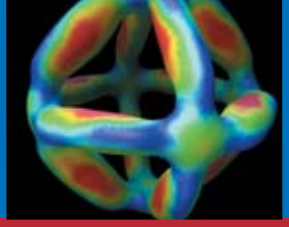
Artificial Molecules Deliver Drugs
p. 72

DNA Building Blocks
p. 30



Nanotech in Sci-Fi
p. 80

SCIENTIFIC AMERICAN
REPORTS
SPECIAL EDITION ON NANOTECHNOLOGY
www.SciAm.com

THE RISE OF NANOTECH

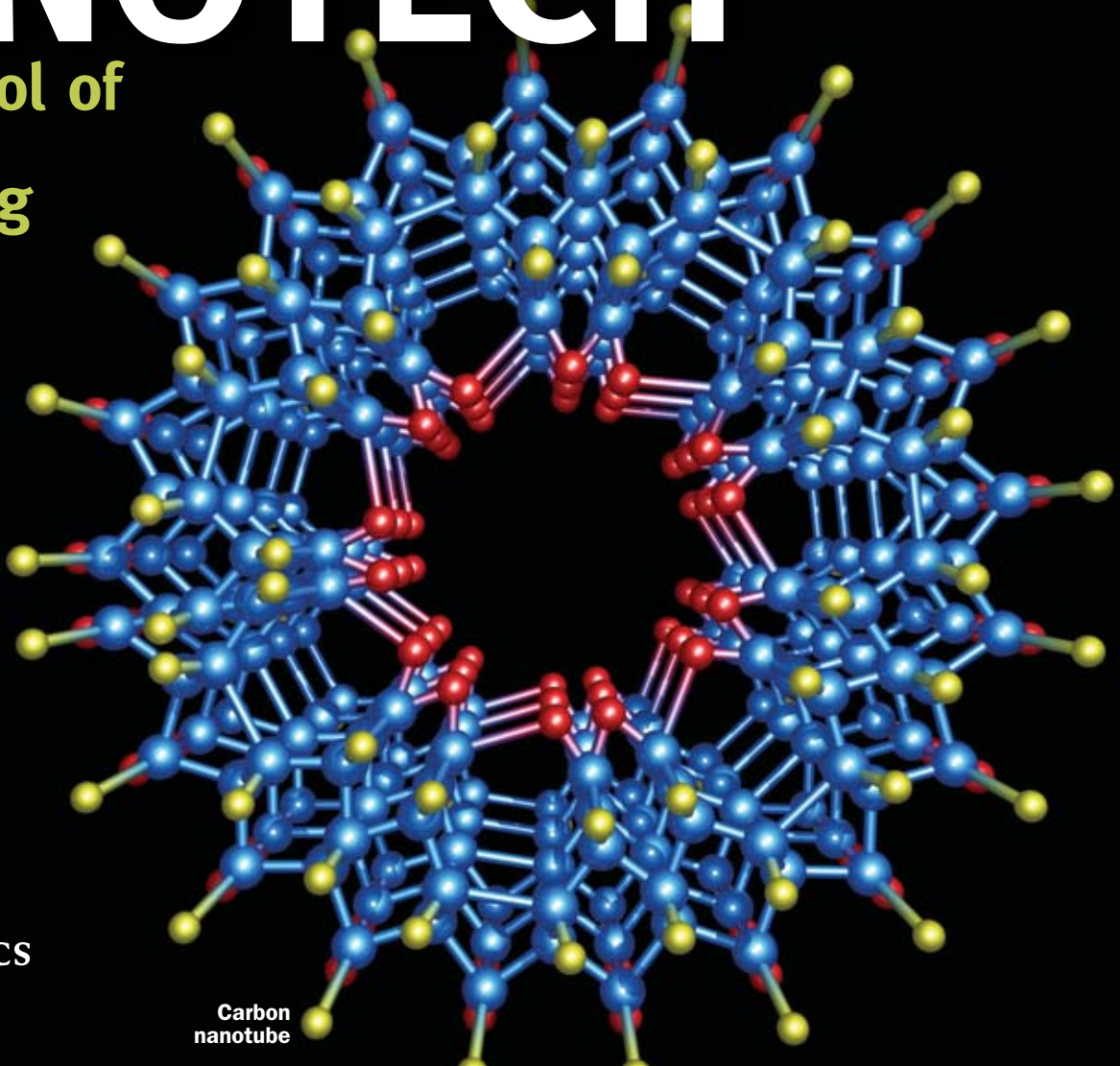
How control of molecules is changing the world

Invisibility Cloak?
Bending Light with Plasmonics

Computing with DNA
Tiny Machines Speak to Cells

Carbon Nanonets
Electronics Gets a Boost

Weird Physics
Small Size, Big Differences



Carbon nanotube

Page Intentionally Blank

SCIENTIFIC AMERICAN
REPORTS

Scientific American Reports is published by the staff of SCIENTIFIC AMERICAN, with project management by:

EDITOR IN CHIEF: John Rennie
EXECUTIVE EDITOR: Mariette DiChristina
ISSUE EDITOR: Dawn Stover

ART DIRECTOR: Patricia Nemoto
PHOTOGRAPHY EDITORS: Emily Harrison, Smitha Alampur
PRODUCTION EDITOR: Richard Hunt

COPY DIRECTOR: Maria-Christina Keller
COPY CHIEF: Daniel C. Schlenoff
ASSISTANT COPY CHIEF: John Matson
COPY AND RESEARCH: Rachel Dvoskin, Aaron Shattuck, Kenneth Silber, Kevin Singer, Michelle Wright

EDITORIAL ADMINISTRATOR: Jacob Lasky
SENIOR SECRETARY: Maya Harty

ASSOCIATE PUBLISHER, PRODUCTION: William Sherman
MANUFACTURING MANAGER: Janet Cermak
ADVERTISING PRODUCTION MANAGER: Carl Cherebin
PREPRESS AND QUALITY MANAGER: Silvia De Santis
PRODUCTION MANAGER: Christina Hippeli
CUSTOM PUBLISHING MANAGER: Madelyn Keyes-Milch

ASSOCIATE PUBLISHER, CIRCULATION: Simon Aronin
CIRCULATION DIRECTOR: Christian Dorbandt
RENEWALS MANAGER: Karen Singer
FULFILLMENT AND DISTRIBUTION MANAGER: Rosa Davis

VICE PRESIDENT AND PUBLISHER: Bruce Brandfon
SALES DEVELOPMENT MANAGER: David Tirpack
SALES REPRESENTATIVES: Jeffrey Crennan, Stephen Dudley, Stan Schmidt

ASSOCIATE PUBLISHER, STRATEGIC PLANNING: Laura Salant
PROMOTION MANAGER: Diane Schube
RESEARCH MANAGER: Aida Dadurian
PROMOTION DESIGN MANAGER: Nancy Mongelli

GENERAL MANAGER: Michael Florek
BUSINESS MANAGER: Marie Maher
MANAGER, ADVERTISING ACCOUNTING AND COORDINATION: Constance Holmes

DIRECTOR, SPECIAL PROJECTS: Barth David Schwartz

MANAGING DIRECTOR AND VICE PRESIDENT, ONLINE: Mina C. Lux
DIRECTOR, WEB TECHNOLOGIES, ONLINE: Vincent Ma
SALES REPRESENTATIVE, ONLINE: Gary Bronson

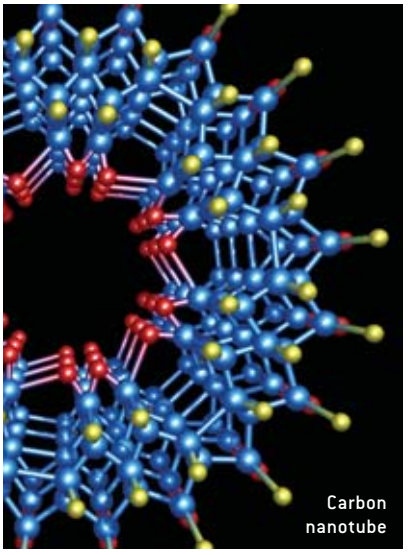
DIRECTOR, ANCILLARY PRODUCTS: Diane McGarvey
PERMISSIONS MANAGER: Linda Hertz

CHAIRMAN: Brian Napack
CHAIRMAN EMERITUS: John J. Hanley

VICE PRESIDENT AND MANAGING DIRECTOR, INTERNATIONAL: Dean Sanderson
VICE PRESIDENT: Frances Newburg

letter from the editor

Small World


Carbon nanotube

ALFRED PASIEKA *SPL/Photo Researchers, Inc.*

IF YOU HAVE heard about nanotechnology at all, you may be aware of its science-fiction-sounding hype. Proponents picture a future in which tiny bots would magically repair tissue to prolong our life span. On the dark side is the disturbing vision of "gray goo," where self-replicating nanodevices destroy the planet. The reality of the burgeoning field of nanotech, however, is hardly less startling in its transformative potential. Some have proclaimed it "the next industrial revolution."

"Nanotechnology" broadly applies to control of materials and components only a few billionths of a meter in size. Already manufacturers sell several hundred products that use nanotech, mainly skin lotions. Next up are advances in biotechnology and electronics—and a merging of the two.

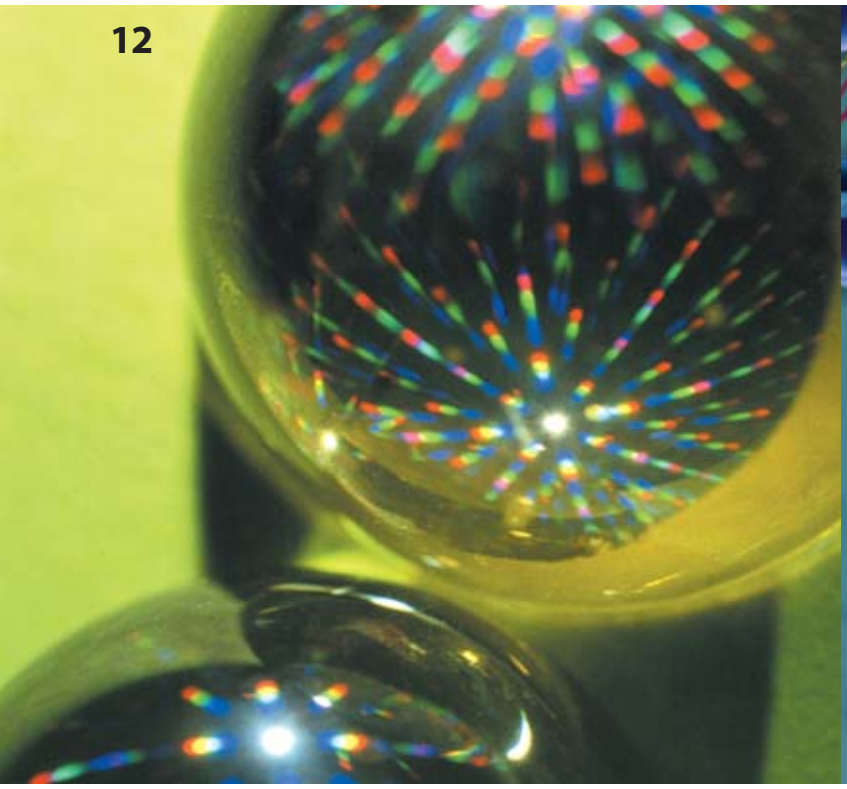
Consider, for instance, molecular building blocks called bis-amino acids, which chemists string together into proteinlike structures, as described by Christian E. Schafmeister in his article, "Molecular Lego," starting on page 22. Applications include medicines, enzymes for catalyzing reactions, sensors, nanoscale valves and computer storage devices. Other researchers are using natural molecular machines to process information: they receive input from other biological molecules and output a tangible result, such as a signal or a therapeutic drug. For more, turn to "Bringing DNA Computers to Life," by Ehud Shapiro and Yaakov Benenson, on page 40.

