

Review

Nanoinformatics: Emerging Databases and Available Tools

Suresh Panneerselvam and Sangdun Choi *

Department of Molecular Science and Technology, Ajou University, Suwon 443-749, Korea;
E-Mail: sureshcibt@gmail.com

* Author to whom correspondence should be addressed; E-Mail: sangdunchoi@ajou.ac.kr;
Tel.: +82-31-219-2600; Fax: +82-31-219-1615.

Received: 17 February 2014; in revised form: 28 March 2014 / Accepted: 9 April 2014 /
Published: 25 April 2014

Abstract: Nanotechnology has arisen as a key player in the field of nanomedicine. Although the use of engineered nanoparticles is rapidly increasing, safety assessment is also important for the beneficial use of new nanomaterials. Considering that the experimental assessment of new nanomaterials is costly and laborious, *in silico* approaches hold promise. Several major challenges in nanotechnology indicate a need for nanoinformatics. New database initiatives such as ISA-TAB-Nano, caNanoLab, and Nanomaterial Registry will help in data sharing and developing data standards, and, as the amount of nanomaterials data grows, will provide a way to develop methods and tools specific to the nanolevel. In this review, we describe emerging databases and tools that should aid in the progress of nanotechnology research.

Keywords: biomedical informatics; databases; molecular simulations; nanoinformatics; nanomedicine; nanotechnology; QSAR; text mining

1. Introduction

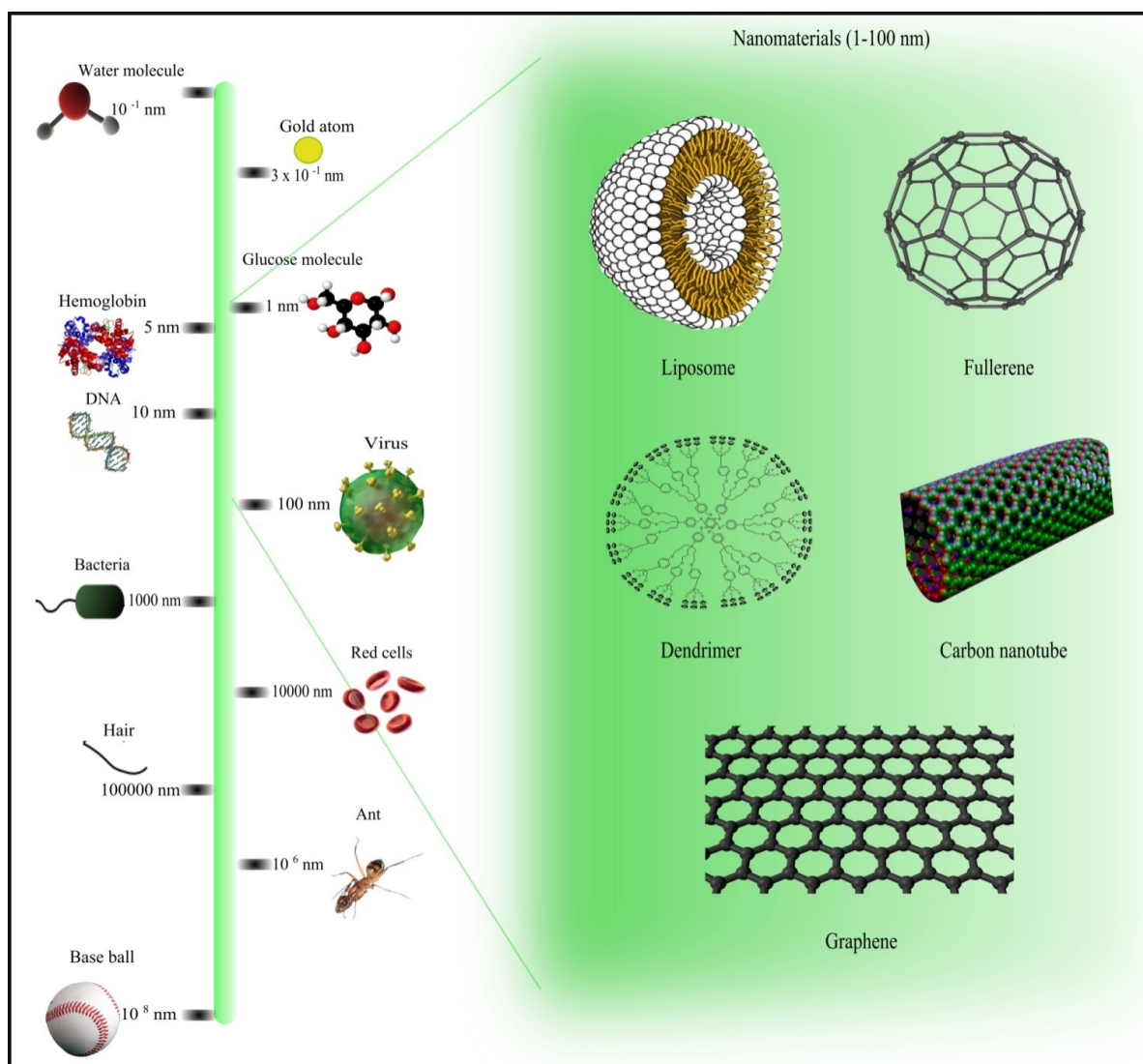
Nanotechnology enables the design and assembly of components with dimensions ranging from 1 to 100 nanometers (nm) [1]. The word nano, which is derived from the Greek word for dwarf, has been part of the scientific nomenclature since 1960 [2]. Macroparticles, which are particles with dimensions greater than 100 nm, have properties that are comparable to those of bulk materials; however, when particles are smaller than 100 nm, their behavior is influenced by atomic, molecular, and ionic interactions. Nanoparticles show novel and changeable properties due to quantum effects [3]. The

material properties at the nanoscale level are unusual for two major reasons. First, nanomaterials have a considerably larger surface area than the equivalent mass of a material in a larger form. For instance, if we cut a one centimeter cube into 10^{21} cubes of one nm each, although the overall mass and volume remains the same, the surface area will increase by one million-fold [3]. Second, at the nanoscale level, quantum effects can begin to dominate the behavior of matter, thereby affecting the optical, electrical, and magnetic properties of materials. This behavior is the result of the actions of the electrons in the nanomaterials. Electrons possess two key characteristics [3]: rotation and a capacity to shift among particular energy levels. In iron oxide magnetic nanoparticles (≤ 20 nm in diameter), electrons rotate in the same direction, whereas in iron oxide macroparticles (> 20 nm in diameter), electrons rotate in opposite directions. When electron spins are aligned in the same direction, the fields are additive, but as soon as the electrons begin to rotate in different directions, the fields eliminate each other. The total magnetic field strength of a nanomaterial is the sum of the magnetic fields of individual electrons. Consequently, nanoparticles comprise greater localized magnetic fields than larger particles [3].

In general, nanoparticles can be classified into three categories: organic, inorganic, and hybrid particles. Simple monomeric building blocks such as polymer-based nanoparticles are organic particles, whereas inorganic nanomaterials include carbon nanotubes, fullerene particles, and metallic nanoparticles (gold and silver nanoparticles). Hybrid nanoparticles comprise engineered nanoparticles such as DNA-carbon nanotube arrays [4]. Because the sizes of nanomaterials are comparable to the sizes of biomolecules such as proteins, they can be designed for biomedical applications such as devices for diagnosis, drug delivery, and therapy [5]. The sizes and shapes of some nanomaterials as compared to those of common materials are shown in Figure 1 [3,6]. Nanoparticles are found in many consumer products such as sheets and clothing (silver nanoparticles), cosmetics and sunscreens (titanium dioxide), bicycles (carbon nanoparticles), and even beer bottles (clay nanoparticles) [7]. Doxil and Abraxane, which are currently available on the market, are Food and Drug Administration-approved nanoformulations for cancer treatment [8]. Although the unique properties of nanomaterials hold promise for many applications, they also raise safety concerns. Nanoparticles can pose risks for human health because of their ability to penetrate into the human body, flow in the bloodstream, and pass through cells, which may cause undesirable side effects [9]. Therefore, knowledge of the biological and toxicological effects of nanomaterials is essential to determine whether the products are safe. However, assessing the safety of nanomaterials through experiments is costly and time consuming; hence, computational methods offer a profitable alternative [10]. Computational approaches can not only serve to build binding models of protein-nanoparticle complexes, but can also be used to simulate the binding processes [11].

Nanoinformatics is described as “the science and practice of determining which information is relevant to the nanoscale science and engineering community, and then developing and implementing effective mechanisms for collecting, validating, storing, sharing, analyzing, modeling and applying that information” [12]. The major setback of the discipline of nanoinformatics is that it is still at an early stage of development and the data are scattered in the form of publications [13–15]. Recently, some initiatives have emerged in the field of nanoinformatics with respect to nanoparticle data management [16–19]. The goal of this review is to describe the emerging databases and tools available for the nanotechnology community. We have categorized the nanoinformatics tools in this review based on relevant examples provided by Maojo *et al.* as well as on the availability of data [13].

Figure 1. Comparison of the sizes of nanomaterials with those of other common materials. Figures were drawn using Inkscape [20] and GNU Image Manipulation Program (GIMP) [21]. All the images in this figure were obtained from Wikimedia Commons.



