Proc. Natl. Acad. Sci. USA
Vol. 78, No. 9, pp. 5275–5278, September 1981
Chemistry

Molecular engineering: An approach to the development of general capabilities for molecular manipulation

(molecular machinery/protein design/synthetic chemistry/computation/tissue characterization)

K. ERIC DREXLER

Space Systems Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Communicated by Arthur Kantrowitz, June 4, 1981

ABSTRACT **Development of the ability to design protein molecules will open a path to the fabrication of devices to complex atomic specifications, thus sidestepping obstacles facing conventional microtechnology. This path will involve construction of molecular machinery able to position reactive groups to atomic precision. It could lead to great advances in computational devices and in the ability to manipulate biological materials. The existence of this path has implications for the present.**

Feynman's 1959 talk entitled "There's Plenty of Room at the Bottom" (1) discussed microtechnology as a frontier to be pushed back, like the frontiers of high pressure, low temperature, or high vacuum. He suggested that ordinary machines could build smaller machines that could build still smaller machines, working step by step down toward the molecular level; he also suggested using particle beams to define two-dimensional patterns. Present microtechnology (exemplified by integrated circuits) has realized some of the potential outlined by Feynman by following the same basic approach: working down from the macroscopic level to the microscopic.

Present microtechnology (2) handles statistical populations of atoms. As the devices shrink, the atomic graininess of matter increasingly creates irregularities and imperfections, so long as atoms are handled in bulk, rather than individually. Indeed, such miniaturization of bulk processes seems unable to reach the ultimate level of microtechnology—the structuring of matter to complex atomic specifications. In this paper, I will outline a path to this goal, a general molecular engineering technology. The existence of this path will be shown to have implications for the present.

Although the capabilities described may not prove necessary to the achievement of any particular objective, they will prove sufficient for the achievement of an extraordinary range of objectives in which the structuring and analysis of matter are concerned. The claim that devices can be built to complex atomic specifications should not, however, be construed to deny the inevitability of a finite error rate arising from thermodynamic effects (and radiation damage). Such errors can be minimized through the use of free energy in error-correcting procedures (including rejection of faulty components before device assembly); the effects of errors can be minimized through fault-tolerant design, as in macroscopic engineering.

The emphasis on devices that have general capabilities should be taken in the spirit of early work on the theoretical capabilities of computers, which did not attempt to predict such practical embodiments as specialized or distributed computation systems. The present argument, however, will proceed from step to step by close analogies between the proposed steps and past developments in nature and technology, rather than by mathematical proof. We commonly accept the feasibility of new devices without formal proof, where analogies to existing systems are close enough: consider the feasibility of making a clock from zirconium. The detailed design of many specific devices to render them describable by dynamical equations would be a task of another order (consider designing a clock from scratch) and appears unnecessary to the establishment of the feasibility of certain general capabilities.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Protein design

Biochemical systems exhibit a "microtechnology" quite different from ours: they are not built down from the macroscopic level but up from the atomic. Biochemical microtechnology provides a beachhead at the molecular level from which to develop new molecular systems by providing a variety of "tools" and "devices" to use and to copy. Building with these tools, themselves made to atomic specifications, we can begin on the far side of the barrier facing conventional microtechnology.

What can be built with these tools? Gene synthesis (3) and recombinant DNA technology can direct the ribosomal machinery of bacteria to produce novel proteins, which can serve as components of larger molecular structures. One might think assembly of such components into complex systems would require a preexisting technology able to handle molecules and assemble them; fortunately, biochemistry demonstrates that intermolecular attraction between complementary surfaces can assemble complex structures from solution. For example, the complex machinery of the ribosome self-assembles from more than 50 different protein molecules and can do so *in vitro* (4).

At present, the design of protein systems as complex as a ribosome seems an awesome task. Indeed, chemists cannot yet predict the three-dimensional conformation of a natural protein from its amino acid sequence, an ability that might seem requisite to the design of new proteins. Two considerations suggest that this obstacle is surmountable: first, the continuing improvement in protein science and, second, the difference between natural science and design engineering.

Regarding the first, computer simulation of protein molecules in solution (5) shows promise. As computer technology and chemical knowledge improve, simulations will increase in accuracy, speed, and size. Improvement promises new insight into protein behavior and may permit the designer to modify (simulated) molecules quickly and to observe their behavior directly.