Nanoscience advances are pushing traditional electronics in new directions as well. In "Carbon Nanonets Spark New Electronics" (*page 48*), George Gruner describes applications that encompass sensors, solar cells, electronic paper and bendable touch screens. Imagine a morning "paper" with headlines that change as news breaks.

Or how about an invisibility cloak? In "The Promise of Plasmonics" (*page 56*), Harry A. Atwater explains how optical signals squeeze through minuscule wires, producing so-called plasmons. Plasmonic circuits could help to move lots of data and improve the resolution of microscopes, the efficiency of light-emitting diodes, and the sensitivity of detectors. Such materials could alter the electromagnetic field around an object to such an extent that it would become invisible. The nanoregime offers enormous promise indeed.

Mariette DiChristina
Executive Editor
Scientific American
editors@SciAm.com

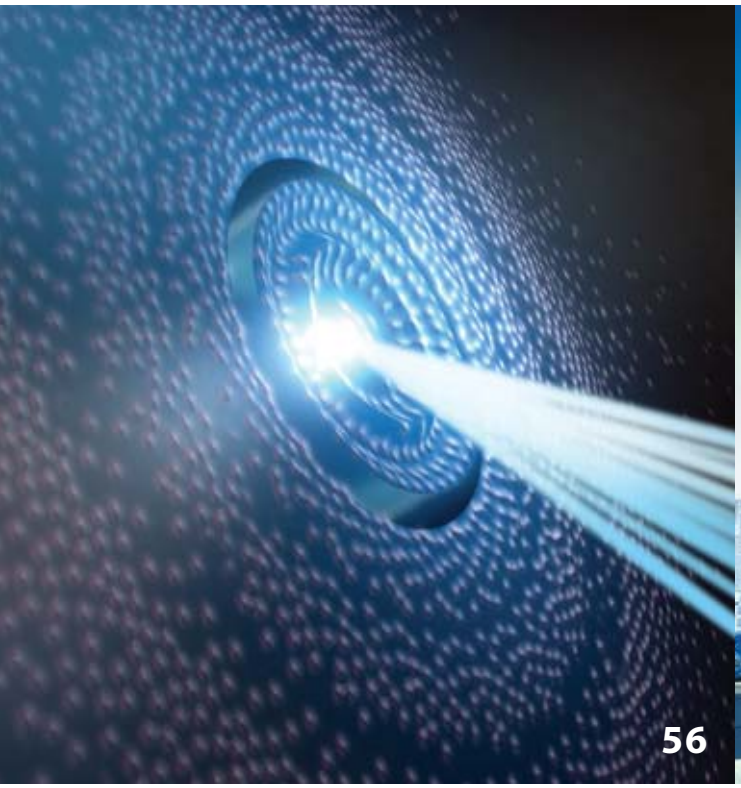
Contents

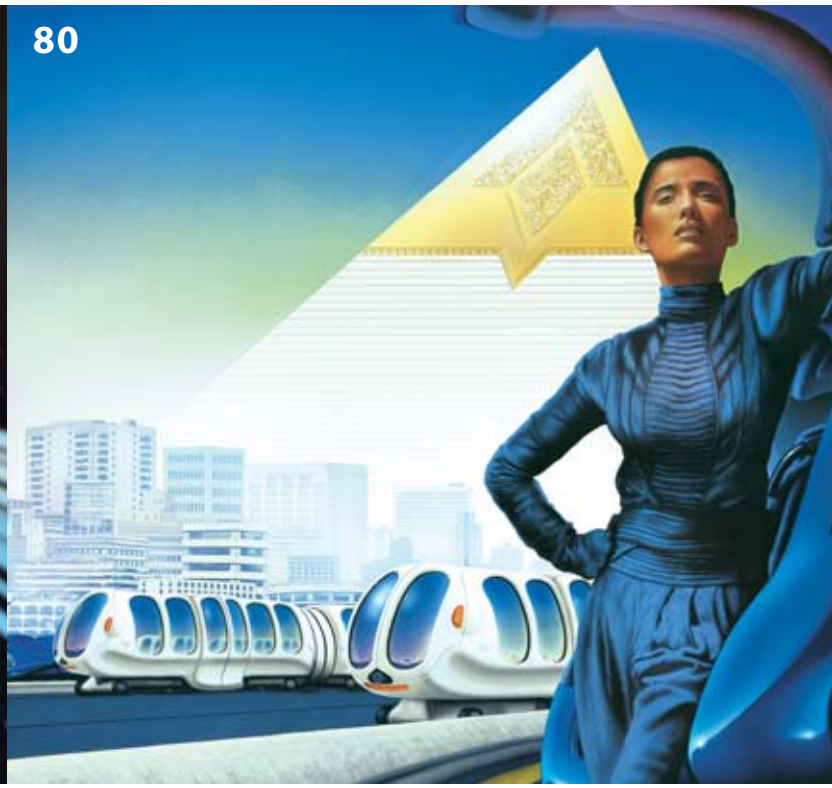


Cover illustration by Alfred Pasieka, SPL/Photo Researchers, Inc.

The articles in this special edition are updated from previous issues of *Scientific American*.

SCIENTIFIC AMERICAN
REPORTS
Volume 17, Number 3, 2007





THE SMALLEST CIRCUITS

FANTASTIC VOYAGE

Scientific American Special (ISSN 1936-1513), Volume 17, Number 3, 2007, published by Scientific American, Inc., 415 Madison Avenue, New York, NY 10017-1111. Canadian BN No. 127387652RT; QST No. Q1015332537. To purchase additional quantities: U.S., $10.95 each; elsewhere, $13.95 each. Send payment to Scientific American, Dept. NT2007, 415 Madison Avenue, New York, NY 10017-1111. **Inquiries: fax 212-355-0408 or telephone 212-451-8442.** Printed in U.S.A.

Plenty of Room Indeed

There is plenty of room for practical innovation at the nanoscale. But first, scientists have to understand the unique physics that governs matter there

By Michael Roukes

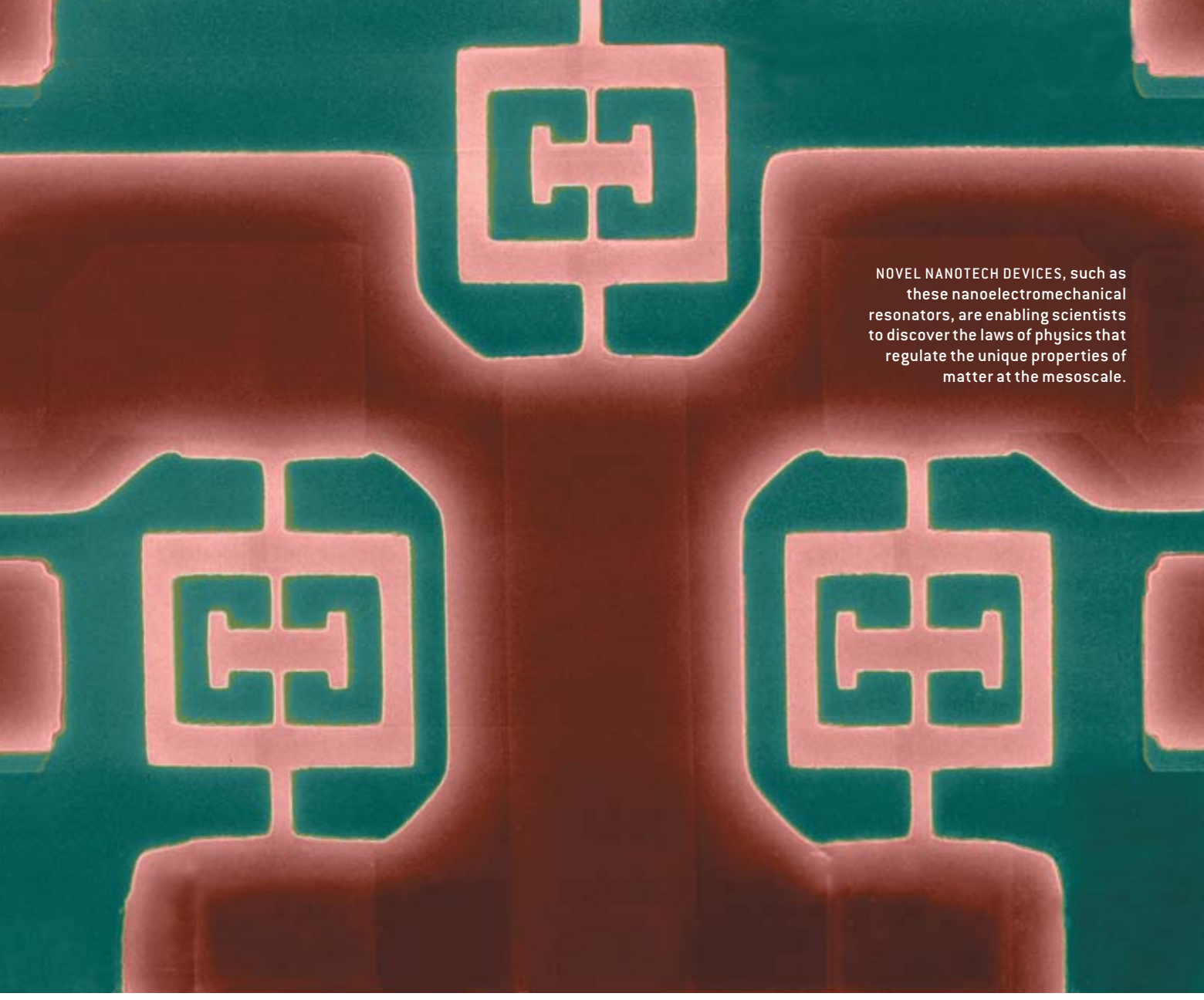
Back in December 1959, future Nobel laureate Richard Feynman gave a visionary and now oft-quoted talk entitled "There's Plenty of Room at the Bottom." The occasion was an American Physical Society meeting at the California Institute of Technology, Feynman's intellectual home then and mine today. Although he didn't intend it, Feynman's 7,000 words were a defining moment in nanotechnology, long before anything "nano" appeared on the horizon.

