8] and proved to be effective for the compensation of scanner's nonlinearity. The nonlinearity [13] of AFMs can be reduced within 10% by software solutions and to less than 1% by hardware solutions with external sensors [14–16] to positioning the scanner. A 1% deviation may be unnoticed or can be ignored by a common user who examines the topograph of micro/nanostructures under AFM. However, the deviation must be taken into account for the application of digital image correlation to the accurate deformation measurement. In fact, the scanner drifts can easily be observed by digital image correlation of AFM images acquired at the same load condition.

The digital image correlation (DIC) is applied to obtain the displacements of the center point of a subimage on the reference image taken before deformation,

when the subimage matches another subimage on the deformed image taken after deformation. The match of a subimage is to be found by maximizing the cross correlation coefficient between intensity patterns of two subimages [11]:

$$C(\mathbf{p}) = \frac{\sum_{y=y_0-N/2}^{y_0+N/2} \sum_{x=x_0-M/2}^{x_0+M/2} g(x,y)h(x + u(x,y,\mathbf{p}), y + v(x,y,\mathbf{p}))}{\sqrt{\sum_{y=y_0-N/2}^{y_0+N/2} \sum_{x=x_0-M/2}^{x_0+M/2} g^2(x,y) \sum_{y=y_0-N/2}^{y_0+N/2} \sum_{x=x_0-M/2}^{x_0+M/2} h^2(x + u(x,y,\mathbf{p}), y + v(x,y,\mathbf{p}))}} \quad (1)$$

where the coordinate (x_0, y_0) represents the center point of a subimage with its size $M \times N$; $g(x, y)$ is the light intensity at the point (x, y) in the reference image; $h(x + u(x, y, \mathbf{p}), y + v(x, y, \mathbf{p}))$ is the light intensity in the deformed image at the match position of the point (x, y) . The intensity at a sub-pixel position is obtained by using bicubic spline interpolation scheme: $I(x,y) = \sum_{n=0}^3 \sum_{m=0}^3 \alpha_{mn} x^m y^n$. The coefficients α_{mn} are determined from the intensity values of pixel locations. $u(x, y, \mathbf{p})$ and $v(x, y, \mathbf{p})$ are displacement functions which define how each of the subimage points located in the reference image at (x, y) is mapped to the deformed image. First-order displacement functions are often used:

$$\begin{aligned} u(x,y,\mathbf{p}) &= p_0 + p_2(x - x_0) + p_4(y - y_0) \\ v(x,y,\mathbf{p}) &= p_1 + p_3(x - x_0) + p_5(y - y_0) \end{aligned} \quad (2)$$

The parameter vector $\mathbf{p}(p_0, p_1, p_2, p_3, p_4, p_5)$ can be explicitly denoted $\mathbf{p}(u(x_0, y_0), v(x_0, y_0), \frac{\partial u}{\partial x}|_{x_0, y_0}, \frac{\partial v}{\partial x}|_{x_0, y_0}, \frac{\partial u}{\partial y}|_{x_0, y_0}, \frac{\partial v}{\partial y}|_{x_0, y_0})$ (where $u(x_0, y_0)$ and $v(x_0, y_0)$ are the displacements of the subimage center in x and y directions respectively). To find the resultant vector \mathbf{p} that maximizes the correlation coefficient (Eq. (1)), the Newton–Raphson method is used to iteratively find the solution with a given approximation \mathbf{p}_{old} to the solution. The next improved approximation \mathbf{p}_{new} can be obtained using the following Newton–Raphson iteration formula:

$$\nabla \nabla C(\mathbf{p}_{\text{old}})(\mathbf{p}_{\text{new}} - \mathbf{p}_{\text{old}}) = -\nabla C(\mathbf{p}_{\text{old}}) \quad (3)$$

Newton–Raphson method is known for its fast convergence capability. Usually, two or three iterations are enough to solve the parameter vector \mathbf{p} , which represents the displacements and displacement gradients at the subimage center. It is practical that the displacement gradients above are used for the direct calculation of strain components in uniform deformation measurements instead of additional operations of displacement field smoothing and derivative calculation.

In what follows, Image correlation results use u to represent the displacement in x /horizontal direction of images, v in y /vertical direction of images.

Figure 1 shows the displacement field obtained from the correlation of two AFM images of size $512 \times 512 \text{ pixel}^2$ or $2 \times 2 \mu\text{m}^2$ from two continuous scans without

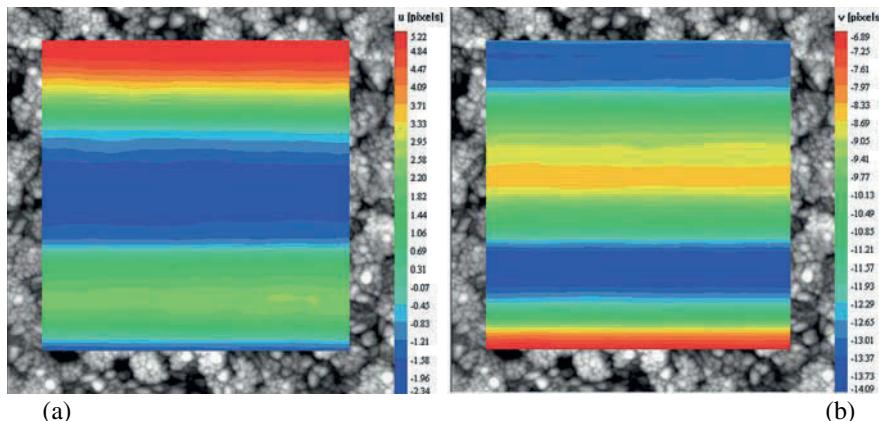


Fig. 1 Artificial displacement fields between two AFM images scanned at the same condition: (a) u field, (b) v field

any load. Although the drift effect or the poor reproducibility has also been mentioned in literature, the method to compensate the AFM nonlinearity and thus derive the accurate displacement field has not been given out.

2 Atomic Force Microscope (AFM)

The AFM is a typical member in the scanning probe microscope family, the powerful modern research tools to investigate the morphology and the local properties of the sample surface from the atomic to the micron level. The experimental system (see Fig. 2) studied in this work is a Digital Instrument-3000 AFM. A sharp tip, typically less than $5 \mu\text{m}$ and often less than 10 nm in diameter at the apex, is located at the free end of a cantilever that is $100\text{--}500 \mu\text{m}$ long.

The interactive forces between the tip and the sample surface cause the cantilever to deflect. A position-sensitive photo-detector (see Fig. 3) measures the cantilever deflections as the sample is scanned under the tip. The measured cantilever deflections allow a computer to generate a map of surface topography.

The interactive force primarily contributed to the deflection of an AFM cantilever is the van der Waals force. The relationship between the van der Waals force and the tip-to-sample separation [13] are shown in Fig. 4. Depending on the distance from the sample surface that the probe tip is held, an AFM can be operated at contact, non-contact, and intermittent contact modes. The selection of AFM modes is dependent on specific applications. Detailed information is available on the operation manual from AFM vendors.

Scanning a sample surface in an AFM is performed by the piezoelectric scanner (Fig. 5). The scanner acts as a fine positioning stage to move the sample under the probe tip and thus makes a high spatial resolution of AFM possible. Similar to

Fig. 2 Schematic of the experimental AFM system

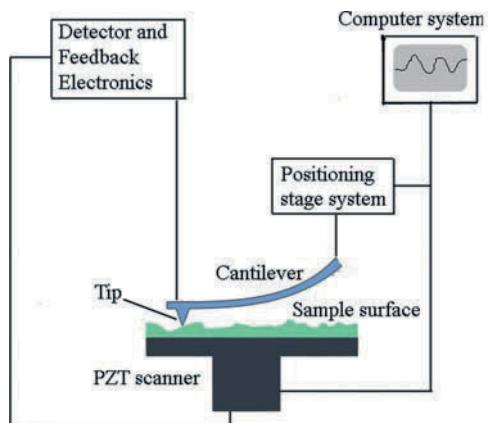


Fig. 3 Position-sensitive photo-detector system in AFM

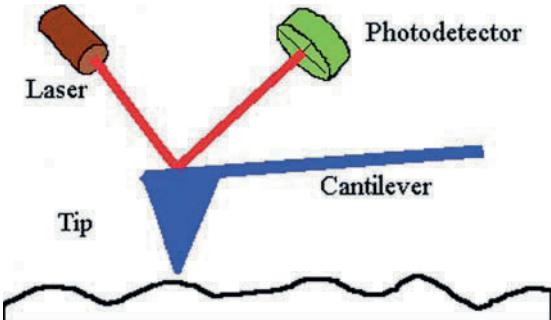


Fig. 4 Curve of van der Waals force vs. distance and AFM modes

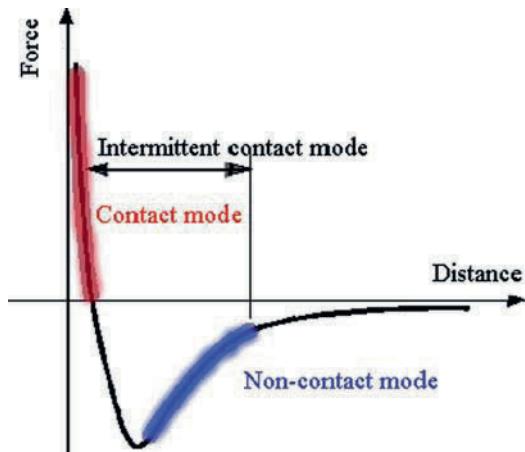


Fig. 5 Tube-shaped piezosscanner in AFM.
Voltage applied to the x - and y -axis produces the bend of the tube. Voltage applied to the z -axis produces the extension of the tube

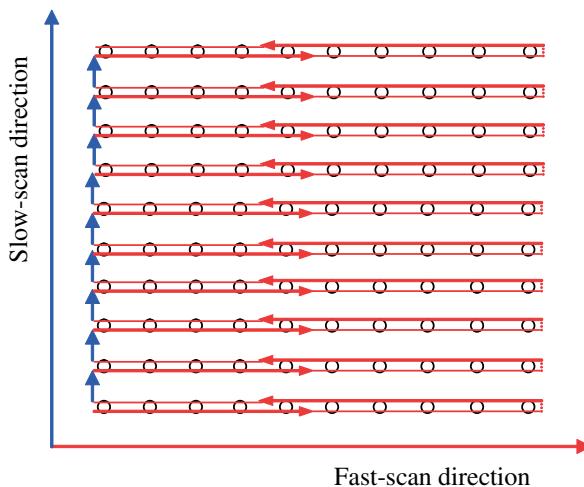
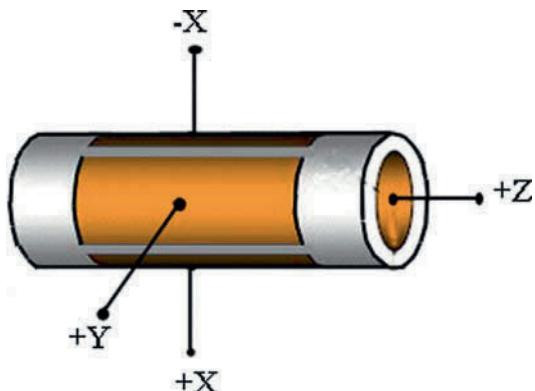


Fig. 6 Schematic description of scanner motion and point sampling

moving an electronic beam on the screen in the cathode ray tube of a television, the AFM control unit drives the scanner in a type of raster pattern as shown in Fig. 6. The tip goes along the first line in forward (fast-scan direction), and then back, then moves across the next line (slow-scan direction perpendicular to the fast scan direction), and so forth. During the scanning, data can be collected separately as the probe moves from left to right (the “trace”) and from right to left (the “retrace”). The data are sampled digitally at equally spaced intervals scheduled in the fast-scan direction and the slow-scan direction. The distance between the sample points is called the step size. The step size is calculated as the full-scan size divided by the number of sample points per line. As a rule, the number of scan lines equals the number

of sample points per line. Thus, the data set from a square grid of measurements, called the AFM frame, consists of a square matrix of data elements, whose size is typically 256×256 or 512×512 . The data elements are the cantilever deflection that reflects the surface topography of the scanned area.

The AFM frame can be visualized on a computer as a three-dimensional topography or two-dimensional color/grayscale image (Fig. 7a, b), where the frame data elements fill image pixels.

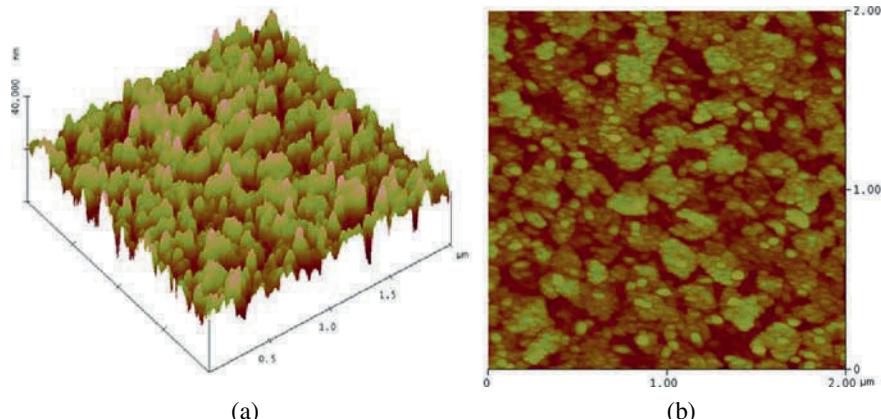


Fig. 7 AFM frame visualized as (a) a 3D topography or (b) 2D color/grayscale image

3 AFM Scanner Imperfections

While an ideally square grid of measurements is scheduled for the scanner, virtually there are a lot of difficulties in achieving perfect tip-over-sample motions in the plane of the scanned sample surface, as well as in the out-of-plane direction. Besides external influences, such as electronic noises, vibrations, acoustic noises, and instability of tip-sample contact, the imperfections of a piezoelectric scanner itself cause the distortions in the AFM images. Here scanner imperfections, including intrinsic nonlinearity, hysteresis, creep, and cross coupling, are discussed in the following sections.

Figure 8 shows the intrinsic nonlinearity and hysteretic behavior of a piezoelectric scanner. The intrinsic nonlinearity describes a nonlinear extension responding to the applied voltage, and the hysteretic behavior of a piezosensor displays differences in forward and reverse motions. To make things worse, the nonlinearity and hysteresis are also dependent on the range and even the rate of change of applied voltage. The intrinsic nonlinearity causes the distortion of the measurement grid shown in Fig. 8.

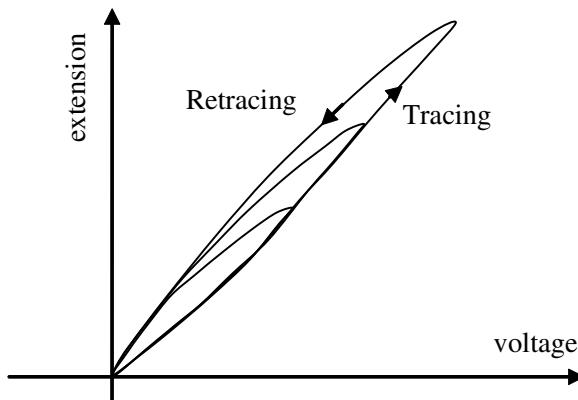


Fig. 8 Intrinsic nonlinearity and hysteresis curves of a piezosscanner for various peak voltages or scan sizes. The hysteresis loop is related to the scan size

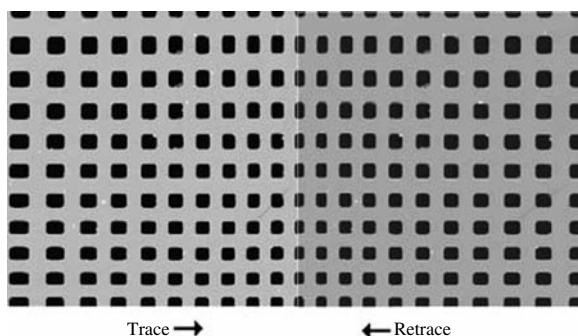


Fig. 9 $100 \times 100 \mu\text{m}$ scans in the forward (trace) and reverse (retrace) directions of a two-dimensional $10 \mu\text{m}$ pitch calibration grating without linearity correction [17]

The sample points are not equally spaced when the scanner does not move linearly with applied voltage. The hysteretic behavior causes the forward and reverse scan directions to behave differently. A surface with periodic structures, such as gratings, is generally used to calibrate an AFM. AFM images of such a structure are shown in Fig. 9. Both scans are in the down direction. Notice the differences in the spacing, size, and shape of the pits between the bottom and the top of each image. The effect of the hysteresis loop on each scan direction in xy plane is demonstrated.

As for this nonlinear relationship, the correction solution in the AFM is the use of a calibration routine by applying a nonlinear voltage (see Fig. 10) in real time to produce a linear scan along x - and y -axis in both tracing and retracing scan directions. The feature distortion in Fig. 9 can be reduced or eliminated by the application of such a solution, as shown in Fig. 11a.

A quantitative investigation of the effect of hysteresis correction can be done by correlating two AFM images scanned in the tracing and retracing directions.

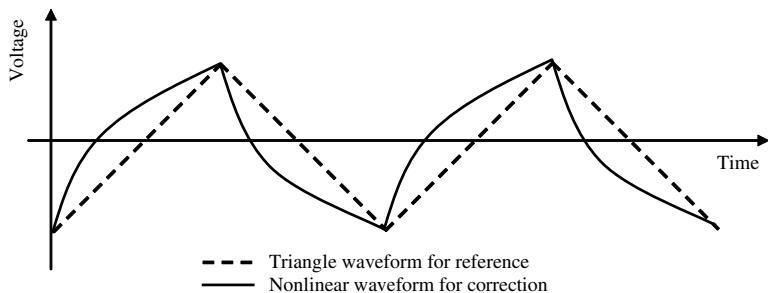


Fig. 10 Nonlinear voltage waveform applied to the piezo electrodes to produce linear scanner movement

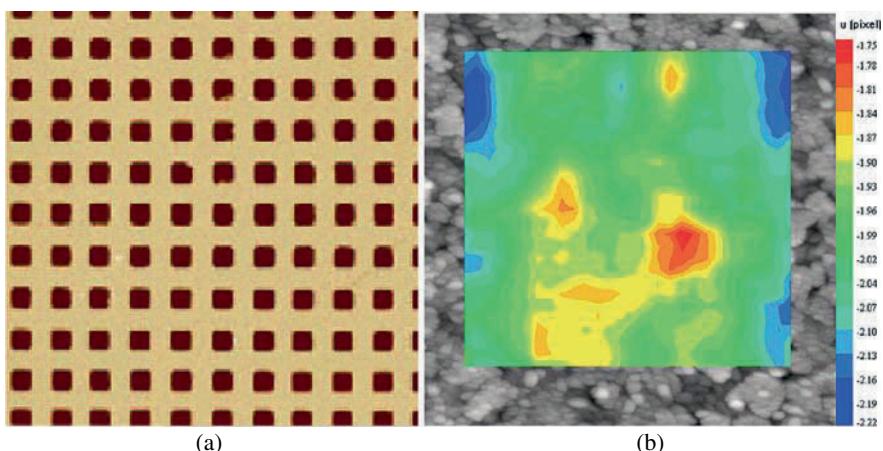


Fig. 11 Effect of hysteresis correction: (a) $100 \times 100 \mu\text{m}$ scans of the same two-dimensional $10 \mu\text{m}$ pitch calibration grating with linearity correction, (b) correlation between the images in the tracing and retracing directions

Figure 11b shows the displacement field in the tracing/retracing direction from the correlation of such a pair of images. The result indicates good linearity of scanner motions along fast-scan direction. However, in fact the procedure does not work well for the linearity correction in slow-scan direction, which can be proved by the results (Fig. 1) from correlating two “trace” images of the same area. The almost horizontal contour bands shown in Fig. 1b show large artificial deformation in slow-scan/vertical direction, but the contour bands of the same type in Fig. 1a denote tiny artificial deformation within a fast-scan/horizontal line. The effectiveness disparity of nonlinearity correction between fast-scan and slow-scan directions ascribed to the disparity of scanner motions in both directions, which is shown in Fig. 6.

Another drawback of a piezosscanner is the so-called creep or drift, a delay in the response to sudden change of the applied voltage. The schematic diagram of applied voltage change versus the scanner motions in xy plane is shown in Fig. 12.

