
CHAPTER 102

Toward Integrated Nanosystems: Fundamental Issues in Design and Modeling

K. Eric Drexler

Molecular Engineering Research Institute, Los Altos, California, USA

CONTENTS

1.	Introduction	1
2.	The Leading Role of Computational Modeling	2
3.	The Diverse Goals of Computational Modeling	4
3.1.	Science and Design Differ Fundamentally	4
3.2.	Protein Design Required the Recognition of a Design Problem	5
3.3.	System Design Presents Distinct Challenges	6
4.	The Strategic Role of Fabrication Technologies	8
4.1.	Nanofabrication Directs the Motion of Atoms	8
4.2.	Self-Assembly Enables Precise Fabrication of Nanosystems	9
4.3.	Productive Nanosystems Offer a Strategic Goal	10
4.4.	Productive Nanosystems Invite Multistage Exploratory Design	10
5.	Conclusion	12
	References	13

1. INTRODUCTION

As technological development moves from scientific investigation to system building, the task of design becomes steadily more important. Successful design of a technological system requires knowledge that enables designers to describe, with adequate reliability, how the components of the system can be fabricated and assembled and how they will behave and interact. This is as true for designing nanosystems as it is for designing the macrotechnologies of aircraft, computers, or factories. But how much knowledge is necessary for successful design?

To ask “how much knowledge?” is, of course, simplistic—scientific knowledge is not a scalar quantity, but a map of the many-dimensional world of nature and artifacts. Even the notion of “successful design” is simplistic—in every area of technology, some systems are simple and robust enough to be designed with confidence, and other systems are so complex or so dependent on unknowns that every design becomes a genuine experiment. To advance nanotechnologies from scientific investigation to system building, we must be careful in formulating both our questions and our standards for judging answers.

This handbook explores diverse modeling methods applied to diverse physical systems with diverse applications. The modeling methods range from *ab initio* quantum chemistry (based directly on fundamental physics), through molecular dynamics (keeping atoms but omitting wavefunctions), to continuum approximations (omitting both atoms and wavefunctions). The physical systems and their state variables range from crystals with magnetic degrees of freedom through disordered polymers with mechanical degrees of freedom. Applications are at least as diverse, ranging from processing molecules using enzymatic catalysts to processing data using superpositions of quantum states.

This chapter explores issues that arise in all areas of theoretical and computational nanotechnology as questions advance from the purely scientific to the technological. It pays particular attention to the author’s own areas of contribution—in protein engineering, molecular machines, and productive nanosystems—examining their relevance to the broader enterprise of nanotechnology and using them to illustrate universal issues at the junction of science, technology, modeling, and design.

Section 2 reviews the role of computational modeling in developing nanotechnologies, showing how computational modeling can, to an unprecedented degree, play a leading role in shaping strategies for research and system development. Section 3 examines the diverse goals of computational modeling, showing why effective research programs must distinguish among the differing goals of science, exploratory design, and design for implementation. Section 4 examines the role of fabrication technologies in the development of nanotechnologies, showing the strategic role of advances in fabrication and the role of computational modeling in facilitating that advance.

2. THE LEADING ROLE OF COMPUTATIONAL MODELING

Theoretical and computational methods will play a leading role in the development of nanosystems, guiding and speeding development. Modeling based on these methods can indicate targets for laboratory development, can reduce experimental failures by rejecting unworkable ideas, and in some instances can enable reliable prediction and design. Because modeling is not constrained by the current limits of our ability to fabricate nanostructures, it can explore and examine potential targets for development well in advance of their physical feasibility. Thus, modeling can help to identify fruitful lines of development for research and investment.

Nanofabrication technologies today can produce only a small subset of the stable, technologically useful structures that are physically possible. For example, in the realm of atomically precise covalent structures (other than regular structures, such as crystals), current technologies can produce little but polymers with controlled sequences of monomers and relatively small organic molecules. Precise construction of extended structures with three-dimensional covalent bonding is easy to conceive and could readily be accomplished, but only by using tools that do not yet exist. Developing fabrication technologies that enable the construction of a wider range of atomically precise structures is a fundamental goal for nanotechnology. By simulating currently inaccessible structures and molecular transformations, modeling can explore applications and development strategies for advanced fabrication processes.

In assessing the adequacy of current modeling capabilities for the design of advanced systems and processes, it is important to remember that structures that are less accessible to fabrication can be more accessible to modeling. For example, the flexible polymeric chains that can be made today explore, through thermal motion, a set of conformations that scales exponentially with chain length. The probability of different conformations depends on an often-delicate balance of energies resulting from London dispersion forces, electrostatic

forces, torsional potentials, and solvent interactions. In some instances (e.g., folded proteins, DNA junctions), these weak forces cooperate to determine a single preferred shape for the molecule. In contrast, extended structures with three-dimensional covalent bonding (e.g., large organic molecules with diamond-lattice frameworks, among many other classes) can form stiff, elastic objects having shapes that are perturbed by weak forces but are not determined by them. The properties of these stiff covalent structures increase the simplicity and reliability of modeling and design for several reasons.

First, the equilibrium geometry of the structure is typically the sole minimum of a nearly parabolic potential energy surface. There is no need to search among different conformations to compare multiple minima having similar energies.