"What I want to talk about," he said, "is the problem of manipulating and controlling things on a small scale.... What I have demonstrated is that there is room—that you can decrease the size of things in a practical way. I now want to show that there is plenty of room. I will not now discuss how we are going to do it, but only what is possible in principle....We are not doing it simply because we haven't yet gotten around to it."

The breadth of Feynman's vision is staggering. In that lecture 48 years ago he anticipated a spectrum of scientific and technical fields that are now well established, among them electron-beam and ion-beam fabrication, molecular-beam epitaxy, nanoimprint lithography, projection electron microscopy, atom-by-atom manipulation, quantum-effect electronics, spin electronics (also called spintronics) and microelectromechanical systems (MEMS). The lecture also projected what has been called the "magic" Feynman brought to everything he turned his singular intellect toward. Indeed, it has profoundly inspired my more than two decades of research on physics at the nanoscale.

Today there is a nanotechnology gold rush. Nearly every major funding agency for science and engineering has its own thrust into the field. Scores of researchers and institutions are scrambling for a piece of the action. But in all honesty, I think we have to admit that much of what invokes the hallowed prefix "nano" falls a bit short of Feynman's mark.

We've only just begun to take the first steps toward his grand vision of assembling complex machines and circuits atom by atom. What can be done now is extremely rudimentary. We're certainly nowhere near being able to



NOVEL NANOTECH DEVICES, such as these nanoelectromechanical resonators, are enabling scientists to discover the laws of physics that regulate the unique properties of matter at the mesoscale.

M. J. MURPHY, D. A. HARRINGTON AND M. L. ROUKES
California Institute of Technology; COLORIZATION BY FELICE FRANKEL

commercially mass-produce nanosystems—integrated multicomponent nanodevices that have the complexity and range of functions readily provided by modern microchips. But there is a fundamental science issue here as well. It is becoming increasingly clear that we are only *beginning* to acquire the detailed knowledge that will be at the heart of future nanotechnology. This new science concerns the properties and behavior of aggregates of atoms and molecules, at a scale not yet large enough to be considered macroscopic but far beyond what can be called microscopic. It is the science of the mesoscale, and until we understand it, practical devices will be difficult to realize.

Scientists and engineers readily fashion nanostructures on a scale of one to a few hundred nanometers—small indeed, but much bigger than simple molecules. Matter at this mesoscale is often awkward to explore. It contains too many atoms to be easily understood by the straightforward application of quantum mechanics (although the fundamental laws still apply). Yet these systems are not so large as to be completely free of quantum effects; thus, they do not simply obey the classical physics governing the macroworld. It is precisely in this intermediate domain, the mesoworld, that unforeseen properties of collective systems emerge.

Researchers are approaching this transitional frontier using complementary top-down and bottom-up fabrication methods. Advances in top-down nanofabrication techniques, such as electron-beam lithography (used extensively by my own research group), yield almost atomic-scale precision, but achieving success, not to mention reproducibility, as we scale down to the single-digit-nanometer regime becomes problematic. Alternatively, scientists are using bottom-up techniques for self-assembly of atoms. But the advent of preprogrammed self-assembly of arbitrarily large systems—with complexity comparable to that built every day in microelectronics, in MEMS and (of course) by Mother Nature—is nowhere on the horizon. It appears that the top-down approach will most likely remain the method of choice for building re-

It is becoming increasingly clear that we are only beginning to acquire the detailed knowledge that will be at the heart of future nanotechnology.

ally complex devices for a good while.

Our difficulty in approaching the mesoscale from above or below reflects a basic challenge of physics. Lately, the essence of Feynman's "Plenty of Room" talk seems to be taken as a license for laissez-faire in nanotechnology. Yet Feynman never asserted that "anything goes" at the nanoscale. He warned, for instance, that the very act of trying to "arrange the atoms one by one the way we want them" is subject to fundamental principles: "You can't put them so that they are chemically unstable, for example."

Accordingly, today's scanning probe microscopes can move atoms from place to place on a prepared surface, but this ability does not immediately confer the power to build complex molecular assemblies at will. What has been accomplished so far, though impressive, is still quite limited. We will ultimately develop operational procedures to help us coax the formation of individual atomic bonds under more general conditions. But as we try to assemble complex networks of these bonds, they certainly will affect one another in ways we do not yet understand and, hence, cannot yet control.

Feynman's original vision was clearly intended to be inspirational. Were he observing now, he would surely be alarmed when people take his projections as some sort of gospel. He delivered his musings with characteristic playfulness as well as deep insight. Sadly for us, the field that would be called nanotechnology was just one of many that intrigued him. He never really continued with it, returning to give but one redux of his original lecture, at the Jet Propulsion Laboratory in 1983.

New Laws Prevail

IN 1959, AND EVEN in 1983, the complete physical picture of the nanoscale was far from clear. The good news for researchers is that, by and large, it still is! Much exotic territory awaits exploration. As we delve into it, we will uncover a panoply of phenomena that we must understand before practical nanotechnology will become possible. The past few decades have seen the elucidation of entirely new, fundamental physical principles that govern behavior at the mesoscale. Let's consider three important examples.

In the fall of 1987 graduate student Bart J. van Wees of the Delft University of Technology and Henk van Houten of the Philips Research Laboratories (both in the Netherlands) and their collaborators were studying the flow of electric current through what are now called quantum-point contacts. These are narrow conducting paths within a semiconductor, along which electrons are forced to flow [*see box on page 8*]. Late one evening van Wees's undergraduate assistant, Leo Kouwenhoven, was measuring the conductance through the constriction as he varied its width systematically. The research team was expecting to see only subtle conductance effects against an otherwise smooth and unremarkable background response. Instead there appeared a very pronounced, and now characteristic, staircase pattern. Further analysis that night revealed that plateaus were occurring at regular, precise intervals.

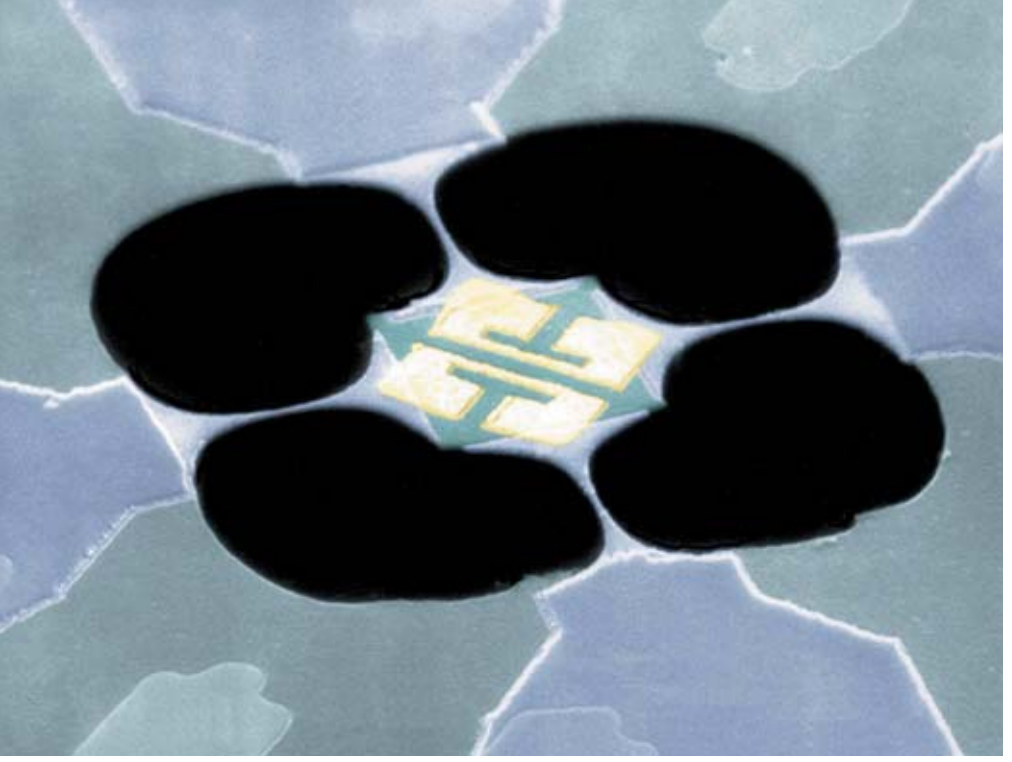
David Wharam and Michael Pepper of the University of Cambridge observed similar results. The two discoveries represented the first robust demonstrations of the quantization of electrical conductance. This is a basic property of small conductors that occurs when the wavelike properties of electrons are coherently maintained from the "source" to the "drain"—the input to the output—of a nanoelectronic device.

Feynman anticipated, in part, such odd behavior: "I have thought about some of the problems of building electric circuits on a small scale, and the problem of resistance is serious...." But the experimental discoveries pointed out something truly new and fundamental: quantum mechanics can completely govern the behavior of small electrical devices.

Direct manifestations of quantum mechanics in such devices were envisioned back in 1957 by Rolf Landauer, a theoretician at IBM who pioneered ideas in nanoscale electronics and in the physics of computation. But only in

Overview/*Nanophysics*

- Smaller than macroscopic objects but larger than molecules, nanotechnological devices exist in a unique realm—the mesoscale—where the properties of matter are governed by a complex and rich combination of classical physics and quantum mechanics.
- Engineers will not be able to make reliable or optimal nanodevices until they comprehend the physical principles that prevail at the mesoscale.
- Scientists are discovering mesoscale laws by fashioning unusual, complex systems of atoms and measuring their intriguing behavior.
- Once we understand the science underlying nanotechnology, we can fully realize the prescient vision of Richard Feynman: that nature has left plenty of room in the nanoworld to create practical devices that can help humankind.



NANOBRIDGE DEVICE allowed Caltech physicists to first observe the quantization of thermal conductance—a fundamental limit to heat flow in minute objects. Four holes (*black*) etched into a silicon nitride membrane defined an isolated thermal reservoir (*central green square*) suspended by four narrow bridges. One gold transducer (*yellow*) electrically heated this reservoir; the second measured its temperature. Thin superconducting films (*blue*) on top of the bridges electrically connected the transducers to off-chip instrumentation but carried no heat. The reservoir therefore cooled only through the silicon nitride bridges, which were so narrow that they passed only the lowest-energy heat waves.

K. C. SCHWAB, E. A. HENRIKSEN AND M. L. ROUKES *Caltech*; COLORIZATION BY M. L. ROUKES

the mid-1980s did control over materials and nanofabrication begin to provide access to this regime in the laboratory. The 1987 discoveries heralded the heyday of "mesoscopia."

A second significant example of newly uncovered mesoscale laws that have led to nascent nanotechnology was first postulated in 1985 by Konstantin Likharev, a young physics professor at Moscow State University working with postdoctoral student Alexander Zorin and undergraduate Dmitri Averin. They anticipated that scientists would be able to control the movement of single electrons on and off a "coulomb island," a conductor weakly coupled to the rest of a nanocircuit. This could form the basis for an entirely new type of device, called a single-electron transistor. The physical effects that arise when putting a single electron on a coulomb island become more robust as the island is scaled downward. In very small devices, these single-electron charging effects can completely dominate the current flow.