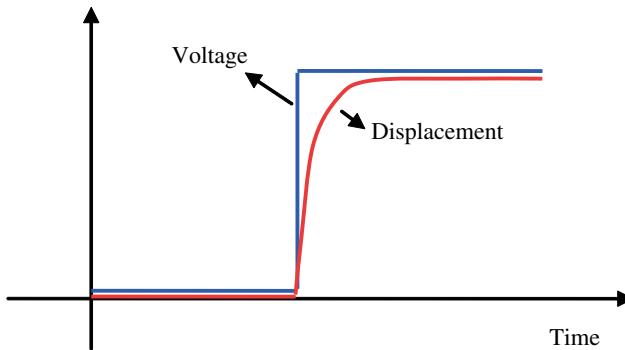


Fig. 12 Schematic time diagram of applied voltage change (blue line) vs. the scanner motion along x-axis (red line)

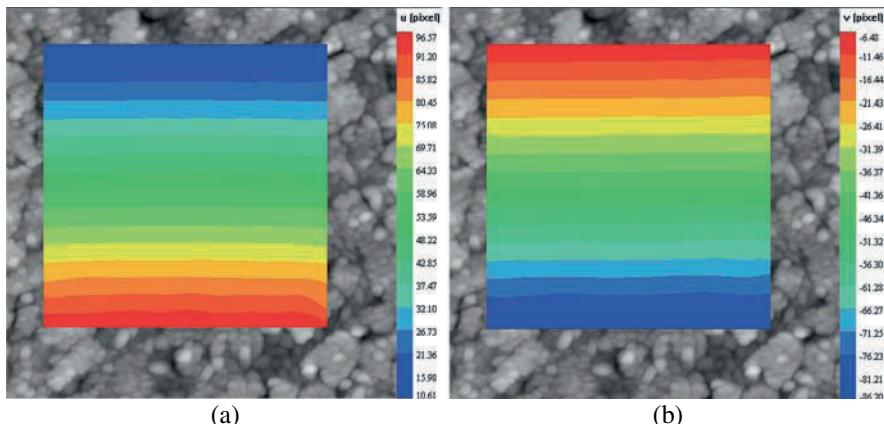
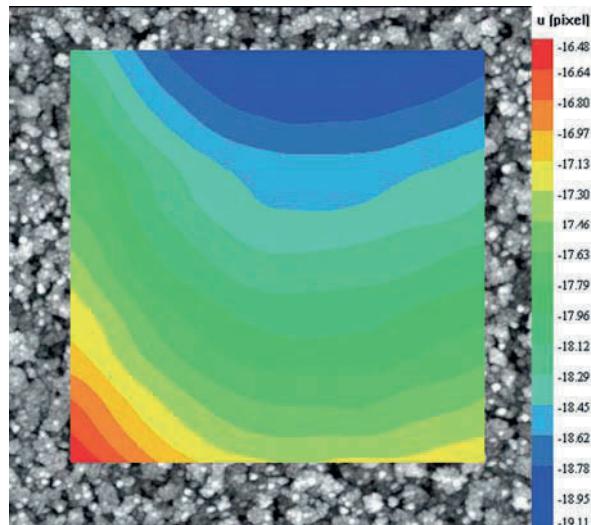


Fig. 13 Artificial displacement fields in AFM images due to creep: (a) u field in the fast-scan direction, (b) v field in the slow-scan direction

Strong influence [18] of the creep occurs on the initial stage of the scanning process (see Fig. 13), when making a change in x and y offsets (see Fig. 14), or when using the Frame Up and Frame Down commands to bring along abrupt voltage change applied to the scanner.

Last but not the least, cross coupling exerts influence during AFM scanning. The cross coupling refers to the interaction between x -, y -, and z -axis scanner movements. The basis of geometric cross coupling is that the scanner is constructed by combining operated piezo electrodes for x , y , and z into a single tube, and the movements of the scanner in xy plane are achieved by bending the tube at the free end. The influences of cross coupling increase with scan size and surface roughness. Higher surface smoothness can reduce the effect of the cross coupling on in-plane deformation measurements by AFM image correlation.

Fig. 14 Effect of creep from 300 nm offsets in x - and y -axis



4 Reconstruction of AFM Images

Considering that Figs. 1 and 13 show the accuracy of the scan along fast-scan/horizontal direction is much better than that in slow-scan/vertical direction, it occurred to the author that the AFM image would have better quality if the image data in two directions were both from the accurate scan along fast-scan direction, and such an image can be produced by combining two AFM images of the same area scanned at the angle of 0° and 90°, respectively. Exactly speaking, new image can be generated from the reconstruction of the image scanned at the angle of 0° whose distortions are determined by digital image correlation with the image scanned at the angle of 90°. The detailed procedure for AFM image reconstruction is introduced as follows:

- (1) Acquire one AFM image at the scan angle of zero degree at some load condition.
- (2) Acquire the other AFM image at the scan angle of 90° at the same load condition, and rotate the image 90° clockwise.
- (3) With the image from step 1 as the reference image, and the image from step 2 as the deformed image, correlate the two AFM images by digital image correlation technique and obtain the displacement field.
- (4) Determine the spurious displacements in the reference image based on the displacement field obtained from step 3.
- (5) Reconstruct the reference AFM image according to the spurious displacements calculated from step 4.

In the procedure, steps 4 and 5 are the core of the method and are illustrated in the following paragraphs.

Considering that the u - and v -displacement contour bands are both in almost horizontal and parallel pattern in Figs. 1 and 13, it is reasonable to assign common u - and v -displacements to all points on a same horizontal/scan line. With this understanding, it can be imagined that the normal grid (Fig. 15a) scanned horizontally will become distorted as schematically shown in Fig. 15b; likewise, the normal grid scanned vertically will become distorted as schematically shown in Fig. 15c. When Figs. 15a, b overlap, the offsets of horizontal scan lines will be exposed as shown in Fig. 15d, where u_j^r and v_j^r denote the horizontal and vertical offsets of horizontal scan line (row) j scanned in the reference image from its ideal position. Likewise, the overlapping of Fig. 15a, c will show up the offsets of vertical scan lines, as shown in Fig. 15e, where u_i^d and v_i^d denote the horizontal and vertical offsets of vertical scan line (column) i scanned in the deformed image from its ideal position. Now if a point is located both on the horizontal line j and the vertical line i in the normal grid, the point will have the position $(i + u_j^r j + v_j^r)$ in the reference image, and have the position $(i + u_i^d j + v_i^d)$ in the deformed image. If $U_{x,y}$ and $V_{x,y}$ denote the horizontal and vertical displacements between a point at the location (x, y) in the reference image and its match point in the deformed image, the displacements of the point at $(i + u_j^r j + v_j^r)$ in the reference image can be denoted as $U_{i+u_j^r j+v_j^r}$ in horizontal direction and $V_{i+u_j^r j+v_j^r}$ in vertical direction, and expressed as:

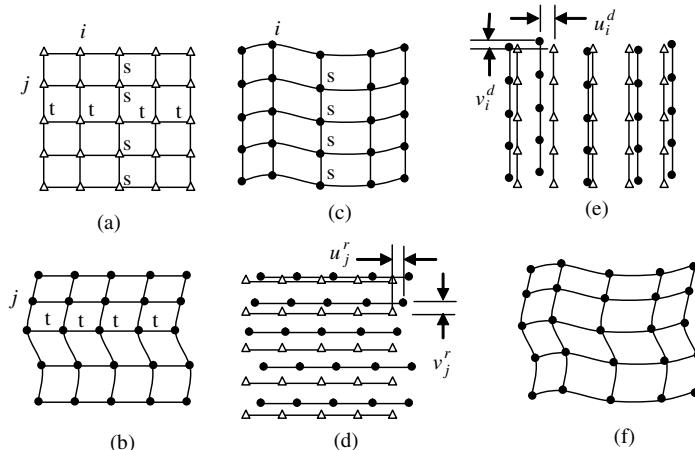


Fig. 15 Schematic diagram for AFM image reconstruction: (a) normal grid to be scanned by the AFM, the grid with row spacing s and column spacing t , j as row index, i as column index, (b) distorted grid scanned at the scan angle of 0° , and used as reference image in DIC, (c) distorted grid scanned at the scan angle of 90° , used as deformed image in DIC, (d) overlapping (a) and (b) shows displacements between corresponding rows, (e) overlapping (a) and (c) shows displacements between corresponding columns, (f) correlation between the two distorted grids (b) and (c)

$$\begin{aligned} U_{i+u_j^r, j+v_j^r} &= (i + u_i^d) - (j + v_j^r) = u_i^d - u_j^r \\ V_{i+u_j^r, j+v_j^r} &= (j + v_i^d) - (j + v_j^r) = v_i^d - v_j^r \end{aligned} \quad (4)$$

If the number of horizontal lines and of vertical lines are n ($j = 1, 2, \dots, n$) and m ($i = 1, 2, \dots, m$) respectively, the number of the unknowns (terms on the right-hand side of Eq. (4)) is $2n + 2m$. Compared with the number of the knowns (terms on the left-hand side of Eq. (4)) mn , we know that there is redundancy for the solution of Eq. (4). The redundancy can be used to reduce the deviation effect of measured displacements by image correlation. Consider random errors that can usually be minimized with the use of the average approach. Sums of the displacements of all m points on scan line j lead to:

$$\begin{aligned} \sum_{i=1}^m U_{i+u_j^r, j+v_j^r} &= \sum_{i=1}^m u_i^d - mu_j^r \\ \sum_{i=1}^m V_{i+u_j^r, j+v_j^r} &= \sum_{i=1}^m v_i^d - mv_j^r \end{aligned} \quad (5)$$

Equation (5) can be rewritten as:

$$\begin{aligned} u_j^r &= -\frac{1}{m} \sum_{i=1}^m U_{i+u_j^r, j+v_j^r} + \frac{1}{m} \sum_{i=1}^m u_i^d \\ v_j^r &= -\frac{1}{m} \sum_{i=1}^m V_{i+u_j^r, j+v_j^r} + \frac{1}{m} \sum_{i=1}^m v_i^d \end{aligned} \quad (6)$$

The subtraction of displacements between two horizontal lines $j + k$ (k is an integer) and j will cancel out the unknowns u_i^d and v_i^d , and is shown as follows:

$$\begin{aligned} u_{j+k}^r - u_j^r &= -\frac{1}{m} \sum_{i=1}^m (U_{i+u_{j+k}^r, j+k+v_{j+k}^r} - U_{i+u_j^r, j+v_j^r}) \\ v_{j+k}^r - v_j^r &= -\frac{1}{m} \sum_{i=1}^m (V_{i+u_{j+k}^r, j+k+v_{j+k}^r} - V_{i+u_j^r, j+v_j^r}) \end{aligned} \quad (7)$$

To reconstruct the reference image, a reference/anchor line is needed. Provided that the central horizontal line jc is selected as the reference line, then $u_{jc}^r = 0$ and $v_{jc}^r = 0$. The displacements of horizontal line $jc + 1$ in the reference image can be rewritten based on Eq. (7):

$$\begin{aligned} u_{jc+1}^r &= u_{jc+1}^r - u_{jc}^r = -\frac{1}{m} \sum_{i=1}^m (U_{i+u_{jc+1}^r, jc+1+v_{jc+1}^r} - U_{i,jc}) \\ v_{jc+1}^r &= v_{jc+1}^r - v_{jc}^r = -\frac{1}{m} \sum_{i=1}^m (V_{i+u_{jc+1}^r, jc+1+v_{jc+1}^r} - V_{i,jc}) \end{aligned} \quad (8)$$

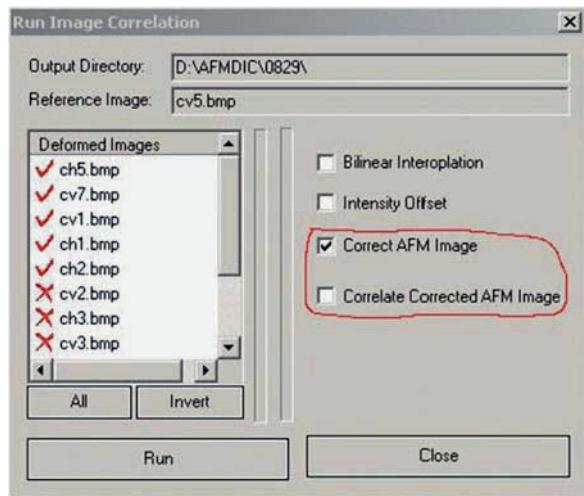
u_{jc+1}^r and v_{jc+1}^r in Eq. (8) can be solved by iteration technique by taking the initial guesses of them as zeros. The solutions are found when the iteration converges. The convergence is reached when the u_{jc+1}^r and v_{jc+1}^r differences between the current and the last estimates of them respectively are both lower than their convergence limits, 0.01 pixel. Likewise, we can obtain u_{jc-1}^r and v_{jc-1}^r , u_{jc+2}^r and v_{jc+2}^r , and so on. When the displacements of all horizontal lines scanned in the reference image are determined, the reference image can be reconstructed to remove its distortion. The reconstruction of the reference image is done based on the following expression:

$$\bar{f}(i,j) = f\left(i + u_j^r, j + v_j^r\right) \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n.) \quad (9)$$

where \bar{f} denotes the reconstructed image profile, and f denotes the reference image profile.

In the implementation of the above algorithms for reconstruction of AFM images, the spurious displacements at a grid of control points located on the image are obtained from DIC. The values of the rest points are calculated using interpolation schemes, such as bilinear interpolation, which is to obtain the values of a point by linear interpolation between its four nearest neighbors. As shown in Fig. 16, this correction function has been added in the original DIC software.

Fig. 16 Software implementation for AFM image correction



5 Verification and Buckling Test

Figure 17 shows the displacement field between two AFM images scanned at the angles of 0° and 90°, respectively but under the same conditions. Either the displacement u field or v field displays the same pattern along a scan line, which agrees well with the supposed (see Fig. 15d).

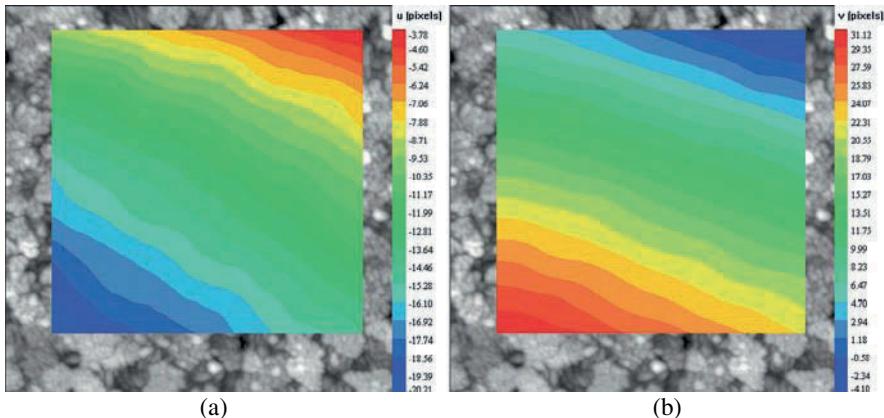


Fig. 17 Displacement (3.906 nm/pixel) fields between two AFM images scanned at 0° and 90°: (a) u field, (b) v field

The second verification case is zero deformation measurement. Two images reconstructed from two pairs of AFM images scanned at the angle of 0° and 90° are correlated using digital image correlation technique. The displacement field between one pair of images scanned at the angle of 0° and 90° is shown in Fig. 17. The displacement field between both images scanned at the angle of 0° is shown in Fig. 18a, b, and the displacement field between two reconstructed images in Fig. 18c, d. The sample ranges of displacements u and v are 65.7 and 68.8 nm before image correction, 2.1 and 2.7 nm after image correction. In this instance, the maximum image distortion was decreased from 4.3% to 0.16% (correlated image area is $1.6 \times 1.6 \mu\text{m}$). Therefore, the correction method to AFM images is effective to cancel out the distortion due to the AFM scanner drifts. In addition, a comparison of measured strains before and after image correction has been listed in Table 1. After image correction, the average remaining strains, ε_x , ε_y , and ε_{xy} , are 76.6, -124, and $-2.83 \mu\varepsilon$, and their standard deviations 0.00148, 0.00180, and 0.00124 respectively, which shows the promise of accurate deformation measurements by AFM image correlation.

A buckling test was conducted as an example of deformation measurements by AFM image correlation. The buckling specimen (see Fig. 19) is a thin PCB (printed circuit board) with cross copper lines located in the centerlines of the plate. The size of the specimen is $56 \times 48 \times 0.42 \text{ mm}$. The center area of the plate to be scanned was sputtered with nanoparticles to obtain AFM images with high contrast pattern. Nanoindentation marks of $1 \mu\text{m}$ size were shaped for the assistance in the positioning of scan areas between two loading states. A simple fixture was designed and the schematic diagram of loading fixture was shown in Fig. 20. The fixture was required to be aligned with the AFM scan direction and be tightly fixed on the AFM stage. The specimen was first scanned at the lengthwise buckling state because of the ease to unload the specimen buckling during the test. An area of $10 \times 10 \mu\text{m}^2$ in the center of the plate surface was selected to be scanned considering that the

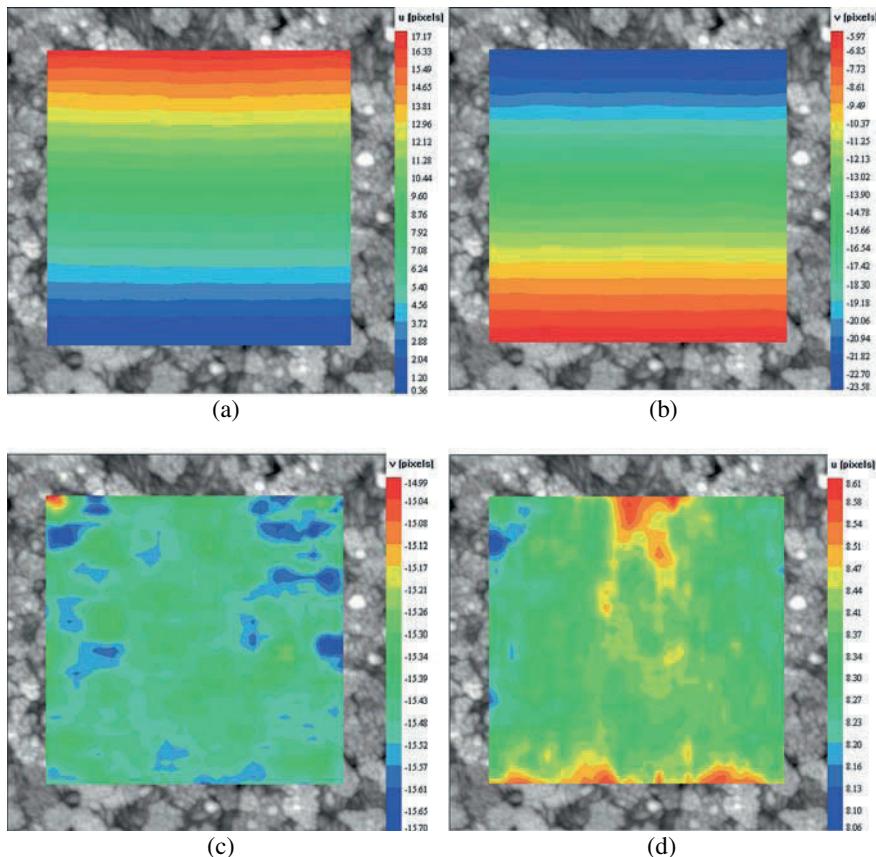


Fig. 18 u and v Displacement (3.906 nm/pixel) fields in the zero deformation test: (a–b) u , v fields before image correction, (c–d) u , v fields after image correction

Table 1 Remaining strains in the zero deformation test

Strains	Before correction			After correction		
	ε_x	ε_y	ε_{xy}	ε_x	ε_y	ε_{xy}
mean(ε)	/	/	/	7.66e-5	-1.24e-4	-2.83e-6
min(ε)	-3.78e-3	3.01e-2	-3.18e-2	-4.92e-3	-7.69e-3	-4.25e-3
max(ε)	5.33e-3	5.48e-2	-1.16e-2	5.50e-3	6.78e-3	3.97e-3
$\sigma(\varepsilon)$	/	/	/	1.48e-3	1.80e-3	1.24e-3

smaller area will increase the difficulty in the search for the same area to be scanned at two separate scans. It is noted that the operation of probe offset was not used in this test because false variational displacements along fast scan axis have been found between AFM images scanned with probe offset and without probe offset. This nonlinearity of AFM scan has been calibrated and shown earlier.

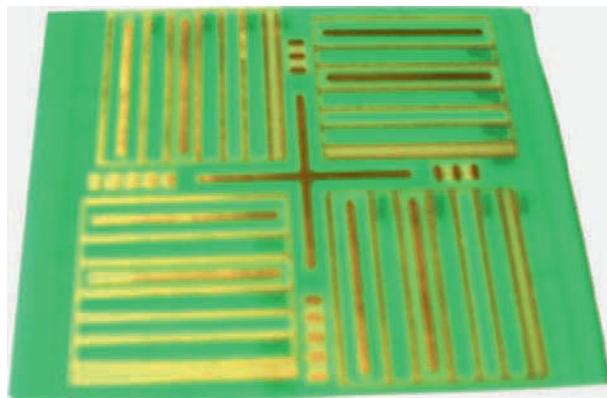
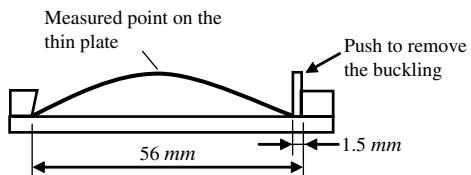


Fig. 19 Thin plate (printed circuit board) for the buckling test

Fig. 20 Schematic diagram of the fixture for the buckling test



The area of interest was scanned by AFM images at the scan angle of 0° and 90° when the specimen was pushed 1.5 mm at one end to buckling deformation and was then unloaded. Acquired images were reconstructed by following the procedure mentioned earlier and digital image correlation was conducted on the reconstructed images. The displacements u and v in the direction of specimen length and width have been shown in Fig. 21a, b.