Second, structural stiffness facilitates free-energy calculations. In flexible polymers, these calculations typically require estimating integrals of the potential energy over a complex surface in a high-dimensional space, typically requiring evaluation of the energy in many quite-different configurations. The free energy of stiff structures, in contrast, can be estimated accurately knowing only the normal modes of the structure, which are often available analytically from a single-point calculation.

Finally, both geometries and (differential) free energies are relatively insensitive to errors in the model potential. The difference between the actual physical potential $E(\mathbf{r})$ and a model potential $E'(\mathbf{r})$ is equivalent to a perturbing force field $\Delta\mathbf{F}'(\mathbf{r}) = -\nabla[E(\mathbf{r}) - E'(\mathbf{r})]$ applied to the model. For reasonably accurate models within their range of applicability, the perturbing forces $\Delta\mathbf{F}'(\mathbf{r})$ are small. In modeling stiff structures, these perturbing forces are associated with small elastic deformations and correspondingly small errors. In flexible polymers, however, the same forces can be associated with shifts among conformations with grossly different geometries. Thus, a modeling technique may prove unreliable when applied to soft structures that can be made today, yet succeed when applied to stiff structures that can be made only with more advanced fabrication systems.

Figure 1 illustrates the design of a nanomachine (a molecular planetary gear) based on stiff covalent structures related to diamond, silicon, and silicon-carbide lattices; these structures place it well beyond reach of current fabrication technologies. The shape and function of its parts have proved insensitive to differences among several molecular mechanics packages.

The principle that advances in fabrication can reduce modeling difficulties has broad application. For example, the drive for lower defect rates in fabrication reduces the set of physical structures that must be modeled to describe products. When products are atomically precise, this set collapses to a singleton. For another example, the drive to fabricate components with better performance can (when holding the desired system-level performance constant) provide a margin of safety that makes system performance relatively insensitive to modeling errors. These observations hold true whether the state variables of functional interest are chemical, electrical, magnetic, or quantum mechanical.

A molecular planetary gear:

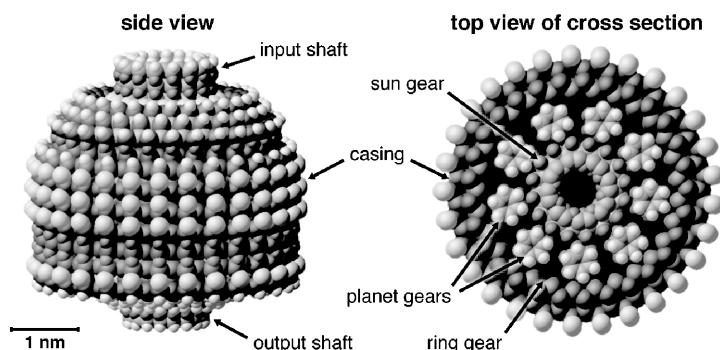


Figure 1. A molecular planetary gear. In mechanical systems, planetary gears convert shaft rotation from high speed and low torque to low speed and high torque. This highly symmetric design is based on stiff covalent structures that cannot be fabricated today but that can easily be modeled using standard molecular mechanics methods. Under operating conditions, the shaft angular speed would be on the order of 10^9 Hz, and the power density (because of scaling laws that favor small mechanical systems) would be in excess of 10^9 W/cm³.

3. THE DIVERSE GOALS OF COMPUTATIONAL MODELING

In nanotechnology, as elsewhere, computational modeling addresses a broad range of questions. These questions differ in surface content—addressing a particular physical system, seeking a particular numerical accuracy—but also differ in their deeper purposes. Distinguishing among these purposes is crucial to effective formulation of problems; misunderstood purposes can lead to misdirected questions, misallocation of research effort, and mistaken judgments regarding the adequacy of answers. Perhaps the deepest distinction among purposes for computational modeling is between advancing science and analyzing a design. Further, within the realm of design there is an important distinction between refining and validating a particular design intended for direct implementation and exploring a space of design choices to describe regions that contain particular, workable designs. Because different purposes place different demands on theoretical and computational studies, it is worth reviewing the implications of these familiar distinctions.

3.1. Science and Design Differ Fundamentally

Science and design are closely related yet in are, in some respects, opposites. Although both may deal with the same sorts of physical systems, even describing them using identical physical models, they nonetheless place their models in quite different relationships to physical systems (Fig. 2).

Scientists take a physical system as given (often after a lengthy search for a fruitful system to study), then seek to develop a model that describes the system. Designers, by contrast, take a desired functional behavior as given, then seek to specify a physical system that exhibits that behavior. Thus, science and design have opposite notions of dependent and independent variables—one develops a model to fit the system, and the other develops a system to fit the model. Roughly speaking, science and design address inverse problems.

This difference between science and design corresponds to different criteria for successful modeling. In science, the inherent goal of precise description establishes the ideal of exact equality between physical state variables and corresponding model state variables. In design, however, the goal of reliable prediction of function establishes the ideal of reliably predicting inequalities between physical state variables and design constraints—for example, predicting that an applied force will not exceed the strength of a component, or that a computation will complete before the result is needed. Design constraints typically define tolerance bands, and models with substantial but well-bounded errors often can reliably predict that a state variable of a physical system will fall within a given tolerance band. Moreover, designers routinely compensate for modeling errors by choosing designs with correspondingly broad tolerance bands. Thus, the adequacy of a model as a basis for design can be evaluated only in the context of a design effort that recognizes and seeks to accommodate the errors in the model.