Regarding the second consideration, natural scientists seek a more general understanding than design engineers require. Science seeks the ability to predict the conformations of all natural polypeptides. In attempting this, protein chemists can search for a minimum-energy chain conformation (in hope that the protein assumes not a local but a global minimum-energy conformation) (6) or can attempt to follow the chain-folding mechanism to find the final conformation (7). Prediction will be easier if the natural conformation has outstanding stability

or if its folding mechanism proceeds in a sequence of strongly preferred steps. Unfortunately, natural selection accepts polypeptides that have natural conformations of low stability (in energetic terms) so long as they exhibit long lifetimes on the cellular time scale (or renature readily). Similarly, natural selection accepts any folding process so long as the chain reaches its natural conformation with essentially 100% yield. Moreover, random mutations are unlikely to enhance the stability of a particular conformation (or the predictability of its folding mechanism). Thus, natural proteins tend to accumulate disruptive changes until they reach the threshold of poor stability or reduced yield of the natural conformation; only then does natural selection come into play. Thus, it is little wonder that chemists cannot yet predict the conformations of natural proteins; they are not designed to fold predictably.

Engineers (in contrast to scientists) need not seek to understand all proteins but only enough to produce useful systems in a reasonable number of attempts. An engineer designing a protein that has 1000 amino acids may choose among some 10^{1300} different amino acid sequences. It might be that only one in 10^9 (or even 10^{700}) randomly selected sequences would yield a predictable conformation, yet this tiny fraction represents a vast number of proteins. Through use of strategically placed charged groups, polar groups, disulfide bonds, hydrogen bonds, and hydrophobic groups, the engineer should be able to design proteins that not only fold predictably to a stable structure (sometimes) but that serve a planned function as well. Even a low success rate will lead to an accumulation of successful designs. Thus, the difficulties encountered in predicting the conformations of natural proteins do not seem insurmountable obstacles to protein engineering.

Computer modeling and chemical understanding of biological targets have already found use in pharmaceutical design (8), and an artificial 34-residue polypeptide designed to interact with RNA has been synthesized and found active (9). It has been proposed to give microcircuitry special sensitivities by adsorbing engineered proteins onto selected surfaces (10). The promise of enzyme design in chemical engineering is evident. As protein science has great promise and difficulties in understanding natural proteins need not block engineering, the substantial payoffs for improved capabilities should lead to development of protein design technology. It would be foolish to underestimate the time and effort that will be required to develop basic design capabilities and then a broad family of working molecular devices; still, the path seems clear to achieving the capabilities exhibited by existing biochemical systems, by copying their features if need be.

Molecular machinery

A comparison of biochemical to macroscopic components will show the possibilities of the former by analogy to the latter (Table 1). With structural members, moving parts, bearings, and motive power, versatile mechanical systems can be built. Molecular assemblages of atoms can act as solid objects, occupying space and holding a definite shape. Thus, they can act as structural members and moving parts. Sigma bonds that have low steric hindrance can serve as rotary bearings able to support $\approx 10^{-9}$ N. A line of sigma bonds can serve as a hinge. Conformation-changing proteins (such as myosin) can serve as sources of motive power for linear motion; the reversible motor of the bacterial flagellum can serve as a source of motive power for rotary motion. The existence of this range of components in nature indicates that power-driven mechanical systems can be constructed on a molecular scale.

By analogy with macroscopic devices, feasible molecular machines presumably include manipulators able to wield a variety of tools. Thermal vibrations in typical structures are a modest fraction of interatomic distances; thus, such tools can be positioned with atomic precision. As present microtechnology (2) can lay down conductors on a molecular scale (10 nm) and molecular devices can respond to electric potentials (through conformation changes, etc.), such devices can be controlled by human operators or macroscopic machines. Further, by analogy with biological sensors, molecular scale instruments can evidently produce macroscopic signals, indicating the feasibility of feedback control in molecular manipulations.

Together, these arguments indicate the feasibility of devices able to move molecular objects, position them with atomic precision, apply forces to them to effect a change, and inspect them to verify that the change has indeed been accomplished. It would be foolish to minimize the time and effort that will be required to develop the needed components and assemble them into such complex and versatile systems. Still, given the components, the path seems clear.

Ordinary chemical synthesis relies on thermal agitation to bring reactant molecules in solution together in the correct orientation and with sufficient energy to cause the desired reaction. Enzyme-like molecular machines can hold reactants in the best relative positions as bonds are strained or polarized. Like some enzymes, they can do work on reactant molecules to drive reactions not otherwise thermodynamically favored.

