

# 材料基因组计划概述

赵朝阳<sup>1)</sup>

1) 北京科技大学……国家材料服役安全科学中心, 北京 100083

✉通信作者, E-mail:2199474541@qq.com

**摘 要** 材料是社会发展的物质基础, 材料创新是各种颠覆性技术革命的核心<sup>[1]</sup>。为了加快新材料的研发速度, 总目标是“将先进材料的发现、开发、制造和使用的速度提高一倍”的“材料基因组计划”应运而生。材料基因组计划创新基础设施包括三个平台: 高通量计算工具平台、高通量实验工具平台和数字化数据平台<sup>[2]</sup>。本文主要对材料基因组计划的三个板块进行初步的介绍。

**关键词** 材料设计; 材料基因组计划; 高通量实验平台; 高通量计算平台; 材料数据库;

**分类号** UNKNOWN

## Material genome project overview

ZHAO Zhaoyang<sup>1)</sup>

1) National Center for Material Service Safety, University of Science and Technology Beijing, Beijing 100083, China

✉Corresponding author, E-mail:2199474541@qq.com

### ABSTRACT

Material is an important material basis of social development, and material innovation is the core of various disruptive technology revolution<sup>[1]</sup>. In order to accelerate the pace of research and development of new material, a “Material Genome Initiative(MGI)” with overall goal of “doubling the speed of discovery, development, manufacture and use of advanced material” comes on board.

The innovative infrastructure of MGI includes three platforms: high-throughput computing platform, high-throughput experimental tool platform and digital data platform<sup>[2]</sup>. This passage mainly focuses on the brief introduction of three platforms of MGI.

**KEY WORDS** material design; material genome initiative; high-throughput computing platform; high-throughput experimental tool platform; material database

## 1 材料基因工程

材料是社会发展的物质基础, 材料创新是各种颠覆性技术革命的核心, 以至于人类文明的进程经常以所使用的材料的类型作为一个时代的标志, 如石器时代、青铜器时代等。新材料技术是体现一个国家科技发展水平的关键标志之一。

长期以来, 欧美发达国家在新材料领域从未停止发展的脚步, 一直处于领先地位。进入 21 世纪以来, 他们越发地意识到依赖于科学直觉与试错的传统材料研究方法, 已跟不上当今技术快速发展的需求, 成为限制社会进步的瓶颈。革新材料研发方法, 加速材料从研究到应用的进程成为各国共同的需求。作为美国政府“先进制造伙伴计划”(Advanced Manufacturing Partnership, AMP)的重要组成部分, 美国总统奥巴马在 2011 年 6 月提出了“材料基因组计划”(Materials Genome Initiative, MGI), 其目的是利用近年来在材料模拟计算、高通量实验和数据挖掘方面取得的突破, 将材料从发现到应用的速度至少提高一倍, 成本至少降低

一半, 发展以先进材料为基础的高端制造业, 从而继续保持美国在核心科技领域的优势。MGI 展现了未来先进材料开发的崭新模式。材料基因组(Materials Genome)这个名词的出现有感于人类基因组计划的成功, 但迄今为止并无特定的科学定义, 仅作为这种新型材料研发模式的代称。

### 1.1 变革传统材料研发模式是国际新材料研究的趋势

根据 MGI 白皮书<sup>[3]</sup>, 在过去的数十年中, 美国持续对于先进材料实验技术和方法的研究作了巨大投入, 发展出一系列先进材料研发实验工具, 包括高通量材料制备和表征设备与技术。MGI 意在对美国已有的材料研究设备和力量进行有效整合, 充分利用超级计算机和大数据技术, 对材料研究的模式进行全面改革, 使这些先进的实验工具发挥更大的效能。MGI 的具体措施包括:

- (1) 发展计算工具和方法, 减少耗时费力的实验, 加快材料设计和筛选;
- (2) 发展和推广高通量材料实验工具, 对候选材料进行筛选和验证, 快速、大量、准确地取得材料计算所需的关键数据;
- (3) 发展和完善材料数据库/信息学工具, 有效管理材料从发现到应用全过程数据链;
- (4) 改革多年来材料界形成的一家一户式的封闭型工作方式, 培育开放、协作的新型合作模式。MGI 的终极目标是通过理论模拟和计算完成先进材料的“按需设计”并实现全程数字化制造。

MGI 是近年来美国发展先进制造技术数字化、智能化趋势在材料科学与工程上的具体体现。2008 年, 美国国家研究理事会出版的一份报告中, 提出要建立一门新学科——集成计算材料工程(Integrated Computational Material Engineering, ICME), 通过将材料计算工具与其他工程领域中计算与分析工具获得的信息相集成, 使材料预测进入产品设计流程, 从而解决材料开发周期日益赶不上产品开发周期的矛盾。而 MGI 将 ICME 的理念扩展到了整个材料科学、技术与工程链条, 贯穿于从新材料发现到应用的全部过程。MGI 力图打破材料科学与材料工程长期脱节的现状, 改变材料行业数据信息相互封闭的传统。这些传统文化被视为造成新材料从发现到应用耗时冗长的重要原因。应该说, MGI 在文化转变方面所面临的挑战丝毫不亚于技术层面。