The duality of technology and science:

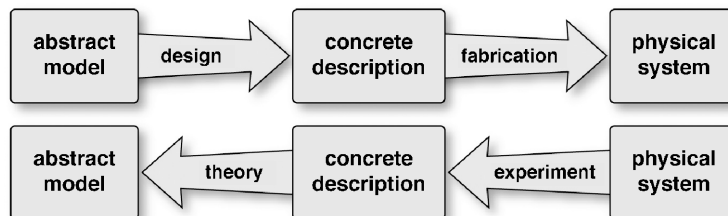


Figure 2. Models and physical systems in technology and science. In science, a physical system has preexisting properties, measurement yields a concrete description, and the concrete description constrains and thus shapes an abstract model (theory) that relates cause to effect. In technology, an abstract model (system concept) relates desired effects to required causes, design yields a concrete description, and the concrete description guides fabrication to produce a physical system with the desired properties. Thus, information in science and design flows in opposite directions. (Of course, practitioners of each generally do both: scientists design experiments, and designers measure their products.)

Science and design also have different criteria for the basis and generality of models. Science ideally aims to describe and explain the whole of physical reality as simply as possible. This grand objective drives a search for models that apply uniform fundamental principles to all questions. In computational modeling of molecular dynamics, for example, a natural goal is to start from quantum laws and accurately describe the behavior of any configuration of atoms, from helium gas, to enzymes, to metallic glasses. Limits on algorithms and computational capacity, of course, force compromises, resulting a patchwork of methods describing different structures at different scales of space and time, retreating from *ab initio* quantum methods to atomistic mechanical models to continuum and lumped-component descriptions.

The design perspective differs in treating generality as a convenience and fundamental physical principles as but one among many sets of useful tools. A designer starts with an objective and seeks to define a physical structure that will achieve it. Typically, the set of possible structures is highly constrained—consider the limited library of micro- and nanoscale structures used to implement silicon-based digital systems, in contrast to the broad range of micro- and nanoscale structures in the world as a whole. Moreover, large designed systems are typically partitioned into components, each described (from the external, system-level perspective) by a small set of geometrical and operational parameters (e.g., a motor must fit within this space, producing torque $> x$ while consuming current $< y$ and producing heat at a rate $< z$). A design process typically requires convenient descriptions of the relationship between structure and behavior for a narrow set of functional components. Whether these descriptions are based on broad fundamental principles or narrow empirical generalizations does not affect their utility.

Science is judged by the quality of its predictive models, but predictive models in design are mere tools: provided the systems in question can be built and then tested to identify and correct errors, even unreliable predictive models can be tolerated. Indeed, if predictive models were perfect, testing would be unnecessary. In reality, extensive testing is the norm in most areas of design practice, and the basic development sequence—design, model, fabricate, test—often forms a cycle that repeats before a product passes its acceptance test. The role of modeling is to enable more frequent recognition of design defects at the modeling stage, and hence less frequent rejections after expending time and cost for fabrication and testing.

These considerations indicate why scientific advances often have broad utility for design and technology: by improving the accuracy and generality of the description of physical systems, these considerations not only aid the design of familiar technological systems but help to open broad new areas to technological development. Equally, these considerations indicate why technology so often exploits phenomena that are reproducible, but poorly understood: limited and defective predictive models are better than none, and (as biological evolution demonstrates) systems can be developed with no predictive models at all by sheer trial and error. Organic synthesis provides a good example of a classic engineering capability—constructing specified objects—that has made great progress without strong, quantitative predictive models. Organic synthesis illustrates another pattern, as well: when predictive models are poor, experiment plays a guiding role and the pursuit of engineering goals merges with scientific investigation.

3.2. Protein Design Required the Recognition of a Design Problem

Protein design is emerging as a useful tool in nanotechnology. Its intellectual history illustrates the value of distinguishing between problems of science and problems of design.

To review, globular protein molecules contain flexible polypeptide chains that fold to form solid objects of nanometer scale. Their shapes and surface properties are determined by the sequence of amino acids that joined to form them; these amino acids are typically selected from the 20 encoded by genes and translated by the molecular machinery of ribosomes. The mechanical stiffness of folded proteins is comparable to that of engineering polymers such as nylon and epoxy (but only about 0.01 that of a stiff covalent solid). Naturally occurring proteins self-assemble to form molecular machines such as enzyme, flagellar motors, and (with RNA) the ribosome itself. If suitable design techniques can be developed, proteins clearly can serve as a basis for artificial self-assembling molecular machine systems. Further,

proteins shaped to be complementary to nonprotein molecular objects could serve as atomically precise “glue” for self-assembled nanosystems containing other active or structural components.

Protein scientists have long pursued solutions to the “protein-folding problem:” given the monomer sequence of a natural protein, predict its three-dimensional folded structure. Solutions are valuable because sequence information has been easy to obtain, whereas determining folded structures—and hence the physical basis of protein function—has been difficult.