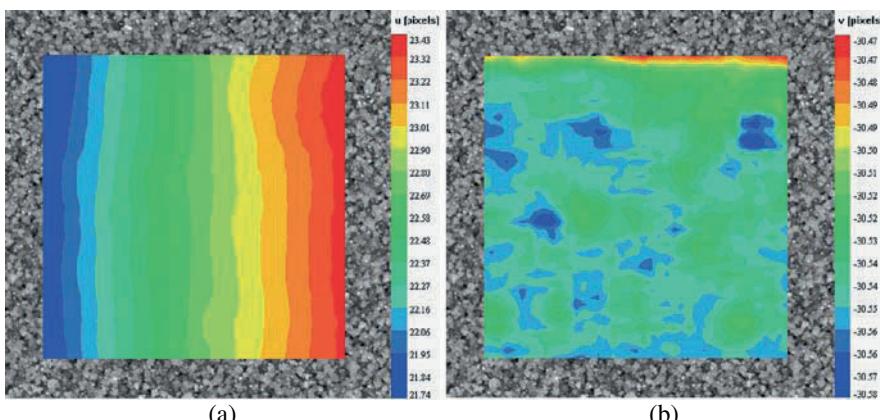


Fig. 21 u (a) and v (b) displacement (19.53 nm/pixel) fields in the buckling test

Table 2 Strains measured by DIC and from FEA σ

Strains		ε_x	ε_y	ε_{xy}
Measured by DIC	mean(ε)	3.96e-3	1.36e-5	2.35e-5
	$\sigma(\varepsilon)$	6.96e-4	3.91e-4	3.58e-4
		3.88e-3	6.59e-8	2.20e-9

The tensile deformation shown in Fig. 21a is almost uniform. Finite element analysis (FEA) of the buckling test using ABAQUS shows that the strains in the center area of the buckling plate are uniform. Measured strains by digital image correlation and calculated strains by FEA are listed in Table 2. The results show the average strain in x direction measured by DIC agrees well with that calculated by FEA. The standard deviations of measured strains in Table 2 are less than those in Table 1. The difference can be attributed to the different scan sizes in two tests. Figure 21b shows a displacement deviation of 0.11 pixels or 2.2 nm which is close to that in the zero deformation test. The minimum strain that can be identified in this instance is estimated at hundreds of micro-strain.

6 AFM-DIC Experiments

It is a challenge to conduct thermo-mechanical tests under AFM. Thermal loading can be exerted on a specimen readily by conducting a thermo-mechanical test. The issues of thermal drift of AFM scanner must be resolved and this led to the design of the thermo-mechanical experiments on a micro thermoelectric cooler under AFM.

A micro thermoelectric cooler (TEC) was used in this work. A micro TEC is a tiny solid-state heat pump, which is a critical component in many applications of integrated electronic packaging assemblies as it provides precise cooling and temperature stabilization. A typical TEC module comprises of two ceramic substrates that serve as a foundation and electrical insulation for p-type and n-type bismuth telluride (Bi_2Te_3) pellets that are connected electrically in series and thermally in parallel between the substrates (Fig. 22). Metallized pads are attached to the ceramic substrates to maintain the electrical connections inside the module. Solder is typically used at the connection joints to enhance the electrical connections and hold the module together. The prototype of the micro TECs with outer dimension of $L3.8 \times W3.0 \times H1.0$ mm is shown in Fig. 23. The dimension details of the micro TEC are listed in Table 3.

After the specimen check, the hot junction side of the good TEC module was attached to a copper heat sink and thermal grease was filled between them (see Fig. 24). Two electric wires were soldered to each terminal of the micro TEC, and then connected to a direct-current power supply. Before the deformation measurement by digital image correlation with the aid of AFM, a calibration check of the assembly performance was done to obtain a configuration of input current, the surface temperature or temperature difference for use in AFM-DIC experiments. For

Fig. 22 Structure of solid-state thermoelectric module

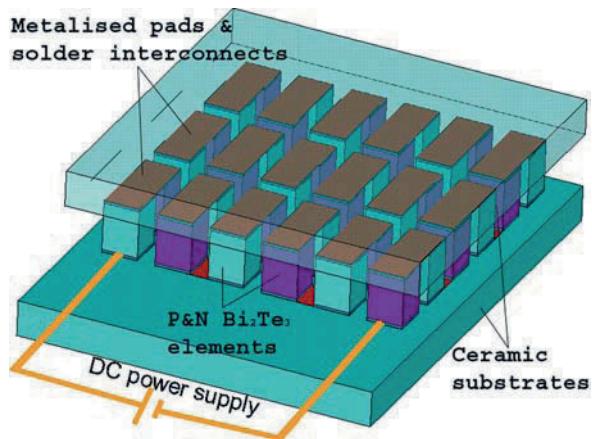


Fig. 23 Prototype of micro thermoelectric cooler



Table 3 Dimensions of the micro TEC module

Components	Dimensions
Top ceramic substrate	3.0 × 3.0 × 0.28 mm
Bottom ceramic substrate	3.8 × 3.0 × 0.28 mm
TE elements	0.3 × 0.3 × 0.4 mm
TE element pitch	0.45 × 0.45 mm
Solder interconnections	0.3 × 0.3 × 0.02 mm

the calibration, a fine thermal couple was attached to the top surface of the TEC module. An input current of 0.8 A and a temperature difference of 50°C under an ambient temperature of 25°C were determined for the following AFM–DIC experiments. In other words, in our experiments, the device working at its cooling stage was subjected to a temperature gradient of 25 to -25°C from the bottom substrate

to the top substrate, while working at its heating stage was subjected to a reverse temperature gradient of 25–75°C from the bottom substrate to the top substrate. The multimode AFM (Fig. 25) of Dimension 3000 produced by Digital Instrument was used in this work. The assembly was then fixed on the stage of AFM with the region of interest facing the AFM scanner tip. The electric wires of the TEC module was fastened on the stage by the tape, squeezed tightly by the AFM shelter, and connected to a DC power supply.

The scanning area (shown in Fig. 24) was located at the right bottom corner of the solder interconnection, which is connected to the bottom substrate. During the experiments, the bottom substrate maintained at the temperature of 25°C, and so the scanning area remained at a temperature of nearly 25°C, which allowed neglecting the thermal effect on the AFM scanner. The tapping mode of AFM was used to scan the specimen. All acquired AFM images were $5 \times 5 \mu\text{m}$ in size with 512 pixel resolution along each scan line. In the present test method, the topography of the area of interest was recorded via AFM at both the horizontal scan and vertical scan directions. The two AFM images were used to reconstruct the topography of the scanned area in order to reduce the drift of AFM scanning. The detail of AFM image correction [11] for deformation measurements was presented in the previous section.

Once the scan parameters of AFM scan were determined, the region of interest in the solder interconnection of the sample was positioned and focused under the AFM. Without the current input, the $5 \times 5 \mu\text{m}$ region of interest was initially scanned in horizontal and vertical directions with the tapping mode, and thus a pair of AFM images before deformation was captured and later correlated to obtain a corrected AFM image as the reference image for DIC analysis. Next, a constant current of 0.8 A was applied to the micro TEC, which was thus operating at its cooling stage. When the steady-state condition was reached after approximately 2 min, the region

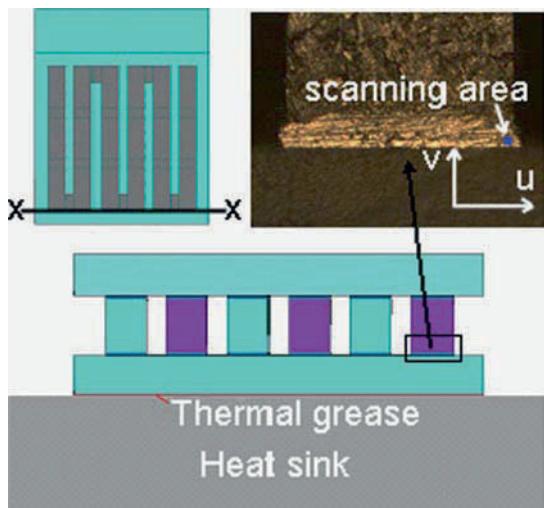
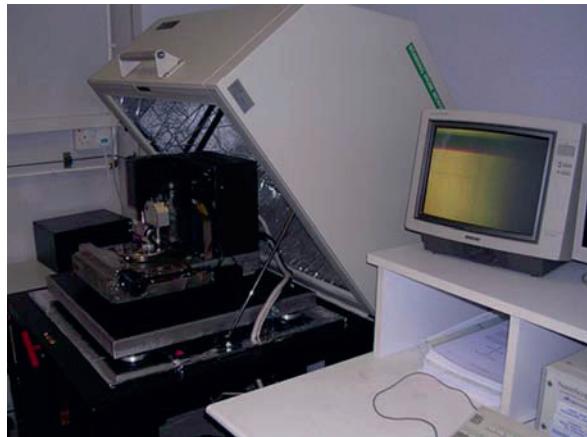


Fig. 24 Sample preparation and experimental test assembly

Fig. 25 AFM used for nanoscale deformation measurement



of interest was scanned again with the same AFM settings and a pair of AFM images after deformation was captured. The pair of AFM images was later correlated to obtain a corrected AFM image, which carried the deformation information due to the TEC's operating at its cooling stage. After that, the input current polarity was changed and so the micro TEC was operating at its heating stage. In other words, the hot side and the cool side of the TEC changed. Under this condition, the region of interest was scanned twice in the horizontal and vertical directions to obtain another pair of AFM images, which carried the deformation information due to the TEC's operating at its heating stage. A typical AFM image of the scanned region captured in this test is shown in Fig. 26.

The AFM images captured before and after deformation were correlated to obtain the deformation using the developed AFM–DIC software developed. The resultant displacement fields due to the TEC's operating at its cooling stage and at its heating stage are shown in Figs. 27 and 28, respectively, where u and v represent the displacement fields in x /horizontal and y /vertical direction of images, respectively (see Fig. 24). Figure 27 shows the range of u -displacements, which is 3.29 nm, is less than that of v -displacements, which is 5.83 nm, and the u -displacement contour bands lean to the right while the v -displacement contour bands are almost horizontal.

Thus, it can be estimated that the average strain in vertical direction $\bar{\varepsilon}_y$ is several times larger than the average horizontal strain component $\bar{\varepsilon}_x$ and the shear strain component $\bar{\varepsilon}_{xy}$. Similarly, Fig. 28 shows the range of u -displacements, which is 4.07 nm, is less than that of v -displacements, which is 7.12 nm, and the u -displacement contour bands incline from the vertical more than the v -displacement contour bands incline from the horizontal.

The heterogeneity of the displacement fields shown in Figs. 27 and 28 did not completely result from the actual deformation. The displacement errors from AFM–DIC analysis can contribute much to the heterogeneous displacement fields when the actual deformation is small. The gradients of strains on the small scanned area

Fig. 26 AFM image of $5 \times 5 \mu\text{m}$ size captured for image correlation analysis

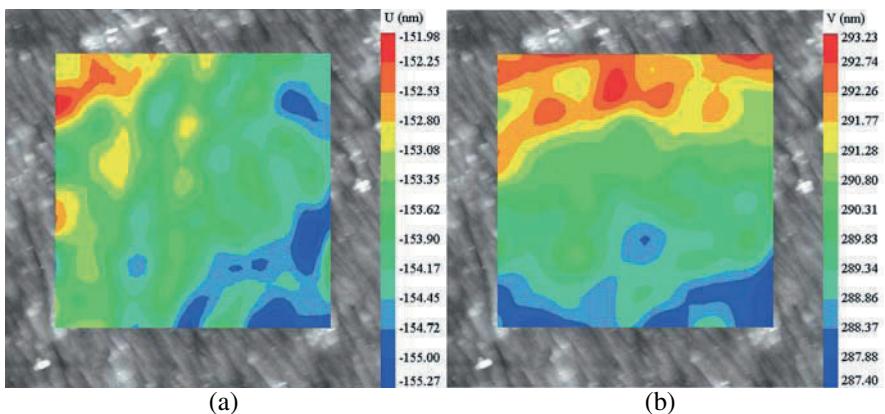
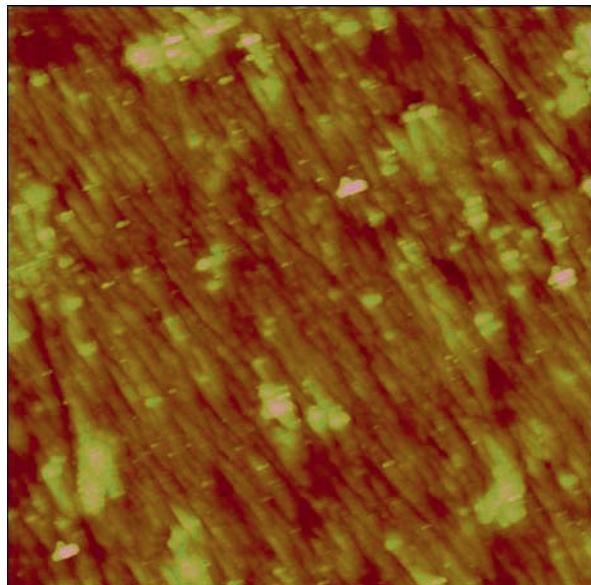


Fig. 27 Displacement fields of the observed region in the solder interconnection when the TEC is working at its cooling stage: (a) u field; (b) v field

of $5 \times 5 \mu\text{m}$ are expected to be small. The average strains derived from linear fitting the displacement fields can be very accurate. The calculation of average strains shows that $\bar{\varepsilon}_x$, $\bar{\varepsilon}_y$, and $\bar{\varepsilon}_{xy}$ are 0.042%, 0.126%, and 0.005%, respectively when the TEC was operating at its cooling stage. The calculated average strain $\bar{\varepsilon}_x$, $\bar{\varepsilon}_y$, and $\bar{\varepsilon}_{xy}$ are 0.035, -0.155, and -0.07%, respectively, when the TEC was operating at its

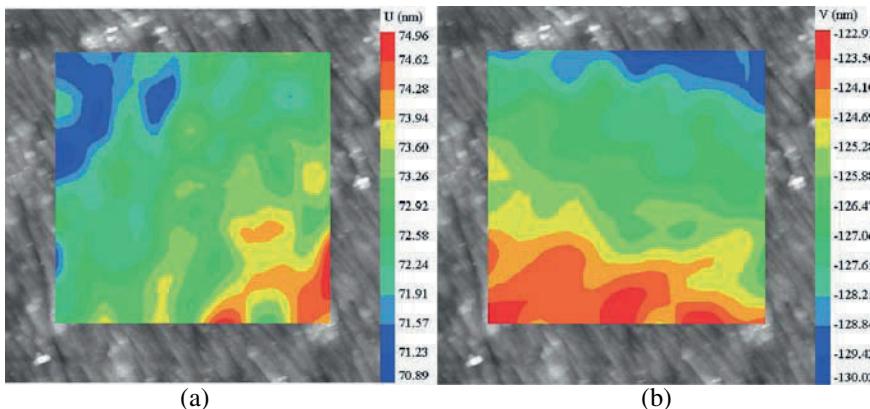


Fig. 28 Displacement fields of the observed region in the solder interconnection when the TEC is working at its heating stage: (a) u field; (b) v field

heating stage. Clearly, compared with $\bar{\varepsilon}_x$ and $\bar{\varepsilon}_{xy}$, the vertical strain component $\bar{\varepsilon}_y$ is dominant in the observed region. When the operation of the TEC module was switched from the cooling mode to the heating mode, $\bar{\varepsilon}_y$ changed from positive to negative, which means that the observed region of solder interconnection was subjected to cyclic stresses between tension and compression.

7 Conclusions

A method was developed to reconstruct AFM images for nanodeformation measurements by digital image correlation. The developed method is to generate a corrected digital image from two AFM images scanned at the angle of 0° and 90° , respectively. The proposed method was shown valid for the correction of distorted AFM images by the calibration test, which shows the displacement deviation in a reconstructed image is less than 0.16%, compared with 4.3% in the original AFM image. The AFM-DIC method was applied for in-situ thermo-mechanical deformation measurements on solder interconnection of a micro thermoelectric cooler. The AFM-DIC measurements can verify finite-element analysis, which was used to study the behavior of the whole device. It was shown that cyclic strains occurred at the top solder interconnections when the TEC worked at its switching mode between cooling stage and heating stage.

The improvement of AFM-DIC technique is desirable considering that the continuous miniaturization of microelectronic devices and interconnections more and more demand experimental strain/stress analysis of micro- and nanoscale components for material characterization and structure reliability analysis. The advance of AFM, calibration algorithm, and DIC algorithm will all contribute toward the robustness of AFM-DIC technique.

References

1. W. H. Peters and W. F. Ranson, Digital imaging techniques in experimental stress analysis, *Opt. Eng.* (1982), **21**, 427–432.
2. M. A. Sutton, W. J. Wolters, W. H. Peters, W. F. Ranson, and S. R. McNeil, Determination of displacements using an improved digital image correlating method, *Image. Vis. Comput.* (1983), **1**, 133–139.
3. T. C. Chu, W. F. Ranson, M. A. Sutton, and W. H. Peters, Applications of digital image correlation techniques to experimental mechanics, *Exp. Mech.* (1985), **25**, 234–244.
4. S. Murray, C. J. Gillham, and A. H. Windle, Characterization and correction of distortions encountered in scanning electron micrographs, *J. Phys. E: Sci. Instrum.* (1973), **6**, 381–384.
5. I. Suganuma, A novel method for automatic measurement and correction of astigmatism in the SEM, *J. Phys. E: Sci. Instrum.* (1987), **20**, 67–73.
6. J. Garnaes, L. Nielsen, K. Dirscherl, J. F. Jorgensen, J. B. Kasmussen, P. E. Lindelof, and C. B. Sorensen, Two-dimensional nanometer-scale calibration based on one-dimensional gratings, *Appl. Phys. A.* (1998), **66**, S831–S835.
7. K. Dirscherl, J. F. Jorgensen, and M. P. Sorensen, Modeling the hysteresis of a scanning probe microscope, *J. Vac. Sci. Technol. B* (2000), **18**, 621–624.
8. E. P. Stoll, Correction of geometrical distortions in scanning tunneling and atomic force microscopes caused by piezo hysteresis and nonlinear feedback, *Rev. Sci. Instrum.* (1994), **65**, 2864–2869.
9. I. Chasiotis and W. G. Knauss, A new microtensile tester for the study of MEMS materials with the aid of atomic force microscopy, *Exp. Mech.* (2002), **42**, 51–57.
10. S. Chang, C. S. Wang, C. Y. Xiong, and J. Fang, Nanoscale in-plane displacement evaluation by AFM scanning and digital image correlation processing, *Nanotechnology* (2005), **16**, 344–349.
11. Y. F. Sun and J. H. L. Pang, AFM image reconstruction for deformation measurements by digital image correlation, *Nanotechnology* (2006), **17**, 933–939.
12. Y. F. Sun, J. H. L. Pang, and W. Fan, Nanoscale deformation measurement of microscale interconnection assemblies by a digital image correlation technique, *Nanotechnology* (2007), **18**, 1–8.
13. R. Howland and L. Benatar, A practical guide to scanning probe microscopy, <http://web.mit.edu/cortiz/www/AFMGallery/PracticalGuide.pdf> (2000).
14. R. C. Barrett and C. F. Quate, Optical scan-correction system applied to atomic force microscopy, *Rev. Sci. Instrum.* (1991), **62**, 1393–1399.
15. J. E. Griffith et al., A scanning tunneling microscope with a capacitance-based position monitor, *J. Vac. Sci. Technol. B* (1990), **8**, 2023–2027.
16. H. Yamada, T. Fujii, and K. Nakayama, Linewidth measurement by a new scanning tunneling microscope, *Jpn. J. Appl. Phys.* (1989), **28**, 2402–2404.
17. D. I. V. M. Group, AFM/LFM instruction manual (Version 4.22ce), Digital Instruments (1999).
18. J. Helena and H. A. Bruck, A new method for characterizing nonlinearity in scanning probe microscopes using digital image correlation, *Nanotechnology* (2005), **16**, 1849–1855.

Nano-Scale and Atomistic-Scale Modeling of Advanced Materials

Ruo Li Dai, Wei-Hsin Liao, Chun-Te Lin, Kuo-Ning Chiang,
and Shi-Wei Ricky Lee

Abstract The studies on advanced materials at sub-micron scales have become a popular subject in engineering sciences and technology development in the past decade. In general, experimental investigations at this scale are very difficult and may be unreliable. Therefore, analytical and computational modeling appears to be rather critical and important in such studies. The objective of this chapter is to review two state-of-the-arts approaches for modeling advanced materials at the nano- and the atomistic scales, respectively.

The first part of this chapter introduces a unit cell model for the analysis of composites with carbon nanotubes (CNTs) and epoxy resin using finite element method (FEM). Varying CNT orientation is considered in order to describe the behaviors of the randomly oriented CNTs inside the epoxy matrix. Composite loss factors are calculated based on the average ratio of the unit cell energy loss to the unit cell energy input. Calculated loss factors under different strain levels are compared to experimental data. With the validated model, parametric study is thereafter performed. Parameters such as CNT dimension and CNT alignment orientation are studied. Those factors lead to higher composite damping capacity are identified.

The second part of this chapter proposes a novel atomistic–continuum mechanics which is a hybrid method for coupling continuum mechanics and interatomic potential function to predict the mechanical behavior of nano-scale single crystal silicon under uniaxial tensile loading. The atomistic–continuum mechanics is based on the transformation of chemical bonds between atoms in molecular mechanics into appropriate elements in finite element method and continuum mechanics. There are many methods used in nano-structure analysis such as molecular dynamics (MD) and ab initio calculation; the molecular dynamics simulation is one of the most promising methods for investigating the mechanical properties of structure and material at nano-scale. However, it requires extensive computing time and cost. Furthermore, in traditional continuum mechanics, the finite element method (FEM)

R.L. Dai (✉)

Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong,
Shatin, NT, Hong Kong
e-mail: ruoli.dai@gmail.com

is widely used to model and simulate the mechanical behaviors of solid/discrete body; it is a mature technology after decades of development. Compared with MD simulation, the continuum mechanics can be very efficient as it can provide results quickly in an accurate range. In the hybrid approach, the total energy of the nano-structure is formed by each individual potential energy of bi-atom, once the total potential is assembled the FEM could take the minimization of the total energy, which includes the potential energy and external works, and to solve the structure by conventional way. Therefore, this study employs FEM to explore the mechanical properties of nano-scale single crystal silicon. This research also utilizes an interatomic potential function to describe the interaction of each atom. A general form of Stillinger–Weber potential function was used for interaction between the silicon atoms in the simulations. Based on atomistic–continuum simulation results, the Young’s modulus of single crystal silicon in various crystallographic planes could be estimated. The results obtained from the present modeling approach are in reasonable agreement with the experiment and simulation results reported in the literature.