2. Growth of Bioinformatics and Applications for Nanoinformatics

In the 1980s, it became mandatory to deposit all published DNA sequences in a central repository. At the same time, a standard was adopted by journals to make gene and protein sequences freely accessible to all and to compile all data in the form of openly available databases. The open accessibility of DNA sequences in databases such as GenBank [22] and the availability of protein structures in databases such as the Protein Data Bank [23] have motivated many researchers to develop powerful methods, tools, and resources for large-scale data analysis [24]. Simultaneously, growth in computational speed and memory storage capacity has led to a new era in the analysis of biological data. Bioinformatics has emerged as a powerful discipline and almost 1000 databases are currently publicly available [25] along with a large number of bioinformatics tools [26]. The growth and development of bioinformatics can provide valuable lessons for applying the same practice to benefit nanoinformatics development [27].

3. Databases

An overview of nanomaterial-related databases is provided in Table 1. The emergence of recent databases such as Investigation Study Assay (ISA) tab-delimited (TAB) format (ISA-TAB-Nano) [19], Cancer Nanotechnology Laboratory (caNanoLab) [17], and Nanomaterial Registry seems encouraging and is discussed below.

Table 1. Summary of nanotechnology-related databases.

No.	Name	Description	Url
1	Nanomaterial Biological Interactions Knowledgebase	The knowledgebase serves as a repository for annotated data on nanomaterial characterization (purity, size, shape, charge, composition, functionalization, and agglomeration state), synthesis methods, and nanomaterial-biological interactions.	[28]
2	InterNano	InterNano supports the information needs of the nanomanufacturing community by bringing together resources related to advances in applications, devices, metrology, and materials that will facilitate the commercial development and/or marketable application of nanotechnology.	[29]
3	Nano-EHS Database Analysis Tool	This web tool provides a quick and thorough synopsis of the Environment, Health and Safety Database.	[30]
4	Nanoparticle Information Library	The goal of the NIL is to help occupational health professionals, industrial users, worker groups, and researchers organize and share information on nanomaterials, including their health and safety-associated properties.	[31]
5	Nanomaterials Registry	The Registry was established to provide researchers with a convenient data management and sharing plan.	[32]
6	National Toxicology Program Database	The goal of this database is to develop and apply modern toxicology and molecular biology tools to identify substances in the environment that may affect human health.	[33]
7	Nano-HUB database	Nano-HUB offers a searchable online database of nanoBIO tools.	[34]
8	National Center for Biomedical Ontology Biportal	A major focus involves the use of biomedical ontologies to aid in the management and analysis of data derived from complex experiments.	[35]
9	ISA-TAB-Nano	The ISA-TAB-Nano standard specifies the format for representing and sharing information about nanomaterials, small molecules, and biological specimens along with their assay characterization data (including metadata and summary data) using spreadsheet or TAB-delimited files.	[36]
10	caNanoLab	caNanoLab is a data-sharing portal designed to facilitate information sharing in the biomedical nanotechnology research community to expedite and validate the use of nanotechnology in biomedicine.	[37]
11	Toxicology Data Network	Includes databases on toxicology, hazardous chemicals, environmental health, and toxic releases.	[38]
12	Nanotechnology Characterization Laboratory	The Nanotechnology Characterization Laboratory performs and standardizes the preclinical characterization of nanomaterials intended for cancer therapeutics and diagnostics developed by researchers from academia, government, and industry.	[39]
13	Collaboratory for Structural Nanobiology	A nanoinformatics service dedicated to the collection, curation, and correlation of structural, physicochemical, and biological and biomedical data.	[40]
14	Toxicology Literature Online	Provides references from the toxicology literature.	[41]
15	OECD	Database on research into the safety of manufactured nanomaterials.	[42]

3.1. ISA-TAB-Nano

Motivated by an absence of standardization in representing and distributing nanomaterial data, the ISA-TAB-Nano was established for the public development of data sharing [19]. The ISA-TAB-Nano file format offers a common framework to record and integrate nanomaterial descriptions and uses the ISA-TAB open-source framework file format. ISA-TAB-Nano defines four file formats for the distribution of the data: (1) the Investigation file; (2) the Study file; (3) the Assay file; and (4) the Material file. The Investigation file provides reference data about each investigation, study, assay, and protocol. The Study file contains the names and attributes of procedures applied for preparing samples for examination (sources and characteristics of biospecimens). The Assay file contains the values of measured endpoint variables and references to external files for each analyzed sample. The Material file describes the material sample and its structural and chemical components. File format templates are available [36], which also provide information on using ontology terms to support standardized descriptions and assist searches for and incorporation of information. Ontology is a formal way of representing knowledge abstracted from the growing body of biological science in a coded form, which can be translated into a programming language [43,44]. A benefit of ontologies is to facilitate the description logic that can be utilized together for querying a dataset. Moreover, the data sets are analyzed which are not usually accessible to searching and comparing. Ontologies are used in databases in various ways. For example, ontologies are characterized in a computer-readable form that can be interpreted by computers as well as domain experts. Thomas *et al.* have provided a review of ontology terms used for standardized descriptions [45]. A list of ontology-related resources for nanotechnology is shown in Table 2.

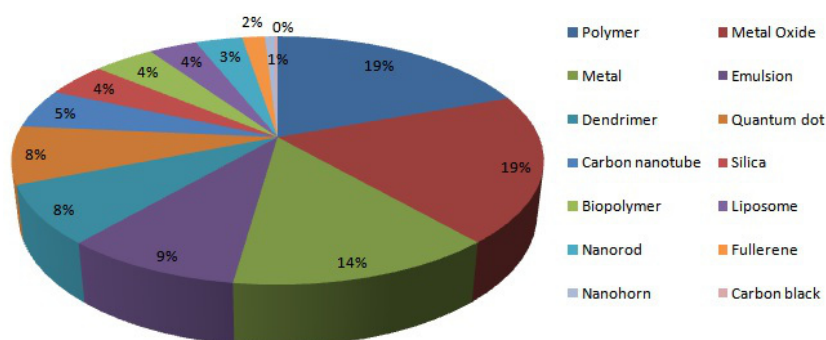
Table 2. List of nanoparticle-related ontology resources.

No.	Name	Description	Url
1	NanoParticle Ontology	The NPO was developed to provide knowledge regarding the description, preparation, and characterization of nanomaterials in cancer nanotechnology research.	[46]
2	BioPortal	BioPortal provides access to commonly used biomedical ontologies and tools for their analysis.	[47]
3	The Open Biological and Biomedical Ontologies	A collaborative experiment involving developers of science-based ontologies who are establishing a set of principles for ontology development with the goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain.	[48]
4	Cancer Open Biomedical Resource	A tool for indexing cancer nanotechnology informatics knowledge with open biomedical resources.	[49]
5	Chemical Entities of Biological Interest	A database and ontology of chemical entities of biological interest.	[50]
6	Gene Ontology	The Gene Ontology project is a major bioinformatics initiative with the aim of standardizing the representation of gene and gene product attributes across species and databases.	[51]
7	Foundational Model of Anatomy ontology	The Foundational Model of Anatomy Ontology is an evolving computer-based knowledge source for biomedical informatics.	[52]
8	Zebrafish Anatomy and Development ontology	A structured, controlled vocabulary of the anatomy and development of the zebrafish.	[53]
9	Ontology Lookup Service	Provides a web service interface to query multiple ontologies from a single location with a unified output format.	[54]
10	Phenotype Quality ontology	Phenotype Quality provides ontology of phenotypic qualities intended for use in a number of applications, primarily defining composite phenotypes and phenotype annotation.	[55]

3.2. caNanoLab

caNanoLab is a data repository that researchers can use for the submission and retrieval of information on nanoparticles, including their composition, function (e.g., therapeutic, targeting, diagnostic imaging), physical (e.g., size, molecular weight), and *in vitro* experimental characterizations (e.g., cytotoxicity, immunotoxicity) [17]. It also encourages data sharing and analysis across the cancer community to accelerate and authenticate the use of nanoparticles in biomedicine. To facilitate data submission, web-based forms are offered so that submitters can restrict the visibility of their records to be private, to be distributed to particular collaboration groups, or to be public. In addition, caNanoLab can be utilized for discovery purposes by using it to search publicly accessible physical and *in vitro* characterizations while providing access to the associated publications. Moreover, the query results are downloadable in spreadsheet-based formats. The caNanoLab project team is also currently engaged with ISA-TAB-Nano to facilitate the submission and exchange of data across the breadth of nanotechnology information. Future features of caNanoLab may include support for the validation, import, and export of ISA-TAB-Nano files by means of customized ISA-TAB tools. A summary of data available from caNanoLab is presented in Figure 2.

Figure 2. Summary of data available from caNanoLab.