美国 MGI 的核心内容是由数家联邦机构牵头构建新材料创新基础设施, 包括高通量材料计算平台、高通量材料制备与检测平台和材料数据库。将这些基础设施与现有的集成计算材料工程无缝衔接, 实现多学科交叉和融合, 从而加速新材料从发现到应用的进程。MGI 实施 3 年以来, 获得了美国材料界的积极响应, 取得了较快的进展。截至 2014 年 6 月, 美国联邦政府用于材料创新基础设施建设的财政投入累计已达 2.5 亿美元, 地方政府、大学、企业及其他非政府机构累计资金投入已超过 3 亿美元, 新建 MGI 协同创新中心逾 20 个, 大型协同创新合作计划近 10 项。在 2014 年 12 月出版的一份报告中, 白宫正式将 MGI 定义为“战略计划”, 充分表明 MGI 在美国国家发展战略中的地位及作用(图 1)。

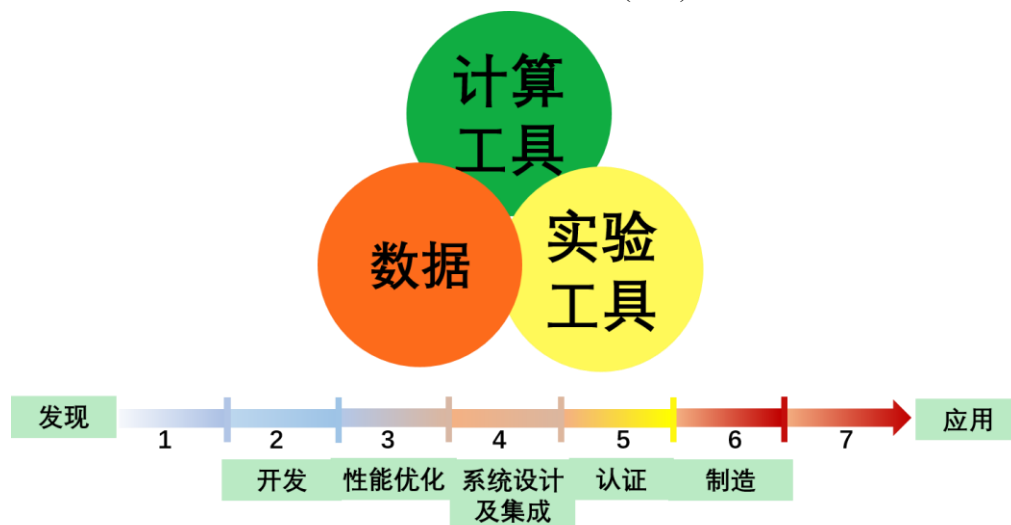


图 1 材料基因组技术三要素协同融合贯穿材料从发现到应用的各个环节

Fig.1 Three elements of MGI run through all aspects of the process from discovery to application synergistically

在美国启动 MGI 的同时, 欧盟以轻量、高温、高温超导、热电、磁性及热磁、和相变记忆存储六类高性能合金材料需求为牵引, 于 2011 年启动了第 7 框架项目“加速冶金学”(Accelerated Metallurgy, ACCMET) 计划。考虑到计算材料学目前尚不具有预测所有材料性能的能力, 项目组织了包括材料需求与制造企业、仪器设备商、政府机构、大学、大科学装置(如欧洲同步辐射光源 ESRF) 等几十家单位参与, 共同开发以激光沉积技术为基础的适用于块体合金材料研发的高通量组合材料制备与表征方法, 对数以万计的合金成份进行自动化筛选、优化与数据积累, 旨在将合金配方研发周期由传统冶金学方法所需的 5~6 年缩短至 1 年以内。2012 年, 欧洲科学基金会又推出总投资超过 20 亿欧元的 2012~2022 欧洲冶金复兴计划, 将高通量合成与组合筛选技术列为其重要内容, 以加速发现与应用高性能合金及新一代其他材料。

日本等国也启动了类似的科学计划。如日本计划建立玻璃、陶瓷、合金钢等领域材料数据库、知识库, 还尝试通过专家系统的建立促进其协同创新能力。俄罗斯联邦纳米技术研发中心 Ulnanotech, 通过引进美国 Intermolecular 公司完整的高通量研发技术平台, 较快地完成了其高通量材料研发能力的建设。

## 1.2 中国新材料领域面临形势严峻

中国的新材料产业虽然在多年攻关中取得了令人瞩目的成绩, 但是起步较晚、起点较低, 与先进国家相比, 整体水平仍存在较大差距。一些急需的新材料找不到, 能找到的又做不出, 能做出的又做不好, 核心竞争力令人堪忧。国家相关机构对 30 余家大型骨干企业调查结果显示: 在所需的 130 种关键材料中有 32% 国内完全空白, 54% 国内虽能生产, 但性能稳定性较差, 只有 14% 左右国内可以完全自给。关键新材料的缺乏长期成为中国国防、能源、信息、环境、交通、医药等重大战略领域发展的软肋, 对国家安全和社会经济发展构成巨大风险。例如, 中国高性能航空发动机的发展受到高温合金材料的严重制约, 大飞机、大型舰艇和尖端武器的制造依赖特种钢材与合金的进口, 在新能源领域, 功率电子器件、太阳能电池和储能等先进技术的发展也受限于关键核心材料。中国 2012 年在进口半导体芯片上的费用接近 2000 亿美元, 超过进口石油的花费近一倍。