These are clearly techniques of great power, yet the synthetic capabilities of systems based on polypeptide chains might seem limited by amino acid properties. However, enzymes show that

Table 1. Comparison of macroscopic and microscopic components

Technology	Function	Molecular example(s)
Struts, beams, casings	Transmit force, hold positions	Microtubules, cellulose, mineral structures
Cables	Transmit tension	Collagen
Fasteners, glue	Connect parts	Intermolecular forces
Solenoids, actuators	Move things	Conformation-changing proteins, actin/myosin
Motors	Turn shafts	Flagellar motor
Drive shafts	Transmit torque	Bacterial flagella
Bearings	Support moving parts	σ bonds
Containers	Hold fluids	Vesicles
Pipes	Carry fluids	Various tubular structures
Pumps	Move fluids	Flagella, membrane proteins
Conveyor belts	Move components	RNA moved by fixed ribosome (partial analog)
Clamps	Hold workpieces	Enzymatic binding sites
Tools	Modify workpieces	Metallic complexes, functional groups
Production lines	Construct devices	Enzyme systems, ribosomes
Numerical control systems	Store and read programs	Genetic system

other molecular structures bound to the polypeptide (such as metal ions and complex ring structures) (11) can extend protein capabilities. The range of such tools is large and greater than found in nature. Thus, the synthetic capabilities of enzymes set only a lower bound on the capabilities of engineered protein systems. Indeed, as tool-wielding protein systems can control the chemical environment of a reaction site completely, they should be able, at a minimum, to duplicate the full range of moderate-temperature synthetic steps achieved by organic chemists. Further, where chemists must resort to complex strategies to make or break specific bonds in large molecules, molecular machines can select individual bonds on the basis of position alone. Conventional organic chemistry can synthesize not only one-, two-, and three-dimensional covalent structures but also exotic strained and fused rings. With the addition of controlled site-specific synthetic reactions, a broad range of large complex structures can doubtless be built.

Still, the synthetic abilities of protein machines will be limited by their need for a moderate temperature aqueous environment (although applied forces can sometimes replace or exceed thermal agitation as a source of activation energy and reaction sites and reactive groups can be protected from the surrounding water, as in some enzymatic active sites). These limits may be sidestepped by using the broad synthetic capabilities outlined above to build a second generation of molecular machinery whose components would be not coiled hydrated polypeptide chains but compact structures having three-dimensional covalent bonding. There is no reason why such machines cannot be designed to operate at reduced pressure or extreme temperatures; synthesis can then involve highly reactive or even free radical intermediates, as well as the use of mechanical arms wielding molecular tools to strain and polarize existing bonds while new molecular groups are positioned and forced into place. This may be done at high or low temperature as desired. The class of structures that can be synthesized by such methods is clearly very large, and one may speculate that it includes most structures that might be of technological interest.

Firmness of the argument

The development path described above should lead to advanced molecular machinery capable of general synthesis operations. As the results of this path can be shown to have consequences for the present, it is of interest to discuss the degree of confidence that should be placed in its feasibility.

It might be argued that complex protein or nonprotein machines are impossible or useless, on the grounds that, if they were possible and useful, organisms would be using them. A similar argument would, however, conclude that bone is a better structural material than graphite composite, that neurons can transmit signals faster than wires, and that technology based on the wheel is impossible or useless. Nature has been constrained less by what is physically possible than by what could be evolved in small steps. Thus, the absence of a proposed kind of molecular machinery in organisms in no way suggests its infeasibility.

To deny the feasibility of advanced molecular machinery, one must apparently maintain either (*i*) that design of proteins will remain infeasible indefinitely, or (*ii*) that complex machines cannot be made of proteins, or (*iii*) that protein machines cannot build second-generation machines.

In light of the expected improvements in computation, the simplified task of design engineers (compared with scientists), the possibilities offered by sheer trial-and-error modification of natural proteins, and the progress already made in protein design, the first seems difficult to maintain. Further, even if protein design were to prove intractible (because of difficulties in predicting conformations), this would in no way preclude developing an alternative polymer system with predictable coiling and using it as a basis for further development.

In light of the presence of the needed components for mechanical devices in the cell, the second seems difficult to maintain. Indeed, the cytoskeleton provides a fair counterexample.

In light of the results of synthetic organic chemistry and the ability of molecular machines to make reactions site specific, it seems difficult to maintain than nonprotein machine components cannot be built and assembled.