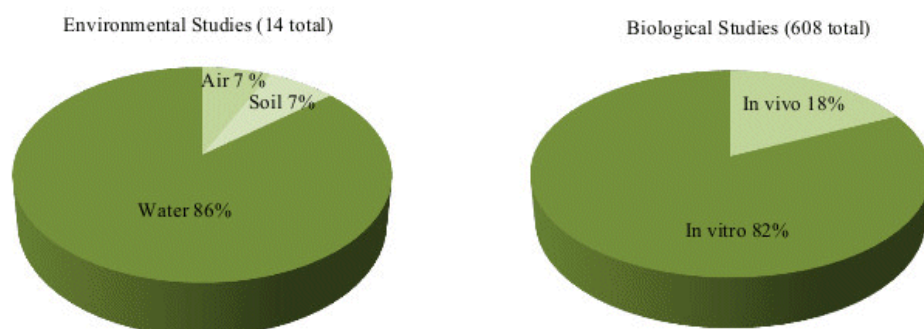


3.3. Nanomaterial Registry

For data to be useful to a wide community of researchers, the data must be easily available in a usable form. However, providing usable data is a task that becomes increasingly difficult as the quantity of information grows. Curation is particularly important for nanomaterials because the experimental conditions and process of sample preparation can affect the actual data measurement. Data curation is best achieved through increasing proficiency in assessing a specific feature of a specific dataset. To increase the efficiency of the information curation method, the Nanomaterial Registry developed a method to assist the combinatorial analysis of a variety of datasets [16]. The Nanomaterial Registry is a web-based database that offers curated data for characterized nanomaterials. The web-based curation tool has a dual functionality that maintains (a) manual data entry and (b) programmatic data upload. The curation tool resides in a protected environment that contains user accounts, roles, and permissions. User roles are supported by an administration tool that offers access, controls, and permissions for curators. As the curation technique improves and the dataset grows, these resources will become increasingly available to researchers and information analysts. Data available from the Nanomaterial Registry are shown in Figure 3.

Figure 3. Data available from the Nanomaterial Registry.

Characteristic	Gold	Silver	Bronze	Total
Aggregation / agglomeration state	70	98		168
Composition	945	230		1175
Particle size	330	322	174	834
Purity	6	33		39
Shape	343	3	23	369
Size distribution	112	3		119
Solubility	38	149		187
Stability	17	28	56	101
Surface area	90	9		99
Surface charge	296	2	30	328
Surface chemistry	162	1		163
Surface reactivity	6			6



4. Text Mining

Text mining is a computerized process for utilizing the enormous amount of knowledge existing in the literature [56,57]. Three important approaches to text mining are currently prevalent in the field of biology: (1) co-occurrence-based methods; (2) rule-based or knowledge-based approaches; and (3) statistical or machine learning-based approaches. Co-occurrence-based methods scan for specific terms or ideas that occur in the same sentence but are sometimes abstract, and then hypothesize an association between them. Rule-based systems make use of common knowledge about how language is structured, specific knowledge about how biologically significant information is stated in the literature, and the relationships that these two types of knowledge can have with one another. In contrast, statistical or machine learning-based systems function by building classifiers that may operate at any level. However, free full-text access for a large proportion of scientific journals is still unavailable. In several disciplines, such as chemistry, it is still hard to find enough abstracts for large-scale analysis [56,57]. Although the application of text mining to nanotechnology is at an early stage, some recent studies have been published [58–60]. The few available text-mining tools are as follows: (1) the Google Scholar search engine [61]; (2) the GoPubMed engine [62]; (3) Textpresso [63]; and (4) BioRAT [64]. When literature is available for analysis, it is possible to predict interactions via text mining, for example, using iHOP [65], which is an online service through which the names of genes and proteins in sentences from abstracts in PubMed can be hyperlinked and assembled into a conceptual network. Likewise, it might be possible to use text mining to predict nanoparticle-protein interactions for nanotechnology.

5. Molecular Modeling

The structures of nucleic acids and many proteins have been determined by crystallography, nuclear magnetic resonance, electron microscopy, and many other techniques. Structural biology provides information on the static structures of biomolecules. However, in reality, biomolecules are highly dynamic, and their motion is important to their function. Different experimental techniques are available to help study the dynamics of biomolecules. Computational power continues to increase, and the development of new theoretical methods offers hope of solving scientific problems at the molecular level. All the theoretical methods and computational techniques that are used to model the behavior of molecules are defined as molecular modeling. Macromolecules can be studied only with molecular mechanics since other quantum methods based on the Schrodinger equations such as *ab initio*, semi-empirical, and density functional theory (DFT) methods require extensive computational time. Molecular mechanics uses classical physics and relies on force field. *Ab initio* and semi-empirical methods are based on approximate solutions of the Schrodinger equation. In addition, semi-empirical methods need additional empirical parameters (*i.e.*, using parameterization) for the calculations. DFT methods are based on electron density. Novel molecules are best studied with *ab initio* or possibly DFT calculations, since the parameterization intrinsic to molecular mechanics or semi-empirical methods makes them unreliable for molecules that are different from those used in the parameterization. The energies of molecules such as proteins or DNA can be calculated using molecular mechanics; small molecule energies can be calculated using *ab initio* and semi-empirical methods [66,67]. There is an extensive body of three-dimensional structures of molecules essential to molecular modeling that can be visualized using software freely available for users for this purpose. A list of software for visualizing macromolecules is presented in Table 3.

Table 3. List of visualization tools.

No.	Name	Operating System	License	Url
1	PyMOL	Windows/GNU Linux	Free and commercial	[68]
2	RasMol	Windows/GNU Linux	GNU General Public License	[69]
3	Raster3D	GNU Linux/Unix/Mac OS X	Artistic License	[70]
4	UCSF Chimera	Windows/GNU Linux/Mac OS X	Free for non-commercial use	[71]
5	VMD	Windows/GNU Linux/Mac OS X/Solaris	VMD License/Free to use	[72]
6	Cn3D	Windows/Mac OS X	Free to use	[73]
7	BALLView	Windows/GNU Linux/Unix/Mac OS X	GNU Lesser General Public License	[74]
8	Jmol	Windows/GNU Linux/Mac OS X	GNU Lesser General Public License	[75]
9	ICM-Browser	Windows, Mac, and GNU Linux	Commercial software	[76]
10	Chemkit	GNU Linux	BSD licenses (free to use)	[77]
11	OpenStructure	GNU Linux/Mac OS X	GNU Lesser General Public License	[78]
12	VEGA ZZ	Windows/GNU Linux/Mac OS X/ Silicon Graphics/Amiga OS	Copyright (free for non-profit academic)	[79]
13	Molekel	Windows/GNU Linux/Mac OS X	GNU General Public License	[80]
14	Avogadro	Windows/GNU Linux/Mac OS X	GNU General Public License	[81]
15	Xeo	Windows/GNU Linux/Mac OS X	GNU General Public License	[82]
16	ChemSketch	GNU Linux/Mac OS X	Commercial software (free for academic and personal use)	[83]

5.1. Docking

Docking is an important computational tool in the drug discovery process and is used to specifically predict protein-ligand interactions. The two basic features of docking software are docking accuracy and scoring reliability [84]. Docking accuracy indicates how similar the prediction of ligand binding is to the ligand conformation that is determined experimentally, whereas scoring reliability ranks ligands based on their affinities. Docking accuracy and scoring reliability are used to assess the searching algorithm and the scoring functions, respectively, of docking software. The numerous searching algorithms used in docking software vary with respect to randomness, speed, and the area covered. Most of the searching algorithms perform well when tested against the known structure. In contrast, scoring functions are rarely successful, and a number of issues needed to be addressed to improve the docking features. Many types of docking software are currently available; AutoDock is a software that is highly accessed and freely available [85]. At present, as protein-nanoparticle complexes are difficult to study using experimental approaches, computational tools show promise. Structural models for carbon nanomaterials such as carbon nanotubes and fullerenes are available, and thus many protein-nanoparticle interactions can be studied computationally. Fullerenols, which are derivatives of fullerenes, are currently in trial for diagnostic and therapeutic uses, although insufficient information is available about the structural interactions and toxicity of fullerenols in biosystems. Yang *et al.* computationally studied the interactions in the fullereneol-lysozyme complex and compared the computational results with experimental results [86]. The structure of fullereneol was constructed using Chem3D Ultra and subjected to blind docking in the active site of chicken lysozyme. The fullereneol bound close to Tryptophan 62, and a π - π stacking interaction was observed. The predicted computational binding result was consistent with all known experimental results. As a result, new proteins that can interact with fullereneol can be identified for future applications as drug targets [87]. The addition of molecular dynamic (MD) simulations to docking studies can also assist in studying interactions. A list of well-known docking software is provided in Table 4.

Table 4. List of tools available for docking.

No.	Name	Operating System	License	Url
1	Autodock 4	Windows/GNU Linux/Mac OS X	GNU General Public License	[88]
2	AutoDockVina	Windows/GNU Linux/Mac OS X	Apache license	[89]
3	Dock	Windows/GNU Linux/Mac OS X	Academic software license	[90]
4	GOLD	Windows/GNU Linux	Commercial software	[91]
5	FlexX	Windows/GNU Linux	Commercial software	[92]
6	Surflex-dock	Windows/GNU Linux	Commercial software	[93]

5.2. Quantitative Structure-Activity Relationships (QSAR)

QSAR is based on the idea that similar structural properties generate similar biological effects [94]. It is used to establish associations between the structural and electronic properties of potential drug candidates along with their binding affinities for common macromolecular targets that are widely used in drug discovery. QSAR studies were originally based on a single physicochemical property such as the solubility or the pKa (acid dissociation constant) value to elucidate the biological effect of a

molecule (one-dimensional QSAR). Two-dimensional QSAR further examines the connectivity of a compound by considering the physicochemical properties of single atoms and functional groups and their roles in biological activity. Current models contain three-dimensional structural descriptors such as the length or width of a substituent. With a suitable force field, it is theoretically possible to determine the feasible binding mode of any given existing molecule to a target protein. Most three-dimensional QSAR models depend on the superposition of ligands, the identification of which is impossible without sufficient structural information for the target protein. The results of QSAR can be used to understand the interactions occurring between functional groups in the molecules with those of their target. QSAR methods are fast and can deal with real complex biological systems. QSAR models provide an alternative to animal testing and offer the advantages of reduced time for experimentation and minimized costs [95]. A few QSAR studies related to nanoparticles have been published to date [96–100]. QSAR can also predict the toxicity of nanoparticles. In addition, it can support targeting and filling gaps in knowledge for known nanoparticles [101]. QSAR tools are presented in Table 5. Recent reviews provide clear future directions for the field [102,103].