改革开放 30 年来, 中国在很多行业广泛采取了“引进—消化吸收—再创新”的策略, 使得一些通用及民用技术在短期内上了一个台阶, 较快地缩短了与国外的差距。而那些国家安全和经济发展所亟需的核心材料仍长期受到西方发达国家的封锁, 无法通过市场渠道来获得。英国《经济学人》杂志在一篇文章中提到:

“在有些情况下, 西方公司会把一部分生产转包给正试图提升自己工业能力的国家的公司; 但有些东西是不可能共享的, 因为它们对于维持一个产品的竞争优势太重要了。”事实证明, 要打破核心技术封锁, 只能自力更生, 自行研发。然而现代技术发展具有“一代材料, 一代产业”的特点, 材料的每一次更新换代, 都会对原先的制造体系, 包括设备、技术、标准等造成颠覆性的改变。因此如果按部就班地发展, 从零起步, 即使一切顺利, 经过 10~20 年即便研发出了某种高端材料并达到了应用水平, 那时国外很可能已经发展出更先进的材料。中国的材料可能依然落后, 不得不进入新一轮的追赶。

## 1.3 中国开展材料基因组计划迫在眉睫

在此国内外背景之下, 中国材料界面临双重挑战: 一是如何在较短的时间内对先进国家进行追赶, 缩小现有差距; 二是在欧美突然发力加速的情况下, 不被落下更远。因此尽快启动中国版的“材料基因组计划”非常有必要, 变革以“炒菜法(试错法)”为基础的材料研发传统模式, 实现新材料领域的超常规速度发展。实际上, 中国材料界一直在关注加速材料研发的新方法。1999 年 6 月, 召开了主题为“发现和优化新材料的集成组合方法”的第 118 次香山科学会议。此后, 国内数家单位也尝试了高通量组合材料实验研究工作, 但由于多种原因, 并未得到普及开展。美国宣布材料基因组计划后, 中国科学家立即敏锐地捕捉到该计划所释放出的重要信息。在多位专家学者建议下, 中国科学院和中国工程院于 2011 年 12 月 21~23 日在北京联合主办了以“材料科学系统工程”为主题的香山科学会议。中国工程院与中国科学院分别设立了重大咨询项目, 并在 2014 年向国务院提交了各自的咨询报告。自 2012~2014 的 3 年中, 中国工程院、中国科学院, 中国硅酸盐学会, 许多大学、研究所等单位组织了多次以材料基因组计划为主题的研讨会、报告会。经过充分的研讨咨询活动, 中国材料界对材料基因组技术的认识不断深入, 已形成基本共识。

材料基因组技术是材料科学技术的一次飞跃, 是新材料研发的“加速器”。在中国实施材料基因组计划, 就是要根据国情构建将先进实验工具、模型计算手段与数据无缝衔接的新型材料创新技术框架体系, 用高

通量并行迭代替代传统试错法中的多次顺序迭代, 逐步由“经验指导实验”向“理论预测、实验验证”的材料研究新模式转变, 以加速中国关键新材料的“发现—开发—生产—应用”进程。这是中国新材料产业实现跨越式发展的必然选择。也只有这样, 才能实现习近平主席提出的“推动中国制造向中国创造转变、中国速度向中国质量转变、中国产品向中国品牌转变”的目标。

实施中国版材料基因组计划为中国材料领域快速追赶并超越国际先进水平提供了机遇。抓住了这个机遇, 就有望在短期内突破一系列国防和国民经济发展亟需解决的材料技术瓶颈, 较快地降低中国关键新材料的对外依存度。随着计划的实施, 中国先进材料科学技术将日趋成熟, 为中国中长期实现在新材料产业领域及其支撑的先进制造业跨越式发展, 赶超美、欧等发达国家提供坚实基础和有力支撑。如果把握不住这一机遇, 就可能会被甩得更远。在材料基因组技术方面, 欧美各国已经走在了前面。因此, 尽快制定并实施中国版材料基因组计划不仅极其必要, 而且异常紧迫。

## 2 材料基因组计划的内容

### 2.1 材料基因组计划的实质

材料基因组是以市场与应用为导向的材料研发新理念, 根本是要通过“多学科融合”实现“高通量材料设计与试验”, 其实质在于“加速”。值得强调指出的是材料基因组理念的运用范畴不仅适用于材料的发现与性能优化, 也适用于制造工艺开发和产品设计阶段, 并需要在材料“发现—研发—生产—应用”的过程中有意识地强化各阶段之间的信息交流与协同, 以总体缩短时间, 降低研发总成本。

#### 2.1.1 材料基因组计划的三要素及其协同关系

材料基因组技术包括高通量材料计算、高通量材料实验和材料数据库 3 个要素, 每一要素都可以在特定阶段起到加速材料研发过程的作用, 应根据实际情况不拘一格, 灵活地组合使用, 才能得到最好的效果。它们之间的协同关系如图 2 所示。