Such considerations are becoming increasingly important technologically. Projections from the International Technology Roadmap for Semiconductors, prepared by long-range thinkers in the industry, indicate that by 2014 the minimum feature size for transistors in computer chips will decrease to 20 nanometers. At this dimension, each switching event will involve the equivalent of only about eight electrons. Designs that properly account for single-electron charging will become crucial.

By 1987 advances in nanofabrication allowed Theodore A. Fulton and Gerald J. Dolan of Bell Laboratories to construct the first single-electron transistor [*see box on page 10*]. The single-electron charging they observed, now called the coulomb blockade, has since been seen in a wide array of structures. As experimental devices get smaller, the coulomb blockade phenomenon is becoming the rule, rather than the exception, in weakly coupled nanoscale devices. This is especially true in experiments in which electric currents are passed through individual molecules. These molecules can act like coulomb islands by virtue of their weak coupling to electrodes leading back to the macroworld. Using this effect to advantage and obtaining robust, reproducible coupling to small molecules (in ways that can actually be engineered) are among the important challenges in the new field of molecular electronics.

In 1990, against this backdrop, I was at Bell Communications Research studying electron transport in mesoscopic semiconductors. In a side project, my colleagues Larry M. Schiavone and Axel Scherer and I began developing techniques that we hoped would elucidate the quantum nature of *heat* flow. The work required much more sophisticated nanostructures than the planar devices used to investigate mesoscopic electronics. We needed freely suspended devices, structures possessing full three-dimensional relief. Ignorance was bliss; I had no idea the experiments would be so involved that they would take almost a decade to realize.

The first big strides were made after I moved to Caltech in 1992, in a collaboration with John M. Worlock of the University of Utah and two successive postdocs in my group. Thomas S. Tighe developed the methods and devices that generated the first direct measurements of heat flow in nanostructures. Subsequently, Keith C. Schwab revised the design of the suspended nanostructures and put in place ultrasensitive superconducting instrumentation to interrogate them at ultralow temperatures, at which the effects could be seen most clearly.

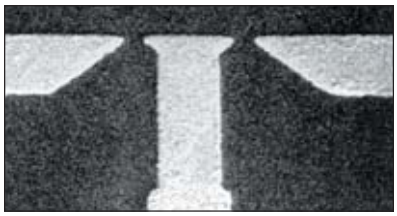
In the late summer of 1999 Schwab finally began observing heat flow through silicon nitride nanobridges [*see illustration above*]. Even in these first data the fundamental limit to heat flow

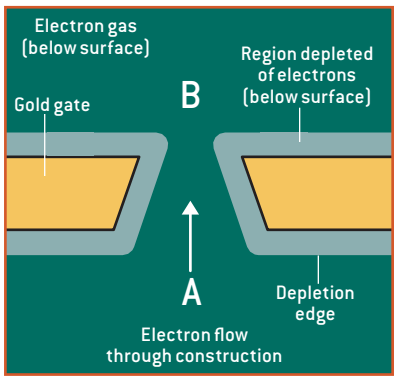
THE AUTHOR

MICHAEL ROUKES, professor of physics, applied physics and bioengineering at the California Institute of Technology, heads a large, cross-disciplinary group studying nanoscale systems. He was recently the founding director of Caltech's Kavli Nanoscience Institute. Among the holy grails his team is chasing are nanodevices to weigh every protein in a single cell and nanodevices to watch the metabolic fluctuations of individual cells in real time through direct measurement of their heat output.

ONE STEP AT A TIME

QUANTIZATION OF ELECTRICAL CONDUCTANCE





In 1987 Bart J. van Wees and his collaborators at the Delft University of Technology and Philips Research Laboratories (both in the Netherlands) built a novel structure (*micrograph*) that revealed a basic law governing nanotech circuits. Gold gate electrodes (*bright areas*) were placed atop a semiconductor substrate (*dark background*). Within the substrate, a planar sheet of charge carriers, called a two-dimensional electron gas, was created about 100 nanometers below the surface. The gates and the gas acted like the plates of a capacitor.

When a negative voltage bias was applied to the gates, electrons within the gas underneath the gates, and slightly beyond the gates' periphery, were pushed away. (The diagram shows this state.) When increasing negative voltage was applied, this "depletion edge" became more pronounced. At a certain threshold, carriers on either side of the constriction (*between points A and B*) became separated, and the conductance through the device was zero. From this threshold level, conductance did not resume smoothly. Instead it increased in stepwise fashion, where the steps occurred at values determined by twice the charge of the electron squared, divided by Planck's constant. This ratio is now called the electrical conductance quantum, and it indicates that electric current flows in nanocircuits at rates that are quantized.

in mesoscopic structures emerged. The manifestation of this limit is now called the thermal conductance quantum. It determines the maximum rate at which heat can be carried by an individual wavelike mechanical vibration, spanning from the input to the output of a nanodevice. It is analogous to the electrical conductance quantum but governs the transport of heat.

This quantum is a significant parameter for nanoelectronics; it represents the ultimate limit for the power-dissipation problem. In brief, all "active" devices require a little energy to operate, and for them to operate stably without overheating, we must design a way to extract the heat they dissipate. As engineers try continually to increase the density of transistors and the clock rates (frequencies) of microprocessors, the problem of keeping microchips cool to avoid complete system failure is becoming monumental. This will only become further exacerbated in nanotechnology.

Considering even this complexity, Feynman said, "Let the bearings run dry; they won't run hot because the heat escapes away from such a small device very, very rapidly." But our experiments indicate that nature is a little more restrictive. The thermal conductance quantum can place limits on how effectively a very small device can dissipate heat. What Feynman envisioned can be correct only if the nanoengineer designs a structure so as to take these limits into account.

From the three examples above, we can arrive at just one conclusion: we are only starting to unveil the complex and wonderfully different ways that nanoscale systems behave. The discovery of the electrical and thermal conductance quanta and the observation of the coulomb blockade are true discontinuities—abrupt changes in our understanding. Today we are not accustomed to calling our discoveries "laws." Yet I have no doubt that electrical and thermal conductance quantization and single-electron-charging phenomena are indeed among the universal rules of nanodesign. They are new laws of the nanoworld. They do not contravene but augment and clarify some of Feynman's original vision. Indeed, he seemed to have anticipated their emergence: "At the atomic level, we have new kinds of forces and new kinds of possibilities, new kinds of effects. The problems of manufacture and reproduction of materials will be quite different."

We will encounter many more such discontinuities on the path to true nanotechnology. These welcome windfalls will occur in direct synchrony with advances in our ability to observe, probe and control nanoscale structures. It would seem wise, therefore, to be rather modest and circumspect about forecasting nanotechnology.

The Boon and Bane of Nano

THE NANOWORLD is often portrayed by novelists, futurists and the popular press as a place of infinite possibilities. But as you've been reading, this domain is not some ultraminiature version of the Wild West. Not everything goes down there; there are *laws*. Two concrete illustrations come from the field of nanoelectromechanical systems (NEMS), in which I am active.

Part of my research is directed toward harnessing small mechanical devices for sensing applications. Nanoscale structures appear to offer revolutionary potential: the smaller a device, the more susceptible its physical properties to alteration. One example is resonant detectors, which are frequently used for sensing mass. The vibrations of a tiny mechanical element, such as a small cantilever, are intimately linked to the element's mass, so the addition of a minute amount of foreign material (the "sample" being weighed) will shift the resonant frequency. Work in my lab by then postdoc Kamil Ekinci shows that nanoscale devices can be made so

B. J. VAN WEES *Delft University of Technology (micrograph)*; NINA FINKEL (*diagram*)

The difficulties in communication between the nanoworld and the macroworld represent a central issue in the development of nanotechnology.

sensitive that "weighing" individual atoms and molecules becomes feasible.

But there is a dark side. Gaseous atoms and molecules constantly adsorb and desorb from a device's surfaces. If the device is macroscopic, the resulting fractional change in its mass is negligible. But the change can be significant for nanoscale structures. Gases impinging on a resonant detector can change the resonant frequency randomly. Apparently, the smaller the device, the less stable it will be. This instability may pose a real disadvantage for various types of futuristic electromechanical signal-processing applications. Scientists might be able to work around the problem by, for example, using arrays of nanomechanical devices to average out fluctuations. But for individual elements, the problem seems inescapable.

A second example of how "not everything goes" in the nanoworld relates more to economics. It arises from the intrinsically ultralow power levels at which nanomechanical devices operate. Physics sets a fundamental threshold for the minimum operating power: the ubiquitous, random thermal vibrations of a mechanical device impose a "noise floor" below which real signals become increasingly hard to discern. In practical use, nanomechanical devices are optimally excited by signal levels a thousandfold or a millionfold greater than this threshold. But such levels are still a millionth to a billionth the amount of power used for conventional transistors.

The advantage, in some future nanomechanical signal-processing system or computer, is that even a million nanomechanical elements would dissipate only a millionth of a watt, on average. Such ultralow power systems could lead to wide proliferation and distribution of cheap, ultraminiature "smart" sensors that could continuously monitor *all* the important functions in hospitals, in manufacturing plants, on aircraft, and so on. The idea of ultraminiature devices that drain their batteries extremely slowly, especially ones with sufficient computational power to function autonomously, has great appeal.

But here, too, there is a dark side. The regime of ultralow power is quite foreign to present-day electronics. Nanoscale devices will require entirely new system architectures that are compatible with amazingly low power thresholds. This prospect is not likely to be received happily by the computer industry, with its overwhelming investment in current devices and methodology. A new semiconductor processing plant today costs more than $1 billion, and it would probably have to be retooled to be useful. But I am certain that the revolutionary prospects of nanoscale devices will eventually compel such changes.

JOE MUNROE *Ohio Historical Society*, COURTESY OF THE ARCHIVES, CALIFORNIA INSTITUTE OF TECHNOLOGY

RICHARD FEYNMAN predicted the rise of nanotechnology in a landmark 1959 talk at Caltech. "The principles of physics," he said, "do not speak against the possibility of maneuvering things atom by atom." But he also anticipated that unique laws would prevail; they are finally being discovered today.

Monumental Challenges

CERTAINLY A HOST of looming issues will have to be addressed before we can realize the potential of nanoscale devices. Although each research area has its own concerns, some general themes emerge. Two challenges fundamental to my current work on nanomechanical systems, for instance, are relevant to nanotechnology in general.

Challenge 1: Communication between the macroworld and the nanoworld. NEMS are incredibly small, yet their motion can be far smaller. For example, a nanoscale beam clamped on both ends vibrates with minimal harmonic distortion when its vibration amplitude is kept below a small fraction of its thickness. For a 10-nanometer-thick beam, this amplitude is only a few nanometers. Building the requisite, highly efficient transducers to transfer information from such a device to the macroworld involves reading out information with even greater precision.