Table 5. List of tools available for QSAR.

No.	Name	Operating System	Url
1	ChemProp	Windows	[104]
2	OECD ToolBox	Windows/GNU Linux/Mac OS X	[105]
3	RmSquare	Web based application	[106]
4	T.E.S.T	Windows/GNU Linux/Mac OS X	[107]
5	Virtual Computational Chemistry Laboratory	Few online servers and commercial software	[108]
6	CORAL	Windows	[109]

High-Throughput Screening Data Analysis Tools (HDAT)

HDAT is a set of web-based high-throughput screening (HTS) data analysis tools [110,111]. The use of HTS in toxicity studies of engineered nanomaterials requires tools for quick and consistent processing and analyses of large HTS datasets. To facilitate this, a web-based platform was developed that presents statistical methods appropriate for toxicity data on engineered nanomaterials. It also offers different plate normalization methods, different HTS summarization statistics, self-organizing map (SOM)-based clustering analysis, and visualization of raw and processed information using heat maps and SOMs. HDAT has been successfully applied in the analysis of a number of HTS studies on the toxicity of engineered nanomaterials, thus enabling analysis of toxicity mechanisms and the development of data-driven structure-activity relationships for nanomaterial toxicity. HDAT offers online instructions and video demonstrations of basic operations such as data formatting and data uploading.

Table 6. List of tools available for molecular modeling.

No.	Name	Operating System	License	Url
Molecular Dynamic Simulations				
1	Amber	Unix-like systems	Academic/Non-profit/Government: \$400	[112]
2	NAMD	Windows/GNU Linux/Mac OS X/Solaris	Free for academics	[113]
3	Gromacs	Solaris, Linux, OS X, Windows by Cygwin, any other Unix variety	GNU Lesser General Public License for version	[114]
4	CHARMM	Unix-like	Academic research groups for a \$600 licensing fee	[115]
5	Desmond	Linux/Windows	Academic and Commercial	[116]
6	LAMMPS	Windows/GNU Linux/Mac OS X	GNU General Public License	[117]
7	TINKER	Windows/GNU Linux/Mac OS X	Available free of charge	[118]
8	MDynaMix	-	GNU Lesser General Public License	[119]
9	CP2K	GNU Linux	GNU Lesser General Public License	[120]
DFT/Ab initio/Monte Carlo/Semi-empirical calculations				
10	ABINIT	Windows/GNU Linux/Mac OS X	GNU General Public License	[121]
11	ACES	GNU Linux/SGI	GNU General Public License	[122]
12	BigDFT	Cross-platform	GNU General Public License	[123]
13	CASINO	Linux related operating system/ Windows by Cygwin	Charge-free license agreements for academics	[124]
14	CASTEP	-	Commercial software	[125]
15	COLUMBUS	GNU Linux	Obtained free of charge	[126]
16	CFOUR	GNU Linux	Charge-free license agreements for academics	[127]
17	CONQUEST	GNU Linux	GNU General Public License	[128]
18	CPMD	GNU Linux	CPMD License (free for non-profit)	[129]
19	CRYSTAL	Windows/GNU Linux/Mac OS X	Commercial software	[130]
20	DACAPO	GNU Linux	CAMP Open Software project	[131]
21	DALTON	GNU Linux	Charge-free license agreements for academics	[132]
22	DFTB+	GNU Linux	Free for academic, educational, and non-profit research	[133]
23	DIRAC	Windows/GNU Linux/Mac OS X	Charge-free license agreements	[134]
24	Elk	GNU Linux	GNU General Public License	[135]
25	ErgoSCF	GNU Linux	GNU General Public License	[136]
26	ERKALE	Windows/GNU Linux	GNU General Public License	[137]
27	FLEUR	GNU Linux	Academic	[138]
28	FreeON	GNU Linux	GNU General Public License	[139]
29	GAMESS	Windows/GNU Linux/Mac OS X	Charge free license agreements	[140]
30	GPAW	GNU Linux/Mac OS X	-	[141]
31	Gaussian	Windows/GNU Linux/Mac OS X	Commercial software	[142]
32	JDFTx	GNU Linux	GNU General Public License	[143]
33	MADNESS	Windows/GNU Linux/Mac OS X	GNU General Public License	[144]
34	MOPAC	Windows/GNU Linux/Mac OS X	Free and Commercial	[145]
35	MPQ	GNU Linux	GNU Lesser General Public License	[146]
36	NWChem	Linux, FreeBSD, Unix and like operating systems, Microsoft Windows, Mac OS X	Educational Community License	[147]
37	Octopus	GNU Linux	GNU General Public License	[148]
38	ONETEP	GNU Linux	Commercial software	[149]

Table 6. Cont.

No.	Name	Operating System	License	Url
DFT/Ab initio/Monte Carlo/Semi-empirical calculations				
39	OpenAtom	GNU Linux/Mac OS X	Free for academic	[150]
40	OpenMX	GNU Linux	GNU General Public License	[151]
41	ORCA	GNU Linux	Free for academic	[152]
42	PSI	GNU Linux/Windows, Mac OS X	GNU General Public License	[153]
43	PyQuante	-	BSD licenses (free for academic)	[154]
44	Quantum ESPRESSO	GNU Linux	Open source distribution	[155]
45	VASP	GNU Linux	Commercial software	[156]
46	Yambo Code	GNU Linux	GNU General Public License	[157]

5.3. MD Simulations

The most broadly used computational technique is MD simulation, which elucidates the structure-function relationships in biomolecules. In MD simulation, atoms and molecules are allowed to interact over time at a given temperature and are assumed to follow the laws of classical mechanics. MD simulation offers a detailed description of atomic motion, and the forces acting on the atoms are computed using a model known as a force field [158]. Such a simulation may serve as a “computational microscope” [159] that reveals interactions that may be difficult to observe experimentally. Dror *et al.* explained MD simulation as follows: “Static structural information might be likened to a photograph of a football game; to understand more readily how the game is played, we want a video recording” [159]. Computer simulation of biomolecule-nanomaterial interactions is gaining popularity as a complement to experimental techniques. When entering human bodies, nanoparticles interact with biomolecules and begin a series of nanoparticle/biological interactions that rely upon on forces as well as dynamic biophysicochemical interactions [11]. Many interactions such as hydrogen bonds, Van der Waals (VDW) forces, and electrostatic, hydrophobic, and π – π stacking interactions contribute to these protein-nanoparticle complexes [11]. The VDW attraction increases with increase in the proximity of the atoms at the interface, within a distance close to the total of the VDW radii of the two atoms. The VDW force acts at a short range and reduces significantly as the contacting atoms depart (Lennard-Jones potential or 6–12 potential). Hydrogen bonding is another important player in intermolecular interactions. Although the hydrogen bond is much stronger than the VDW interaction, the number of hydrogen bonds connecting proteins and nanoparticle is much smaller than that of VDW interactions. Hydrophobic interactions are an entropic effect originating from the repulsion of ordered water molecules from a nonpolar surface. The π – π stacking is an attraction involving aromatic rings. However, as only a few amino acids contain aromatic rings and the rings are often buried within the hydrophobic core of the proteins, only specific interactions are possible. A recently published review of molecular modeling in structural nanotoxicology discusses methods for predicting interactions between nanomaterials and biomolecules [160]. The review also discusses how nanoparticles of different sizes, shapes, structures, and chemical properties are able to influence the organization and function of nanomachinery in cells. In summary, computer modeling studies have

become a dominant tool in risk assessment studies. The molecular modeling tools currently available are shown in Table 6.

Table 7. Summary of image-related databases and tools. All tools listed are freely available for use.

No.	Name	Operating System	License	Url
1	Micro-manager	Windows and Mac	Copyright (distributed free of charge)	[161]
2	OMERO	Windows, Mac and Linux	GNU General Public License and few commercial licenses	[162]
3	Bisque	Windows, Mac and Linux	Free and open source	[163]
4	Bio-Formats	GNU Linux	GNU General Public License	[164]
5	ImageJ	Windows, Mac and GNU Linux (any java based)	Public domain	[165]
6	Fiji	Windows, Mac OS X, GNU Linux and Unix	GNU General Public License (the plugin interface is excluded from the license; some plugins have different licenses)	[166]
7	FARSIGHT TOOLKIT	Windows, Mac OS X and GNU Linux	Apache License, Version 2.0	[167]
8	ICY	Windows, Mac OS X and GNU Linux	GNU General Public License	[168]
9	CellProfiler	Windows, Mac OS X, and GNU Linux	GNU General Public License and some portions are also BSD-licensed	[169]
10	Micro-Pilot	Windows	Free for academic users	[170]
12	Cell ID	Windows and GNU Linux	GNU Lesser General Public License	[171]
13	bioView3D	Windows, Mac OS X, and GNU Linux	Open source (free for non-commercial)	[172]
14	V3D	Windows, Mac and GNU Linux	Free for non-profit research	[173]
15	IMOD	Windows, Mac and GNU Linux	Free to use	[174]
16	iCluster	Windows, Mac OS X, and GNU Linux	GNU General Public License	[175]
17	Image Surfer	Windows and Mac	Free to use	[176]

6. Imaging

Medical imaging technologies are used for diagnosis as well as for many other biomedical applications. The use of nanomaterials in imaging applications is growing rapidly. Molecular imaging can assist in early diagnosis and also provides information on pathological processes. Compared with currently used fluorescent proteins and small-molecule dyes, nanotechnology imaging probes can offer signals that are several-fold brighter and more stable. Nanoparticles, predominantly those made of precious metals, enhance signals in imaging approaches. For example, semiconductor quantum dots are tiny light-emitting particles that are used as fluorescent probes for molecular imaging and medical diagnostics. However, there are still many disadvantages associated with imaging technologies, such as tissue specificity and systemic toxicity. Progress in nanotechnology may overcome these challenges and offer more sensitive and specific information [177]. *In vivo* and *in vitro* imaging applications of nanomaterials have been reviewed [3,178]. The new informatics methods link tissue banks with histology data in order to offer enhanced image annotation down to the nano level [13]. Due to the

increase in the use of image analysis tools and methods in medical applications, well-supported tools and databases that can be used for nanotechnology have become increasingly available; these are shown in Table 7 [179]. Image databases and tools used in fluorescence microscopy can also be used to enhance image annotation with novel methods as the amount of experimental data at the nanolevel grows.