Keywords Nano-scale · Atomistic scale · Modeling · Materials

1 Modeling of Carbon Nanotube Composites for Vibration Damping

Damping materials play an important role in vibration damping and control systems. Vibration can be reduced via damping materials by transferring the kinetic energy into heat. One of the widely used damping materials is viscoelastic material (VEM). In general, the VEM damping effects are significant for being used in structures due to the energy dissipation when the structures are vibrating. However, the stiffness of traditional viscoelastic materials is low; thus large shear strain could be caused while having significant energy dissipation. The softness property may constrain the VEMs from being used in the structures with high strength-to-weight ratio, where the strength and/or transmissibility issues are crucial. On the other hand, epoxy resin (which is also a VEM, but has a highly crosslinked 3D network) is much stiffer than the VEMs and could be used to enhance structural stiffness; but it dissipates little energy from the structure. Considering the above situations, it is desirable to develop alternative means for significant vibration damping while the stiffness and strength of the material/structure are high. It could be possible to make it when composites made up of carbon nanotubes and epoxy resins are used.

People have expected that carbon nanotubes would bring us a revolution in scientific research since the multiwalled carbon nanotubes were discovered [1]. Carbon nanotubes (CNTs) are cylinder allotropes of carbon with diameter in the scale of nanometers. This results in a nano-structure where the length-to-diameter ratio might exceed 10,000. Carbon nanotubes are one of the strongest and stiffest materials known, in terms of tensile strength and elastic modulus, respectively.

A multiwalled carbon nanotube was tested to have a tensile strength of 63 GPa. For comparison, high-carbon steel has a tensile strength of approximately 1.2 GPa. CNTs have very high elastic moduli, on the order of 1 TPa [2]. Since carbon nanotubes have a low density of 1.3–1.4 g/cm³, its specific strength of up to 48,462 kN·m/kg is the best of known materials, compared with high-carbon steel's 154 kN·m/kg.

It was recognized that CNTs might be the ultimate reinforcing materials for developing composites with excellent properties, such as exceptionally high stiffness, strength, and resilience, as well as superior electrical and thermal properties [3]. Researches have been conducted in the field of nano-structures enhanced polymer resin composites. Most of those researches focused on the strength, toughness, thermal conduction, and electrical conduction enhancement [4–7].

Recently, researchers have started to pay attention to the damping enhancement of the composites due to the CNT addition. The unique damping properties of CNT-based structures, such as densely packed CNT thin films [8–11] and polymeric composites with dispersed CNT fillers [12–16], have been of great interest. Multiwalled carbon nanotube (MWNT) thin films were fabricated using catalytic chemical vapor deposition (CVD) of xylene and ferrocene mixture precursor [10, 11]. The nanotube films were employed as inter-layers within the composite to reinforce the interfaces between composite plies, enhance laminate stiffness as well as structural damping. Zhou et al. [17] conducted experimental and analytical investigations on the energy dissipation of polymeric composites containing carbon nanotube fillers. Their specimens were fabricated by directly mixing single-walled carbon nanotubes into two types of polymers. The loss factors of the composites were characterized while the loading was applied in the normal direction. It was also reported that poor load transfer between nanotubes and surrounding polymer chains may detriment to high stiffness and strength, but greatly enhance the structural damping [14]. Dai and Liao [18] experimentally studied the damping effect caused by CNT addition. General procedures for fabricating CNT/epoxy composites were given as follows. A set of material testing equipment was constructed. The fabricated composite specimens were evaluated under different loading conditions by varying excitation frequencies and amplitudes. Storage modulus and loss factors were measured for both pure epoxy specimen and CNT composite specimens. In particular, CNT addition was randomly dispersed in the composite fabrication and shear loading was applied during the damping measurement. Experimental results showed that the CNT additives greatly enhanced the damping capacity of pure epoxy resins. This fabricated CNT/epoxy composite has the same level of storage modulus as epoxy resin and a similar level of loss factor as compared to VEMs while strain is over a certain level. It was concluded by Dai and Liao that for sufficiently high strain level, CNT composite could provide more damping than traditional damping materials. The observed phenomena matched the anticipation and the results were promising for potential applications when both damping and stiffness are desired.

Zhou et al. [17] first developed a “stick-slip” damping model for composites containing well-dispersed single-wall carbon nanotubes (SWNTs) and performed

damping measurements on CNT-based composites. SWNTs were assumed to be in parallel bundles referred to as “nanotube ropes” or “nanoropes.” In each bundle, the SWNTs were held together by relatively weak van der Waals forces. Based on this assumption, a micromechanical model for the composites with aligned nanoropes was developed. Analytical results showed strain-dependent damping enhancement due to “stick-slip” motion at the interfaces of the SWNTs and the resin. Zhou et al. [17] and Liu et al. [15, 16] presented an analysis on the structural damping characteristics of polymeric composites containing dilute, randomly oriented nanoropes. The SWNT rope was modeled as a closely packed lattice consisting of seven nanotubes in hexagonal array. The resin was described as a viscoelastic material using two models: Maxwell model and three-element standard solid model. The composite was modeled as a three-phase system consisting of a resin, a resin sheath acting as a shear transfer zone, and SWNT ropes. The analytical results indicated that the loss factor of the composite is sensitive to stress magnitude. It was illustrated that the “stick-slip” friction is the main contribution for the total loss factor of CNT-based composites even with a small amount of nanotubes/ropes. Odegard et al. [19] also gave an analytical model to simulate nanotube-reinforced polymer composites. Liu et al. [20] used the boundary element method (BEM) for analyzing carbon nanotube-based composites. It should be noted that the damping performance of CNT composites was not considered in by the last two research groups.

In this chapter, a finite element method (FEM)-based model for CNT composite is developed. A composite unit cell model is built. The FEM-based composite shows the detailed sliding surface status developed between CNT and base epoxy matrix, while none of previous models does. Under the shear loading, simulation is carried out to investigate interfacial sliding under different strain level and CNT orientation. Loss factors can be obtained from the ratio of energy dissipation out of energy input per cycle. The simulation results are compared with experimental data. Based on the FEM model, further parametric studies are performed. Several parameters such as CNT dimension and CNT alignment orientation are discussed with respect to composite damping capacity (loss factor).

2 Nanotube Composite Modeling

2.1 Modeling Platform and Assumptions

The finite element method (FEM) is used for modeling the CNT composites. The chosen platform is commercial FEM software ANSYS 10.0/Multiphysics. Three-dimensional solid element SOLID185 is used. SOLID185 is defined by eight nodes having 3° of freedom at each node: translations in x -, y -, and z -directions [21]. This element can be used in relatively large deformation cases and there is no redundant degree of freedom that brings extra computational load.

It was assumed that the energy dissipation in CNT composite mostly comes from interfacial sliding friction between CNT and base matrix [17]. In addition, the base

matrix (i.e., epoxy resin) itself also contributes to the composite damping, which is assumed to be constant with respect to strain and the loss factor of the base matrix equals to that of the pure epoxy resin. Several assumptions are made in order to model CNT composites:

- (1) Damping consists of two parts: contribution from epoxy resin matrix and contribution from CNT/matrix interfacial sliding.
- (2) CNT and base matrix are perfectly bonded at initial state before loading is applied. As the interfacial strain increases with applied shear loading, sliding occurs when the critical strain is reached.
- (3) Interfacial friction transfers mechanical energy to thermal energy thus dissipates energy out of system.
- (4) The CNTs in composite are assumed to be well aligned.
- (5) CNT segments are well separated and each single CNT segment is isolated within a small epoxy resin cubic, which is called “composite unit cell.” The dimension of this unit cell is calculated from CNT volume ratio. This assumption would be acceptable when the damping mainly comes from CNT/epoxy sliding but not CNT/CNT friction.
- (6) CNTs in matrix are randomly oriented and uniformly distributed so that we can average the energy loss over all CNT segments.

Based on assumptions (4) and (5), the CNTs in composite are modeled as straight segments isolated in composite unit cells. It has been reported that while ultrasonic process is used for dispersing the CNTs into the epoxy resin, the sonication energy breaks the CNTs into shorter segments. The ultrasonic process will decrease CNT aspect ratio while the dispersibility is increased [22]. For CNT composite fabricated in this research, it can be observed from a SEM image given in Fig. 1 that the CNT dispersion quality is good in general. The aspect ratio of CNTs is also found to be decreased. Under such condition, the model with straight CNT segments isolated in composite unit cells seems acceptable.

2.2 Finite Element Model

The composite unit cell (cubic shaped) containing a single CNT segment is modeled in 3D as shown in Fig. 2. Extra blocks are mounted on the top and bottom of this unit cell in the model in order to facilitate applied loading. The dimension of this composite unit cell is obtained from CNT weight ratio. Higher CNT weight ratio gives smaller CNT composite unit cell. Detailed dimension and properties are given in Table 1. The CNT segment is modeled as a hollow tube with rounded ends and a built-in thin sheath tube in order to simulate the contact/bonding surface between CNT and matrix (see Fig. 3). This sheath layer plays a very important role in damping simulation. Some of the over-stressed elements in this layer will be deleted in order to simulate the interfacial sliding mechanism.

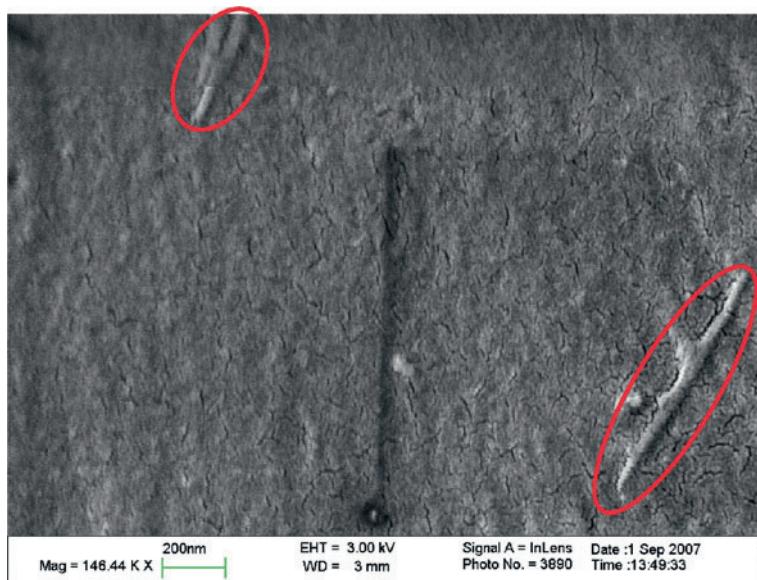


Fig. 1 CNT segments (*circled*) in matrix

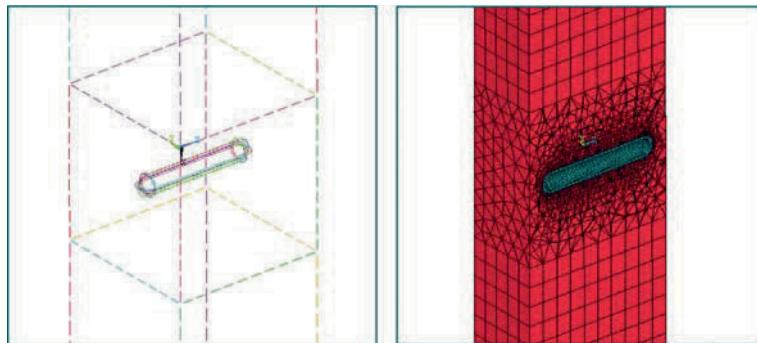


Fig. 2 Finite element model of composite unit cell

Table 1 Model parameters

CNT segment	Epoxy resin unit cell	Sheath layer			
Length	100 nm	Cubic edge length	300 nm	Thickness	2 nm
Outer diameter	20 nm	Young's modulus	3 GPa	Young's modulus	3 GPa
Thickness	4 nm	Poisson's ratio	0.37	Poisson's ratio	0.37
Young's modulus	1000 GPa	Loading	Shear	Critical sliding strain	0.0016
Poisson's ratio	0.3				

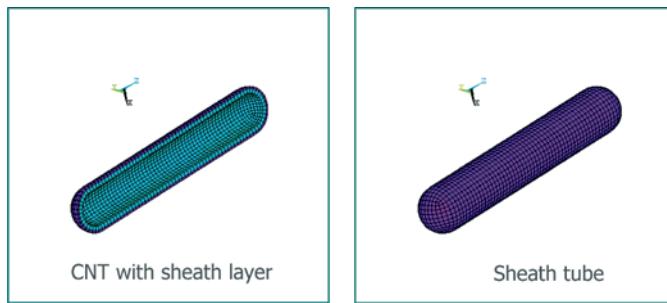


Fig. 3 CNT model with sheath layer

For finite element modeling, hexahedron elements are preferred over tetrahedron elements because fewer well-formed hexahedron elements are required than tetrahedron elements for the same accuracy. In another word, the size of the model and the computational effort could be reduced by using hexahedron elements. However, the mesh generation may be more complicated or even impossible for some of the topologies [21]. In this model, an all-hexahedron mesh is generated for the CNT segment, the sheath layer, and the blocks, which are fixed on the top and bottom sides of the composite unit cell. Tetrahedron and prism-shaped elements are used for meshing the epoxy part of the unit cell. This treatment greatly reduces the calculation time as compared with an all-tetrahedron meshed model. Figure 4 shows the comparison between different meshing strategies. The model shown in the top two figures is generated by an all-tetrahedron meshing, which contains 81,531 elements while the model shown in the bottom two is generated by all-hexahedron meshing, which has 56,968 elements. In addition, the mesh quality is improved by using all-hexa mesh as the element shape is more regular and the mesh density in the epoxy part is even finer with fewer elements. As a result, it takes 4.5 h using the all-tetra meshed model for obtaining a solution while the model with all-hexa mesh runs a little bit longer than 3 h (the results of the two models are close to each other). The solver convergence of hexahedral versus tetrahedral meshes has been studied. Cifuentes et al. [23] concluded that the results obtained with quadratic tetrahedral elements, compared to bilinear hexahedral elements, were equivalent in terms of both accuracy and CPU time. Bussler et al. [24] reported that better accuracy can be achieved by using the same order hexahedral elements over tetrahedrons. Therefore, the model containing less tetrahedron elements would have better solver convergence and accuracy when material nonlinearity or large deformation is significant.

The Young's modulus and Poisson's ratio of the sheath layer elements are set to be the same as the regular epoxy resin but a failure strain (0.0016) is set in order to simulate the interfacial sliding mechanism. The value of this critical sliding strain can be determined from experiments and numerical analysis, which will be discussed in the next section.

The flow chart for calculating composite loss factor is given in Fig. 5. The model is first constructed with boundary conditions and initial loading applied. The model

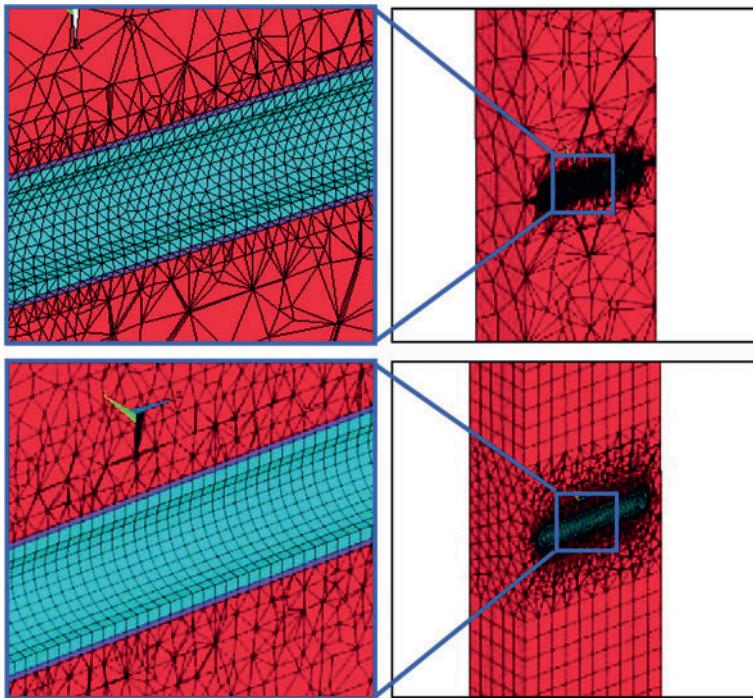


Fig. 4 Comparison between different meshing strategies (*Top*: all-tetra mesh, *Bottom*: all-hexa mesh except the epoxy part of composite unit cell)

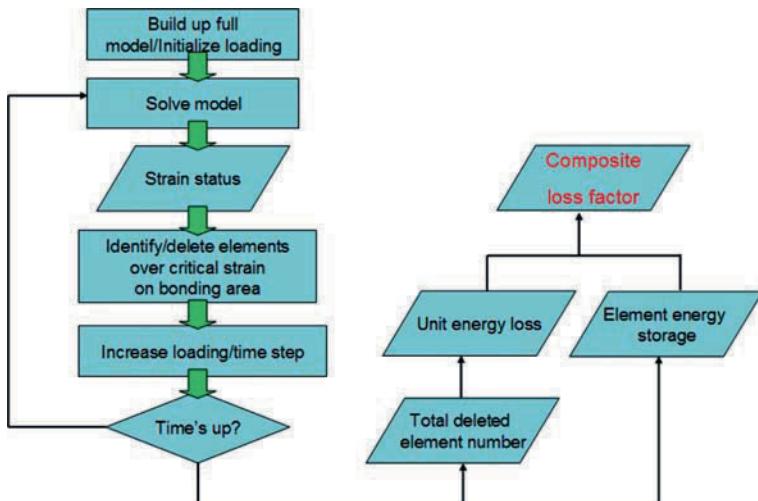


Fig. 5 Flow chart for calculating composite loss factor

is then solved by using ANSYS® implicit sparse solver and DYNA® explicit solver for highly nonlinear case such as large displacement. Observing the results, we can pick up the sheath elements over the critical sliding strain. These elements are deleted from system in order to simulate the debonding between nanotube segment and base epoxy matrix. Then, as clock steps up, another set of loading is applied, the model is solved again and more elements are deleted in the new round of calculation. After one period of cyclic loading applied, the total number of deleted elements is counted. The number of deleted elements is used for calculating energy loss. The total element energy storage (element strain energy) during this time interval is also calculated from the model. By using these two energies, we can calculate composite loss factor. More specifically, from simulated strain/stress distribution, the stored energy (energy input) of a composite unit cell per loading cycle, U_{cell} can be calculated as follows:

$$U_{\text{cell}} = \frac{1}{2} \sum_n \sigma_{\text{cell}} \varepsilon_{\text{cell}} V \quad (1)$$

where V , n , σ_{cell} , and $\varepsilon_{\text{cell}}$ represent the volume of element in composite unit cell, the total number of elements in composite unit cell, the calculated von Mises equivalent stress, and strain, respectively [21].

As CNT is much stiffer than base matrix, it is assumed that the strain near CNT/matrix interface is larger than the average strain in composite unit cell. Once the critical sliding strain is determined, the sheath elements over this critical strain during each loading cycle are detected and removed from the model in order to simulate the debonding process and sliding effect. As mentioned, the critical strain in this simulation is chosen as 0.0016; further discussions on critical strain will be given in the next section. The number of deleted sheath tube elements is counted after each loading cycle in order to calculate the cyclic energy dissipation. The zenith angle and azimuth angle of the CNT segment are denoted by two parameters α and β as shown in Fig. 6 to describe CNT orientation with respect to global coordinate system. After the solver finishes computing, the energy loss during one cycle is calculated by adding the strain energy of all deleted elements

$$U_{\text{loss}} = \frac{1}{2} \sum_m \sigma_{\text{cr}} \varepsilon_{\text{cr}} v \quad (2)$$

where v , m , σ_{cr} , and ε_{cr} are volume of sheath element, number of deleted sheath elements, critical sliding stress, and critical sliding strain, respectively.

2.3 Simulation Results

In this simulation, the interfacial sliding between CNT segments and base epoxy resin is modeled. The model shows when and where the sliding occurs on

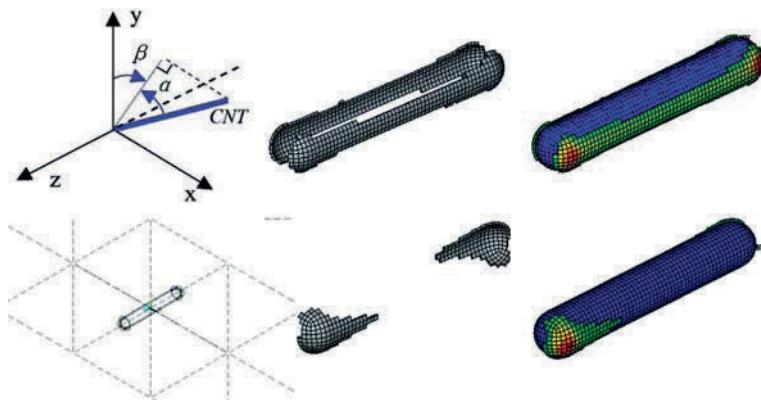


Fig. 6 Deleted sheath elements during CNT/epoxy sliding

CNT/matrix interface and how the sliding develops. These conditions were not considered by previous researches.