Compounding this problem, the natural frequency of the vibration increases as the size of the beam is decreased. So to track the device's vibrations usefully, the ideal NEMS transducer must be capable of resolving extremely small displacements, in the picometer-to-femtometer (trillionth to quadrillionth of a meter) range, across very large bandwidths, extending into the microwave range. These twin requirements pose a truly monumental challenge, one much more significant than those faced so far in MEMS work. A further complication is that most of the methodologies from MEMS are inapplicable; they simply don't scale down well to nanometer dimensions.

These difficulties in communication between the nanoworld and the macroworld represent a generic issue in the development of nanotechnology. Ulti-

In each new regime, some wonderful scientific phenomenon emerges. But then a thorny host of underlying, equally unanticipated problems appear.

mately, the technology will depend on robust, well-engineered information transfer pathways from what are, in essence, individual macromolecules. Although the grand vision of futurists may involve self-programmed nanobots that need direction from the macroworld only when they are first wound up and set in motion, it seems more likely that most nanotechnological applications realizable in our lifetimes will entail some form of reporting up to the macroworld and feedback and control back down. The communication problem will remain central.

Orchestrating such communication immediately invokes the very real possibility of collateral damage. Quantum theory tells us that the process of measuring a quantum system nearly always perturbs it. This can hold true even when we scale up from atoms and molecules to nanosystems comprising millions or billions of atoms. Coupling a nanosystem to probes that report back to the macroworld always changes the nanosystem's properties to some degree, rendering it less than ideal. The transducers required for communication will do more than just increase the nanosystem's size and complexity. They also necessarily extract some energy to perform their measurements and can degrade the nanosystem's performance. Measurement always has its price.

Challenge II: Surfaces. As we shrink MEMS to NEMS, device physics becomes increasingly dominated by the surfaces. Much of the foundation of solid-state physics rests on the premise that the surface-to-volume ratio of objects is infinitesimal, meaning physical properties are always dominated by the physics of the bulk. Nanoscale systems are so small that this assumption breaks down completely.

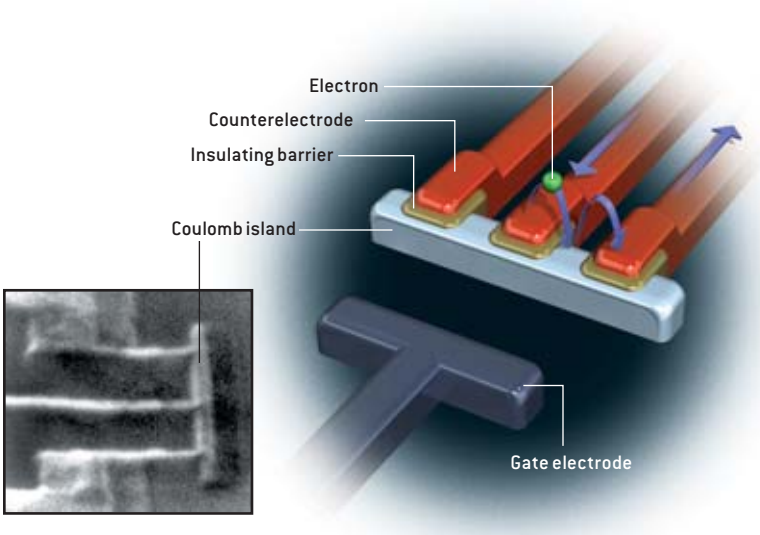
For example, mechanical devices patterned from single-crystal, ultrapure materials can contain very few (even zero) crystallographic defects and impurities. My initial hope was that, as a result, there would be only very weak damping of mechanical vibrations in monocrystalline NEMS. But as we shrink mechanical devices, we repeatedly find that acoustic energy loss seems to increase in proportion to the increasing surface-to-volume ratio. This result clearly implicates surfaces in

TAKING CHARGE

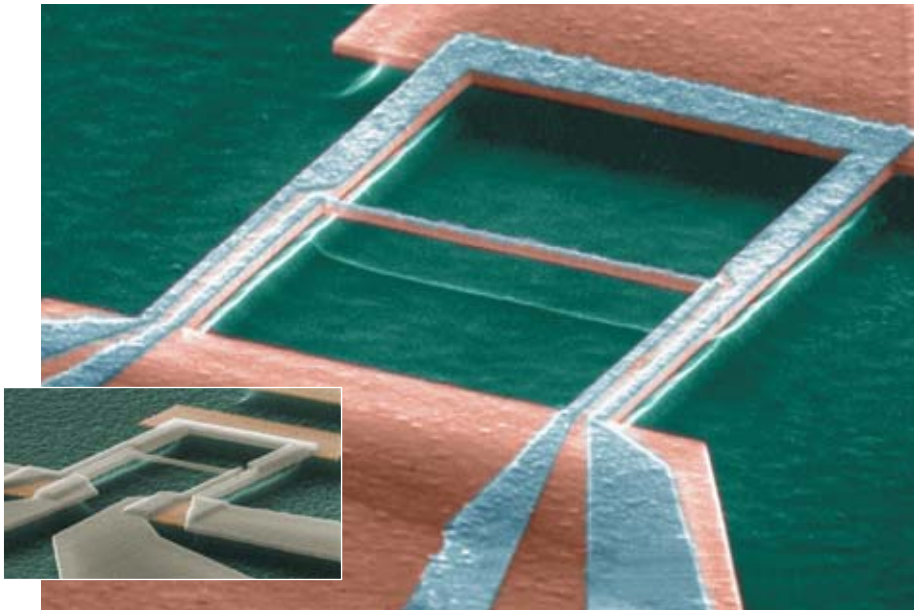
SINGLE ELECTRONICS

Advances in nanofabrication allowed Theodore A. Fulton and Gerald J. Dolan to build a single-electron transistor at Bell Laboratories in 1987 (*micrograph*). In this structure, the controlled movement of individual electrons through a nanodevice was first achieved. At its heart was a coulomb island, a metallic electrode isolated from its counterelectrodes by thin insulating oxide barriers (*diagram*). The counterelectrodes led up to the macroscale laboratory instrumentation used to carry out the experiments. An additional gate electrode (*visible in diagram but not in micrograph*) was offset from the coulomb island by a small gap; it allowed direct control of the charge introduced to the island. Electric current flowed through the device from one counterelectrode to another, as in a conventional circuit, but here it was limited by the stepwise hopping of electrons onto and off the coulomb island.

Fulton and Dolan's experiments demonstrate both the fundamental physics of single-electron charging and the potential of these devices as ultrasensitive electrometers: instruments that can easily detect individual electron charges. Circuits that switch one electron at a time could someday form the basis for an entirely new class of nanoelectronics. The advent of such single electronics, however, also presages problems that will have to be faced as conventional electronic circuits are shrunk to the nanoscale.



T. A. FULTON AND G. J. DOLAN *Bell Laboratories (micrograph)*; BRYAN CHRISTIE DESIGN (*diagram*)



NANOMECHANICAL AMPLIFIER overcomes the vexing problem of communication with the macroworld by providing up to 1,000-fold amplification of weak forces. Two suspended bridges of monocrystalline silicon carbide (*left* and *right*, *main panel*) support the central crossbridge, to which a high-frequency (17 megahertz) signal force is applied. Thin-film electrodes (*silver*) atop these structures provide very sensitive readouts of nanoscale motion. A smaller, second-generation device (*inset*), shown to scale with the first, operates at the very high frequency of 140 megahertz.

the devices' vibrational energy-loss processes. In a state-of-the-art silicon beam measuring 10 nanometers wide and 100 nanometers long, more than 10 percent of the atoms are at or next to the surface. It is evident that these atoms will play a central role, but understanding precisely how will require a major, sustained effort.

In this context, nanotube structures, which are the focus of much current research, ostensibly look ideal. A nanotube is a crystalline, rodlike material perfect for building the miniature vibrating structures of interest to us. And because it has no chemical groups projecting outward along its length, one might expect that interaction with "foreign" materials at its surfaces would be minimal. Apparently not. Although nanotubes exhibit ideal characteristics when shrouded within pristine, ultrahigh-vacuum environments, samples in more ordinary conditions, where they are exposed to air or water vapor, evince electronic properties that are markedly different. Mechanical properties are likely to show similar sensitivity. So surfaces definitely *do* matter. It would seem there is no panacea.

Payoff in the Glitches

FUTURISTIC THINKING is crucial to making the big leaps. It gives us some wild and crazy goals—a holy grail to chase. And the hope of glory propels us onward. Yet 19th-century chemist Friedrich August Kekulé once said, "Let us learn to dream, gentlemen, then perhaps we shall find the truth.... But let us beware of publishing our dreams before they have been put to the proof by the waking understanding."

This certainly holds for nanoscience. While we keep our futuristic dreams alive, we also need to keep our expectations realistic. It seems that every time we gain access to a regime that is a factor of 10 different—and presumably "better"—two things happen. First, some wonderful, unanticipated scientific phenomenon emerges. But then a thorny host of underlying, equally unanticipated new problems appear. This pattern has held true as we have pushed to decreased size, enhanced sensitivity, greater spatial resolution, higher magnetic and electric fields, lower pressure and temperature, and so on. It is at the heart of why projecting forward too many orders of magnitude is usually perilous. And it is what should imbue us with a sense of humility and proportion at this, the beginning of our journey. Nature has already set the rules for us. We are out to understand and employ her secrets.

Once we head out on the quest, nature will frequently hand us what initially seems to be nonsensical, disappointing, random gibberish. But the science in the glitches often turns out to be even more significant than the grail motivating the quest. And being proved the fool in this way can truly be the joy of doing science. The delightful truth is that, for complex systems, we do not, and ultimately probably cannot, know everything that is important.