7. Computational Design of Biomolecular Nanostructures

DNA and protein self-assemblies are used to build nanostructures because of the programmability property of DNA and protein and the surplus of chemical techniques available for their manipulation. For example, DNA is a valuable self-assembly material because, with appropriate design of its base sequences, assembly is enabled through specific interactions among complementary base pairs [180]. Many self-assembly molecules that can be used as vehicles for drug delivery are built using DNA [181]. Some of the software currently available for designing three-dimensional structures of carbon nanotubes and for DNA origami design is listed in Table 8.

Table 8. Programs for DNA origami, nanomaterial design, and carbon nanotube design. All the tools listed are freely available for use.

No.	Name	Operating System	Url
1	Tiamat	Windows	[182]
2	caDNAno	Web-Accessible	[183]
3	SARSE	GNU Linux/Mac	[184]
4	cadnano	Windows, OSX and GNU Linux	[185]
5	Uniquimer3D	Windows	[186]
6	CoNTub	Windows, Mac OS X, GNU Linux and Unix	[187]
7	Fullerene	GNU Linux/Unix and Mac	[188]
8	CS Catalyst	Windows	[189]
9	Ninithi	Windows	[190]
10	TubeGen	Web-Accessible	[191]
11	Wrapping	Windows	[192]

8. Specific Challenges and Opportunities in Nanoinformatics

The databases and tools presented here highlight the increasing resources that are available to users. Their improvement is accompanied by the continuous expansion in the number of databases (Table 1) and various computer programs (Tables 6 and 7). The new database initiatives such as ISA-TAB-Nano, caNanoLab, and Nanomaterial Registry will facilitate data sharing, data standards, and, depending on the growth of nanomaterials data, the development of methods and tools specific to the nanolevel. Moreover, the growth of these fields essentially requires that the ISA-TAB-Nano standard be adopted by journals and other organizations to ensure consistent representation of nanotechnological data [19]. It should be noted that open accessibility and the liberty to use published DNA sequences in databases such as GenBank has encouraged scientists to build powerful methods, tools, and resources that have substantially enriched the field of bioinformatics [24]. It should also be noted that nanomaterial data are much more complex than sequence information or molecular data. An additional challenge is the

lack of availability of scientific literature from subscription-based journals in nanotechnology, particularly in chemistry, for which even abstracts are inaccessible for large-scale analyses such as text mining; this is a big drawback for text mining and nanoinformatics analysis [193]. With respect to QSAR, a recent review provided a summary of the advances and a clear roadmap for future research and also identified major gaps in the field [102]. Molecular modeling and imaging in nanotoxicology appear promising, as many software tools mentioned in this review are free under the General Public License, so that there is freedom to use, study, share, and modify the software. Although the modeling behavior of nanomaterials presents different challenges compared to those for drugs, it is nonetheless possible to adapt new methods and develop specific tools from existing tools. There are also anticipated beneficial side effects of nanoinformatics development, such as the possible minimization of animal studies. Recent initiatives and efforts emerging from the nanotechnology community are encouraging.

9. Conclusions

The databases and tools presented herein will be helpful to the nanotechnology community. We highlight the following recommendations that we hope will assist this community: (i) Consensus standards and ontology terms should be used in nanotechnology-related research; (ii) Journals should adopt formats to ensure uniform representation of information; (iii) Authors should submit their data to the databases; (iv) Similar to the current model of molecular biology databases, free and unrestricted access to nanomaterial data would enable nanoinformatics researchers to create tools that will benefit the broad nanotechnology community; (v) Open access and freely available nanotechnology literature will also benefit the nanotechnology community.

Acknowledgments

This work was supported by the Mid-Career Researcher Program through the National Research Foundation of Korea, funded by the Ministry of Education, Science, and Technology (2012-016803) and partly supported by a grant from the Priority Research Centers Program, Korea (NRF 2012-0006687).

Author Contributions

S.P. obtained and analyzed the data and wrote the manuscript; S.C. supervised the work and wrote the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Albanese, A.; Tang, P.S.; Chan, W.C. The effect of nanoparticle size, shape, and surface chemistry on biological systems. *Annu. Rev. Biomed. Eng.* **2012**, *14*, 1–16.
2. Toumey, C. Why not zwergo-technology? *Nat. Nanotechnol.* **2011**, *6*, 393–394.
3. Kim, B.Y.; Rutka, J.T.; Chan, W.C. Nanomedicine. *N. Engl. J. Med.* **2010**, *363*, 2434–2443.

4. Makarucha, A.J.; Todorova, N.; Yarovsky, I. Nanomaterials in biological environment: A review of computer modelling studies. *Eur. Biophys. J.* **2011**, *40*, 103–115.
5. Nie, S.; Xing, Y.; Kim, G.J.; Simons, J.W. Nanotechnology applications in cancer. *Annu. Rev. Biomed. Eng.* **2007**, *9*, 257–288.
6. Yokel, R.A.; Macphail, R.C. Engineered nanomaterials: Exposures, hazards, and risk prevention. *J. Occup. Med. Toxicol.* **2011**, *6*, 7.
7. Love, S.A.; Maurer-Jones, M.A.; Thompson, J.W.; Lin, Y.S.; Haynes, C.L. Assessing nanoparticle toxicity. *Annu. Rev. Anal. Chem.* **2012**, *5*, 181–205.
8. Bharali, D.J.; Khalil, M.; Gurbuz, M.; Simone, T.M.; Mousa, S.A. Nanoparticles and cancer therapy: A concise review with emphasis on dendrimers. *Int. J. Nanomed.* **2009**, *4*, 1–7.
9. Buzea, C.; Pacheco, I.I.; Robbie, K. Nanomaterials and nanoparticles: Sources and toxicity. *Biointerphases* **2007**, *2*, 17–71.
10. De la Iglesia, D.; Garcia-Remesal, M.; de la Calle, G.; Kulikowski, C.; Sanz, F.; Maojo, V. The impact of computer science in molecular medicine: Enabling high-throughput research. *Curr. Top. Med. Chem.* **2013**, *13*, 526–575.
11. Yang, S.T.; Liu, Y.; Wang, Y.W.; Cao, A. Biosafety and bioapplication of nanomaterials by designing protein-nanoparticle interactions. *Small* **2013**, *9*, 1635–1653.
12. Nanoinformatics. Available online: http://nanoinformatics.org/nanoinformatics/index.php/Main_Page (accessed on 14 March 2014).
13. Maojo, V.; Fritts, M.; de la Iglesia, D.; Cachau, R.E.; Garcia-Remesal, M.; Mitchell, J.A.; Kulikowski, C. Nanoinformatics: A new area of research in nanomedicine. *Int. J. Nanomed.* **2012**, *7*, 3867–3890.
14. NNI Signature Initiative. Available online: http://www.nano.gov/sites/default/files/nki_nsi_white_paper_-_may_14_2012_secured.pdf (accessed on 14 March 2014).
15. The ACTION-Grid White Paper on Nanoinformatics. Available online: <http://www.action-grid.eu/documents/Final%20White%20Paper%20-%20Nano.pdf> (accessed on 14 March 2014).
16. Ostraat, M.L.; Mills, K.C.; Guzan, K.A.; Murry, D. The Nanomaterial Registry: Facilitating the sharing and analysis of data in the diverse nanomaterial community. *Int. J. Nanomed.* **2013**, *8* (Suppl. 1), 7–13.
17. Sharon, G.; George, W.H.; Stephanie, A.M.; Michal, L.; Mervi, H.; Juli, D.K. Cananolab: Data sharing to expedite the use of nanotechnology in biomedicine. *Comput. Sci. Discov.* **2013**, *6*, 014010.
18. De la Iglesia, D.; Cachau, R.E.; García-Remesal, M.; Maojo, V. Nanoinformatics knowledge infrastructures: Bringing efficient information management to nanomedical research. *Comput. Sci. Discov.* **2013**, *6*, 014011.
19. Thomas, D.G.; Gaheen, S.; Harper, S.L.; Fritts, M.; Klaessig, F.; Hahn-Dantona, E.; Paik, D.; Pan, S.; Stafford, G.A.; Freund, E.T.; *et al.* ISA-TAB-Nano: A specification for sharing nanomaterial research data in spreadsheet-based format. *BMC Biotechnol.* **2013**, *13*, 2.
20. Inkscape Is a Professional Vector Graphics. Available online: <http://www.inkscape.org/en/> (accessed on 14 March 2014).
21. GIMP—The GNU Image Manipulation Program. Available online: <http://www.gimp.org/> (accessed on 14 March 2014).