Solving this scientific protein-folding problem was generally assumed to be a prerequisite for successful protein design. On the surface, this dependency seems plausible: solving the scientific problem would enable the prediction of protein folds, and protein design aims to produce proteins that fold predictably. Protein-folding research has, of course, greatly assisted protein design, yet the assumed dependency was mistaken. As the author argued in Ref. [1], the problem of designing a sequence that will fold to a given structure differs fundamentally from the problem of predicting the structure adopted by a given sequence. To distinguish these two “protein-folding problems,” Pabo [2] dubbed the design task identified in Ref. [1] as the “inverse folding problem;” Ponder and Richards then suggested computational approaches to this problem [3]. Protein design based on these and further developments has proved successful [4] and has now become both automated [5] and increasingly reliable and routine [6]. Results include proteins that mimic natural folding patterns (but with different internal and surface structures) [6], and design of a novel fold [7] (Fig. 3). Moreover, these results have repeatedly confirmed the initial prediction, based on a margin-of-safety argument, that designed proteins could be made more stable than natural proteins [1]. Progress has also been made on the original, scientific protein-folding problem (in part through insights stemming from the inverse-folding concept [8]), yet the computational methods used for the two problems remain distinct.

In protein engineering, distinguishing the design problem from a related science problem was a crucial step toward success. This distinction will surely prove fruitful in other areas of nanotechnology.

3.3. System Design Presents Distinct Challenges

A goal of nanotechnology is to enable the development of nanosystems that integrate diverse nanoscale components to form complex, functional products. In the world of macroscopic and microelectronic systems, development projects span wide ranges of time, cost, and complexity. The development of a new computer chip, car, or spacecraft can take years, cost millions to billions of dollars, and require designs comprising millions of functional components, each playing a distinct role. System design requires methods and levels of abstraction that differ from both those of scientific investigation and of component-level design. As nanofabrication capabilities advance, broader areas of nanotechnology will follow nanoscale lithography, advancing into the realm of system design. To understand the role of theoretical and computational methods in the development of nanosystems, it becomes important to recognize the distinctive tasks and questions that arise in the design of complex systems.

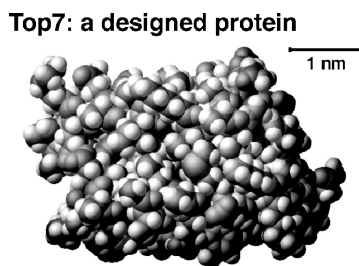


Figure 3. A designed protein with a novel backbone structure. Iterative computational design yielded TOP7, a 93-residue α/β protein. Produced by the ribosomal machinery of bacteria, TOP7 was found to be folded and extremely stable. X-ray crystallography revealed a structure with a root mean square deviation of 0.12 nm from the design model [6]. Note the striking qualitative differences between this protein and the comparably sized structure in Fig. 1.

Fundamental to system design is the concept of hierarchical structure [9]. A familiar hierarchy in descriptions of physical systems starts with the application of quantum theory to small structures and short timescales, moving to atomistic and then continuum and lumped-component models to describe larger structures and longer timescales. System design adopts a different hierarchical perspective—one that focuses on components, sets of components that form subsystems, which form components at a higher level, and so on up to level of the total system under design (embedded in some environment or broader system context). This hierarchy can be deep: a wire designer treats insulation as a component, a motor designer treats a wire as a component, an aircraft designer treats an electric motor as a component, and a transportation-system designer treats an aircraft as a component. As this example shows, at the top level of system design, the very concept of physics-based modeling can lose direct relevance.

A system-design process starts at some upper level of this hierarchy with an overall design concept—a low-resolution sketch of a functional system—then refines this design by adding detail to descriptions of subsystems and their relationships, working down to the level of preexisting components (motors, transistors, atoms). At any level, more detailed description and modeling may reveal difficulties or inconsistencies that force revision of the higher-level design. In a successful design effort, this iterative process of moving between higher and lower levels converges on a detailed, implementable, and attractive design. In early phases of an effort, a designer commonly works with a parameterized, high-level description—one in which a multidimensional range of systems is described by a mathematical model in which variables relate (for example) component volume, power consumption, mass, and so forth. Values for these parameters may be based not on detailed designs of components but on estimates of the properties of well-designed components of the required kind. Because design works chiefly with inequalities and tolerance bands, designers often can compensate for uncertainty in these estimates by including a margin of conservatism to absorb errors. Where large margins of conservatisms are compatible with adequate system performance, early-phase high-level design can proceed with confidence using quite rough approximations in describing components. This enables preliminary system-level design to proceed in advance of detailed component-level design and modeling. This can enable the formulation of long-term goals.

In ambitious, long-term system development projects, the requirement for novel components may generate substantial research projects. The many technological spin-offs of the 1960s moon landing program resulted from this stimulus to research. Long-term goals in nanotechnology can likewise motivate a broad research program, including extensive development of theoretical and computational methods to support nanosystem design.

System-level modeling will be crucial to the design of integrated nanosystems. Analysis at this level, however, will rely on descriptions of component behavior that emerge from—but do not exhibit—the physics of interactions at the nanoscale. This partial independence of higher levels from lower levels is a familiar pattern. For example, from an abstract logic-system design perspective, reliable vacuum tubes, CMOS transistors, and molecular electronic switches are essentially equivalent. From a more concrete computer-system design perspective, their differences in size, speed, voltage, current, power dissipation, and reliability are important, but their differences in molecular structure and electron quantum states are not.