Figure 6 shows interfacial sliding examples between CNT segment and base matrix in two different CNT orientations. The upper two show the sliding status while CNT segment is perpendicular to loading surface ($\alpha = 90^\circ$, $\beta = 90^\circ$). The lower two illustrate the sliding status while CNT segment's zenith angle and azimuth angle are 30 and 90° , respectively. The sheath layer elements shown in the middle of Fig. 6 are the ones over critical strain, which are removed from the model in order to simulate the sliding effect during each loading cycle. More studies on CNT segment's zenith and azimuth alignment angles will be given in next section. Other than CNT orientation, interfacial sliding area is also related to the applied strain level. Figure 7 shows CNT segment sliding status under different strain levels, i.e., 0.0001, 0.0002, 0.0003, 0.0005, 0.0006, 0.0007, 0.0008, and 0.0012, respectively, from (a) to (h). The sliding area is developed along with strain level – higher strain induces larger sliding area.

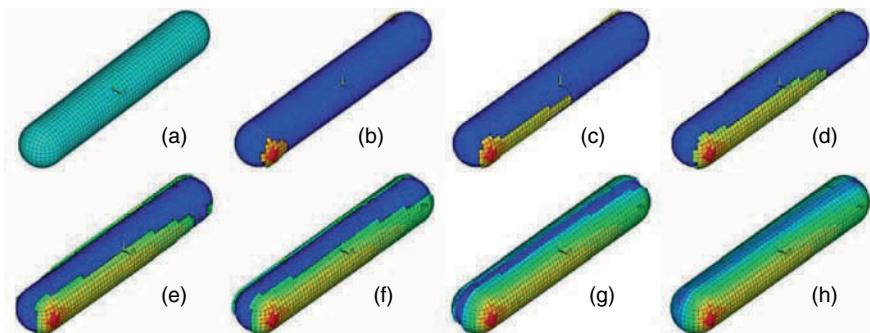


Fig. 7 CNT/epoxy interfacial sliding status under different strain level ($\alpha = 7.5^\circ$, $\beta = 90^\circ$, $\varepsilon_{cell} = [0.0001, 0.0012]$)

The loss factor is the ratio of energy dissipated as the result of damping during one cycle to the total energy of the vibration system under cyclic loading. The loss factor of the unit cell could be obtained as follows:

$$\eta = \frac{U_{\text{loss}}}{U_{\text{cell}}} = \frac{\frac{1}{2} \sum_m \sigma_{\text{cr}} \varepsilon_{\text{cr}} V}{\frac{1}{2} \sum_n \sigma_{\text{cell}} \varepsilon_{\text{cell}} V} = \frac{\sum_m E_{\text{epoxy}} \varepsilon_{\text{cr}}^2 V}{\sum_n E_{\text{cell}} \varepsilon_{\text{cell}}^2 V} \quad (3)$$

Shear loading is applied on the composite unit cell. CNT orientation parameters α and β range from 0 to 90° while the step size is 7.5° , which leads to $13 \times 13 = 169$ different cases built in the first octant for each loading level. The angular range covers one-eighth of the 3D coordinate system so that it is sufficient for our study due to geometric symmetry. The number of deleted sheath elements is recorded in terms of α and β under different strain level. In this model, the total number of elements in the sheath layer $n_{\text{max}} = 1973$, which corresponds to the maximum number of elements involved in interfacial sliding. In other words, at most 1973 elements could be deleted during sliding to result in the largest possible energy dissipation. In general, higher strain level leads to larger sliding area that induces more energy dissipation; different CNT alignment causes different sliding area size and location under the same loading. A 3D graph is given in Fig. 8 to show the number of deleted sheath elements versus CNT orientation when $\varepsilon_{\text{cell}} = 0.0007$. The related data are given in Table 2.

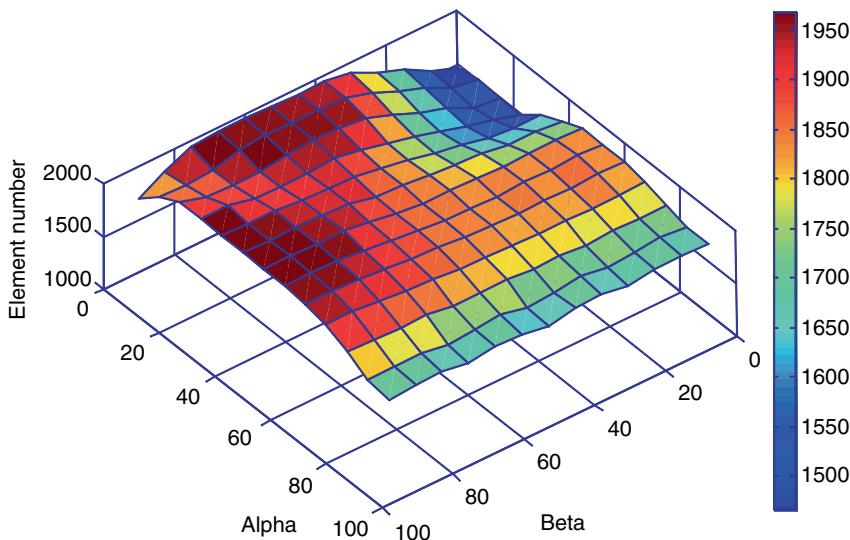


Fig. 8 Number of deleted sheath elements versus CNT orientation while $\varepsilon_{\text{cell}} = 0.0007$

For each loading level, based on simulation results for 169 cases (represent 13 \times 13 different CNT orientations in the first octant), the average number of sliding sheath elements is calculated. The average energy loss of each composite unit

Table 2 Number of deleted sheath elements versus CNT orientation while $\varepsilon_{cell} = 0.0007$

$\beta \setminus \alpha$	0	7.5	15	22.5	30	37.5	45	52.5	60	67.5	75	82.5	90
0	1453	1407	1382	1365	1452	1632	1726	1753	1747	1705	1645	1639	1572
7.5	1445	1423	1406	1394	1505	1600	1694	1746	1721	1700	1642	1604	1604
15	1578	1540	1529	1508	1552	1612	1713	1754	1745	1717	1641	1610	1607
22.5	1707	1668	1616	1579	1588	1650	1718	1736	1718	1707	1643	1623	1573
30	1812	1797	1706	1673	1659	1701	1731	1760	1736	1734	1632	1581	1589
37.5	1938	1898	1818	1750	1716	1729	1747	1750	1750	1721	1648	1612	1574
45	1945	1935	1874	1811	1745	1754	1784	1780	1777	1748	1644	1598	1555
52.5	1953	1938	1899	1827	1777	1775	1810	1810	1805	1760	1671	1573	1566
60	1929	1896	1845	1851	1840	1840	1849	1849	1835	1773	1659	1602	1583
67.5	1887	1783	1796	1852	1882	1924	1911	1896	1870	1813	1661	1582	1566
75	1639	1761	1783	1856	1903	1941	1955	1936	1901	1830	1711	1631	1610
82.5	1622	1637	1776	1887	1955	1961	1957	1957	1926	1844	1689	1617	1609
90	1617	1627	1690	1965	1963	1966	1963	1957	1929	1850	1672	1601	1566

cell can be further obtained. The average number of sliding elements, number of maximum sliding elements, number of minimum sliding elements, and standard deviation of sliding number (w.r.t. CNT segment orientation) associated with strain range [0.0002, 0.001] are given in Table 3. The loss factor of the CNT composite thus can be calculated. The calculated loss factor of the CNT composite is compared to the experimental data as shown in Fig. 9. We can see from this figure that the calculated and measured loss factors of the CNT composite are in the similar trend and level with respect to strain.

Table 3 Number of deleted sheath elements under different strain level

Strain	0.0002	0.0003	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009	0.001
Average	92	441	957	1333	1586	1725	1810	1911	1961
Max	352	746	1174	1755	1945	1966	1973	1973	1973
Min	0	157	707	1033	1220	1365	1463	1577	1694
Std. dev.	95	167	136	159	155	143	123	84	55
Loss factor	0.095	0.141	0.159	0.149	0.133	0.117	0.104	0.096	0.089

In this simulation, the CNT/matrix interfacial critical sliding strain is set to 0.0016. The critical strain is determined from experimental data. However, from available experimental data and other references [2, 12, 16], the critical strain can only be estimated. A range of critical strain is varied to find out which one gives the loss factor curve most fit to experimental data. Simulation results for different critical strain values are shown in Fig. 10, where critical strain value is set to 0.0012, 0.0014, 0.0016, and 0.0018. We can see from this figure that the simulation with $\varepsilon_c = 0.0016$ fits experimental data the best, and thus this value is used for further simulation in this research.

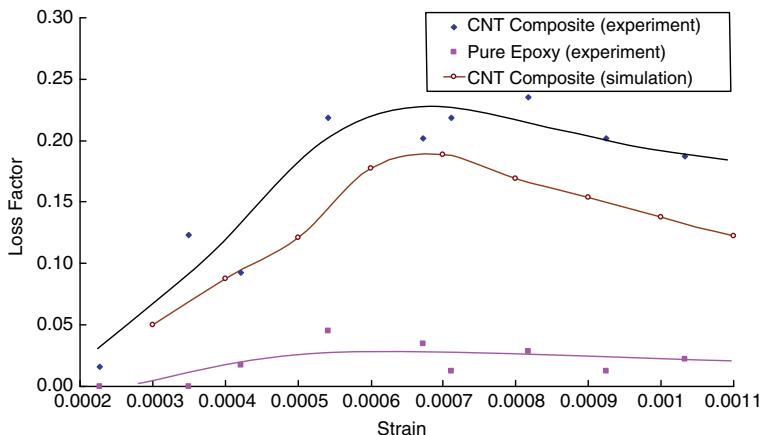


Fig. 9 Comparison of loss factor between simulation and experiment

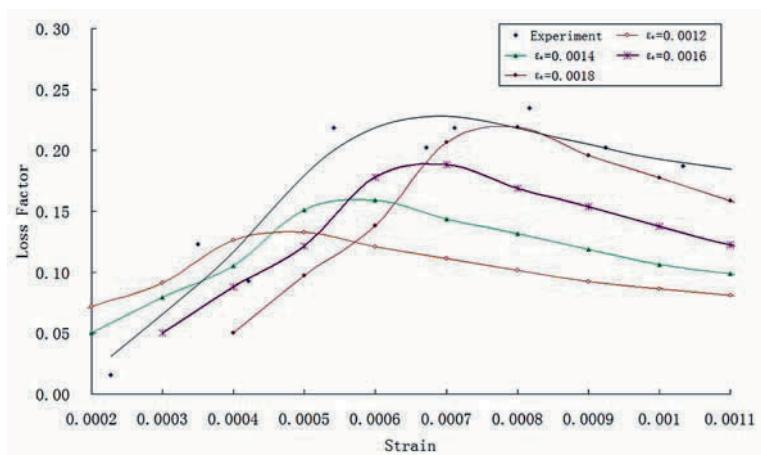


Fig. 10 Simulated composite loss factors with different critical strains

Figure 9 shows that simulated loss factor is smaller than the one measured from experiment for a given strain. It should be noted that only the contributions from epoxy resin matrix and CNT/matrix interfacial sliding were considered when modeling damping, as stated in the first assumption. Some other mechanisms could also dissipate energy. CNT/CNT stick-slip motion could be one of these mechanisms while energy dissipation from bubbles/cavities inside epoxy matrix might be another one. These mechanisms could be further considered for more accurate models.

2.4 Parametric Study on CNT Composites

The validated model could be used for predicting CNT composite dynamic properties and guiding composite fabrication. Several factors could affect the energy dissipation ability of the CNT composite, such as the dimensions of CNTs and the alignment orientations of CNTs inside matrix. These key factors will be studied in order to provide some references for material selection and composite fabrication.

There are several commercially available carbon nanotubes. A brief summary on geometrical properties of some major MWNTs is given in Table 4. As shown in the table, typical diameter of commercially available MWNTs ranges from 5 to 50 nm while the length ranges from 200 nm to 20 μm . For the studies in this research, CNT diameter range [10, 50 nm] and length range [100, 600 nm] are set as they cover most commonly used MWNTs. Another reason for the chosen CNT length range is that an ultrasonic process was performed during CNT dispersion. High-energy sonication could break CNTs down to small segments, thus the selection of the length range [100, 600 nm] would be reasonable.

Table 4 Geometrical properties of commercial MWNTs

Supplier	Type	Length	Diameter
NanoLab	MWNT 1	1–5 μm	15 \pm 5 nm
	MWNT 2	5–20 μm	15 \pm 5 nm
	MWNT 3	1–5 μm	30–15 nm
	MWNT 4	5–20 μm	30–15 nm
	MWNT 1	200 nm	6.5 nm
Rosseter	MWNT 2	300 nm	12 nm
	MWNT 3	500 nm	20 nm
SunNano	MWNT	400 nm–1 μm	10–30 nm
n-Tec	MWNT	2 μm	2–50 nm

CNT models with different CNT length while keeping the same CNT diameter are built. Loss factor versus strain with respect to different CNT length are shown in Fig. 11. Corresponding data are given in Table 5. These simulation results are obtained based on the same epoxy matrix and CNT weight ratio while assuming randomly oriented dispersion of CNTs in composite. It is observed that, given the same CNT diameter, CNT addition with shorter segment length induces slightly higher composite damping while strain is over certain level (around 0.00065 for the case shown in Fig. 11) and longer CNTs bring out higher damping before this strain level. However, composite loss factor curves with different CNT length are close to each other in general. In other words, composite loss factor is not sensitive to CNT length. As it is hard to keep CNTs with high aspect ratio than instead of low one, this observation would keep the cost of composite fabrication relatively low.

Another study is performed with respect to CNT diameter in range [10, 50 nm]. Other geometrical parameters including CNT segment length, diameter, and the dimension of composite unit cell are kept the same. Loss factor versus strain with different CNT diameter are shown in Fig. 12 and corresponding data are given in

Table 5 Calculated loss factors with different CNT length (CNT diameter = 20 nm)

Global strain	Loss factor (L100)	Del. elem. (L100)	Loss factor (L200)	Del. elem. (L200)	Loss factor (L300)	Del. elem. (L300)	Loss factor (L400)	Del. elem. (L400)	Loss factor (L600)	Del. elem. (L600)
0.0001	0.050	0	0.050	0	0.069	2	0.105	8	0.122	16
0.0002	0.112	9	0.152	31	0.160	51	0.165	71	0.163	105
0.0003	0.175	44	0.192	102	0.203	167	0.209	233	0.207	345
0.0004	0.243	131	0.256	284	0.260	436	0.262	589	0.260	874
0.0005	0.270	241	0.269	480	0.262	693	0.261	918	0.259	1360
0.0006	0.265	338	0.256	640	0.252	939	0.251	1243	0.250	1855
0.0007	0.241	397	0.229	731	0.224	1064	0.223	1401	0.221	2078
0.0008	0.213	428	0.201	780	0.197	1130	0.195	1484	0.193	2196
0.0009	0.188	446	0.178	814	0.174	1179	0.172	1546	0.171	2286
0.0010	0.168	460	0.159	835	0.155	1208	0.154	1584	0.152	2341
0.0011	0.152	469	0.143	851	0.140	1231	0.139	1612	0.137	2383
0.0012	0.138	476	0.131	865	0.128	1251	0.127	1641	0.126	2425
0.0013	0.127	483	0.120	877	0.118	1270	0.117	1666	0.116	2458
0.0014	0.118	488	0.112	886	0.110	1284	0.109	1683	0.108	2478
0.0015	0.110	491	0.105	892	0.103	1291	0.102	1690	0.101	2488

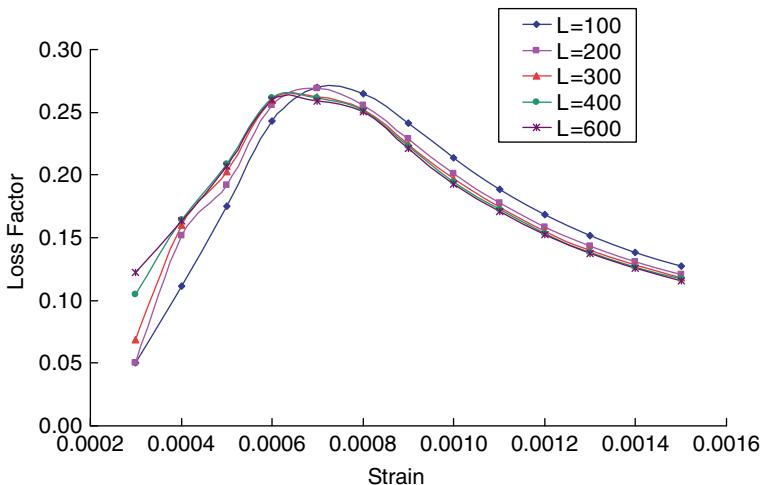


Fig. 11 Loss factor versus strain with respect to different CNT length (CNT diameter $D = 20$ nm; L represents CNT segment length in nanometer)

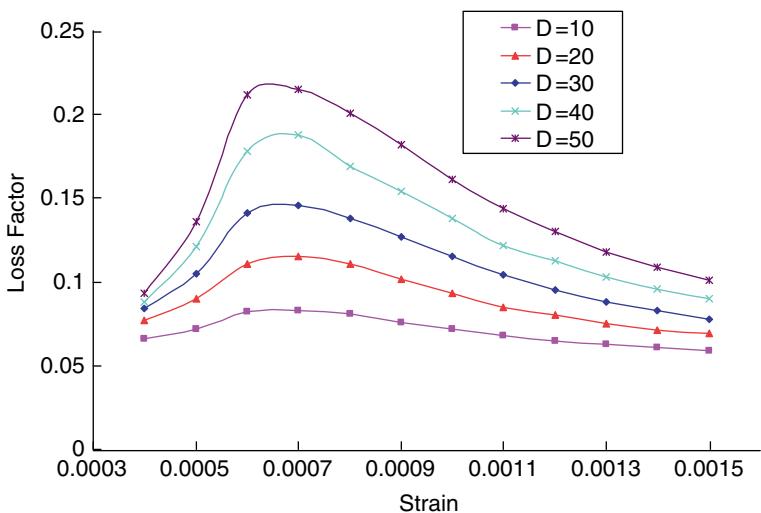


Fig. 12 Loss factor versus strain with respect to different CNT diameters (D represents CNT diameter and CNT segment length $L = 200$ nm)

Table 6. It can be seen that the effect of CNT diameter on the composite loss factor is significant. More specifically, larger CNT diameter gives more energy dissipation for a single composite unit cell as increasing CNT diameter enlarges the surface area of the CNT segment. Larger composite loss factor is thus induced when the number of CNT segments in composite is the same.

Table 6 Calculated loss factors with different CNT diameter (CNT length = 200 nm)

Global strain	Loss factor (D10)	Del. elem. (D10)	Loss factor (D20)	Del. elem. (D20)	Loss factor (D30)	Del. elem. (D30)	Loss factor (D40)	Del. elem. (D40)	Loss factor (D50)	Del. elem. (D50)
0.0004	0.066	18	0.077	31	0.084	39	0.088	43	0.093	49
0.0005	0.072	56	0.090	102	0.105	143	0.121	189	0.136	233
0.0006	0.082	146	0.111	284	0.141	437	0.178	643	0.212	845
0.0007	0.083	235	0.115	480	0.146	726	0.188	1100	0.215	1358
0.0008	0.081	321	0.111	640	0.138	951	0.169	1333	0.201	1755
0.0009	0.076	365	0.102	731	0.127	1123	0.154	1555	0.182	2041
0.0010	0.072	399	0.093	780	0.115	1226	0.138	1688	0.161	2200
0.0011	0.068	412	0.085	814	0.104	1275	0.122	1733	0.144	2303
0.0012	0.065	420	0.080	835	0.095	1305	0.113	1830	0.130	2372
0.0013	0.063	422	0.075	851	0.088	1328	0.103	1866	0.118	2421
0.0014	0.061	423	0.071	865	0.083	1345	0.096	1889	0.109	2459
0.0015	0.059	424	0.069	877	0.078	1357	0.090	1906	0.101	2488

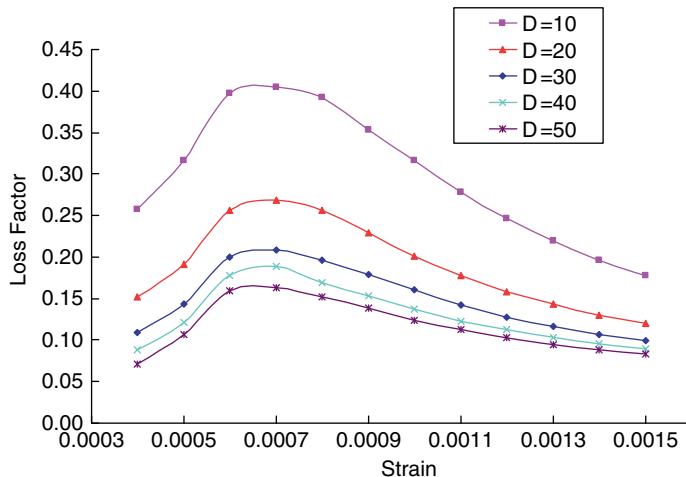


Fig. 13 Loss factor versus strain with respect to different CNT diameters (D represents CNT diameter and CNT segment length $L = 200$ nm, CNT weight ratio is fixed for each case)

Additional analysis on the loss factor with different CNT diameters is performed while CNT weight ratio is fixed at 2%. The results are shown in Fig. 13 and Table 7. It can be observed from the figure that thinner CNTs induce higher composite damping. The maximum composite loss factor with 10 nm CNT additives is 0.4 while the one with 50 nm CNT additives is 0.16. The reason is as follows. Although a smaller diameter gives smaller surface area for a single CNT segment, a larger total sliding surface is induced as the number of CNT segments in composite increases when CNT diameter decreases given the same CNT weight ratio. Therefore, if the weight ratio of CNT additives is the same, CNTs with smaller diameter can be chosen to obtain higher composite damping. If cost is not a big concern, we can even use SWNTs, which are with a much smaller diameter (only a few nanometers).