22. Genbank. Available online: <https://www.ncbi.nlm.nih.gov/genbank/> (accessed on 14 March 2014).
23. Protein Data Bank. Available online: <http://www.rcsb.org/pdb/home/home.do> (accessed on 14 March 2014).
24. Brown, P.O.; Eisen, M.B.; Varmus, H.E. Why PLoS became a publisher. *PLoS Biol.* **2003**, *1*, e36.
25. Fox, J.A.; Butland, S.L.; McMillan, S.; Campbell, G.; Ouellette, B.F. The bioinformatics links directory: A compilation of molecular biology web servers. *Nucleic Acids Res.* **2005**, *33*, W3–W24.
26. Web Server Issue. Available online: <http://nar.oxfordjournals.org/content/41/W1.toc> (accessed on 14 March 2014).
27. González-Nilo, F.; Pérez-Acle, T.; Guínez-Molinos, S.; Geraldo, D.A.; Sandoval, C.; Yévenes, A.; Santos, L.S.; Laurie, V.F.; Mendoza, H.; Cachau, R.E. Nanoinformatics: An emerging area of information technology at the intersection of bioinformatics, computational chemistry and nanobiotechnology. *Biol. Res.* **2011**, *44*, 43–51.
28. Nanomaterial-Biological Interactions Knowledgebase. Available online: <http://nbi.oregonstate.edu/> (accessed on 14 March 2014).
29. Internano—Resources for Nanomanufacturing. Available online: <http://www.internano.org/> (accessed on 14 March 2014).
30. Nano-EHS Database Analysis Tool. Available online: <http://icon.rice.edu/report.cfm> (accessed on 14 March 2014).
31. Nanoparticle Information Library. Available online: <http://nanoparticlelibrary.net/> (accessed on 14 March 2014).
32. Nanomaterial Registry. Available online: <https://www.nanomaterialregistry.org/> (accessed on 14 March 2014).
33. National Toxicology Program Database. Available online: http://tools.niehs.nih.gov/ntp_tox/ (accessed on 14 March 2014).
34. Nano-HUB Database. Available online: <http://nanohub.org/resources/databases> (accessed on 14 March 2014).
35. National Center for Biomedical Ontology. Available online: <http://www.bioontology.org/> (accessed on 14 March 2014).
36. ISA-TAB-Nano. Available online: <https://wiki.nci.nih.gov/display/ICR/ISA-TAB-Nano> (accessed on 14 March 2014).
37. caNanoLab. Available online: <https://cananolab.nci.nih.gov/caNanoLab/> (accessed on 14 March 2014).
38. Toxicology Data Network. Available online: <http://toxnet.nlm.nih.gov/> (accessed on 14 March 2014).
39. Nanotechnology Characterization Laboratory. Available online: http://ncl.cancer.gov/objectives_ncl_obj5.asp (accessed on 14 March 2014).
40. The Collaboratory for Structural Nanobiology. Available online: http://csn.ncifcrf.gov/Advanced_Structure_Analysis/HOME.html (accessed on 14 March 2014).
41. Toxicology Literature Online. Available online: <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE> (accessed on 14 March 2014).

42. OECD Database on Research into the Safety of Manufactured Nanomaterials. Available online: <http://webnet.oecd.org/NANOMATERIALS/Pagelet/Front/Default.aspx> (accessed on 14 March 2014).
43. Schuurman, N.; Leszczynski, A. Ontologies for bioinformatics. *Bioinform. Biol. Insights* **2008**, *2*, 187–200.
44. Bard, J.B.; Rhee, S.Y. Ontologies in biology: Design, applications and future challenges. *Nat. Rev. Genet.* **2004**, *5*, 213–222.
45. Thomas, D.G.; Klaessig, F.; Harper, S.L.; Fritts, M.; Hoover, M.D.; Gaheen, S.; Stokes, T.H.; Reznik-Zellen, R.; Freund, E.T.; Klemm, J.D.; *et al.* Informatics and standards for nanomedicine technology. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **2011**, *3*, 511–532.
46. NanoParticle Ontology for Cancer Nanotechnology Research. Available online: <http://www.nano-ontology.org/> (accessed on 14 March 2014).
47. BioPortal. Available online: <http://bioportal.bioontology.org/> (accessed on 14 March 2014).
48. The Open Biological and Biomedical Ontologies. Available online: <http://www.obofoundry.org/> (accessed on 14 March 2014).
49. caOBR—A Tool for Indexing Cancer Informatics Resources. Available online: <http://www.bioontology.org/caOBR> (accessed on 14 March 2014).
50. The Database and Ontology of Chemical Entities of Biological Interest. Available online: <http://www.ebi.ac.uk/chebi/> (accessed on 14 March 2014).
51. The Gene Ontology. Available online: <http://www.geneontology.org/> (accessed on 14 March 2014).
52. The Foundational Model of Anatomy Ontology. Available online: <http://sig.biostr.washington.edu/projects/fm/AboutFM.html> (accessed on 14 March 2014).
53. Zebrafish Anatomy and Development Ontology. Available online: http://www.obofoundry.org/cgi-bin/detail.cgi?id=zebrafish_anatomy (accessed on 14 March 2014).
54. Ontology Lookup Service. Available online: <http://www.ebi.ac.uk/ontology-lookup/browse.do?ontName=REX> (accessed on 14 March 2014).
55. Phenotypic Quality Ontology. Available online: <http://bioportal.bioontology.org/ontologies/PATO> (accessed on 14 March 2014).
56. Cohen, K.B.; Hunter, L. Getting started in text mining. *PLoS Comput. Biol.* **2008**, *4*, e20.
57. Rzhetsky, A.; Seringhaus, M.; Gerstein, M.B. Getting started in text mining: Part two. *PLoS Comput. Biol.* **2009**, *5*, e1000411.
58. Grulke, C.M.; Goldsmith, M.R.; Vallero, D.A. Toward a blended ontology: Applying knowledge systems to compare therapeutic and toxicological nanoscale domains. *J. Biomed. Biotechnol.* **2012**, *2012*, 308381.
59. Garcia-Remesal, M.; Garcia-Ruiz, A.; Perez-Rey, D.; de la Iglesia, D.; Maojo, V. Using nanoinformatics methods for automatically identifying relevant nanotoxicology entities from the literature. *Biomed. Res. Int.* **2013**, *2013*, 410294.
60. Liu, R.; France, B.; George, S.; Rallo, R.; Zhang, H.; Xia, T.; Nel, A.E.; Bradley, K.; Cohen, Y. Association rule mining of cellular responses induced by metal and metal oxide nanoparticles. *Analyst* **2014**, *139*, 943–953.
61. Google Scholar. Available online: <http://scholar.google.com> (accessed on 14 March 2014).

62. Gopubmed. Available online: <http://www.gopubmed.org/> (accessed on 14 March 2014).
63. Textpresso. Available online: <http://www.textpresso.org/> (accessed on 14 March 2014).
64. BioRAT. Available online: http://bioinf.cs.ucl.ac.uk/software_downloads/biorat/ (accessed on 14 March 2014).
65. iHOP. Available online: <http://www.ihop-net.org/UniPub/iHOP/> (accessed on 14 March 2014).
66. Lewar, E.G. *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, 2nd ed.; Springer: New York, NY, USA, 2010.
67. Molecular Mechanics and Dynamics. Available online: <http://employees.csbsju.edu/hjakubowski/classes/ch331/protstructure/mechdynam2.html> (accessed on 14 March 2014).
68. PyMOL. Available online: <http://www.pymol.org/> (accessed on 14 March 2014).
69. RasMol. Available online: <http://www.bernstein-plus-sons.com/software/rasmol/> (accessed on 14 March 2014).
70. Raster3D. Available online: <http://skuld.bmsc.washington.edu/raster3d/raster3d.html> (accessed on 14 March 2014).
71. UCSF Chimera. Available online: <http://www.cgl.ucsf.edu/chimera/index.html> (accessed on 14 March 2014).
72. Visual Molecular Dynamics. Available online: <http://www.ks.uiuc.edu/Research/vmd/> (accessed on 14 March 2014).
73. Cn3D Is a Macromolecular Structure Viewer. Available online: <http://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml> (accessed on 14 March 2014).
74. BALL—Biochemical Algorithms Library. Available online: <http://www.ball-project.org/> (accessed on 14 March 2014).
75. Jmol. Available online: <http://jmol.sourceforge.net/> (accessed on 14 March 2014).
76. ICM-Browser. Available online: http://www.molsoft.com/icm_browser.html (accessed on 14 March 2014).
77. Chemkit. Available online: <http://sourceforge.net/projects/chemkit/files/chemkit-0.1.tar.gz/download> (accessed on 14 March 2014).
78. OpenStructure. Available online: <http://www.openstructure.org/> (accessed on 14 March 2014).
79. VEGA ZZ. Available online: <http://www.vegazz.net/> (accessed on 14 March 2014).
80. Molekel. Available online: <http://molekel.cscs.ch/wiki/pmwiki.php> (accessed on 14 March 2014).
81. Avogadro. Available online: http://avogadro.cc/wiki/Main_Page (accessed on 14 March 2014).
82. Xeo. Available online: <http://sourceforge.net/projects/xeo/> (accessed on 14 March 2014).
83. ChemSketch. Available online: <http://www.acdlabs.com/resources/freeware/chemsketch/> (accessed on 14 March 2014).
84. Hu, X.; Shelver, W.H. Docking studies of matrix metalloproteinase inhibitors: Zinc parameter optimization to improve the binding free energy prediction. *J. Mol. Graph. Model.* **2003**, *22*, 115–126.
85. Morris, G.M.; Huey, R.; Olson, A.J. Using autoDock for ligand-receptor docking. *Curr. Protoc. Bioinform.* **2008**, doi:10.1002/0471250953.bi0814s24.
86. Yang, S.T.; Wang, H.; Guo, L.; Gao, Y.; Liu, Y.; Cao, A. Interaction of fullerenol with lysozyme investigated by experimental and computational approaches. *Nanotechnology* **2008**, *19*, 395101.
87. Calvaresi, M.; Zerbetto, F. Baiting Proteins with C60. *ACS Nano* **2010**, *4*, 2283–2299.