Accordingly, the chief role of theoretical and computational nanotechnology in the development of integrated nanosystems will be to support the design and fabrication of nanoscale components and subsystems. Here, the leading role of modeling is clear. In system development, a key to launching a project is confidence that a system of the intended sort can function—even if many of its required processes and components are not yet demonstrated. Because modeling can precede fabrication capabilities (and can model the fabrication process itself), modeling-based research and design can build the confidence in success that will motivate the investment necessary to achieve ambitious goals.

The nature of system development highlights the second distinction mentioned above, between refining and validating a particular design intended for direct implementation, and exploring a space of design choices to describe regions that contain particular, workable

designs. The early phase of system development generates designs that cannot directly be built—they are preliminary and exploratory in nature, intended to guide not fabrication, but further research leading to more detailed designs. In this phase, the purpose is not to specify every detail of the system (to enable fabrication) but to define required subsystems and components, to estimate their feasible capabilities, and to determine whether the project is likely to encounter no insurmountable obstacles. In this early phase, the flexibility of parameterized system designs can reduce the pressure on modeling capabilities. If exploratory modeling gives results with large error bars, the designer can readily generate an estimate of system performance based on the corresponding worst-case assumptions. If this worst-case performance is high enough to represent a desirable goal (as will often be the case, given the fundamental advantages of nanosystems in many spheres), then confidence in feasibility can precede precise specification and complete design.

4. The Strategic Role of Fabrication Technologies

Human capabilities in nanotechnology (as in all technologies) are limited by what we can design and by what we can fabricate. As reviewed above, theoretical and computational methods can make direct contributions to the design of nanosystems, helping to select and validate targets for fabrication. But theoretical and computational methods can also help in the design of fabrication processes and systems and hence can make strong indirect contributions to the fabrication of nanosystems.

Among physical technologies, fabrication systems play a special, strategic role. Other systems can compute, display, transport, and so on, but they cannot make either other systems or systems like themselves. Thus, improvements in these areas cannot directly physically advance other technologies. Fabrication is different because fabrication systems are responsible for making all others. Improvements in fabrication can directly, physically advance other technologies and can lay the foundations for further advances in fabrication. From this perspective, fabrication forms the trunk of a tree on which all other technologies are branches.

4.1. Nanofabrication Directs the Motion of Atoms

Nanofabrication systems include ribosomes, scanning probe atom-manipulation systems, organic synthesis apparatus, and billion-dollar chip-production facilities. The physical processes exploited by nanofabrication systems include bond formation in solution, bond formation guided by molecular machines, molecular self-assembly guided by intermolecular forces, and lithographically directed etching and thin-film growth. From a modeling perspective, the fundamental feature common to all these processes is the motion of atoms—an inescapable requirement for processes that restructure matter. This leads to a focus on molecular dynamics, using quantum methods (or special force fields, such as the Brenner hydrocarbon potential [10]) to describe motions that change bonding and a broader class of methods to describe other motions.

Most fabrication processes in wide use today operate on disordered systems. Organic synthesis and molecular self-assembly rely on selective interactions between molecules that collide randomly in solution, yet they can produce fully ordered molecular and supramolecular products. Lithographically directed processes rely on statistical processes of etching, deposition, oxidation, and so forth, typically yielding disordered results. Modeling of these disordered systems requires the exploration of a large configuration space. In many instances, the products of these disorderly physical processes depend on the relative kinetics of competing pathways, or the relative thermodynamic stability of competing end states. In these instances, correctly predicting which pathway or product state will predominate may require correctly calculating differences in transition states or product states on the order of kT . The challenges presented by large configuration spaces and sensitivity to small energy differences currently limits the use of computational methods in the designing and modeling of (for example) novel organic syntheses.

4.2. Self-Assembly Enables Precise Fabrication of Nanosystems

Many nanofabrication processes yield crystalline structures; it would be useful to be able to design molecular building blocks that assemble to form selected structures. In the general case, crystal structure prediction is difficult. Of the many lattice packing arrangements (crystal polymorphs) that a molecule might conceivably form, the one that appears in the actual physical system, may depend on very small differences in the free energy of transition states or end states in the crystal growth process. It has recently been stated that “Because we still lack a general and accurate method for CSP [crystal structure prediction], there is no known example of a designed crystalline molecular material in wide technological use” [11]. This statement, however, exactly parallels the early views regarding protein design discussed above: it conflates a scientific problem of general prediction with a design problem of specifying structures that behave predictably. As with protein engineering, it should be fruitful to distinguish between these fundamentally different problems. A focus on design directs attention to computational methods that take the basic form of a target structure as given (here, a lattice) while seeking molecular structures that will favor its formation.

Crystal structure design is a special case of atomically precise molecular self-assembly (as distinct from the self-assembly of micelles or other disordered structures). Striking examples of precise self-assembly (including small two-dimensional crystals) have been produced through the design and fabrication of branched, double-stranded DNA structures (DNA junctions). These structures have been used to implement simple molecular machines [12] and have been proposed as a means for organizing active components in (for example) molecular electronic systems [13]. The design methodology for these structures uses computational methods to identify DNA sequences in which bases will pair to link particular strand segments while minimizing unwanted pairing between all other segments.