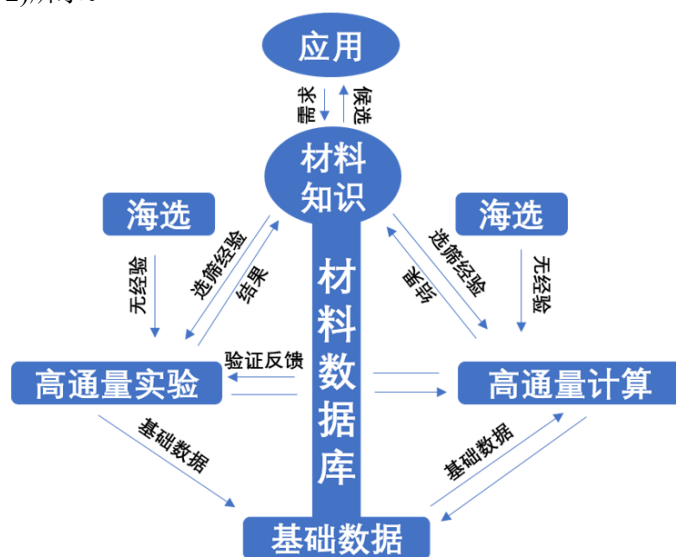


图 2 材料基因组三要素间的协同工作流程

Fig. 2 Collaborative workflow among three elements of the MGI

以材料筛选过程为例, 材料基因组技术面向应用需求, 以已有的材料知识为起点, 可通过高通量材料实验和高通量材料计算两条路径来丰富材料知识, 积累材料基础数据。一般而言, 无论是实验还是计算, 其筛选过程是在既有材料知识(理论或经验)指导下进行。实验获得的结果可以直接丰富材料知识, 而计算获得的结果在现阶段通常还需要经过实验验证才能进入材料知识库。计算结果对实验设计具有指导作用, 有助于大大降低实验工作量, 提高实验效率。实验获得的基础数据可以验证计算的结果, 有助于提高计算的准确性, 丰富材料数据库。三要素协同工作, 可以使得材料研发过程中的理论与实验结合更加紧密, 迅速排解不

确定因素,更快地得到结果。因此,三要素的有机融合和协同工作对加速材料的研发至关重要;组织相应的跨学科的研发团队,协同攻关,是加速材料研发的重要保障。

### 2.1.2 材料基因组计划的核心内容

材料基因组计划的关键在于由研究理念的变革带来的思维与工作模式的转换。带给人类文明迄今辉煌的传统方法经由成百上千年发展形成,已在材料业界深入人心,根深蒂固。因此新理念新文化的树立与普及面临重大考验,成为推行材料基因组计划的首要难题。

建立适应新型研究理念的新型材料研发基础设施是材料基因组计划的核心内容。如果说理念的变革是一场革命,它反映了人类对自然认识积累到一定程度而孕育着升华,即将进入更为主动、可控地改造自然的新阶段,新型平台设施,包括高通量材料计算、制备、表征平台等的建立过程则更像是一次维新,它主要由对既有技术在新理念下的重新集成整合而来,并根据需要在此基础上进行方法与功能的拓展以满足新形势下的新需求,保证新理念在实际工作中得以实施。大量技术层面工作,包括对各要素平台技术和工具的研究正有待开展。

(1) 多尺度材料计算软件及高通量计算驱动引擎的自主开发,如基于第一性原理和分子动力学的原子分子尺度计算,基于热力学/动力学参数的微观尺度计算(应用于化学反应、相变),基于有效介质理论的介观尺度计算(材料混合、复合),基于经验公式、半经验公式的宏观尺度计算,以及高通量计算作业生成、提交、监控算法,计算与实验数据集成等;特别需要指出的是材料计算不仅限于基于基本原理的理论计算,也应包括材料的发现、研发、生产和使用过程中的经验公式、半经验公式的计算。实际上,每个研究单位和企业都有这样的技术秘诀,特别是在工程设计和应用阶段,这些经验、半经验计算公式及专用计算软件起到了理论计算与实际应用间的桥梁作用;

(2) 高通量材料制备及表征方法、技术、仪器装备的自主设计、开发和制造,如基于不同材料形态的高通量组合制备技术及装备,针对材料成分、结构的高通量表征技术及仪器装备,针对材料各种性能的高通量测试技术及仪器装备,以及针对材料工艺过程的高通量原位实时表征测试技术及仪器装备等;

(3) 材料数据库及材料信息学技术的自主开发,如数据标准制定,数据共享协议制定,数据内容建设,数据分析与数据挖掘技术,以及数据库管理和使用机制等。

## 3 计算工具平台

通常我们讲成分—结构—性能关系是材料科学的基础,所以从成分直接预测晶体结构是最基本的,也是很重要的。近几年来开发的方法基本可以做到从成分直接预测晶体结构,特别是二元系和固定成分的化合物相的晶体结构。Ceder 小组开发了把第一性原理计算与信息学(数据挖掘)相结合来预测晶体结构的方法,其可靠性在 80 个二元系中得到了很好的验证<sup>[4]</sup>; Zunger 和他的合作者也开发了一种预测晶体结构的基因算法;中国最近开发的卡里普索(CALYPSO, Crystal Structure Analysis by Particle Swarm Optimization)晶体结构预测方法也有很好的应用前景<sup>[5]</sup>;另外一种叫 PEGS+DFT 的方法已经被成功地用来预测复杂离子晶体的晶体结构,这种方法最近甚至被用来成功地预测了一个非晶态化合物的局部晶体结构。可以看出,最近几年从过去比较简单的用结构图(Structure Maps)预测到现在更准确的晶体结构预测已有很大的进展,但是结构图在预测多组元合金中的相结构可能还是有它的用处<sup>[6]</sup>。下面介绍几款运用较为成熟的计算工具:

### 3.1 VASP

VASP(Vienna Ab-initio Simulation Package)是维也纳大学 Hafner 小组开发的进行电子结构计算和量子力学-分子动力学模拟软件包。它是目前材料模拟和计算物质科学研究中最流行的商用软件之一。

VASP 通过近似求解 Schrödinger 方程得到体系的电子态和能量,既可以在密度泛函理论(DFT)框架内求解 Kohn-Sham 方程(已实现了混合(hybrid)泛函计算),也可以在 Hartree-Fock(HF)的近似下求解 Roothaan 方程。此外, VASP 也支持格林函数方法(GW 准粒子近似, ACFDT-RPA)和微扰理论(二阶 Møller-Plesset)。

VASP 使用平面波基组,电子与离子间的相互作用使用模守恒赝势(NCPP)、超软赝势(USPP)或投影扩充波(PAW)方法描述。VASP 使用高效的矩阵对角化技术求解电子基态。在迭代求解过程中采用了 Broyden 和 Pulay 密度混合方案加速自洽循环的收敛。VASP 可以自动确定任意构型的对称性。利用对称性可方便地设



定 Monkhorst-Pack 特殊点, 可用于高效地计算体材料和对称团簇。Brillouin 区的积分使用模糊方法或 Blöchl 改进的四面体布点-积分方法, 实现更快的  $k$  点收敛。

### 3.2 CASTEP

CASTEP(Cambridge Sequential Total Energy Package)是由剑桥大学研发的第一原理计算软件。

CASTEP 可以用来研究一个系统的表面化学、结构特性、带结构、态密度、光学特性、电荷密度的空间分布及其波函数。另外, CASTEP 可以用来计算晶体的弹性常数及相关的力学特征, 比如: 泊松比、体模量和杨氏模量等。CASTEP 还可以用来计算半导体或其他材料中的点缺陷(空位、杂质原子取代和间隙)和扩展缺陷(晶体界面和断层)。使用线性响应理论, CASTEP 还可以计算固体的振动特性(声子的色散关系、声子的态密度和相关的热学特性)。所有用 CASTEP 计算得到的结果都有非常重要的应用, 如可以用来研究表面吸附物的振动特性, 解释实验得到的中子光谱或振动谱以及研究在高温高压下相的稳定性。通过线性响应理论方法, CASTEP 还可以计算对外加电场的响应: 分子的极化率和固体的电容率。

此外还有由丹麦 QuantuumWise 公司研发的 ATK, 由丹麦技术大学研发的开源工具中间件 ASE 等。近期由德国马普所牵头的地平线 2020 项目 “NoMatD”, 2015 年 10 月 1 日正式获欧盟委员会的资助, 为期 3 年, 其理念与课题组目前已初步研制的 Mat Cloud Lib 很接近, 通过 “Centralized Data Warehouse” 方法让各个研究团队贡献计算数据, 建立 “材料百科全书” 和研发大数据分析工具。此外, 瑞士洛桑联邦工学院(EPFL)也正在研发高通量自动流程计算相关技术。实际上, 英国在美国提出材料基因组计划前, 在当时英国 e-Science 计划的资助下, 也开始了高通量材料计算模拟和材料计算基础数据库的研究, 如 eMinerals 和 “材料网格” 项目。