88. AutoDock 4. Available online: <http://autodock.scripps.edu/> (accessed on 14 March 2014).
89. AutoDock Vina. Available online: <http://vina.scripps.edu/> (accessed on 14 March 2014).
90. Dock. Available online: <http://dock.compbio.ucsf.edu/> (accessed on 14 March 2014).
91. GOLD. Available online: <http://www.ccdc.cam.ac.uk/Solutions/GoldSuite/Pages/GOLD.aspx> (accessed on 14 March 2014).
92. FlexX. Available online: <http://www.biosolveit.de/FlexX/> (accessed on 14 March 2014).
93. Surflex-Dock. Available online: http://www.tripos.com/index.php?family=modules,SimplePage,,,&page=surflex_dock&s=0 (accessed on 14 March 2014).
94. Tropsha, A. Best Practices for QSAR model development, validation, and exploitation. *Mol. Inf.* **2010**, *29*, 476–488.
95. Mays, C.; Benfenati, E.; Pardoe, S. Use and perceived benefits and barriers of QSAR models for reach: Findings from a questionnaire to stakeholders. *Chem. Cent. J.* **2012**, *6*, 159.
96. Toropov, A.A.; Toropova, A.P.; Benfenati, E.; Leszczynska, D.; Leszczynski, J. Smiles-based optimal descriptors: QSAR analysis of fullerene-based HIV-1 pr inhibitors by means of balance of correlations. *J. Comput. Chem.* **2010**, *31*, 381–392.
97. Fourches, D.; Pu, D.; Tassa, C.; Weissleder, R.; Shaw, S.Y.; Mumper, R.J.; Tropsha, A. Quantitative nanostructure—Activity relationship modeling. *ACS Nano* **2010**, *4*, 5703–5712.
98. Fourches, D.; Pu, D.; Tropsha, A. Exploring quantitative nanostructure-activity relationships (QNAR) modeling as a tool for predicting biological effects of manufactured nanoparticles. *Comb. Chem. High Throughput Screen.* **2011**, *14*, 217–225.
99. Puzyn, T.; Rasulev, B.; Gajewicz, A.; Hu, X.; Dasari, T.P.; Michalkova, A.; Hwang, H.M.; Toropov, A.; Leszczynska, D.; Leszczynski, J. Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nat. Nano* **2011**, *6*, 175–178.
100. Liu, R.; Rallo, R.; George, S.; Ji, Z.; Nair, S.; Nel, A.E.; Cohen, Y. Classification NanoSAR development for cytotoxicity of metal oxide nanoparticles. *Small* **2011**, *7*, 1118–1126.
101. Poater, A.; Saliner, A.G.; Solà, M.; Cavallo, L.; Worth, A.P. Computational methods to predict the reactivity of nanoparticles through structure–property relationships. *Expert Opin. Drug Deliv.* **2010**, *7*, 295–305.
102. Winkler, D.A.; Mombelli, E.; Pietroiusti, A.; Tran, L.; Worth, A.; Fadeel, B.; McCall, M.J. Applying quantitative structure-activity relationship approaches to nanotoxicology: Current status and future potential. *Toxicology* **2013**, *313*, 15–23.
103. Epa, V.C.; Burden, F.R.; Tassa, C.; Weissleder, R.; Shaw, S.; Winkler, D.A. Modeling biological activities of nanoparticles. *Nano Lett.* **2012**, *12*, 5808–5812.
104. ChemProp. Available online: <http://www.ufz.de/index.php?en=10684> (accessed on 14 March 2014).
105. OECD. Available online: <http://www.oecd.org/env/ehs/> (accessed on 14 March 2014).
106. RmSquare. Available online: <http://aptsoftware.co.in/rmsquare/> (accessed on 14 March 2014).
107. T.E.S.T. Available online: <http://www.epa.gov/nrmrl/std/qsar/qsar.html> (accessed on 14 March 2014).
108. Virtual Computational Chemical Laboratory. Available online: <http://www.vcclab.org/> (accessed on 14 March 2014).
109. CORAL. Available online: <http://www.insilico.eu/coral/> (accessed on 14 March 2014).

110. Rong, L.; Taimur, H.; Robert, R.; Yoram, C. HDAT: Web-based high-throughput screening data analysis tools. *Comput. Sci. Discov.* **2013**, *6*, 014006.
111. High Throughput Screening Data Analysis Tools (HDAT). Available online: <http://nanoinfo.org/hdat/> (accessed on 14 March 2014).
112. Assisted Model Building with Energy Refinement. Available online: <http://ambermd.org/> (accessed on 14 March 2014).
113. NAMD—Scalable Molecular Dynamics. Available online: <http://www.ks.uiuc.edu/Research/namd/> (accessed on 14 March 2014).
114. Gromacs. Available online: <http://www.gromacs.org/> (accessed on 14 March 2014).
115. Chemistry at HARvard Macromolecular Mechanics. Available online: <http://www.charmm.org/> (accessed on 14 March 2014).
116. Desmond. Available online: http://www.deshawresearch.com/resources_desmond.html (accessed on 14 March 2014).
117. LAMMPS Molecular Dynamics Simulator. Available online: <http://lammps.sandia.gov/> (accessed on 14 March 2014).
118. TINKER—Software Tools for Molecular Design. Available online: <http://dasher.wustl.edu/tinker/> (accessed on 14 March 2014).
119. MDynaMix: A Molecular Dynamics Program. Available online: <http://www.mmk.su.se/~sasha/mdynamix/> (accessed on 14 March 2014).
120. CP2K—Open Source Molecular Dynamics. Available online: <http://www.cp2k.org/> (accessed on 14 March 2014).
121. ABINIT. Available online: <http://www.abinit.org/> (accessed on 14 March 2014).
122. Advanced Concepts in Electronic Structure. Available online: <http://www.qtp.ufl.edu/ACES/> (accessed on 14 March 2014).
123. BigDFT. Available online: http://bigdft.org/Wiki/index.php?title=BigDFT_website (accessed on 14 March 2014).
124. Quantum Monte Carlo. Available online: <http://vallico.net/casinoqmc/> (accessed on 14 March 2014).
125. CASTEP. Available online: <http://www.castep.org/> (accessed on 14 March 2014).
126. The Columbus Quantum Chemistry Programs. Available online: <http://www.univie.ac.at/columbus/> (accessed on 14 March 2014).
127. Coupled-Cluster Techniques for Computational Chemistry. Available online: <http://www.cfour.de/> (accessed on 14 March 2014).
128. Conquest: Linear Scaling DFT. Available online: <http://www.order-n.org/> (accessed on 14 March 2014).
129. CPMD. Available online: <http://www.cpmc.org/> (accessed on 14 March 2014).
130. A Computational Tool for Solid State Chemistry and Physics. Available online: <http://www.crystal.unito.it/index.php> (accessed on 14 March 2014).
131. Dacapo Is a Total Energy Program Based on Density Functional Theory. Available online: <https://wiki.fysik.dtu.dk/dacapo> (accessed on 14 March 2014).
132. Dalton. Available online: <http://daltonprogram.org/> (accessed on 14 March 2014).