In both DNA and protein design, a key requirement for stable and specific binding is that matching surfaces be large enough to display multiple molecular features with distinct properties (of shape, charge, polarity, hydrogen bonding, etc.). These features enable the design of complementary surfaces that exhibit strong cooperative binding while disrupting binding among other, noncomplementary surfaces. Although scientific studies can benefit from a focus on small, simple structures (which better reveal differences in elementary binding interactions), design principles favor larger structures (which better conceal errors in estimating elementary binding interactions). Larger structures with larger interfaces enable a designer to control more features, offering more opportunities for strengthening and disrupting selected binding interactions. Larger interfaces also increase the tolerance for modeling errors: when adding multiple interactions, each expected to be stabilizing, cumulative errors in the total binding energy grow as the square root of area, whereas the expected binding energy increases linearly. This reduces sensitivity to modeling errors and enables more reliable design of strong binding.

All these considerations favor larger structures, motivating the use of computational methods to search a space of design options that tends to grow exponentially with structure size. Scientific intuition may indicate that a larger search space creates a harder problem, but this perception stems from the scientific goal of gaining information, of reducing a set of possibilities toward the single, uniquely correct description of a given system. From a design perspective, however, enlarging a search space—adding design options—typically creates an easier problem. A larger space of possibilities will generally contain larger solution-regions (not just a single, correct point hidden in a larger space). Adding options offers more points of control, more ways to improve a design, and more paths around obstacles. (This is explored in a protein-engineering context in Ref. [14]) Here again, the inverse relationship between problems in science and design gives a superficially similar issue—search-space size—a very different significance.

Unfortunately, fabricating structures large enough to self-assemble into well-ordered structures remains a challenge. Instances constructed via general organic synthesis have been sporadic and specialized—typically, they have been unique structures rather than modular structure-building systems. The flexible and general methods in use today use large biopolymer molecules, made either by organic synthesis (e.g., DNA) or by programmable molecular machine systems (nucleic acids, made by polymerases, or proteins, made by ribosomes).

4.3. Productive Nanosystems Offer a Strategic Goal

Fabrication based on programmable biological molecular machine systems demonstrates striking performance: ribosomes can build proteins at a rate of about 10 monomers per second with an error rate on the order of 10^{-4} , and polymerases can build DNA at a rate of about 10^3 monomers per second with an error rate (via kinetic proofreading) as low as 10^{-9} to 10^{-11} [15]. These productive molecular machine systems are themselves made of proteins and nucleic acids and have been induced to self-assemble *in vitro*. In light of growing capabilities for the design of self-assembling protein and nucleic acid structures, it is natural to consider the design of artificial molecular machine systems, including productive systems. The ubiquitous role of molecular machines in biological sensing, information processing, movement, and molecular transformation indicates that artificial molecular machines could have broad applications in nanotechnology.

A class of molecular machine systems having special strategic importance to the development of nanotechnology would be productive systems that, like ribosomes, can fabricate components suitable for building next-generation productive systems. Productive nanosystems of this class could open a new range of molecular structures to reliable, programmable fabrication. In the past, the application of advances in fabrication to the fabrication of more advanced fabrication systems has sustained a spiral of improvement that led from hammered iron to modern nanostructures. Continuing this spiral in the domain of productive nanosystems holds great promise.

In considering how machines of this sort might be designed, we must distinguish the scientific problem of learning how biological machines work from the design problem of designing machines that perform similar tasks. Nonbiological approaches may prove simpler. For example, artificial machines could exploit the sequential addition of reactants to construct monomer sequences (as in a Merrifield synthesis), rather than translating a sequence encoded by an information molecule. Alternatively, a sequence of monomers could be bound to a DNA strand in a specific sequence via complementary DNA linkers long enough to bind with high reliability and specificity; a system that builds structures using this monomer-bearing strand would resemble a ribosome, but simplified and made more reliable by the use of larger, prebound analogues to tRNA molecules. The advantage of either class of machine with respect to ribosomes and DNA polymerases would be an ability to fabricate novel families of molecules; their advantage with respect to Merrifield synthesis would be greater speed and reliability, enabling the construction of molecules large enough to support reliable self-assembly (of, e.g., next-generation machine components). “Artificial ribosomes” are cited in a chemical industry roadmap as a goal for the 3–10-year time frame [16].

4.4. Productive Nanosystems Invite Multistage Exploratory Design

Theoretical and computational methods can help to guide development because they can describe what has not yet been made. Perhaps most significant, in light of the strategic role of fabrication technologies, modeling can explore and guide the development of potential fabrication processes and their products. Here, a natural objective is to explore potential paths for a multistage system development project in which fabrication systems would be used to produce improved fabrication systems through several generations of technology. As we have seen, in early-stage system design, a modest amount of knowledge can substantially reduce uncertainties regarding even quite large projects.

Preliminary work along these lines [17–19] indicates that a natural, long-term objective for fabrication-system development is to cross the implementation gap diagrammed and discussed in Fig. 4. This gap separates current fabrication technologies able to make polymeric products (like the structure illustrated in Fig. 3) from a more capable technology able to make products based on stiff covalent structures (like the gear illustrated in Fig. 1). Theoretical and computational investigations indicate that these products will have large performance advantages relative to current molecular and macroscale machines in areas such as strength, power density, density and energy cost of computation, and breadth of fabrication capabilities.