Each of the development steps outlined above seems closely analogous to past steps taken by nature or by technology. Each of these steps can be accomplished in many ways. To argue their infeasibility would seem to require some general principle precluding success, and it is difficult to see what such a principle might be like. Thus, the claim that advanced molecular technology can be developed seems well founded.

Although the existence of molecular machinery in cells indicates the feasibility of some sort of artificial molecular machinery, errors in assembly might limit the synthesis of structures of great complexity. In the cell, molecular machinery uses DNA to direct the assembly of DNA and other molecules. In some eukaryotic cells, DNA directs DNA synthesis with an error rate of $\approx 10^{-11}$ per nucleotide added (12). As engineers commonly design systems to function reliably with many more failed components than 1 in 10^{11}, such an error rate seems no barrier to the construction of quite complex devices.

The possibility of low error rates is not surprising. For synthesis systems permitting error detection and correction (such as DNA synthesis), the net error rate in assembly can be reduced to roughly the product of the raw error rate in assembly and the rate at which errors are falsely identified as correct. As no uncertainty principle prohibits accurate discrimination between objects of different kinds (such as correctly and incorrectly assembled molecular structures), no limits to the detection and correction of errors are apparent.

Applications to computation

Molecular technology has obvious application to the storage and processing of information. A crude approach would involve literal "molecular machinery" patterned on the Babbage machine. In a more subtle approach, bits could be represented by protons, bound electrons, reactive groups, or conformation changes and transferred by movement of protons or of well-localized electrons (13), excitons, or phonons. The range of plausible device speeds is suggested by the 10^{-6}-sec turnover time for a fast enzyme, by the 10^{-13}-sec scale of collisional interactions (11), and by the 10^{-16} sec taken for an electron to cross an interatomic distance at a typical Fermi velocity.

It seems highly likely that a cubic cell 0.1 μm on a side (containing some 10^8 optimally arranged atoms) can hold a bit or perform a logic operation and, at the same time, transmit bits through itself to provide communication from cell to cell in a lattice. If so, then computers can be built with at least 10^{15} active elements per cubic centimeter. In a well-designed computer (with elements closer to their true technological limit and not laid out in regular cubical cells), this volume estimate should prove quite conservative. Elements so small will be sensitive to radiation damage; to be reliable, systems will require a large measure of redundancy.

Concern might be raised about the cost of such intricately patterned matter, either because of labor or energy requirements. It seems clear, however, that molecular-scale production systems can be completely automated (what use is there

for hands?). Thus, labor costs of production (including production of additional production equipment) can approach zero. The energy needed to produce molecularly engineered material will generally be greater than the energy needed to produce ordinary materials of similar bulk composition, but analogy suggests that the energy cost need not be vastly greater than for the production of biological materials. In many cases (e.g., advanced computers or any of a number of applications not discussed here), the unique value of the products would make such energy costs unimportant, even if energy costs remained high.

Some biological applications

Molecular devices can interact directly with the ultimate molecular components of the cell and thus serve as probes of unique value in studying processes within the cell. Further, molecular devices can characterize a frozen cell in essentially arbitrary detail by removal and characterization of successive layers of material (atomically thin layers, if desired). Although the amount of data involved is large (a typical cell contains billions of protein molecules), the physical bulk of a device able to store and manipulate this amount of data will be quite small.

The change of temperature and water distribution during freezing modifies cell structures in several ways, primarily by physical displacement of structures by ice crystals and denaturation of proteins by concentration of solutes in the residual liquid (14). With frozen tissue, knowledge of normal structures (membrane geometries, natural protein structures) and analysis of frozen structures (position of ice crystals, position of denatured proteins) should permit quite accurate reconstruction of the nature of the tissue before freezing.

Such procedures would have special utility in analyzing the structure of tissue in the brain. Unlike, say, muscle or liver tissue, the function of brain tissue depends on the detailed three-dimensional structure of intertwined cells and their interfaces. The freezing process is far too slow to stop such dynamic processes as action potentials and synaptic transmission; short-term memory, however, is suspected to involve chemical modification of the neurons, and long-term memory is believed to involve the growth and modification of neuronal structures, particularly synapses (15). At the modest freezing rates possible in substantial pieces of tissue, ice crystals may be expected to nucleate and grow in the intercellular fluid, displacing the cell membranes as they do so (16). Electron micrographs, however, show that synapses (like many intercellular junctions) involve complementary structures on both sides of the intercellular gap, which should provide information enough to reconstruct the pre-freezing configurations of the cells almost regardless of ice crystal locations.