Complex systems are often exquisitely sensitive to a myriad of parameters beyond our ability to sense and record—much less control—with sufficient regularity and precision. Scientists have studied, and in large part already understand, matter down to the fundamental particles that make up the neutrons, protons and electrons that are of crucial importance to chemists, physicists and engineers. But we still cannot predict how complex assemblages of these three elemental components will finally behave en masse. For this reason, I firmly believe that it is on the foundation of the experimental science under way, in intimate collaboration with theory, that we will build the road to true nanotechnology. Let's keep our eyes open for surprises along the way! SA

MORE TO EXPLORE

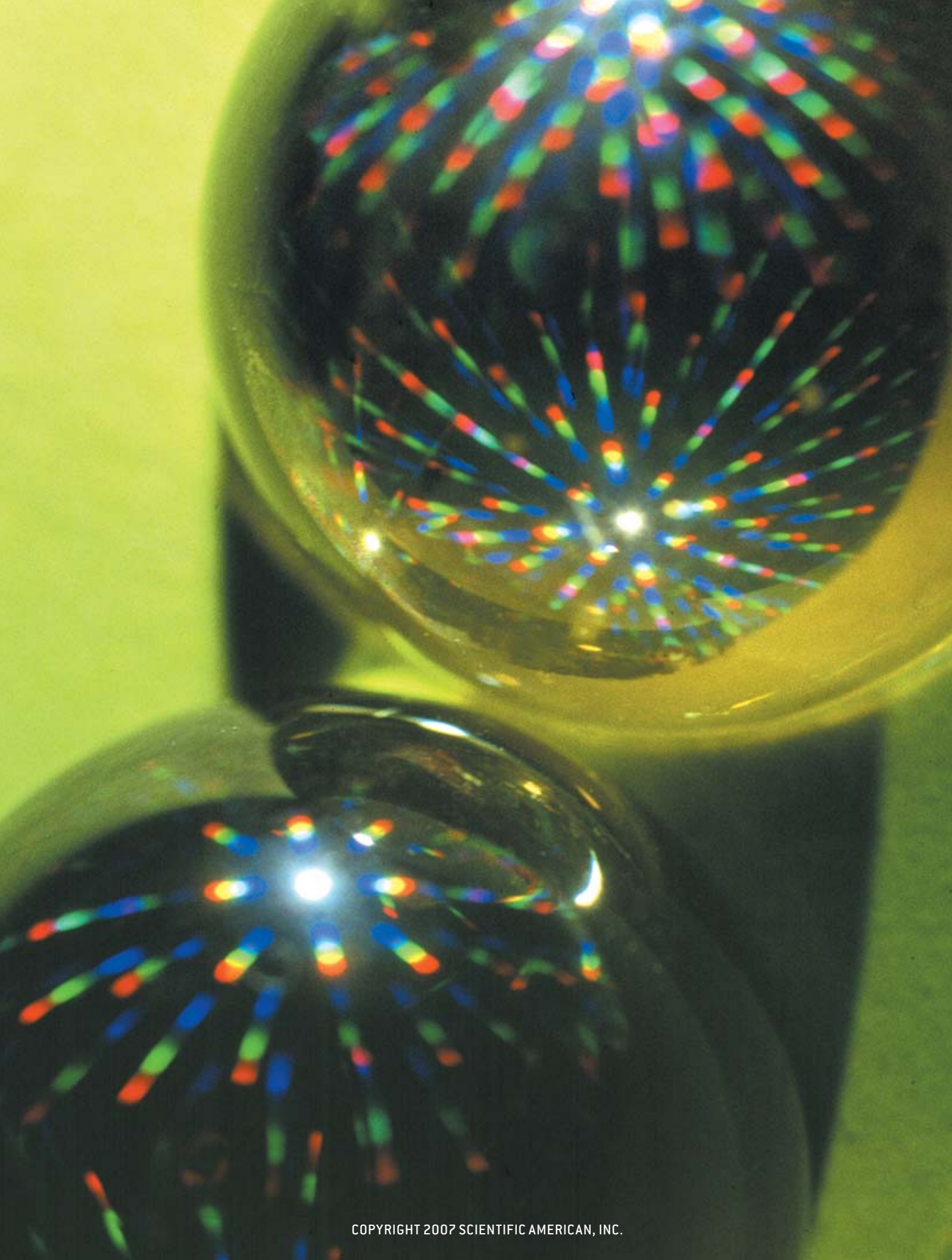
Nanoelectromechanical Systems Face the Future. Michael Roukes in *Physics World,* Vol. 14, No. 2; February 2001. Available at **physicsweb.org/article/world/14/2/8**

Putting Mechanics into Quantum Mechanics. Keith C. Schwab and Michael L. Roukes in *Physics Today,* Vol. 58, No. 7, pages 36–42; July 2005.

The author's group: **http://nano.caltech.edu**

Richard Feynman's original lecture "There's Plenty of Room at the Bottom" can be found at **www.its.caltech.edu/~feynman/plenty.html**

D. A. HARRINGTON AND M. L. ROUKES *Caltech*; COLORIZATION BY M. L. ROUKES (*main panel*); R. B. KARABALIN, X. L. FENG AND M. L. ROUKES *Caltech*; COLORIZATION BY M. L. ROUKES (*inset*)



The art of Building Small

RESEARCHERS ARE DISCOVERING CHEAP, EFFICIENT WAYS TO MAKE STRUCTURES ONLY A FEW BILLIONTHS OF A METER ACROSS

BY GEORGE M. WHITESIDES AND J. CHRISTOPHER LOVE

"Make it small!" is a technological edict that has changed the world. The development of microelectronics—first the transistor and then the aggregation of transistors into microprocessors, memory chips and controllers—has brought forth a cornucopia of machines that manipulate information by streaming electrons through silicon. Microelectronics rests on techniques that routinely fabricate structures smaller than 100 nanometers across (that is, 100 billionths of a meter). This size is tiny by the standards of everyday experience—about one-thousandth the width of a human hair—but it is large on the scale of atoms and molecules. The diameter of a 100-nanometer-wide wire would span about 500 atoms of silicon.

INTRICATE DIFFRACTION PATTERNS are created by nanoscale-width rings (too small to see) on the surface of one-centimeter-wide hemispheres made of clear polymer. Kateri E. Paul, then a graduate student in George M. Whitesides's group at Harvard University, fashioned the rings in a thin layer of gold on the hemispheres using a nanofabrication technique called soft lithography.

The idea of making "nanostructures" that comprise just one or a few atoms has great appeal, both as a scientific challenge and for practical reasons. A structure the size of an atom represents a fundamental limit: to make anything smaller would require manipulating atomic nuclei—essentially, transmuting one chemical element into another. In recent years, scientists have learned various techniques for building nanostructures, but they have only just begun to investigate their properties and potential applications. The age of nanofabrication is here, and the age of nanoscience has dawned, but the age of nanotechnology—finding practical uses for nanostructures—has not really started yet.

The Conventional Approach

RESEARCHERS may well develop nanostructures as electronic components, but the most important applications could be quite different: for example, biologists might use nanometer-scale particles as minuscule sensors to investigate cells. Because scientists do not know what kinds of structures they will ultimately want to build, they have not yet determined the best ways to construct them. Photolithography, the technology used to manufacture computer chips and virtually all other microelectronic systems, has been refined to make structures smaller than 100 nanometers, but the processes are very difficult, expensive and inconvenient. In a search to find better alternatives, nanofabrication researchers have adopted the philosophy "Let a thousand flowers bloom."

First, consider the advantages and disadvantages of photolithography. Manufacturers use this phenomenally productive technology to churn out more than three billion transistors *per second* in the U.S. alone. Photolithography is basically an extension of photography. One first makes the equivalent of a photographic negative containing the pattern required for some part of a microchip's circuitry. This negative, which is called the mask or master, is then used to copy the pattern into the metals and semiconductors of a microchip. As is the case with photography, the negative may be hard to make, but creating multiple copies is easy, because the mask can be used many times. The process thus separates into two stages: the preparation of the mask (a one-time event, which can be slow and expensive) and the use of the mask to manufacture replicas (which must be rapid and inexpensive).

To make a mask for a part of a computer chip, a manufacturer first designs the circuitry pattern on a conveniently large scale and converts it into a pattern of opaque metallic film (usually chromium) on a transparent plate (usually glass or silica). Photolithography then reduces the size of the pattern in a process analogous to that used in a photographic darkroom [*see box on opposite page*]. A beam of light (typically ultraviolet light from a mercury arc lamp) shines through the chromium mask, then passes through a lens that focuses the image onto a photosensitive coating of organic polymer (called the photoresist) on the surface of a silicon wafer. The parts of the photoresist struck by the light can be selectively removed, exposing parts of the silicon wafer in a way that replicates the original pattern.

Why is it challenging to make nanostructures by photolithography? The technology faces two limitations. The first is that the shortest wavelength of ultraviolet light currently used in production processes is about 190 nanometers. Trying to make structures much smaller than half of that spacing is like trying to read print that is too tiny: diffraction causes the features to blur and meld together. Various technical improvements have made it possible to push the limits of photolithography. The smallest structures created in mass production are about 70 nanometers across, and with creative modifications to the optics employed (phase-shifting masks and immersion optics), it is possible to make structures that are only 40 nanometers across. But these structures are still not small enough to explore some of the most interesting aspects of nanoscience.

The second limitation follows from the first: because it is technically difficult to make such small structures using light, it is also very expensive to do so. The photolithographic tools that will be used to make chips with features well below 100 nanometers will each cost tens of millions to hundreds of millions of dollars. This expense may or may not be acceptable to manufacturers, but it is prohibitive for the biologists, materials scientists, chemists and physicists who wish to explore nanoscience using structures of their own design.

Future Nanochips

THE ELECTRONICS industry is deeply interested in developing new methods for nanofabrication so that it can continue its long-term trend of building ever smaller, faster and less expensive devices. It appears that the evolution from microelectronics to nanoelectronics will advance in the near term on incremental modifications to existing photolithographic techniques. But because these adaptations become more

IMAGE BY FELICE FRANKEL, WITH TECHNICAL HELP FROM KATERI E. PAUL; COURTESY OF GEORGE M. WHITESIDES *Harvard University* *(page 12)*

Overview/*Nanofabrication*

- The development of nanotechnology will depend on the ability of researchers to efficiently manufacture structures smaller than 100 nanometers (100 billionths of a meter) across.
- Photolithography, the technology now used to fabricate circuits on microchips, can produce nanometer-scale structures, but the modifications are technically difficult and hugely expensive.
- Nanofabrication methods can be divided into two categories: top-down methods, which carve out or add aggregates of molecules to a surface, and bottom-up methods, which assemble atoms or molecules into nanostructures.
- Two examples of promising top-down methods are soft lithography and dip-pen lithography. Researchers are using bottom-up methods to produce quantum dots that can serve as biological dyes.

difficult as the dimensions of the structures become smaller, manufacturers are exploring alternative technologies for making future nanochips.

One leading contender is electron-beam lithography. In this method, the circuitry pattern is written on a thin polymer film with a beam of electrons. An electron beam does not diffract at atomic scales, so it does not cause blurring of the edges of features. Researchers have used the technique to write lines with widths of only a few nanometers in a layer of photoresist on a silicon substrate. The electron-beam instruments currently available, however, are very expensive and impractical for large-scale manufacturing. Because the beam of electrons is needed to fabricate each structure, the process is similar to the copying of a manuscript by hand, one line at a time.