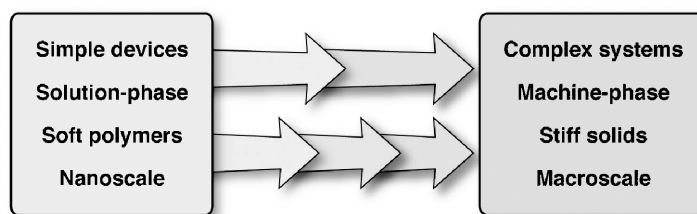
A multi-stage development process:

Figure 4. The implementation gap between present and advanced fabrication systems. Productive molecular machine systems today (ribosomes, polymerases) work in solution to make small polymeric structures with relatively simple functions. Physical considerations and natural directions for technological advance suggest the objective of developing productive systems that can make large, complex products from a wider range of materials. Performing fabrication in machine-phase systems (in which molecular motion results from mechanical conveyance and positioning, rather than from diffusion in a solvent) offers fundamental advantages in the control of molecular encounters, and hence of synthetic reactions.

Preliminary, exploratory design indicates that the implementation gap diagrammed in Fig. 4 can be crossed via a multiphase, multistage development process (diagrammed and discussed in Fig. 5). Current tools and techniques are not directly able to fabricate systems on the far side of the gap, but apparently they can build next-generation tools that are stepping-stones on paths to that ability. Several steps of this sort, followed by scale-up of the resulting high-performance productive nanosystems, apparently can close the implementation gap.

In design work, one first considers goals and then the means for implementing them, working backward from objectives to concrete systems. Here, however, some of the systems contemplated will require new systems to enable their fabrication. The purpose of the early-phase exploration is not to design systems that could be built only years from now but, rather, to build confidence in the feasibility of the objectives and to better understand how current research efforts can be coordinated to move more directly toward long-term, high-payoff goals.

Work to date indicates that achieving these goals would be of great value to many areas of science and technology. Both biological examples and analyses using standard theoretical and computational methods indicate that productive molecular machine systems can enable economical, large-scale fabrication of products (both macro- and nanoscale) built to complex, atomic specifications. The implications of this for technology as a whole are profound.

Perhaps surprisingly, longer-term objectives can be easier to model and analyze than are shorter-term objectives. One reason is that components constructed from stiff covalent solids are (as discussed in Section 2.1) easier to model than devices constructed from experimentally accessible folded polymer structures. Advanced machine systems can be made of these components, and hence inherit their ease of modeling; designs can then avoid problematic conformational degrees of freedom and sensitivity to small energy differentials. Further, in advanced systems, the motions of reactive molecules—and hence their opportunities to react—can be strictly controlled. Mechanical transport and positioning can prevent unwanted molecular encounters and can sequence and (with careful design of stiff structures that limit thermal displacements) can precisely direct desired reactive encounters. This degree of control permits the use of highly reactive species in synthesis (where solution-phase with these species would be unspecific) and can make reaction results insensitive to small energy differentials among transition and product states. Overall, these characteristics of advanced designs can reduce sensitivity to modeling errors. A substantial set of components and reactions has been analyzed, but further work is needed to produce a more detailed picture. Modeling techniques that facilitate the design of reaction trajectories guided by mechanical constraints would be of particular use in this work. The results to date, however, support a system-level analysis that indicates the feasibility of clean, highly productive fabrication systems able to make a wide range of high-performance macro- and nanoscale products [17].

Modeling of productive molecular machine systems can exploit molecular mechanics methods to describe mechanical motions (e.g., in molecular transport and positioning), together

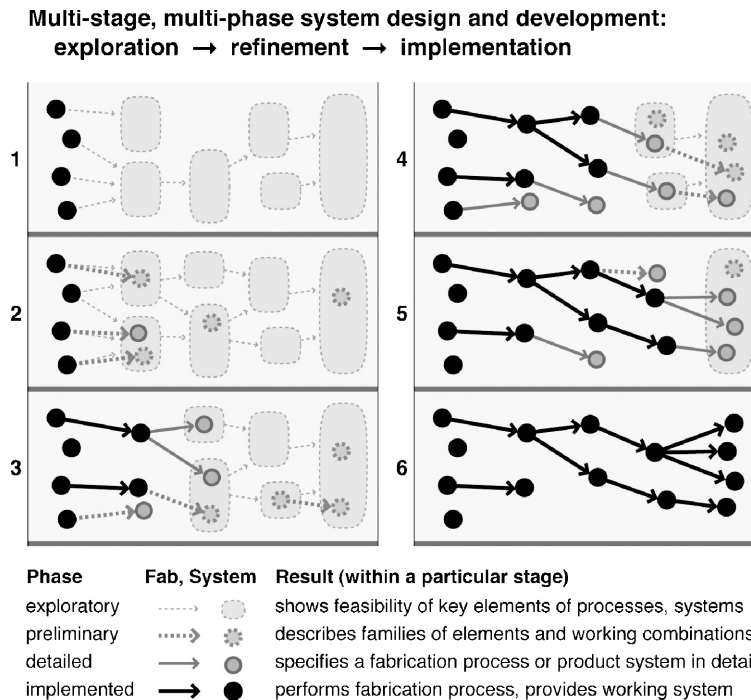


Figure 5. Phases versus time in multistage development of fabrication systems. Closing the implementation gap (Fig. 4) will require development to pass through several intermediate stages, each corresponding to a set of fabrication technologies. This diagram illustrates the natural structure of a multistage development process for fabrication systems. Each panel represents the overall state of development at a particular time. Each region in a panel (reading left to right) represents a specific set of fabrication systems (at some phase in its development), moving from those currently available, to next-stage systems, to those further from implementation. Each arrow represents a fabrication process (at some phase in its development) that uses the set of systems at its left to implement the set at its right. The style of a region or arrow represents the phase of development of the corresponding systems or process (see key). Note that some designs are discarded before being implemented, and that some are developed well before their physical implementation becomes possible. In the first panel, a continuous chain of valid exploratory designs indicates the feasibility of designing and implementing one or more corresponding physical systems and processes for each step. The final panel represents a chain of implemented systems.