## 4 实验工具平台

### 4.1 高通量实验的发展历程

高通量实验的主要发展历程(图 3)<sup>[7]</sup>。

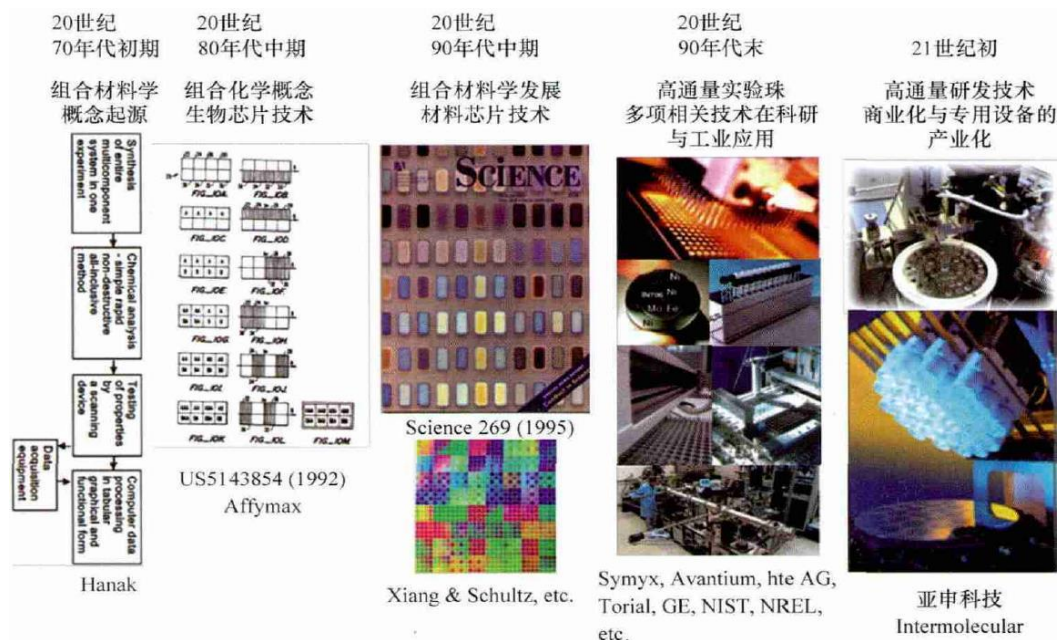


图 3 高通量实验的主要发展历程

Fig. 3 The mainly development course of high-throughput experiment

1970 年, Hanak 首先提出了 “多样品实验” 的概念, 并应用于薄膜形态的二元、三元超导材料研究<sup>[8]</sup>, 其基本思想是通过一次实验合成完整覆盖多组分材料体系中成分组合的样品阵列, 利用高效的测试分析手段快速获取阵列中各样品的成分、结构以及性能数据, 最终通过计算机进行数据处理并以适当的方式呈现。然而, 由于当时计算机等相关支撑性技术水平的限制, 该方法未能得到快速推广。20 世纪 80 年代中期兴起

了组合化学,并派生到高通量新药筛选、高通量基因测序、高通量平行反应器(用于有机材料和催化剂等的合成)等,显著地提高了生物和化学领域的研发效率。20 世纪 90 年代中期,美国劳伦斯伯克利国家实验室的项晓东和 Schultz 发展和完善了现代高通量组合材料实验方法,率先展示了高通量实验的巨大潜力,并随后在多种材料系统上进行了应用与示范推广,取得了一系列新材料成果,并基于此在美国创办了 Symyx Corp 和 Intematix Corp 两家上市公司。

20 世纪 90 年代末期,高通量组合材料实验方法已在较大范围被材料科技工业领域接受<sup>[9]</sup>,应用于金属、陶瓷、无机化合物、高分子等材料的研发与产业化。适用的材料形态从最初的薄膜形态扩展至液体、胶体、块体、粉末等多种形态,并取得了一系列商业上的成功。典型的案例包括:Symyx 公司发展出新型化工催化剂;Intematix 公司开发出突破专利封锁的固体发光器件荧光材料;通用电气公司(GE)开发了高性能的特殊合金材料;康宁公司 PMN-PT 电光陶瓷的发明及光通信元器件产业化;Intermolecular 公司开发出新一代低辐射膜材料;Intel 公司和三星公司用于相变存储合金和高介电材料研究。美国国防部(DOD)、能源部(DOE)、国家标准局(NIST)、劳伦斯伯克利国家实验室(LBNL)及可再生能源国家实验室(NREL)等通过先进材料项目(ATP)、半导体自旋器件(SPINS)等多个研究项目,大力资助高通量组合材料实验研究,成功发展出许多新型材料,如半导体室温伽玛射线和中子探测材料,储氢材料、热电材料,纳米管生长催化剂等。进入 21 世纪后,还出现了专门提供商业化的高通量组合材料实验仪器设备与高通量组合材料实验研发服务的公司,如中国的亚申科技研发中心(上海)有限公司和美国 Intermolecular 公司。

世纪末开始尝试采用高通量组合材料实验方法,如中国科技大学开展了液滴喷射制备技术与同步辐射在组合材料方法中的应用研究,中国科学院上海硅酸盐研究所提高了镀锌汽车板表面抗盐液腐蚀能力及力学性能,清华大学原子分子纳米科学教育部重点实验室和中国科学院大连物理化学研究所分别优化了 CO 氧化催化剂和 NO 还原催化剂,大连中国石化研究院通过引进美国 Symyx 公司的高通量设备展开石化冶炼催化材料的快速筛选等。但在普及与应用规模上与发达国家有相当差距。

## 4.2 高通量实验流程

### (1) 高通量实验流程概述

Hanak 在 1970 年提出的工作流程(图 4)已初步包含了高通量实验最基本的特征:

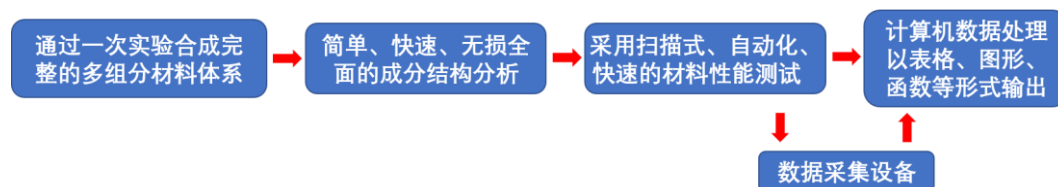


图 4 Hanak 提出的材料高通量实验的初试概念<sup>[9]</sup>

Fig. 4 First concept of high-throughput put forward by Hanak<sup>[9]</sup>

- 1) 高通量合成制备,即在 1 次实验中完成多组分目标材料体系制备,使制备具有高效性、系统性和一致性;
  - 2) 快速分析测试,即采用扫描式、自动化、快速的分析测试技术,原则上 1 天制备的样品 1 天内完成测试分析,避免成为瓶颈;
  - 3) 计算机数据处理输出,即充分利用计算机数据处理和分析功能,以表格、图形等多种形式输出<sup>[10]</sup>。
- 在此基础上,经过多年发展与演化,形成了新型高通量组合材料的实验流程(图 5)。它除保持传统特征外,还具有若干重要的新特点。