As mentioned, CNT/epoxy interfacial sliding status is related to CNT alignment orientation in matrix. A model is used here to study the relationship between CNT orientation and composite damping ability. Figure 14 shows simulation results regarding CNT/matrix interfacial sliding status in different CNT segment orientations, while CNT segments are aligned on $x-z$ plane, swept from negative z toward positive x ($\alpha = [0^\circ, 90^\circ]$, $\beta = 90^\circ$). As we can see from this figure, interfacial sliding occurs along the CNT longitudinal direction while $\alpha = 0^\circ$. The sliding surface location moves to the both ends of the segment while CNT rotating around y axis with $\alpha = 52.5^\circ$.

As observed in the simulation results, the effect of CNT alignment on the composite damping varies under different strain level. The sliding element numbers versus CNT orientation are illustrated in Fig. 15 for strain range $[0, 0.0012]$. As increasing the applied load, interesting phenomena are observed. The CNT/matrix interfacial sliding could be categorized into five different stages with respect to the average strain in composite unit cell.

Table 7 Calculated loss factors with different CNT diameter, fixed CNT weight ratio (CNT length = 200 nm)

Global strain	Loss factor (D10)	Del. elem. (D10)	Loss factor (D20)	Del. elem. (D20)	Loss factor (D30)	Del. elem. (D30)	Loss factor (D40)	Del. elem. (D40)	Loss factor (D50)	Del. elem. (D50)
0.0004	0.258	18	0.152	31	0.109	39	0.088	43	0.072	49
0.0005	0.316	56	0.192	102	0.143	143	0.121	189	0.107	233
0.0006	0.397	146	0.256	284	0.200	437	0.178	643	0.160	845
0.0007	0.404	235	0.269	480	0.208	726	0.188	1100	0.163	1358
0.0008	0.392	321	0.256	640	0.196	951	0.169	1333	0.152	1755
0.0009	0.353	365	0.229	731	0.179	1123	0.154	1555	0.139	2041
0.0010	0.317	399	0.201	780	0.160	1226	0.138	1688	0.124	2200
0.0011	0.279	412	0.178	814	0.143	1275	0.122	1733	0.112	2303
0.0012	0.247	420	0.159	835	0.128	1305	0.113	1830	0.102	2372
0.0013	0.219	422	0.143	851	0.116	1328	0.103	1866	0.095	2421
0.0014	0.196	423	0.131	865	0.107	1345	0.096	1889	0.088	2459
0.0015	0.178	424	0.120	877	0.099	1357	0.090	1906	0.083	2488

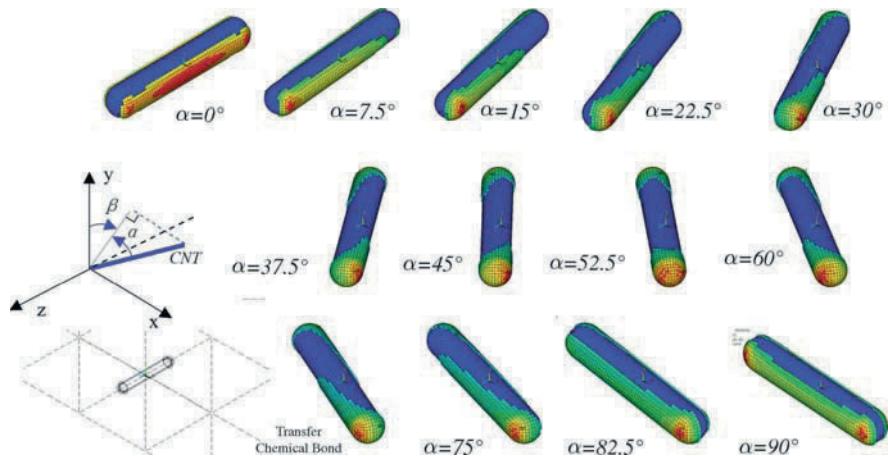


Fig. 14 CNT/epoxy interfacial sliding status by varying α ($\alpha = [0^\circ, 90^\circ]$, $\beta = 90^\circ$)

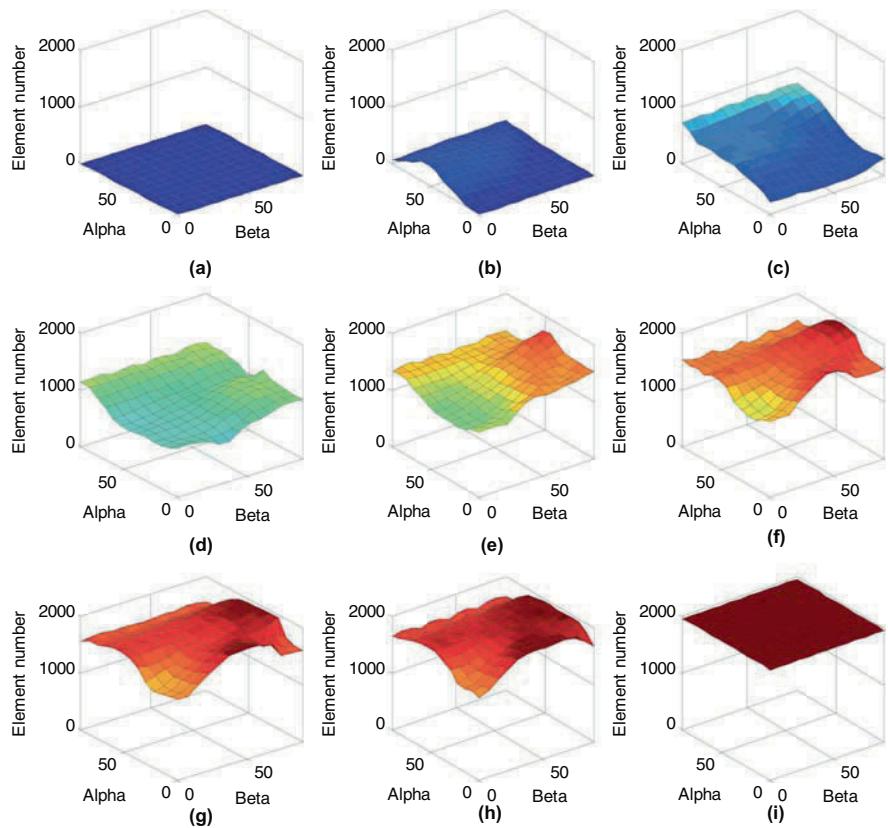


Fig. 15 Sliding element number versus CNT orientation ($\varepsilon_{cell} = [0, 0.0002, 0.0003, 0.0004, 0.0005, 0.0006, 0.0007, 0.0008, 0.0012]$)

- (1) For very low strain level (Fig. 15a), no sliding occurs between CNTs and matrix. At this stage, composite loss factor is small and corresponding to that of pure epoxy.
- (2) For low strain level (Fig. 15b), sheath layer strain could become larger than the critical strain and interfacial sliding occurs only on CNT segments in some orientations. In this case, the orientation with $\alpha \approx 50^\circ$ and $\beta = 0^\circ$ (all CNT segments are in $x-y$ plane) gives higher loss factor.
- (3) For medium strain level (Fig. 15c, d, and e), sliding occurs on all-oriented CNT segments. We can see from these figures that, under strain range $\varepsilon_{cell} = [0.0003, 0.0005]$, CNTs in composite are partially debonded (the numbers of sliding sheath elements are less than 1973). Meanwhile, CNTs with orientation $\alpha = 90^\circ$ dissipate relatively more energy than other orientations.
- (4) For high strain level (Fig. 15f, g, and h), sliding occurs on all oriented CNT segments. For some CNT orientations, the CNT segment is fully debonded. It is observed from these figures that, for some CNT orientations ($\beta = 90^\circ$, $\alpha \approx 50^\circ$ and $\beta \approx 50^\circ$, $\alpha = 90^\circ$), more sheath layer elements are deleted to result in higher damping effect. It should be noted that, although not all the CNTs are fully debonded, composite loss factor could reach its maximum at this stage as observed in Fig. 9 (strain about 0.0007). This is because the relationship between interfacial sliding area and strain is not linear. The energy dissipation (which is related to sliding area) at this stage is already close to the maximum while the total energy input (which is related to strain) can still be further increased for even higher strain level (stage 5).
- (5) For very high strain level (Fig. 15i), all CNT segments are fully debonded, nearly the full sheath layer is deleted (over 1940 elements are deleted). The interfacial sliding is complete. In this case, regardless of the CNT alignment, the energy loss of composite unit cell reaches its upper bound and the composite gives maximum energy loss U_{loss} . However, as total energy input U_{cell} keeps going up as the strain level is further increased beyond the energy loss upper bound, the loss factor η calculated from Eq. (3) would drop.

Based on the aforementioned discussion, we can conclude that if the strain is above a certain level (0.0012 in this case), as all CNTs inside composite are thoroughly debonded, CNT alignment is not a concern as the composite already provides the maximum energy dissipation.

3 Modeling of Nano-scale Single Crystal Silicon with Atomistic–Continuum Mechanics

Single crystal silicon (SCS) is one of the most common structure materials used in microelectromechanical system (MEMS). Rapid development in the field of MEMS, particularly in the design of mechanical devices, requires an accurate value of the mechanical properties that include Young's modulus, Poisson's ratio, and

tensile strength. Tests of MEMS materials have been carried out for evaluating their mechanical properties, such as tensile tests [25–29], bending tests [30–32]. However, the material properties have been estimated only on a micrometer to millimeter scale structure because of difficulties in fabrication of a nano-scale test specimens and problems associated with measuring ultra-small physical phenomena in an experiment.

Although, the material testing techniques can provide much knowledge, they may not be adequate for the accurate determination of the material properties. Even if these testing methods were possible, the equipment cost would be significant in performing nano-scale tensile testing. An alternate method would be to investigate the nature of deformation and the material properties of nano-scale devices by simulation techniques. Recently, in the nano or atomistic regime, the molecular dynamics (MD) simulation technique has been successfully applied to a variety of phenomena including solid fracture, surface friction, crack propagation, material testing [33], and machining [34]. The essence of the MD simulation method is the numerical solution of Newton's equation on motion for an ensemble of atoms. To render atomistic simulation studies practical, a classical or semi-classical potential function from which interatomic force can be derived is necessary. This is accomplished by using an appropriate empirical potential energy function that satisfies material properties criteria that include the lattice constant, compressibility, elastic constant, cohesive energy. The interatomic forces, involve attractive and repulsive, are defined by an appropriate empirical potential energy function, such as Morse potential for metal, Stillinger–Weber potential for single crystal silicon. The MD simulations can be likened to the dynamic response of numerous nonlinear spring-mass systems under an applied load, velocity, or displacement conditions. From this point of view, the mass represents the atoms (ions) linked one-with-other by spring elements to represent the interatomic bonds (interatomic force). However, MD simulations are limited to very small systems over very short times. Furthermore, in traditional continuum mechanics, the FEM is widely used to model and simulate the mechanical behaviors of solid/discrete body [49], it is a mature technology after decades of development. Compared with MD simulation, the continuum mechanics is considered very efficient and can provide results quickly within an accurate range. Moreover, many researchers successfully employed the traditional continuum theory that directly incorporates the interatomic potential or MD algorithmic into the constitutive model of the solid [35–39]. According to their results, the continuum mechanics theory is applicable on the nanometer scale.

Therefore, this investigation will adopt a hybrid method which combines FEM and the interatomic potential function to explore the mechanical properties of nano-scale single crystal silicon (SCS) under uniaxial tensile loading condition. In the present approach, the minimum potential energy criterion will be applied to individual atoms for obtaining the force–displacement relation of atoms, once the force–displacement curve is determined the analysis could follow the conventional FEM steps to complete the solution procedures. This novel approach is called the atomistic–continuum mechanics (ACM, [48]). Mechanical properties such as the

Young's modulus, coefficient of thermal expansion (CTE), and Poisson's ratio are evaluated from the ACM simulation and compare with the experimental data or other simulation results in the literature.

4 Numerical Simulation Strategy of Atomic and Nano-structures

This section briefly reviews several numerical simulation strategies of the nano-scale materials and nano-structures. Similarly, the benefits and disadvantages of the simulation methods would be addressed. The ab initio method, quantum mechanics (QM), Monte Carlo (MC), molecular mechanics (MM), molecule dynamics (MD), and the ACM methods are included.

4.1 Ab Initio Method

The ab initio method is computational methodology that can be used without any experimental parameters. In general, the Schrödinger equation is used to calculate the quantum description of the atomic systems. Except for some special cases, one normally assumes the Bohr–Oppenheimer approximation in which the motions of nuclei are suppressed.

The time-dependent Schrödinger equation would be expressed as

$$\left(-\frac{1}{2} \nabla^2 - \sum_i \sum_a \frac{Z_a}{|r_i - R_a|} + \sum_{i,j} \frac{1}{|r_i - r_j|} \right) |\Psi\rangle = E |\Psi\rangle \quad (4)$$

where r_i and r_j are the positions of the electrons and R_a and Z_a are the positions and the atomic number of the nuclei, E is the energy, and Ψ is the wave function. The energy may be computed by the Schrödinger equation which in the time-dependent Bohr–Oppenheimer approximation is

$$H\Psi(r_1, r_2, \dots, r_N) = E\Psi(r_1, r_2, \dots, r_N) \quad (5)$$

The wave function Ψ depends on the position and spins of the N electrons. The Hamiltonian operator, H , consists of a sum of three terms: the kinetic energy, the interaction with the external potential (V_{ext}), and the electron–electron interaction (V_{ee}). Therefore, to compare Eqs. (4) and (5), the Hamiltonian operator could be expressed as

$$H = -\frac{1}{2} \nabla^2 - \sum_i \sum_a \frac{Z_a}{|r_i - R_a|} + \sum_{i,j} \frac{1}{|r_i - r_j|} \quad (6)$$

To resolve the issue, the Hartree–Fock theory is by far the most frequently quoted ab initio methods. This theory allows for the Pauli exclusion principle while lets each electron have its own orbital.

The ab initio method also exhibits important limitations. The major drawback of this method is its intense computational cost, which limits the modeled systems to within a certain length and time.

4.2 Monte Carlo Simulation

Related to molecular dynamics are Monte Carlo methods which randomly move to a new trajectory. If this conformation has a lower energy it is accepted, if not, an entirely new conformation is generated. This process is continued until a set of low-energy conformers has been generated in a certain number of times.

4.3 Molecular Mechanics

Molecular mechanics is a classical mechanical model that represents a molecule as a group of atoms held together by elastic bonds. Molecular mechanics looks at the bonds as springs, which can be stretched, compressed, and bent at the bond angles. Interactions between non-bonded atoms also are considered. The sum of all these forces is called the force field of the molecule. To make a molecular mechanics calculation, a force field is chosen and suitable molecular structure values are constructed. Molecular mechanics is a valuable tool for predicting geometries of molecules.

4.4 Molecule Dynamics

The molecular dynamics is developed to study of the properties of atoms system with empirical interatomic potential functions. Thus, the interaction between two atoms is represented by a potential energy. Based on the electronic database or by alternatively using experimental measurements of specific properties, an approximate empirical potential function can be constructed. According to classic Newtonian mechanics, the dynamic evolution of all atoms can be determined by numerical integration. In principle, once the positions and velocities of atoms in the finite region within the simulation core are known, all thermodynamic properties can be readily extracted. A successful simulation depends on three major factors:

- (1) The computational implementation of the MD method.
- (2) The construction of accurate interatomic potential energy.
- (3) The analysis of massive data resulting from computer simulation.

On the other hand, how can MD use Newton's law to move atoms when everyone knows that the system at the atomistic level obeys quantum law rather than classic laws? A simple validation of the classic approximation is based on the de Broglie thermal wavelength [40], which is defined as

$$\Lambda = \sqrt{\frac{h^2}{2\pi M k_B T}} \quad (7)$$

where h represents Plank's constant, M is the atomic mass, k_B denotes the Boltzmann's constant, and T is the temperature. The classic approximation is justified if $\Lambda < a$, where a is the mean nearest neighbor separation of atoms. As such, the quantum effect must be considered to demonstrate the behaviors of nano-scale. For example, the de Broglie thermal wavelength of the silicon atom is the 0.19 Å whose value is smaller than the nearest neighbor separation of 2.53 Å. Therefore, classical mechanics is appropriate to study the motion of the Si atom system in this statement.

The following presents the critical concepts in the numerical implementation of molecular dynamics methods. The MD is used to describe the solution of classic equations of motion for a set of atoms. The microscopic coordinates r_j and momentum P_i of the molecular computed over time may then be used to obtain thermodynamic and transport properties through statistical mechanics. It is assumed that the motion of atoms is governed by Newton's equations of motion of classic mechanics, as given by

$$\frac{dr_i}{dt} = \frac{P_i}{m_i} \quad (8)$$

$$\frac{dP_i}{dt} = F_i \quad (9)$$

where m_i and F_i are the mass and force experienced by atom i as given by

$$F_i = -\frac{\partial V}{\partial r_i} = -\nabla V \quad (10)$$

where V is the potential function. Therefore, all the forces between the particles have been calculated using Eq. (10). The Newton's equation of motion can be integrated to calculate a new velocity and consequently a next position for each atom.

For a system of N atoms, Eqs. (8) and (9) represent $6N$ first-order differential equations that require solving. These equations are always solved using finite different methods such as Verlet algorithms and predictor–corrector algorithms which are the two most common methods. The basic Verlet method is based on position $r(t)$, acceleration $a(t)$, and the position $r(t-\delta t)$ from the previous step. The equations for advancing the position, velocity, and acceleration are

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \quad (11)$$

$$a(t) = -(1/m) \nabla(r(t)) \quad (12)$$

$$v(t) = \frac{r(t + \delta t) - r(t - \delta t)}{2\delta t} \quad (13)$$

where acceleration is substituted using Newton's equation of motion to give force divided by mass. The application of the Verlet algorithm (central difference) to MD simulation is completed in two steps. First, the force is calculated using position $r(t)$ and second, positions $r(t)$ and $r(t-\delta t)$, along with the force determined in step one, are used to calculate the position at $r(t + \delta t)$. These steps are repeated to integrate the motion over a particular time length. To ensure accuracy, the Leap-Frog algorithm (modified central difference) is used for a slight modification of the Verlet algorithm

$$\begin{aligned} r(t + \delta t) &= r(t) + v\left(t + \frac{1}{2}\delta t\right)\delta t \\ v\left(t + \frac{1}{2}\delta t\right) &= v\left(t - \frac{1}{2}\delta t\right) + a\delta t \end{aligned} \quad (14)$$

Accordingly, a basic MD simulation is usually a period of about a few femtoseconds (fs) since the intermolecular or interatomic vibration.

Molecular dynamics is a very powerful technique but has limitations [40].

- (1) Classical description of atomic motion: According to the de Broglie thermal wavelength of Eq. (7), the classic approximation is justified if $\Lambda < a$, where a is the mean nearest neighbor separation of atoms. As such, the quantum effect must be considered to demonstrate the behaviors of nano-scale.
- (2) Realism of forces: In the molecular dynamics, atoms interact with each other. These interactions originate forces with act upon atoms, and atoms move under the action of these instantaneous forces. As the atoms move, their positions change and force changes as well. According the above-mentioned, forces are usually obtained as the gradient of a potential energy function, depending on the positions of the particles. The realism of the simulation therefore depends on the ability of the potential chosen to reproduce the behavior of the material under the conditions at which the simulation is run.
- (3) Time and size limitation:
 - Time scale
The maximum time step of integration in MD simulation is defined by the fastest motion in the system. Vibration frequencies in a molecular system are up to 100 THz which corresponds to a period of about 10 fs. Optical phonon frequencies are 10 THz, period of about 100 fs. Therefore, a typical time step in MD simulation is on the order of a femtosecond.

Using modern computers it is possible to calculate 10^6 – 10^8 time steps. Therefore, it can only simulate processes that occur within 1–100 ns.

- Size scale

The size of the computational model is limited by the number of atoms that can be included in the simulation, typically 10^4 – 10^6 atoms. This corresponds to the size of the computational cell on the order of tens of nanometer.

4.5 Atomistic–Continuum Mechanics

ACM is a novel atomistic level numerical methodology based on the classic finite element theory. Figure 16 shows the schematic of the atomistic–continuum mechanics. In this figure, a simple model for a diamond crystal is that of an orderly array of spheres which represents the atoms linked with one another by elastic bonds (non-linear spring elements) to represent the interatomic bonds (interatomic force). To move one atom aside from its equilibrium position entails stretching/compressing or attractive/repulsive behavior. If one atom is pulled to increase the distance between the atoms then the spring will exert an attractive force on the atom and try to attract it back to its equilibrium position. If the atoms are pushed together then the spring will exert a repulsive force on the atom and restore the atom to its equilibrium position.