133. Density Functional Based Tight Binding. Available online: <http://www.dftb-plus.info/> (accessed on 14 March 2014).
134. DIRAC—Program for Atomic and Molecular Direct Iterative Relativistic All-electron Calculations. Available online: <http://www.diracprogram.org/doku.php> (accessed on 14 March 2014).
135. The Elk FP-LAPW Code. Available online: <http://elk.sourceforge.net/> (accessed on 14 March 2014).
136. ErgoSCF Is a Quantum Chemistry Program for Large-Scale Self-Consistent Field Calculations. Available online: <http://www.ergoscf.org/> (accessed on 14 March 2014).
137. ERKALE Is a Quantum Chemistry Program Used to Solve the Electronic Structure of Atoms, Molecules and Molecular Clusters. Available online: <https://code.google.com/p/erkale/> (accessed on 14 March 2014).
138. FLEUR—The Julich FLAPW Code Family. Available online: <http://www.flapw.de/pm/index.php> (accessed on 14 March 2014).
139. FreeON Is a Suite of Programs for Linear Scaling Quantum Chemistry. Available online: https://www.freeon.org/index.php/Main_Page (accessed on 14 March 2014).
140. The General Atomic and Molecular Electronic Structure System (GAMESS) Is a General *ab initio* Quantum Chemistry Package. Available online: <http://www.msg.ameslab.gov/gamess/index.html> (accessed on 14 March 2014).
141. Grid-Based Projector-Augmented Wave Method. Available online: <https://wiki.fysik.dtu.dk/gpaw/> (accessed on 14 March 2014).
142. Gaussian. Available online: <http://www.gaussian.com/> (accessed on 14 March 2014).
143. Joint Density Functional Theory. Available online: <http://sourceforge.net/p/jdftx/wiki/Home/> (accessed on 14 March 2014).
144. Multiresolution Adaptive Numerical Environment for Scientific Simulation. Available online: <http://code.google.com/p/m-a-d-n-e-s-s/> (accessed on 14 March 2014).
145. A Practical Quantum Chemistry Tool for Modeling Biological Systems and Co-Crystals. Available online: <http://cacheresearch.com/mopac/Mopac2002manual/> (accessed on 14 March 2014).
146. The Massively Parallel Quantum Chemistry Program. Available online: <http://www.mppqc.org/> (accessed on 14 March 2014).
147. NWChem—Delivering High-Performance Computational Chemistry. Available online: http://www.nwchem-sw.org/index.php/Main_Page (accessed on 14 March 2014).
148. Octopus Is a Scientific Program Aimed at the *Ab Initio* Virtual Experimentation. Available online: http://www.tddft.org/programs/octopus/wiki/index.php/Main_Page (accessed on 14 March 2014).
149. ONTEP Is a Linear-Scaling Code for Quantum-Mechanical Calculations Based on Density-Functional Theory. Available online: <http://www2.tcm.phy.cam.ac.uk/onetep/> (accessed on 14 March 2014).
150. OpenAtom Is a Highly Scalable and Portable Parallel Application for Molecular Dynamics Simulations at the Quantum Level. Available online: <http://charm.cs.uiuc.edu/OpenAtom/> (accessed on 14 March 2014).
151. Open Source Package for Material Explorer. Available online: <http://www.openmx-square.org/> (accessed on 14 March 2014).

152. The ORCA Program Is to Carry out Geometry Optimizations and to Predict a Large Number of Spectroscopic Parameters at Different Levels of Theory. Available online: <http://www.ccc.mpg.de/forum/> (accessed on 14 March 2014).
153. PSI4—*Ab Initio* Quantum Chemistry Program. Available online: <http://www.psicode.org/> (accessed on 14 March 2014).
154. PyQuante—Python Quantum Chemistry. Available online: <http://pyquante.sourceforge.net/> (accessed on 14 March 2014).
155. Quantum Espresso Is an Integrated Suite of Open-Source Computer Codes for Electronic-Structure Calculations and Materials Modeling at the Nanoscale. Available online: <http://www.quantum-espresso.org/> (accessed on 14 March 2014).
156. Vienna *Ab Initio* Simulation Package. Available online: <http://www.vasp.at/> (accessed on 14 March 2014).
157. Yambo Is a FORTRAN/C Code for Many-Body Calculations in Solid State and Molecular Physics. Available online: <http://www.yambo-code.org/> (accessed on 14 March 2014).
158. Dodson, G.G.; Lane, D.P.; Verma, C.S. Molecular simulations of protein dynamics: New windows on mechanisms in biology. *EMBO Rep.* **2008**, *9*, 144–150.
159. Dror, R.O.; Dirks, R.M.; Grossman, J.P.; Xu, H.; Shaw, D.E. Biomolecular simulation: A computational microscope for molecular biology. *Annu. Rev. Biophys.* **2012**, *41*, 429–452.
160. Yanamala, N.; Kagan, V.E.; Shvedova, A.A. Molecular modeling in structural nano-toxicology: Interactions of nano-particles with nano-machinery of cells. *Adv. Drug Deliv Rev.* **2013**, *65*, 2070–2077.
161. Micro-Manager Open Source Microscopy Software. Available online: <http://www.micro-manager.org/wiki/Micro-Manager> (accessed on 14 March 2014).
162. OMERO. Available online: <http://www.openmicroscopy.org/site/products/omero> (accessed on 14 March 2014).
163. Bisque Database. Available online: <http://www.bioimage.ucsb.edu/bisque> (accessed on 14 March 2014).
164. Bio-Formats Is a Standalone Java Library for Reading and Writing Life Sciences Image file Formats. Available online: <http://loci.wisc.edu/software/bio-formats> (accessed on 14 March 2014).
165. Image Processing and Analysis in Java. Available online: <http://rsb.info.nih.gov/ij/> (accessed on 14 March 2014).
166. Fiji Is an Image Processing Package. Available online: <http://fiji.sc/Fiji> (accessed on 14 March 2014).
167. Quantitative Tool for Studying Complex and Dynamic Biological Microenvironments from 4D/5D Microscopy Data. Available online: http://farsight-toolkit.org/wiki/Main_Page (accessed on 14 March 2014).
168. Icy—An Open Community Platform for Bioimage Informatics. Available online: <http://icy.bioimageanalysis.org/> (accessed on 14 March 2014).
169. Cellprofiler—Cell Image Analysis Software. Available online: <http://www.cellprofiler.org/> (accessed on 14 March 2014).

170. Micropilot—Microscope Automation. Available online: http://www.embl.de/services/core_facilities/almf/services/high-throughput-microscopy/microscope-automation/micropilot/index.html (accessed on 14 March 2014).
171. Cell-ID Is an Open-Source Cell-Finding, Tracking, and Analysis Package. Available online: <http://www.molsci.org/protocols/software.html> (accessed on 14 March 2014).
172. BioView3D Is an Open Source and Cross-Platform Application Intended for Biologists to Visualize 3D Stack Imagery. Available online: <http://www.bioimage.ucsb.edu/downloads/BioView3D> (accessed on 14 March 2014).
173. Vaa3D Is a Swiss Army Knife for Bioimage Visualization & Analysis. Available online: http://penglab.janelia.org/proj/vaa3d/Vaa3D/About_Vaa3D.html (accessed on 14 March 2014).
174. The IMOD Is a Set of Image Processing, Modeling and Display Programs. Available online: <http://bio3d.colorado.edu/imod/> (accessed on 14 March 2014).
175. iCluster SubCellular Localisation Image Visualiser. Available online: <http://icluster.imb.uq.edu.au/> (accessed on 14 March 2014).
176. ImageSurfer Is free 3D Imaging Software to Visualize and Analyze Multi-Channel Volumes. Available online: <http://imagesurfer.cs.unc.edu/> (accessed on 14 March 2014).
177. Rosen, J.E.; Yoffe, S.; Meerasa, A.; Verma, M.; Gu, F.X. Nanotechnology and diagnostic imaging: New advances in contrast agent technology. *J. Nanomed. Nanotechnol.* **2011**, *2*, 115.
178. Kostarelos, K.; Bianco, A.; Prato, M. Promises, facts and challenges for carbon nanotubes in imaging and therapeutics. *Nat. Nano* **2009**, *4*, 627–633.
179. Hamilton, N.A. Open source tools for fluorescent imaging. *Methods Enzymol.* **2012**, *504*, 393–417.
180. Sanderson, K. Bioengineering: What to make with DNA origami. *Nature* **2010**, *464*, 158–159.
181. Zhao, Y.X.; Shaw, A.; Zeng, X.; Benson, E.; Nystrom, A.M.; Hogberg, B. DNA origami delivery system for cancer therapy with tunable release properties. *ACS Nano* **2012**, *6*, 8684–8691.
182. Tiamat Is a Graphical User Interface Program for the Design of DNA Nanostructures and Sequences. Available online: <http://yanlab.asu.edu/Resources.html> (accessed on 14 March 2014).
183. Computer-Aided Engineering for DNA Origami. Available online: <http://cando-dna-origami.org/> (accessed on 14 March 2014).
184. DNA Origami Design Software. Available online: <http://www.cdna.dk/default.asp?getreq=%2Findex.php%2Fsoftware.html> (accessed on 14 March 2014).
185. Cadnano Simplifies and Enhances the Process of Designing Three-Dimensional DNA Origami Nanostructures. Available online: <http://cadnano.org/> (accessed on 14 March 2014).
186. Uniquimer3D Is a User-Friendly Software System for Structural DNA Nanotechnology. Available online: <http://ihome.ust.hk/~keymix/> (accessed on 14 March 2014).
187. Contub Is a Set of Tools Dedicated to the Construction of Complex Carbon Nanotube Structures for Use in Computational Chemistry. Available online: <http://www.ugr.es/~gmdm/java/contub/contub.html> (accessed on 14 March 2014).
188. A Program for Creating Fullerene Structures and for Performing Topological Analyses. Available online: <http://ctcp.massey.ac.nz/index.php?group=&page=fullerenes&menu=fullerenes> (accessed on 14 March 2014).

189. CS Catalyst. Available online: <http://www.camsoft.co.kr/chemnews/catalyst/5/51/bucky.html> (accessed on 14 March 2014).
190. Ninithi Is a Free Modeling Software Used in Nanotechnology. Available online: <http://sourceforge.net/projects/ninithi/files/> (accessed on 14 March 2014).
191. TubeGen Is a Web Accessible Nanotube Structure Generator. Available online: <http://turin.nss.udel.edu/research/tubegenonline.html> (accessed on 14 March 2014).
192. Nanotube Coordinate Generator with a Viewer for Windows. Available online: <http://www.photon.t.u-tokyo.ac.jp/~maruyama/wrapping3/wrapping.html> (accessed on 14 March 2014).
193. Van Noorden, R. Tensions grow as data-mining discussions fall apart. *Nature* **2013**, *498*, 14–15.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).