The ability to reconstruct the prefreezing structure of tissue, when combined with the general synthetic capabilities outlined above, will make feasible the physical restoration of tissue damaged by ordinary freezing through characterization, reconstruction, and restoration of successive segments of frozen material. Although restored to a frozen condition, such tissue would lack the characteristic damage caused by the freezing process. As many tissues can survive the gross insult of ordinary freezing (17), it seems likely that most could survive freezing followed by repair. The remaining mode of damage would seem to be denaturation of proteins sensitive to cold alone during the thawing process. Should cell components of some species prove sensitive to short periods of cold, they could presumably be modified to resemble those of hardier species (hamsters can survive freezing of half their body water; ref. 17) without changing either cell function or DNA.

Implications for the present

The existence of a path to an advanced molecular technology has implications for the present. As with all technologies, long-range promise should tend to increase interest in undertaking the early steps, even beyond the interest springing from more immediate benefits. The longer the expected wait, however, the less the interest.

On the other hand, molecular engineering of materials and devices can extend the capabilities of technology many fold in many areas. The implications of the feasibility of molecular technology are important to present day speculations concerning the probable behavior (and likelihood of existence) of extraterrestrial technological civilizations. Similarly, those concerned with the long-range future of humanity must concern themselves with the opportunities and dangers arising from this technology. Finally, the eventual development of the ability to repair freezing damage (and to circumvent cold damage during thawing) has consequences for the preservation of biological materials today, provided a sufficiently long-range perspective is taken.

Conclusion

Development of the ability to design protein molecules will, by analogy between features of natural macromolecules and components of existing machines, make possible the construction of molecular machines. These machines can build second-generation machines able to perform extremely general synthesis of three-dimensional molecular structures, thus permitting construction of devices and materials to complex atomic specifications. This capability has implications for technology in general and in particular for computation and characterization, manipulation, and repair of biological materials.

I thank C. Peterson, P. Morrison, J. Lettvin, A. Kantrowitz, and C. Walsh for their comments and criticism.

1. Feynman, R. (1961) in *Miniaturization*, ed. Gilbert, H. D. (Reinhold, New York), pp. 282–296.
2. Krumhansl, J. A. & Pao, Y. H. (1979) *Phys. Today* **32** (11), 25–32.
3. Itakura, K. & Riggs, A. D. (1980) *Science* **209**, 1401–1405.
4. Nomura, M. & Held, W. (1974) in *Ribosomes*, eds. Nomura, M., Tissiera, A. & Lengyel, P. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 193–203.
5. McCammon, J. A., Gelin, B. R. & Karplus, M. (1977) *Nature (London)* **267**, 585–590.
6. Scheraga, H. A. (1978) in *Versatility of Proteins*, ed. Li, C. H. (Academic, New York), pp. 119–132.
7. Karplus, M. & Weaver, D. L. (1976) *Nature (London)* **260**, 404–406.
8. Gund, P., Andose, J. D., Rhodes, J. B. & Smith, G. M. (1980) *Science* **208**, 1425–1431.
9. Gutte, B., Dannigen, M. & Wittschieber, E. (1979) *Nature (London)* **281**, 650–655.
10. Anonymous (1980) *Semicond. Int.* **3** (5), 10.
11. Walsh, C. (1979) *Enzymatic Reaction Mechanisms* (Freeman, San Francisco), pp. 33, 38.
12. Drake, J. (1969) *Nature (London)* **221**, 1132.
13. Chance, B., Mueller, P., DeVault, D. & Powers, L. (1980) *Phys. Today* **33** (10), 32–38.
14. Fennema, O. R. (1973) in *Low-Temperature Preservation of Foods and Living Matter*, eds. Fennema, O. R., Powrie, W. D. & Marth, E. H. (Dekker, New York), pp. 476–503.
15. Entingh, D., Dunn, A., Glassman, E., Wilson, J. E., Hogan, E. & Damstra, T. (1975) in *Handbook of Psychobiology*, eds. Gazzinga, M. S. & Blakemore, C. (Academic, New York), pp. 201–238.
16. Fennema, O. R. (1973) in *Low-Temperature Preservation of Foods and Living Matter*, eds. Fennema, O. R., Powrie, W. D. & Marth, E. H. (Dekker, New York), pp. 150–239.
17. Fennema, O. R. (1973) in *Low-Temperature Preservation of Foods and Living Matter*, eds. Fennema, O. R., Powrie, W. D. & Marth, E. H. (Dekker, New York), pp. 436–475.