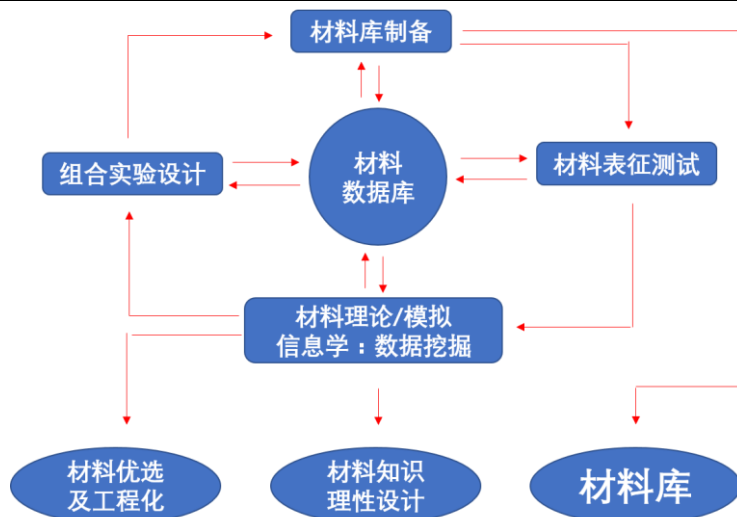


图 5 现代材料高通量实验流程示意图<sup>[9]</sup>

Fig. 5 The flow diagram of modern material high-throughput experiment<sup>[9]</sup>

- 1) 强调实验设计的重要性, 合理的实验设计减少工作量, 提高筛选速度和成功率;
- 2) 明确材料数据库在流程中的轴心位置, 材料数据库兼具实验管理、数据处理、信息存储、数据挖掘等多项功能;
- 3) 注重材料计算模拟与实验的互动, 相互验证, 便于及时优化方向, 快速收敛。

#### (2) 高通量实验的设计

高通量实验设计的原则是首先要将实验的目的、目标、待解决的问题、欲获得的信息等定义清晰, 找出关键的科学参数。在设计中充分利用现有经验和知识, 积极采用理论计算模型作为指导, 尽量缩小筛选范围。现代高通量组合材料实验强调理论和实验的融合与协同, 力图通过理论模型对可能的组合进行预先的计算仿真, 避免盲目组合, 从而提高实验的效率与成功几率。在高通量组合材料实验方案的设计中还需要充分考虑目标材料体系的制备和表征能力, 包括样品的制备方法、表征工具的空间分辨率等, 从而确定样品的空间密度和组合形式。(图 6)给出样品组合方案设计的参考因素。

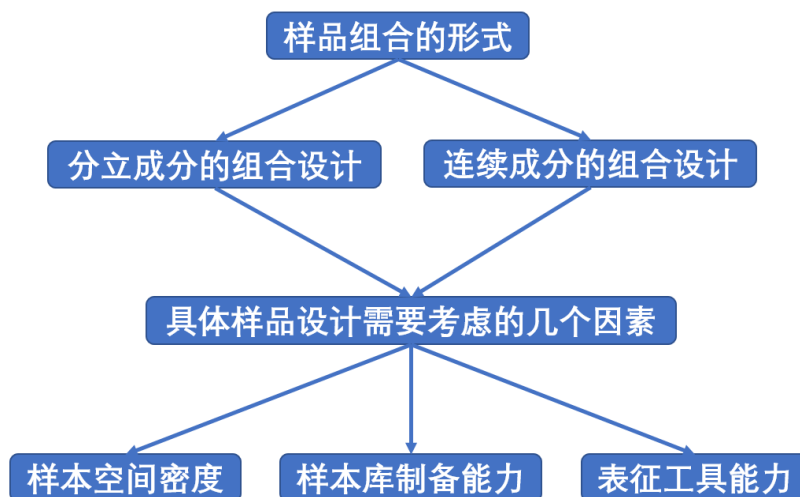


图 6 样品组合方案设计的参考因素<sup>[9]</sup>

Fig. 6 Reference factors of sample portfolio design

#### (3) 数据管理在高通量实验中的轴心作用

数据管理是现代高通量实验技术发展中的主要挑战之一。现代材料高通量制备与表征技术能够以较快的速度产生海量的数据。使海量的实验数据真正发挥加速材料研发和应用进程的作用, 还需对其进行