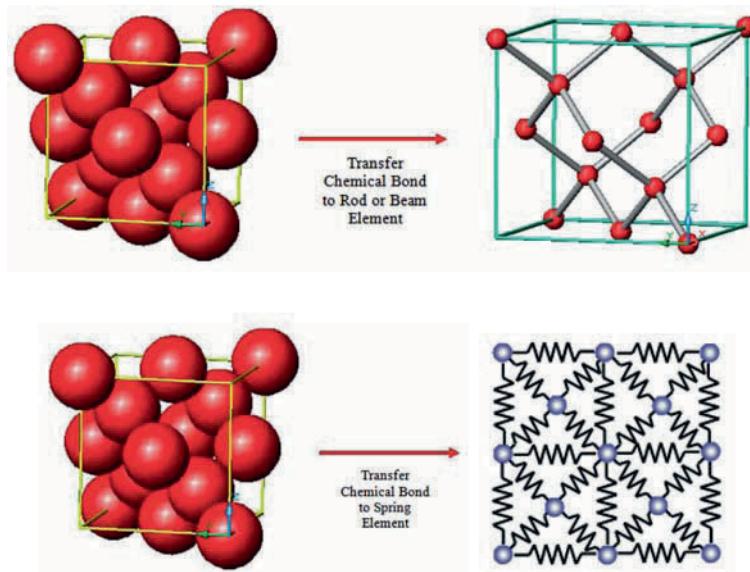


Fig. 16 Schematic of the atomistic–continuum mechanics (ACM)

5 Concept of Atomistic–Continuum Mechanics

Potential functions are developed to simplify the complexity of quantum mechanics-based computation such as ab initio calculations. These functions return a value of energy based on the conformation of the molecules or atoms. Some potentials such as Morse potential [35] and embedded-atom method (EAM [19, 50]) have been proposed to describe the potential energy of metal atoms and they have been successfully applied to estimate the material properties of metals at room temperature in many articles [41–43]. For covalent bonds the most commonly used potential functions are the reactive empirical bond order (REBO) potential [44] for carbon and hydrocarbon, Cornell potential [45] for bio-structure, Stillinger–Weber potential and Tersoff potential [46] for silicon, germanium family. The Stillinger–Weber potential is selected as a benchmark for silicon model in this research.

In general the force–elongation (or so-called force–displacement) curve between two atoms in a crystal is considered to have the form shown in Fig. 17.

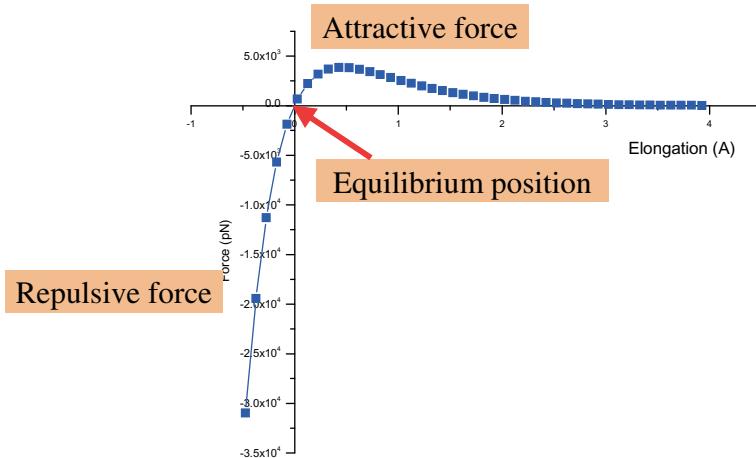


Fig. 17 The force–elongation curve (interatomic potential function for silicon)

The minimization of the total potential energy is considered due to the ACM. Therefore, this section also introduces the principle of minimum potential energy. The total potential energy of an elastic body Π is defined as

$$\Pi = U - W \quad (15)$$

where U is the strain energy, and W is the energy of the external loads, respectively. The strain energy of a linear elastic body is defined as

$$U = \frac{1}{2} \int_V \{\varepsilon\}^T [E] \{\varepsilon\} dV$$

The energy of the external loads can be stated as

$$W = \int_V \{u\}^T \{F\} dV + \{u\}^T \{P\} \quad (17)$$

where the $\{\varepsilon\}$ represents the strain field, $[E]$ represents the material matrix, $\{u\}$ represents the displacement field, $\{F\}$ is the body force, $[D]$ represents the nodal of degree-of-freedom (D.O.F) of the structure, and $\{P\}$ is the load applied to D.O.F by external agencies.

Equations (16) and (17) which could be rewritten as the finite element form for a 3D analysis were given by

$$\begin{aligned} U &= \frac{1}{2} \int_V [D]^T [k] [D] dV \\ W &= \int_V [D]^T [N]^T \{F\} dV + [D]^T \{P\} \end{aligned} \quad (19)$$

where $[k]$ is the element stiffness matrix ($[k] = \int_V [B]^T [E] [B] dV$), the $[B]$ is the strain–displacement matrix, and $[N]$ represents the shape function matrix. The minimization of total potential energy with respect to the nodal displacement requires that

$$\frac{\partial \Pi}{\partial [D]} = 0 \quad (20)$$

According to Eq. (20), the overall structure can now be obtained as

$$[k][D] = \{R\} \quad (21)$$

where $[k] = \sum_{i=1}^n [k_i]$ and $\{R\} = \{P\} + \sum_{i=1}^n [N_i] \{F\}$. The required solution for the nodal displacement and reaction force can be obtained after solving Eq. (21).

Therefore, the atomistic–continuum mechanics is a methodology that the forces are usually obtained as the gradient of a potential energy function, depending on the positions of the particles. The realism of the simulation therefore depends on the ability of the potential chosen to reproduce the behavior of the material under the conditions at which the simulation is run.

The feasible atomistic–continuum mechanics should be established, making it capable of including the chemical bond force in the analysis of the single crystal silicon finite element model. In this investigation, the Stillinger–Weber and Tersoff potentials were used to describe the chemical bond force between the silicon atoms.

Compared with the ab initio method and the molecular dynamics, the ACM based on the classical finite element theory has several benefits. First, the ACM model could predict accurate physical and chemical behaviors of atom and molecule systems with the million or billion atoms. Second, the ACM numerical method based

on the finite element theory could reduce the computational time of the simulation. However, the ACM method has two disadvantages. First, the potential function is based on known properties. If one is interested in the properties of a new type of molecules, an appropriate force field probably may not be available for that type of molecules. Second, because ACM models set at molecules as a group of springs, it cannot be used to predict electronic properties of molecules. This novel simulation method to investigate nano-scale materials is not limited to any materials; it can also be applied to other nano-structured materials once the interatomic potential and the atomic structure of the material are known. Table 8 lists the characteristic comparison of ab initio method, Monte Carlo simulation, molecular mechanics, molecular dynamics, and atomistic–continuum mechanics.

Table 8 The characteristic comparison of many-particles system simulations

	Characteristic	Quantum mechanics included	Degree of freedom of particles	The max. size of many particle systems
Ab initio (QM)	To solve the Schrödinger equation for many particle systems	Yes	High	30–100 atoms
MC	New particle configurations are created randomly step by step	No	Case by case	Case by case
MM	With potential functions, to solve static molecular configuration near equilibrium position	No	Medium	10^5 – 10^6 atoms
MD	Works through the solution of Newton's equations of motion	No	Medium	10^4 – 10^6 atoms
ACM	Transfer interatomic relationship to continuum mechanics, and then solve problem by the finite element method	No	Low	10^5 – 10^7 atoms

6 Interatomic Potential Energy for Silicon

The potential energy function considered for the interaction between the silicon atoms in this investigation is a general form of Stillinger–Weber (S–W) potential. The Stillinger–Weber potential is a combination of two-body and three-body potentials, f_2 and f_3 . The two-body potential describes the formation of a chemical bond between two atoms. The three-body potential represents the angle between two bonds. The two-body potential has a Lennard-Jones type form:

$$f_2(r) = \begin{cases} A(Br^{-p} - r^{-q}) \exp [(r - a)^{-1}], & r < a \\ 0, & r \geq a \end{cases} \quad (22)$$

Note that $f_2(r)$ is a function of the rescaled bond length r only, and it vanishes without discontinuities at $r = a$.

The three-body potential f_3 is given by w

$$f_3(r_i, r_j, r_k) = h(r_{ij}, r_{ik}, \theta_{jik}) + h(r_{ji}, r_{jk}, \theta_{ijk}) + h(r_{ki}, r_{kj}, \theta_{ikj}) \quad (23)$$

where θ_{jik} is the angle between r_j and r_k bond at vertex atom i . Provided that both r_j and r_k are less than the previously introduced cutoff a , it has the following form:

$$h(r_{ij}, r_{ik}, \theta_{jik}) = \lambda \exp \left[\gamma (r_{ij} - a)^{-1} + \gamma (r_{ik} - a)^{-1} \right] \cdot \left(\cos \theta_{jik} + \frac{1}{3} \right)^2 \quad (24)$$

Original the parameters of the S–W potential function were fitting to the simulation results of the total energy and the lattice constant to the experiment data. The energy and the distance units are deduced from the observed atomic energy and lattice constant at 0 K in the silicon diamond structure: $\sigma = 0.21$ nm. The parameters A , B , p , q , a , λ , and γ are constant used in the Stillinger–Weber potential function for single crystal silicon (see Table 9 for detail) [36].

Table 9 Parameters used in the Stillinger–Weber potential function for silicon [36]

Parameter	Parameter		
A	7.049556277	a	1.8
B	0.602224558	Λ	21
p	4	Γ	1.2
q	0		

7 Methodology of Atomistic–Continuum Mechanics

Single crystal silicon forms a covalently bonded structure, the diamond cubic structure. Silicon, with its four covalent bonds, coordinates itself tetrahedrally, and these tetrahedrons make up the diamond-cubic structure. This structure can also be represented as two interpenetrating face-centered cubic (FCC) lattices, one displaced (1/4, 1/4, 1/4) with respect to the other, as shown in Fig. 18. Figure 19a is a schematic of the model used in the ACM simulations of uniaxial tensile testing of SCS. The top and bottom surface are free from any prescribed loads (displacement), as in conventional tensile testing. Furthermore, this study also investigated the different crystallographic plane, whose prescribed loading direction axes were aligned in plane of (100), (110), and (111), respectively. Figure 19b displayed the different crystallographic plane of SCS.

According to the fundamentals of ACM simulation, the covalent bond (interatomic force) between the two atoms was replaced by the nonlinear spring element. Moreover, in order to describe completely the two-body and three-body interaction

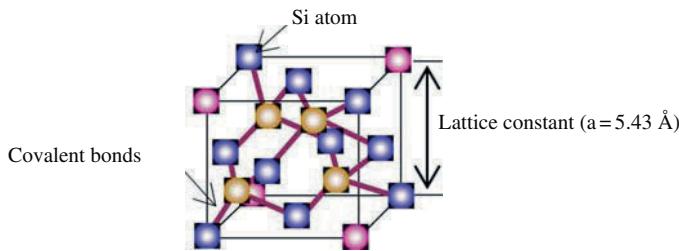


Fig. 18 Single crystal silicon structure

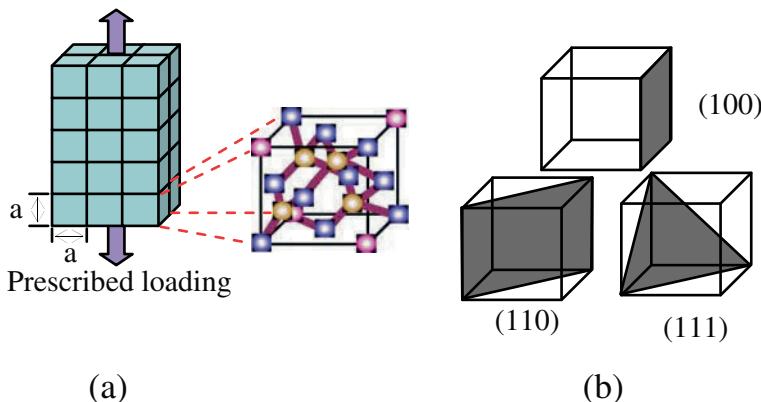
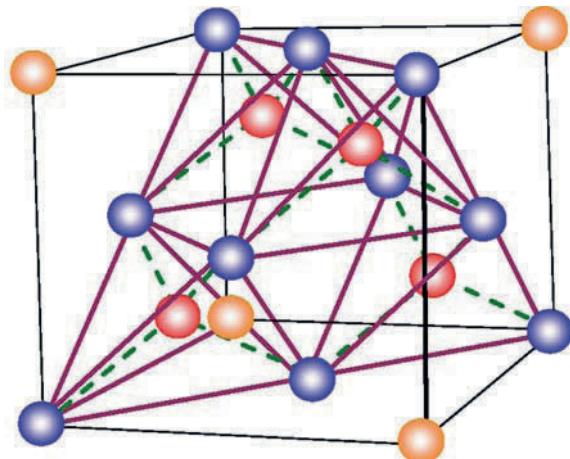


Fig. 19 (a) Schematic of the model used in the ACM simulations of uniaxial tensile testing of SCS ($a = 5.43 \text{ \AA}$: lattice constant). (b) The different crystallographic planes include (100), (110), and (111)

in ACM simulation, this investigation was set to two nonlinear spring elements of two kinds of the different behaviors. One was illustrated the stretch motion, another was defined the angle variation behavior between the two atoms and two covalent bonds, i.e., two-body and three-body term of the Stillinger–Weber potential function, respectively. Therefore, the original single crystal silicon structure in Fig. 18 was reconstructed the new model in ACM simulation, as shown in Fig. 20. Figure 21 showed that the force–elongation curves of the nonlinear spring elements of the two-body and three-body potential function. The relationship between the force and elongation was derived from the gradient of the potential function with respect to the position of atom, i.e.,

$$F(r) = -\nabla f(r) \quad (25)$$

Fig. 20 The modification of single crystal silicon model.
 (Solid line: three-body term of S-W potential function;
 Dotted line: two-body term of S-W potential function)



8 Atomistic–Continuum Mechanics Models

In order to study the mechanical properties of the nano-scale single crystal silicon, the finite element software was adopted to simulate its behavior. In this investigation, the commercial software ANSYS[®] was applied to analyze the mechanical performance of the nano-scale single crystal silicon.

The half of finite element model of nano-scale single crystal silicon of the (110) crystallographic plane were established in Fig. 22, since the single crystal silicon is a half symmetry. The boundary conditions of the top surface were the prescribed displacement along y -direction, the x - and z -direction fixed and the y -direction fixed on the bottom surface. Table 10 gives the dimension of ACM models, loading, etc.

9 Results and Discussion

In the following, the mechanical behaviors of the nano-scale single crystal silicon on different crystallographic plane under various prescribed displacement will be presented.

9.1 Young's Modulus Estimation

Table 11 listed the simulated Young's modulus compared with those values in the literature measured for bulk silicon [25, 26, 30, 47]. It could be observed that the estimated Young's modulus of SCS along the (100), (110), and (111) plane are in agree with the values reported in the literature based on macro- and microscale

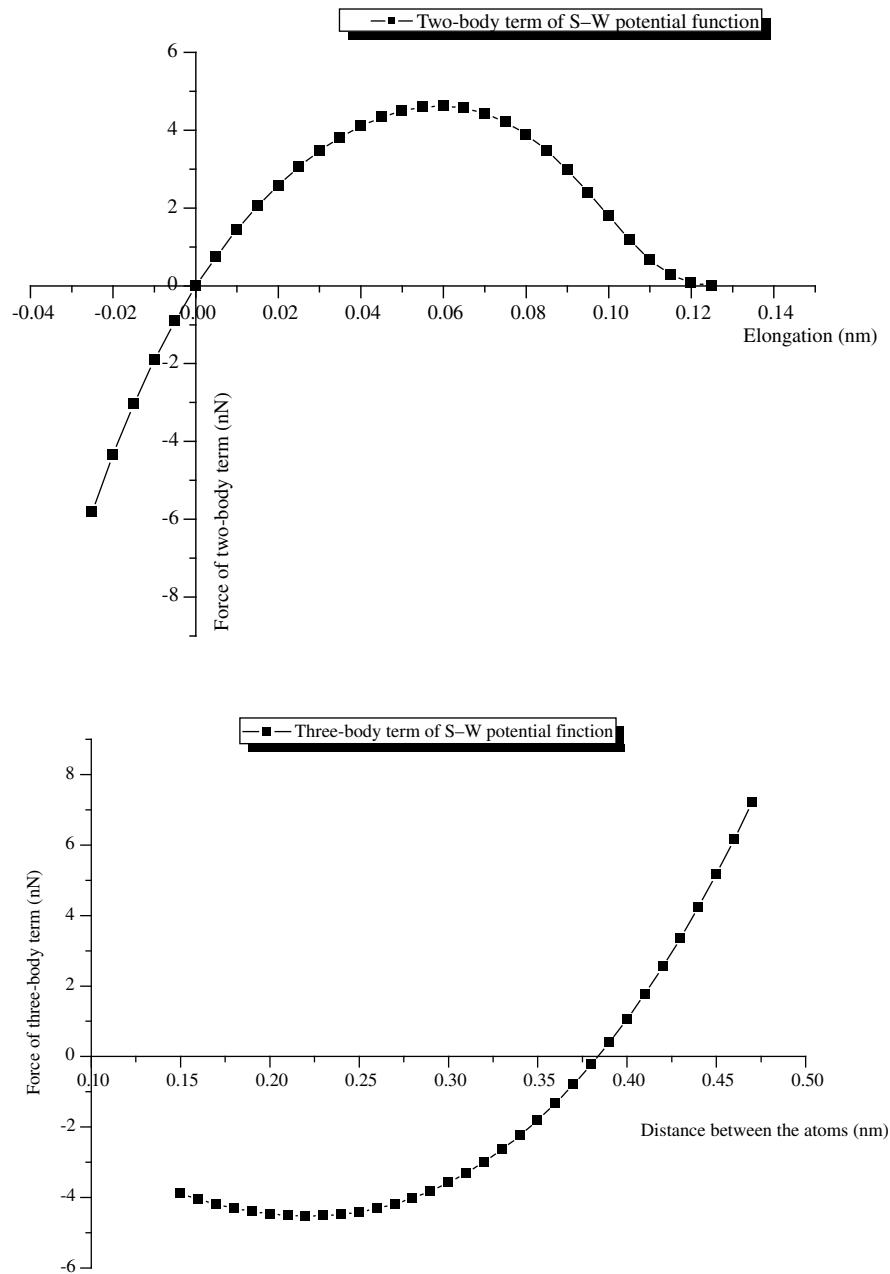


Fig. 21 Force–elongation curves of the nonlinear spring elements of the two-body and three-body term of S–W potential function

Fig. 22 The half of finite element model of nano-scale single crystal silicon of the (110) crystallographic plane

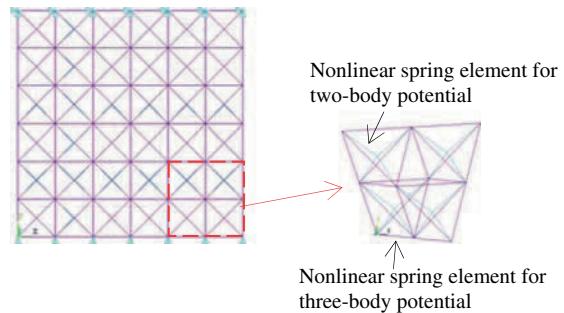


Table 10 ACM simulations conditions used for the tensile loading

Materials	Single crystal silicon
Potential used	Stillinger–Weber potential
Simulation model	Three dimension
Simulation size	$3 \times 5 \times 3a$ (a is the lattice constant)
Prescribed displacement	0.000543–0.00543 nm
Strain	0.1–1%
Crystallographic plane	(100), (110), and (111)

experimental work. The errors of the Young's modulus on the different crystal plane were 5, 4, and 5%, respectively. There are two main reasons to explain why the Young's modulus of simulation results agreed with the experiment data. First, the single crystal silicon was a defect free, no impurity structure as well as its atoms were also arranged regular and perfect. Therefore, the silicon material was identical in behavior no matter in simulation method or experiment testing. Second, the parameters of Stillinger–Weber potential function (listed in Table 11) were accurate, since the parameters fitted and deduced from the experiment data.

Table 11 Young's modulus estimation of SCS (GPa)

Direction	(100)	(110)	(111)
Brantley [47]	130.2	169.9	187.5
Wilson and Beck [30, 31]	130	161	—
Yi et al. [26]	—	169.2	—
Sato [25]	125	140	180
Present study	121.8	153	174.6

9.2 Effect of Prescribed Displacement (Strain) on the Tensile Response

ACM simulations of uniaxial tension were conducted at various prescribed displacement for SCS to investigate the effect of strain on the mechanical properties.

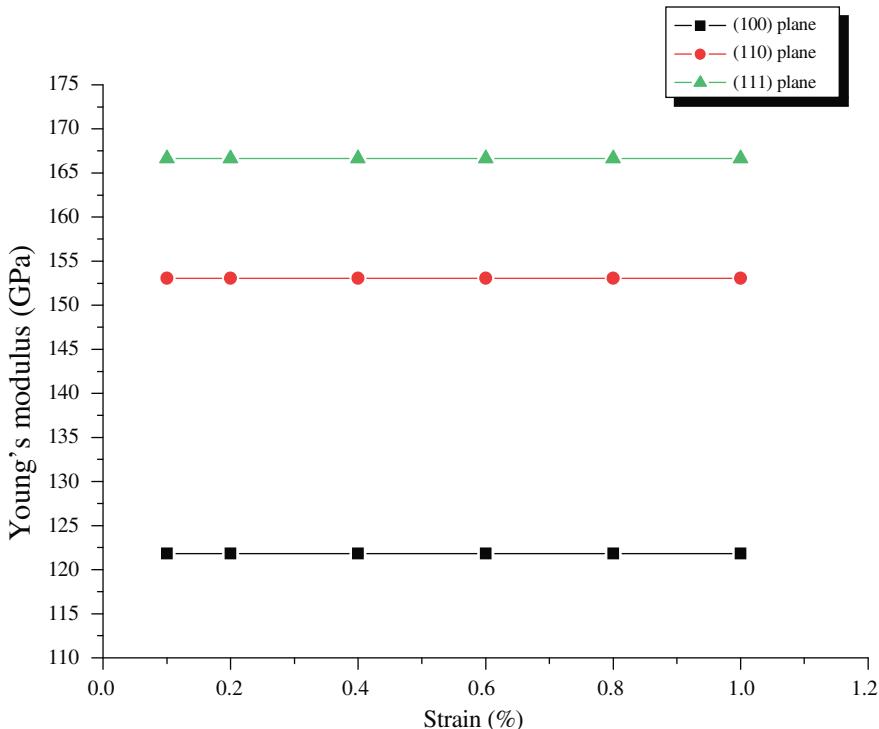


Fig. 23 Effect of various prescribed displacement (strain) on the Young's modulus for SCS

Figure 23 showed that the effect of various prescribed displacement (strain) on the Young's modulus for nano-scale SCS. It could be observed that the Young's modulus was nearly constant and relatively insensitive to the strain. According to the literature [33] also revealed that the elastic modulus for silicon was not affected by the loading rate. The foregoing results were laid on the strain lower than 1%. However, while the strain exceeded 1%, the SCS would probably have reached the fracture strain and then the SCS structure would be directly fractured. Since the SCS was a brittle material, there was no necking effect before failure, unlike the metal materials.