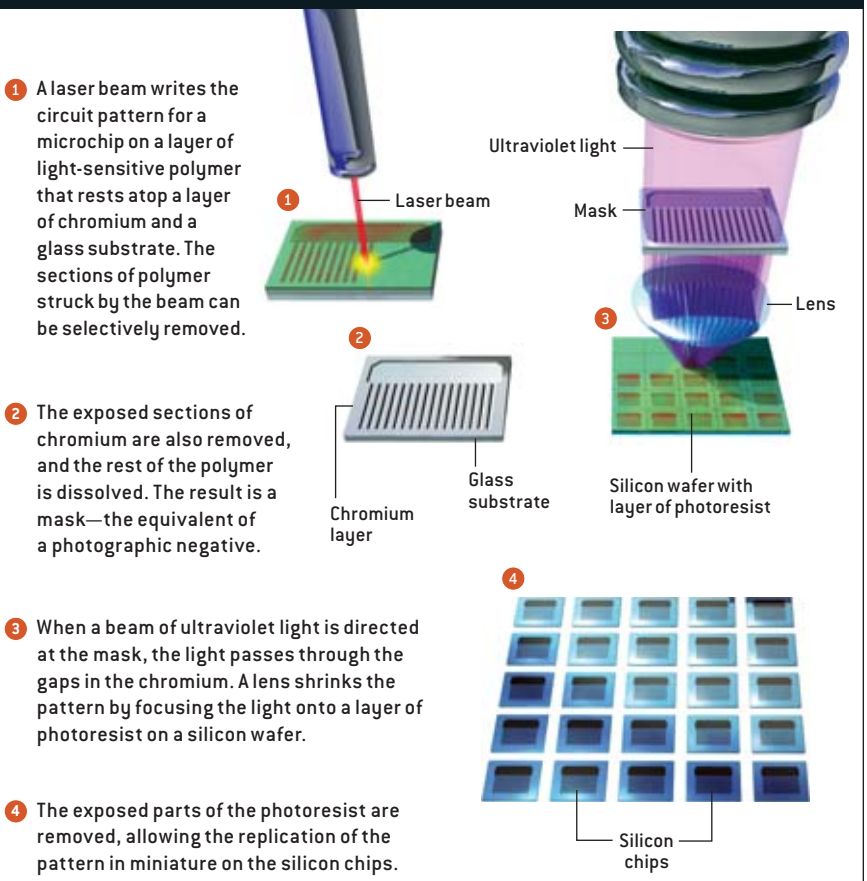
If electrons are not the answer, what is? Another contender is lithography using x-rays with wavelengths between 0.1 and 10 nanometers or extreme ultraviolet light with wavelengths between 10 and 70 nanometers. Because these forms of radiation have much shorter wavelengths than the ultraviolet light currently used in photolithography, they minimize the blurring caused by diffraction. These technologies face their own set of problems, however: conventional lenses are not transparent to extreme ultraviolet light and do not focus x-rays. Furthermore, the energetic radiation rapidly damages many of the materials used in masks and lenses. But the microelectronics industry clearly would prefer to make advanced chips using extensions of familiar technology, so these methods are being actively developed. Some of the techniques (for example, advanced ultraviolet lithography for chip production) will probably become commercial realities. They will not, though, make inexpensive nanostructures and thus will do nothing to open nanotechnology to a broader group of scientists and engineers.

The need for simpler and less expensive methods of fabricating nanostructures has stimulated the search for unconventional approaches that have not been explored by the electronics industry. We first became interested in the topic in the 1990s when we were engaged in making the simple structures required in microfluidic systems—chips with channels and chambers for holding liquids. This lab-on-a-chip has myriad potential uses in biochemistry, ranging from drug screening to genetic analysis. The channels in microfluidic chips are enormous by the standards of microelectronics: 50 microns (or 50,000 nanometers) wide, rather than 100 nanometers. But the techniques for producing those channels are quite versatile. Microfluidic chips can be made quickly and inexpensively, and many are composed of organic polymers and gels—materials not found in the world of electronics. We discovered that we could use similar techniques to create nanostructures.

The methods represented, in a sense, a step backward in technology. Instead of using the tools of physics—light and electrons—we employed mechanical processes that are familiar in everyday life: printing, stamping, molding, embossing and cutting. The techniques are called soft lithography because the tool they have in common is a block of polydimethylsiloxane (PDMS)—the rubbery polymer used to caulk the leaks around bathtubs. (Physicists often refer to such organic chemicals as "soft matter.")

To carry out reproduction using soft lithography, one first makes a mold or a

CONVENTIONAL PHOTOLITHOGRAPHY



BRYAN CHRISTIE DESIGN

THE AUTHORS

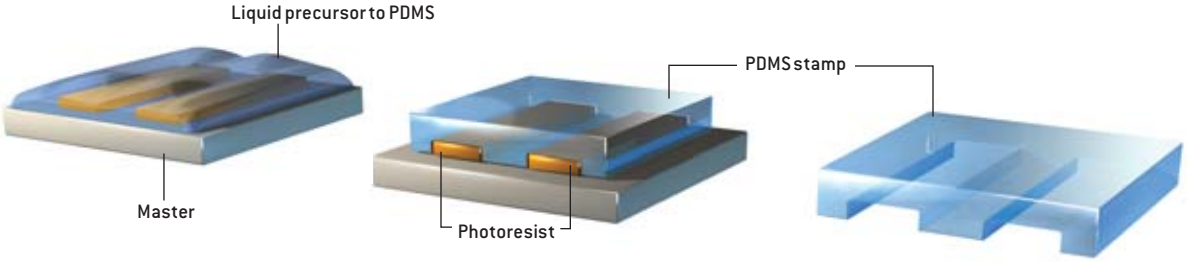
GEORGE M. WHITESIDES and *J. CHRISTOPHER LOVE* collaborate on developing unconventional methods of nanofabrication. Whitesides, a professor of chemistry at Harvard University, received his Ph.D. from the California Institute of Technology in 1964 and joined the Harvard faculty in 1982. Love is an assistant professor of chemical engineering at the Massachusetts Institute of Technology. He received his Ph.D. in chemistry from Harvard in 2004 under the guidance of Whitesides.

SOFT LITHOGRAPHY

Printing, molding and other mechanical processes carried out using an elastic stamp can produce patterns with nanoscale features. Such techniques can fabricate devices that might be used in optical communications or biochemical research.

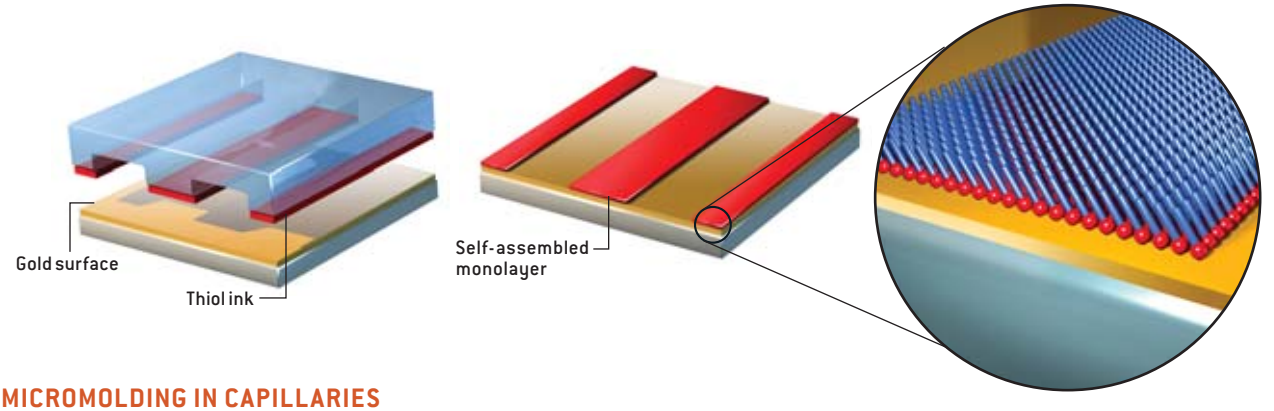
MAKING AN ELASTIC STAMP

1. A liquid precursor to polydimethylsiloxane (PDMS) is poured over a bas-relief master produced by photolithography or electron-beam lithography.
2. The liquid is cured into a rubbery solid that matches the original pattern.
3. The PDMS stamp is peeled off the master.

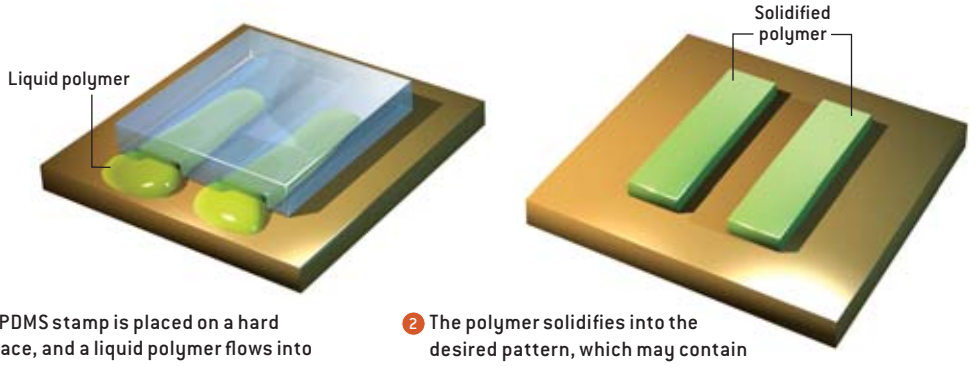


MICROCONTACT PRINTING

1. The PDMS stamp is inked with a solution consisting of organic molecules called thiols and then pressed against a thin film of gold on a silicon plate.
2. The thiols form a self-assembled monolayer on the gold surface that reproduces the stamp's pattern; features in the pattern are as small as 50 nanometers.



MICROMOLDING IN CAPILLARIES



1. The PDMS stamp is placed on a hard surface, and a liquid polymer flows into the recesses between the surface and the stamp.
2. The polymer solidifies into the desired pattern, which may contain features smaller than 10 nanometers.

BRYAN CHRISTIE DESIGN

stamp. The most prevalent procedure is to use photolithography or electron-beam lithography to produce a pattern in a layer of photoresist on the surface of a silicon wafer. This process generates a bas-relief master in which islands of photoresist stand out from the silicon [*see top illustration in box on opposite page*]. Then a chemical precursor to PDMS—a free-flowing liquid—is poured over the bas-relief master and cured into the rubbery solid. The result is a PDMS stamp that matches the original pattern with astonishing fidelity: the stamp reproduces features from the master as small as a few nanometers. Although the creation of a finely detailed bas-relief master is expensive because it requires electron-beam lithography or other advanced techniques, copying the pattern on PDMS stamps is cheap and easy. And once a stamp is in hand, it can be used in various inexpensive ways to make nanostructures.

The first method—originally developed by Amit Kumar, then a postdoctoral student in our group at Harvard University—is called microcontact printing. The PDMS stamp is "inked" with a reagent solution consisting of organic molecules called thiols [*see middle illustration in box on opposite page*]. The stamp is then brought into contact with an appropriate sheet of "paper"—a thin film of gold on a glass, silicon or polymer plate. The thiols react with the gold surface, forming a highly ordered film (a self-assembled monolayer, or SAM) that replicates the stamp's pattern. Because the thiol ink spreads a bit after it contacts the surface, the resolution of the monolayer cannot be quite as high as that of the PDMS stamp. But when used correctly, microcontact printing can produce patterns with features as small as 50 nanometers.

In another method of soft lithography, called micromolding in capillaries, the PDMS stamp is used to mold patterns. The stamp is placed on a hard surface, and a liquid polymer flows by capillary action into the recesses between the surface and the stamp [*see bottom illustration in box on opposite page*]. The polymer then solidifies into the desired pattern. This technique can replicate structures smaller than 10 nanometers.