有效的分析、组织和呈现。现代信息学已经建立了较为先进的大数据理论,并已经开发出一系列的计算机软件工具辅助大数据的管理和分析。材料信息学正是现代信息学技术与材料科学研究相结合的产物。基于材料信息学的实验数据管理是高通量实验流程中的中心环节,它包含了实验过程管理、多变量表征的追踪、信息存储、材料数据库信息检索以及大数据挖掘等 5 项主要功能,如(图 7)所示,是始于实验设计开始,涉及原始实验数据采集、存储、管理、处理、分析、挖掘等各个步骤的高通量实验全流程管理。



图 7 组合材料试验信息学的内容<sup>[9]</sup>

Fig.7 The content of composites experimental informatics

其中,大数据挖掘技术包括聚类分析(cluster analysis)、预测模型(predictive modeling)、关联分析(correlation)、异常检测(abnormality test)等方法,对海量材料数据进行挖掘,快速寻找材料“工艺-成分-结构-性能”之间的内在规律,完善材料科学知识和基础数据库。

## 5 材料数据库

与传统选材方法相比,通过材料数据库选材具有如下优势:

### (1) 快且准

利用材料数据库选材,能按照用户的需求快速准确地找到所需材料的相关信息。利用材料数据库进行选材,充分利用计算机的快速数据处理功能,简化数据处理的任务,快速而准确查找与定位所需数据和信息。

### (2) 全且易

利用材料数据库进行选材避免了选材者四处查找手册的麻烦,一个人机对话的、菜单式的数据库系统,用户使用起来十分方便。友好的选材界面使用户非常容易地从浩瀚的材料数据中轻松地选出所需材料,而且有效地避免了漏选错选。另外,利用计算机可以对数据库方便地进行数据维护,保证数据的统一性,规范用户的操作。

### (3) 智且新

利用材料数据库在得到最佳用材的同时,可以给出多种相近或相似的可用推荐材料,让用户可以结合自己的实际情况进行必要的分析、判断,从而得出最终选择。另外,若将数据库与专家系统结合,不仅可以在库中已存储的数据中找到所需信息,而且还可以在经验公式的基础上进行一定程度的推理和预测,使选材过程更智能化。如果通过网络将数据进行共享还可提高数据的利用水平并使数据库的数据更全更新。

## 6 结论

高通量实验技术在过去 40 年中取得了长足的发展,其有效性在材料科学的不同领域、多个研究案例中得到验证。随着材料基因组方法被广泛接受,高通量实验作为一种要素工具技术进入到一个新的发展阶段。

(1) 高通量实验技术与材料模拟计算、数据库的充分融合,将使各自的效果倍增。研究基于理论计算-数据库-实验一体化以及组合(制造)-表征实验一体化的新技术路线,实现与材料特性相关的优化,以介观尺度为基点,探索微观-介观-宏观材料跨尺度组合与表征实验系统。以海量数据为基础构建材料组成-组织结构-性能统计相关性的模型,引入现代信息学理论中关于大数据处理和挖掘的成果,开展新材料发现、材

料改性及工艺优化的新技术与新方法研究, 加速材料研究进程。同时, 基于先进信息技术建立高通量实验与材料模拟计算、数据库的融合接口, 也是材料基因组技术的本征需求。

(2) 高通量实验的日益普及呼唤制备与表征仪器装备技术的创新和发展。材料基因组方法是材料研发思维的重大变革, 对实验方法与仪器装置也提出了更新的要求。开发一系列普适、精准、快速的制备表征仪器对推广应用高通量实验技术至关重要。高通量实验仪器的开发需集中物理、化学、材料、电子、仪器、机械和计算机软件等多个领域的顶尖人才共同参与才能实现。

## 参 考 文 献

- [1] 汪洪, 向勇, 项晓东, 等. 材料基因组——材料研发新模式[J]. 科技导报, 2015, 33(10):13-19.
- [2] 赵继成. 材料基因组计划简介[J]. 自然杂志, 2014, 36(2):89-104.
- [3] National Research Council. Integrated computational materials engineering: a transformational discipline for improved competitiveness and national security[M]. National Academies Press, 2008.
- [4] Curtarolo S, Morgan D, Ceder G. Accuracy of ab initio methods in predicting the crystal structures of metals: A review of 80 binary alloys[J]. Calphad, 2005, 29(3): 163-211.
- [5] Wang Y, Lv J, Zhu L, et al. Crystal structure prediction via particle-swarm optimization[J]. Physical Review B, 2010, 82(9): 094116.
- [6] Seiser B, Drautz R, Pettifor D G. TCP phase predictions in Ni-based superalloys: Structure maps revisited[J]. Acta Materialia, 2011, 59(2): 749-763.
- [7] 王海舟, 汪洪, 丁洪, 等. 材料的高通量制备与表征技术[J]. 科技导报, 2015, 33(10): 31-49.
- [8] Xiang X D, Sun X, Briceno G, et al. A combinatorial approach to materials discovery[J]. Science, 1995, 268(5218): 1738.
- [9] Potyrailo R, Rajan K, Stoewe K, et al. Combinatorial and high-throughput screening of materials libraries: Review of state of the art[J]. ACS combinatorial science, 2011, 13(6): 579-633.
- [10] Hanak J J. The “multiple-sample concept” in materials research: Synthesis, compositional analysis and testing of entire multicomponent systems[J]. Journal of Materials Science, 1970, 5(11): 964-971.