with quantum chemistry methods to describe reactive transformations in the synthesis of products. This holds true both for early, polymeric systems and for later systems based on stiff covalent solids, provided that these systems are small. Larger systems call for multiscale modeling techniques in which atomistic descriptions support continuum and lumped-component models that describe extended structures and the interaction of subsystems. Bridging these levels of description in a consistent modeling environment suited to system design will be a key challenge.

5. CONCLUSION

As nanotechnology advances, ever more areas of practice will move from relatively experimental, science-intensive work on simple systems to relatively predictable, design-intensive work on complex systems (a transition already seen in nanoscale lithography for integrated circuits). In this advance, theoretical and computational methods will speed scientific progress by narrowing the range of unknowns that require experimental testing, and will speed system development by enabling more effective exploration and design. By describing systems that have not or cannot yet be made, computational modeling can help to guide progress in nanotechnology by identifying attractive goals well in advance of their achievement.

To fulfill their potential, applications of theoretical and computational methods must take account of the fundamental difference between scientific and design problems. Although physical principles are identical in both spheres, the relationship between physical systems

and models is essentially reversed. Both the questions and the required methods differ when the problem is to design a system to meet the requirements of an abstract design model, rather than to fit a theoretical, scientific model to a given system. Mistaking a design problem for a scientific problem can obstruct progress (as it did when the protein-fold design was conflated with a protein-fold prediction). To design a complex, hierarchical system, rather than a simple material or device, requires a further shift in perspective. In system design, early development work produces not concrete, directly implementable designs but abstract, parameterized design models that depend on estimates of the performance of components that have not yet been designed in detail.

Because technological advances are paced both by advances in what can be designed and in what can be made, a particularly strategic role for computational and theoretical methods is in the design of systems that can extend fabrication capabilities. Here, the productive capabilities of biological molecular machine systems indicate a rich set of opportunities. A natural goal is to use biopolymers as building blocks in the construction of improved polymer-building systems. Turning to longer-term developments, preliminary studies have produced both detailed designs for stiff, covalent components and exploratory designs of systems that would use and produce such components. These studies indicate that productive nanosystems can eventually achieve the cleanliness, low cost, and productivity of biological molecular machinery, while employing different means and producing a wider range of products. Because of the greater stiffness and reduced configuration space available to stiff, covalent components, advanced systems are easier to model than those that can be built today. This invites the use of standard modeling techniques to explore component and system designs well in advance of their implementation.

Computational and theoretical methods can both facilitate laboratory work today and aid in the identification of paths to more advanced capabilities. They thus are positioned to play a central role in advancing and guiding developments in many areas within the broad field of nanotechnology.

REFERENCES

1. K. Drexler, *Proc. Natl. Acad. Sci. USA* 78, 5275 (1981).
2. C. Pabo, *Nature* 301, 200 (1983).
3. J. Ponder and F. Richards, *J. Mol. Biol.* 193, 775 (1987).
4. H. Hellinga, *Nat. Struct. Mol. Biol.* 5, 525 (1998).
5. B. Dahiyat and S. Mayo, *Science* 278, 82 (1997).
6. G. Dantas, B. Kuhlman, D. Callender, M. Wong, and D. Baker, *J. Mol. Biol.* 332, 449 (2003).
7. B. Kuhlman, G. Dantas, G. Ireton, G. Varani, B. Stoddard, and D. Baker, *Science* 302, 1364 (2003).
8. A. Godzik, A. Kolinski, and J. Skolnick, *J. Comput. Aided Mol. Des.* 4, 397.
9. H. Simon, "The Sciences of the Artificial." MIT Press, Cambridge, MA, 1969.
10. D. Brenner, O. Shenderova, J. Harrison, S. Stuart, B. Ni, and S. Sinnott, *J. Phys. Condens. Matter* 14, 783 (2002).
11. R. Desiraju, *Nat. Mater.* 1, 77 (2002).
12. Seeman, *Biochemistry* 42, 7259 (2003).
13. B. Robinson and N. Seeman, *Prot. Engin.* 1, 295 (1987).
14. K. Drexler, *Annu. Rev. Biophys. Biomol. Struct.* 23, 377 (1994).
15. J. Watson, N. Hopkins, J. Roberts, J. Steitz, and A. Weiner, "Molecular Biology of the Gene," 4th Edn. Benjamin/Cummings Publishing Company, 1987.
16. American Chemical Society, "Technology Roadmap for Materials." American Chemical Society, Washington, DC, 2000.
17. K. Drexler, "Nanosystems: Molecular Machinery, Manufacturing, and Computation." Wiley/Interscience, 1992.
18. R. Merkle, *Nanotechnology* 8, 149 (1997).
19. R. Merkle, *Trends Biotechnol.* 17, 271 (1999).