10 Summary and Conclusions

10.1 Modeling of Carbon Nanotube Composites for Vibration Damping

In this research, an FEM composite unit cell model was developed, and the loss factor of the CNT composite was calculated and compared with the experimental results. The simulation curve fits well in general with experimental data. It could be observed from simulation results that a composite loss factor is related to the strain

level, CNT dimension, CNT weight ratio, and alignment of CNT segments. As the location, size, and shape of CNT/epoxy interfacial sliding area could be simulated by using this model, before composite fabrication, CNT alignment for higher damping under different loading level could be identified. Furthermore, the upper bound of the CNT composite loss factor could be estimated by using the developed model. In parametric studies, several factors are discussed while composite energy dissipation ability is taken into consideration. Parameters such as CNT length, diameter, and alignment orientation are investigated. It is found that the composite loss factor is insensitive to CNT segment length but highly sensitive to CNT diameter. The interfacial sliding developed between CNT and matrix could be categorized into five different stages under different strain level while the CNT alignment orientation for larger composite loss factor was identified.

10.2 Atomistic–Continuum Mechanics Models

This research presents a novel mechanics for coupling continuum mechanics and interatomic potential function (Stillinger–Weber potential) to predict the mechanical behavior of nano-scale single crystal silicon under uniaxial tensile loading. According to this methodology, once the interatomic potential and the atomic structure of the material are known, a systematic approach, which does not involve any parameter fitting, has been proposed. The measured Young's modulus of SCS were 121.8, 153, and 174.6 GPa along the (100), (110), and (111) crystallographic plane, respectively, which are in reasonable agreement with the results reports in the literature based on macro- and microscale experiment. The effect of strain through the variation of prescribed displacement on the measured mechanical properties was investigated. It could be observed that the Young's modulus were nearly constant and relatively insensitive to the strain (0.1–1%).

This novel simulation mechanics to investigate the nano-scale materials was not limited to single crystal silicon; it can also be applied to other nano-structured materials once the interatomic potential and the atomic structure of the material are known. In the future, the author would explore the size effects of the nano-scale single crystal silicon using the atomistic–continuum mechanics with Stillinger–Weber potential function.

Acknowledgments The work presented in this chapter was sponsored by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK4144/07E), the Research Boost Program of National Tsing Hua University (Project No. 98N2914E1), and the National Nano Science and Technology Project of National Science Council, Taiwan (Project No. NSC98-2120-M-194- 002). The authors wish to acknowledge these supports.

References

1. Iijima S. Helical microtubules of graphitic carbon. *Nature* 1991;354:56–58.
2. Yu M.F., O. Lourie, M.J. Dyer. Strength and breaking mechanism of multiwalled carbon nanotubes under tensile load. *Science* 2000;287:637–640.

3. Thostenson E.T., Z.F. Ren, T.W. Chou. Advances in the science and technology of carbon nanotubes and their composites: a review. *Composites Science & Technology* 2001;61: 1899–1912.
4. Ajayan P.M., O. Stephan, C. Colliex, D. Trauth. Aligned carbon nanotube arrays formed by cutting a polymer resin-nanotube composite. *Science* 1994;256:1212–1214.
5. Qian D., E. Dickey, C.R. Andrews, T. Rantell. Load transfer and deformation mechanisms in carbon nanotube-polystyrene composites. *Applied Physics Letters* 2000;76:2868–2870.
6. Ruan S.L., P. Gao, X.G. Yang, T.X. Yu. Toughening high performance ultrahigh molecular weight polyethylene using multiwalled carbon nanotubes. *Polymer* 2003;44:5643–5654.
7. Lau K.T., S.Q. Shi, L.M. Zhou, H.M. Cheng. Micro-hardness and flexural properties of randomly-oriented carbon nanotube composites. *Journal of Composite Materials* 2003;37:365–376.
8. Koratkar N., B. Wei, P.M. Ajayan. Carbon nanotube films for damping applications. *Advanced Materials* 2002;14:997–1000.
9. Koratkar N., B. Wei, P.M. Ajayan. Multifunctional structural reinforcement featuring carbon nanotube films. *Composites Science & Technology* 2003;63:1525–1531.
10. Lass E.A., N.A. Koratkar, P.M. Ajayan, B.Q. Wei, P. Kebinski. Damping and stiffness enhancement in composite systems with carbon nanotubes films. *Proceedings of Materials Research Society Symposium* 2002;740:371–376.
11. Lass E., P. Ajayan, N. Koratkar. Engineered connectivity in carbon nanotube films for damping applications. *Proceedings of SPIE* 2003;5052:141–150.
12. Wong M., M. Paramsothy, X.J. Xu, Y. Ren, S. Li, K. Liao. Physical interactions at carbon nanotube-polymer interface. *Polymer* 2003;44:7757–7764.
13. Zhou X., E. Shin, K.W. Wang, C.E. Bakis. Interfacial damping characteristics of carbon nanotube based composites. *Composites Science & Technology* 2004;64:2425–2437.
14. Suhr J., N. Koratkar, P. Kebinski, P. Ajayan. Viscoelasticity in carbon nanotube composites. *Nature Materials* 2005;4:134–137.
15. Liu A., J.H. Huang, K.W. Wang, C.E. Bakis. Effects of interfacial friction on the damping characteristics of composites containing randomly oriented carbon nanotube ropes. *Journal of Intelligent Material Systems & Structures* 2006;17:217–229.
16. Liu A., K.W. Wang, C.E. Bakis, J.H. Huang. Analysis of damping characteristics of a viscoelastic polymer filled with randomly oriented single-wall nanotube ropes. *Proceedings of SPIE International Symposium on Smart Structures and Materials* 2006;6169:275–291.
17. Zhou X., K.W. Wang, C.E. Bakis. Effects of interfacial friction on composites containing nanotube ropes. *Proceedings of 44th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference* 2004;AIAA–2004–1784.
18. Dai R.L., W.H. Liao. Experimental studies on carbon nanotube composites for vibration damping. *Proceedings of 17th International Conference on Adaptive Structures and Technologies* 2006.
19. Odegard G.M., T.S. Gates, K.E. Wise, C. Park, E.J. Siochi. Constitutive modeling of nanotube-reinforced polymer composites. ICASE Report 2002;NASA/CR-2002-211760:No. 2002–2027.
20. Liu Y.J. X.L. Chen. Continuum models of carbon nanotube-based composites using the boundary element method. *Electronic Journal of Boundary Elements* 2003;1:316–335.
21. Inc ANSYS Theory Reference 2005;10.0.
22. Andrews R., M.C. Wiesenberger. Carbon nanotube polymer composites. *Current Opinion in Solid State & Materials Science* 2004;8:31–37.
23. Cifuentes A.O., A. Kalbag. A performance study of tetrahedral and hexahedral elements in 3-D finite element structural analysis. *Finite Elements in Analysis & Design* 1992;12:313–318.
24. Bussler M., A. Ramesh. The eight-node hexahedral element in FEA of part designs. *Foundry Management & Technology* 1993;121(11):26–28.
25. Sato K., T. Yoshioka, T. Ando, M. Ahilida, T. Kawabata. Tensile testing of silicon film having different crystallographic orientation carrier out on a silicon chip. *Sensor & Actuator A* 1998;70:148–152.

26. Yi T., L. Lu, C.J. Kim. Microscale material testing of single crystalline silicon: process effects on surface morphology and tensile strength. *Sensor & Actuator A* 2000;**83**:172–178.
27. Sharpe W.N., B. Yuan, R.L. Edwards. A new technique for measuring the mechanical properties of thin films. *Journal of Microelectromechanical System* 1997;**6**:193.
28. Li X., T. Kasai, S. Nakao, T. Ando, M. Shikida, K. Sato, H. Tanaka. Anisotropy in fracture of single crystal silicon film characterized under uniaxial tensile condition. *Sensor & Actuator A* 2005;**117**:143–150.
29. Tsuchiya T., O. Tabata, J. Sakata, Y. Taga. Specimen size effect on tensile strength of surface-micromachined polycrystalline silicon thin film. *Journal of Microelectromechanical System* 1998;**7**:106.
30. Wilson C.J., P.A. Beck. Fracture testing of bulk silicon microcantilever beams subjected to a side load. *Journal of Microelectromechanical System* 1996;**5**:142–150.
31. Wilson C.J., A. Oremggi, M. Narbutovskih. Fracture testing of silicon microcantilever beams. *Journal of Applied Physics* 1996;**79**:2386.
32. Namazu T., Y. Isono, T. Tanaka. Evaluation of size effect on mechanical properties of single crystal silicon by nanoscale bending test using AFM. *Journal of Microelectromechanical System* 2000;**9**:450–459.
33. Komanduri R., N. Chandrasekaran, L.M. Raff. Molecular dynamics simulation of uniaxial tension at nanoscale of semiconductor materials for micro-electro-mechanical systems (MEMS) application. *Materials Science and Engineering A* 2003;**34**:58–67.
34. Komanduri R., L.M. Raff. A review on the molecular dynamics simulation of machining at the atomic scale. *Proceedings of Institution Mechanical Engineers Part B* 2001;**215**:1639–1672.
35. Girifalca L.A., V.G. Weizer. Application of morse potential function to cubic metals. *Physical Review* 1959;**114**:687–690.
36. Stillinger F.H., T.A. Weber. Computer simulation of local order in condensed phases of silicon. *Physical Review B* 1985;**31**:5262–5271.
37. Zhang P., Y. Huang, P.H. Geubelle, P.A. Klin, K.C. Hwang. The elastic modulus of single-wall carbon nanotubes: a continuum analysis incorporating interatomic potentials. *International Journal of Solid and Structures* 2002;**39**:3893–3906.
38. Xiao S.P., T. Belytschko. A bridging domain method for coupling continua with molecular dynamics. *Computational Methods for Applied Mechanical Engineering* 2004;**193**:1645–1669.
39. Wagner G.J., W.K. Liu. Coupling of atomistic and continuum simulations using a bridging scale decomposition. *Journal of Computational Physics* 2003;**190**:249–274.
40. Ercollessi F.A., P. Dynamics (website: <http://www.sissa.it/furio/>)
41. Chiang K.N., C.Y. Chou, C.J. Wu, C.A. Yuan. Prediction of the bulk elastic constant of metals using atomic-level single-lattice analytical method. *Applied Physics Letters* 2006;**88**:171904.
42. Jeng Y.R., C.M. Tan. Computer simulation of tension experiments of a thin film using atomic model. *Physics Review B* 2002;**65**:174107.
43. Foiles S.M., M.S. Daw. Calculation of the thermal expansion of metals using the embedded-atom method. *Physics Review B* 1988;**38**:12643.
44. Brenner D.W., O.A. Shenderova, J.A. Harrison, S.J. Stuart, B. Ni, S.B. Sinnott. A second-generation reactive empirical bond order potential energy expression for hydrocarbons. *Journal of Physics: Condensed Matter* 2002;**14**:783–802.
45. Cornell W.D., P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of American Chemical Society* 1995;**117**:5179–5197.
46. Tersoff J. New empirical model for the structure properties of silicon. *Physical Review Letters* 1986;**56**:632–635.
47. Brantly W.A. Calculated elastic constants for stress problems associated with semiconductor devices. *Journal of Applied Physics* 1973;**44**:534–535.

48. Chiang K.N., C.A. Yuan, C.N. Han, C.Y. Chou, Y. Cui. Mechanical characteristic of ssDNA/dsDNA molecule under external loading. *Applied Physics Letters* 2006;88: 023902–023903.
49. Chiang K.N., C.H. Chang, C.T. Peng. Local-strain effects in Si/SiGe/Si islands on oxide. *Applied Physics Letters* 2005;87:191901–191903.
50. Daw M.S., M.I. Baskes. Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Physics Review B* 1984;29:6443.

Index

A

- AFM, 60, 62, 76, 186–196, 204–205, 209–210, 324–325, 568, 580, 660, 695–717
Anisotropic conductive adhesives (ACAs), 5–6, 21, 23–24, 32–42
Atomic force microscope, 186, 322, 325, 565, 580, 637, 660, 696, 698–701
Atomistic scale, 14, 719–755

B

- Bio-assay, 49, 78–80
Biocompatibility, 147, 506, 514, 518, 522, 531, 542–550, 603, 605
BioMEMS, 505–550, 567
Biomimetic nanosurface, 47
Bio packaging, 1–15
Biosensor, 150, 510, 513, 521–522, 584, 587, 613–661
Bonding, 3, 7, 21, 23–24, 89, 102, 105–106, 118, 339–343, 464–465, 531, 534, 567–568, 584, 589, 592, 596–597

C

- Capacitors, 9, 158–161, 250, 255–256, 258–265, 271
Carbon nanotubes (CNT), 1–3, 29–30, 90–98, 282, 290, 302–308, 640, 644–645, 649–651, 687–689, 720–722, 754–755
Carbon nanotubes interfaces, 1
CHEMFET, 568
Chemical sensor, 567–593, 607
Chip-to-world connection, 505
Conductive adhesives, 5, 19–42, 148, 150–152, 278, 280, 286, 293–294, 297, 678
Current-carrying capability, 5–6, 24, 33, 41, 124

D

- Depression, 2, 4–5, 13, 32, 220–222, 225, 229, 232, 236, 244

Digital image correlation (DIC), 695–717

1D Nanowire electrode, 167–182
Disposable, 507–508, 511–513

E

- Electrical interconnect, 2, 538–540
Electrically conductive adhesives (ECAs), 20–22, 41, 280, 297, 678
Electrical and thermal, 2–3, 13, 26, 29, 35, 417, 468, 721
Electronic cooling, 358
Electronic packaging, 3, 7, 97–98, 113, 115, 244, 249, 258, 297, 509, 666, 689, 712
Electronics cooling, 431
Experimental techniques, 13, 315, 448–463

F

- Fast charge, 153, 168
Fine-pitch interconnection, 20
Flow boiling, 358, 418, 436, 438, 440–442, 447
Flow condensation, 445
Fluid handling, 507, 511, 529, 531
Fluidic interconnect, 468–469, 528–529, 540, 602
Fluorescence resonance energy transfer, 482, 484, 486, 592

G

- Gas sensor, 2, 168, 567–568, 574, 578–581, 594, 596, 645
Graphene, 2, 19, 88–92, 119–120, 282, 303, 316, 322, 644

H

- Hairpin probe, 482–486, 490–491, 498–499
Heat transfer, 118, 279, 282, 286, 296–297, 300, 354, 356–357, 377–378, 401–402, 417–423, 431–472, 684, 687

Hot spots, 351, 355, 357, 360, 374, 377, 385, 387, 397, 417, 416, 443

I

Inductor, 158, 255–256, 265–269

Integrated passives, 247–273

Integration, 7, 112, 212, 255–256, 259, 273, 357, 454, 472, 508, 510–511, 516–517, 526, 528–530, 549, 567, 589, 594, 603, 617–618, 658, 661, 744

Interconnection, 2, 11, 19–42, 124, 222, 269, 468–469, 529, 536, 541, 677–680, 714, 716–717

Interface, 2, 12, 34, 50–54, 87–124, 153, 194–196, 211, 278, 306, 315–346, 361, 376, 390–391, 413–414, 450, 471, 512, 531, 538–539, 596, 618, 656, 680–683, 728

ISFET, 568–574, 605

Isotropically conductive adhesives (ICA), 21, 25–32

L

Lab-on-a-chip, 400, 446, 614

Large-scale integration, 529–530

Laser-induced fluorescence, 449, 457, 459–461

Li-ion, 152–153, 156, 158, 167–168, 177

Lithium-ion batteries, 148, 152–158, 167

Live cell, 479–499

Lotus surface, 10–12

M

Materials

biological, 624–628

bioMEMS, 544–548

1D nanowire electrode, 167–182

electrically conductive contact surface, 649

low surface energy, 52–53

nanocomposite-embedded capacitor, 258–265

nanocomposite-embedded inductor, 265–269

nanoparticle-based embedded resistor, 269–273

nano-scale and atomistic-scale modeling of advanced, 719–755

nonlinear optical, 161–164

packaging, 542–543

sealing, 600–601

silicon thermoelectric, 372–376

thermal interface, 87–124, 315–346

Medical diagnostics, 480, 522

Melting point, 2–5, 13, 29, 104, 219–229, 232–233, 235–238, 278, 677

MEMS, 11, 49, 76–77, 252, 355, 366, 509, 514–516, 533, 536–540, 585, 594, 607, 739–740

Microcantilever, 579–582, 586

Microchannel, 88, 355–357, 404, 418, 431–472, 510, 590

Microchannel flows, 471

Microcoolers, 359–401

Microelectrode, 522, 621–623

Microelectronics, 1–3, 5, 20, 40, 57, 167–182, 278, 370, 372, 471, 527, 536, 539, 619

Microfabrication, 356, 373, 464–470, 534, 536, 543–544, 601, 661

Microfluidics, 49, 77–80, 446, 456, 508, 531, 535, 540, 598, 617, 657–658

Microhotplate, 574–579

Microneedle, 514, 543, 588–689, 592

Micro-PIV, 449–457, 457–459

Modeling, 13, 106, 120–121, 283, 317, 378, 438, 447, 690, 719–755

Molecular beacon, 481–499

Molecular dynamics, 96, 307, 447–448, 665–690, 740, 742–744, 747–748

Molecular simulations, 672

N

Nanocomposites, 2, 7–10, 120–121, 148, 158–161, 258–273, 277–308

Nanoconductive fillers, 19

Nanoelectronics, 41–42, 90, 119–120

Nano-embedded capacitors, 1

Nanogenerator, 186, 195–209

Nano interconnect, 87

Nanomaterials, 1–15, 87–124, 139–165, 265, 370, 617, 649

Nanoparticles, 3–5, 9, 13, 26, 28, 102–103, 220–222, 237–244, 247–273, 447, 654, 709

Nanoparticle synthesis, 225

Nanopiezotronics, 203

Nanopowders, 141–143, 146–147, 150–151, 154–155, 164

Nano-scale, 13, 318, 462, 719–755

Nanosilver, 19, 25, 281, 297, 300

NanoSpray, 139–165

Nanostructure, 11, 71, 141, 168, 720

Nanotechnology, 1–2, 20, 24, 41, 104, 149, 256–273, 567, 622, 640, 660

Nano texture, 60, 78

Nanowires, 19, 24, 38, 167–182, 185–215, 307–308, 508, 576–577, 645

Nccvc, 141–150, 155, 157, 164

NCCVD, 142–143, 149–150, 164

Nonlinear optical materials, 161–164
Numerical techniques, 447–448

O

Oligonucleotide probe, 481–483, 485–486,
490–491, 494, 510, 621, 648
Optical transparency, 61, 67, 524

P

Packaging, 1–15, 23–24, 41, 49, 75, 88, 99,
103, 123, 201, 253, 258, 259, 297,
505–550, 557, 564, 565–607, 613–661,
665–690
PDMS, 59, 462–463, 508, 512, 521–522,
529–530, 535–537, 540, 545–546, 572,
584, 696–601, 657, 675–676
Percolation threshold, 20–21, 25–27, 30, 39,
159, 278, 281, 304
Piezoelectricity, 191, 636
Point-of-care diagnostic, 522

R

Resistor, 158, 249–250, 255–256, 269–273,
526, 577, 587
RNA detection, 482–487, 490, 495, 498

S

Self-assembled monolayer (SAM), 2, 34, 41,
116, 634, 654, 682
Simulation, 13–14, 38, 228, 283, 307, 329,
380, 391, 418–419, 447–448, 665–666,
669–670, 676–677, 684, 687, 690,
719–720, 727–731, 740–741, 744–745
Sintering, 26, 29, 31–33, 35, 147, 298, 365
Soft lithography, 510, 529, 589,
597–600, 654
Solar cell, 10–11, 49, 69, 77–78
Solar cell packaging, 49

Stacked microchannels, 432–433, 471
Stereolithography, 586, 601–605
Superhydrophobic surface, 47, 49–52, 57–77,
80–81
Surface modification, 7–8, 64, 70, 143, 149,
512, 516, 543–549, 618, 649–651

T

Thermal conductivity, 7, 21, 30, 35, 89, 96–98,
103–108, 119–120, 278–284, 286–298,
351, 353–356, 359, 404–406, 409,
414–415, 417, 420–421, 436, 471, 524,
580, 684–687

Thermal conductivity measurement, 286–298,
305

Thermal interface materials (TIM), 2, 29,
87–124, 150, 278, 281, 296, 315–346,
357, 390, 401, 417–418, 687

Thermally conductive composites, 280–282,
286, 291, 298, 301, 303, 306

Thermal management, 7, 88–89, 115, 316,
349–423, 471–472, 603

Thermal property measurements, 87

Thermo-compression bonding (TCB), 23–24

Thermoelectric coolers, 358–361, 365–372,
386, 389, 423

Tin oxide, 39, 144, 169, 524, 531, 569, 574,
577–578, 596

Total internal reflection fluorescence, 461

Transparent electrodes, 19

Two-phase flow regimes, 436–438

Two-phase flows, 355, 357, 431, 436–438,
442, 447

Z

ZnO, 144, 186–191, 193–197, 199, 201–203,
205, 209–214