A related extension of micromolding, called nanoskiving, produces arrays of metallic nanostructures by cutting cross sections of the molded patterns. A PDMS stamp is first used to mold a hard plastic such as an epoxy—with cylindrical posts, for example. The molded epoxy is then coated with a thin (about 50 nanometers) film of metal and covered with more epoxy. Cutting parallel to the plane of this sandwiched structure, like slicing meat at the deli counter, creates a thin sheet of plastic containing nanostructures shaped like the cross section of the original molded structures—for cylindrical posts, the resulting shapes are rings. The thickness of the slice determines their height and the deposited film their thickness. These techniques are particularly well suited for producing subwavelength optical devices, waveguides and optical polarizers, all of which could be used in optical fiber networks and eventually perhaps in optical computers. Other possible applications are in the field of nanofluidics, an extension of microfluidics that would involve producing chips for biochemical research with channels only a few nanometers wide. At that scale, fluid dynamics may allow new ways to separate materials such as fragments of DNA.

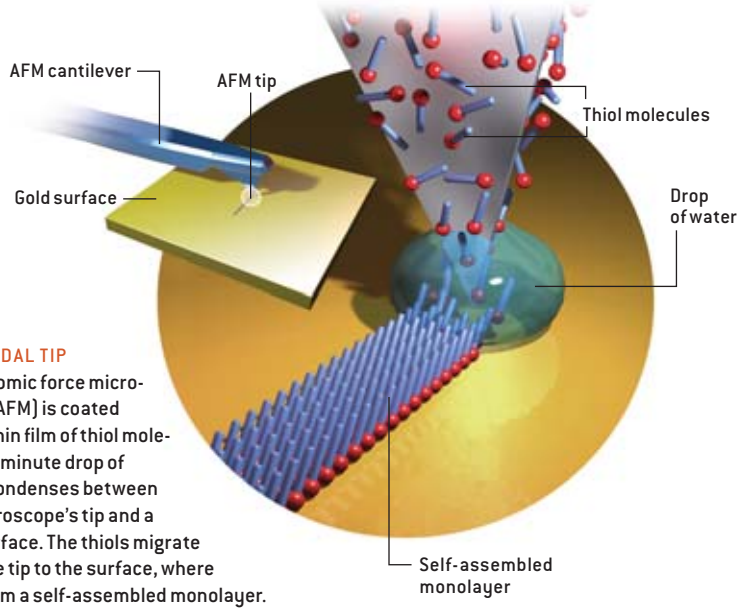
The majority of these methods require no special equipment and in fact can be carried out by hand in an ordinary laboratory. Conventional photolithography must take place in a cleanroom facility devoid of dust and dirt; if a piece of dust lands on the mask, it will create an unwanted spot on the pattern. As a result, the device being fabricated (and sometimes neighboring devices) may fail. Soft lithography is generally more forgiving because the PDMS stamp is elastic. If a piece of dust gets trapped between the stamp and the surface, the stamp will compress over the top of the particle but maintain contact with the rest of the surface. Thus, the pattern will be reproduced correctly except for where the contaminant is trapped.

Moreover, soft lithography can produce nanostructures in a wide range of materials, including the complex organic molecules needed for biological studies. And the technique can print or mold patterns on curved as well as planar surfaces. But the technology is not ideal for making the structures required for complex nanoelectronics. Currently all integrated circuits consist of stacked layers of different materials. Deformations and distortions of the soft PDMS stamp can produce small errors in the replicated pattern and a misalignment of the pattern with any underlying patterns previously fabricated. Even the tiniest distortions or misalignments can destroy a multilayered nanoelectronic device. Therefore, soft lithography is not well suited for fabricating structures with multiple layers that must stack precisely on top of one another.

Researchers have found ways, however, to correct this shortcoming—at least in part—by employing a rigid stamp instead of an elastic one. In a technique called step-and-flash imprint lithography, developed by C. Grant Willson of the University of Texas at Austin, photolithography is used to etch a pattern into a quartz plate, yielding a rigid bas-relief master. Willson eliminated the step of making a PDMS stamp from the master; instead the master itself is pressed against a thin film of liquid polymer, which fills the master's

These methods require no special equipment and in fact can be carried out by hand in an ordinary lab.

DIP-PEN LITHOGRAPHY



PYRAMIDAL TIP of an atomic force microscope (AFM) is coated with a thin film of thiol molecules. A minute drop of water condenses between the microscope's tip and a gold surface. The thiols migrate from the tip to the surface, where they form a self-assembled monolayer.

recesses. Then the master is exposed to ultraviolet light, which solidifies the polymer to create the desired replica. A related technique called nanoimprint lithography, developed by Stephen Y. Chou of Princeton University, also employs a rigid master but uses a film of polymer that has been heated to a temperature near its melting point to facilitate the embossing process. Both methods can produce two-dimensional structures as small as 20 nanometers with good fidelity, and it appears likely that they will contend with photolithography for manufacturing next-generation ultrabright LEDs, flash drives and disk drives. One significant advantage for imprint lithography over standard photolithography is the ability to pattern three-dimensional topographies in a single step: such topography is critical for connecting different layers on integrated circuits, and imprint methods should save many steps (and thus costs) during the fabrication of microchips.

Pushing Atoms Around

THE CURRENT REVOLUTION in nanoscience started in 1981 with the invention of the scanning tunneling microscope (STM), for which Heinrich Rohrer and Gerd K. Binnig of the IBM Zurich Research Laboratory received the Nobel Prize in Physics in 1986. This remarkable device detects small currents that pass between the microscope's tip and the sample being observed, allowing researchers to "see" substances at the scale of individual atoms. The success of the STM led to the development of other scanning probe devices, including the atomic force microscope (AFM). The operating principle of the AFM is similar to that of an old-fashioned phonograph. A tiny probe—a fiber or a pyramid-shaped tip that is typically between two and 30 nanometers wide—is brought into direct contact with the sample. The probe is attached to the end of a cantilever, which bends as the tip moves across the sample's surface. The deflection is measured by reflecting a beam of laser light off the top of the cantilever. The AFM can detect variations in vertical surface topography that are smaller than the dimensions of the probe.

But scanning probe devices can do more than simply allow scientists to observe the atomic world—they can also be used to create nanostructures. The tip on the AFM can be used to physically move nanoparticles around on surfaces and to arrange them in patterns. It can also be used to make scratches in a surface (or, more commonly, in monolayer films of atoms or molecules that coat the surface). Similarly, if researchers increase the currents flowing from the tip of the STM, the microscope becomes a very small source for an electron beam, which can be used to write nanometer-scale patterns. The STM tip can also push individual atoms around on a surface to build rings and wires that are only one atom wide.

An intriguing new scanning probe fabrication method is called dip-pen lithography. Developed by Chad A. Mirkin of Northwestern University, this technique works much like a goose-feather pen [*see box at left*]. The tip of the AFM is coated with a thin film of thiol molecules that are insoluble in water but react with a gold surface (the same chemistry used in microcontact printing). When the device is placed in an atmosphere containing a high concentration of water vapor, a minute drop of water condenses between the gold surface and the microscope's tip. Surface tension pulls the tip to a fixed distance from the gold, and this distance does not change as the tip moves across the surface. The drop of water acts as a bridge over which the thiol molecules migrate from the tip to the gold surface, where they are fixed. Researchers have used this procedure to write lines a few nanometers across.

Although dip-pen lithography is relatively slow, it can use many different types of molecules as "inks" and thus brings great chemical flexibility to nanometer-scale writing. Parallel arrays of independent scanning probes have improved the throughput of dip-pen lithography significantly and may catalyze successful commercialization of the method. Researchers have not yet determined the best applications for the technique, but two ideas being pursued are the precise repair of damaged photomasks or electrical circuits and the creation of anticounterfeit labels for pharmaceuticals or other products.

An interesting cousin to these techniques involves another kind of nanostructure, called a break junction. If you break a thin, ductile metal wire into two parts by pulling sharply, the process seems abrupt to a human observer,

BRYAN CHRISTIE DESIGN

but it actually follows a complex sequence. When the force used in breaking the wire is first applied, the metal begins to yield and flow, and the diameter of the wire decreases. As the two ends move apart, the wire gets thinner and thinner until, in the instant just before breaking, it is a single atom in diameter at its narrowest point. This process of thinning a wire to a break junction can be detected easily by measuring the current that flows through the wire. When the wire is slender enough, current can flow only in discrete quantities (that is, current flow is quantized).

The break junction is analogous to two STM tips facing each other, and similar physical rules govern the current that flows through it. Mark A. Reed of Yale University has pioneered a particularly inventive use of the break junction. He built a device that enabled a thin junction to be broken under carefully controlled conditions and then allowed the broken tips to be brought back together or to be held apart at any distance with an accuracy of a few thousandths of a nanometer. By adjusting the distance between the tips in the presence of an organic molecule that bridged them, Reed was able to measure a current flowing across the organic bridge. This experiment was an important step in the development of technologies for using single organic molecules as electronic devices such as diodes and transistors.

Top-Down and Bottom-Up

ALL THE FORMS of lithography we have discussed so far are called top-down methods—that is, they begin with a pattern generated on a larger scale and reduce its lateral dimensions (often by a factor of 10) before carving out nanostructures. This strategy is required in fabricating electronic devices such as microchips, whose functions depend more on their patterns than on their dimensions. But no top-down method is ideal; none can conveniently, cheaply and quickly make nanostructures of any material. So researchers have shown growing interest in bottom-up methods, which start with atoms or molecules and build *up* to nanostructures. These methods can easily make the smallest nanostructures—with dimensions between two and 10 nanometers—and do so inexpensively. But these structures are usually generated as simple particles in suspension or on surfaces, rather than as designed, interconnected patterns.

Two of the most prominent bottom-up methods are those used to make nanotubes and quantum dots. Scientists have made long, cylindrical tubes of carbon by a catalytic growth process that employs a nanometer-scale drop of molten metal (usually iron) as a catalyst. The most active area of research in quantum dots originated in the laboratory of Louis E. Brus (then at Bell Laboratories) and has been developed by A. Paul Alivisatos of the University of California, Berkeley, Moungi G. Bawendi of the Massachusetts Institute of Technology, and others. Quantum dots are crystals containing only a few hundred atoms. Because the electrons in a quantum dot are confined to widely separated energy levels, the dot emits only one wavelength of light when it is excited. This property makes the quantum dot useful as a biological marker.

Nanofabrication: Comparing the Methods

Researchers are developing an array of techniques for building structures smaller than 100 nanometers. Here is a summary of the advantages and disadvantages of four methods.

Photolithography

Advantages: The electronics industry is already familiar with this technology because it is currently used to fabricate microchips. Manufacturers can modify the technique to produce nanometer-scale structures by employing electron beams, x-rays or extreme ultraviolet light.

Disadvantages: The necessary modifications will be expensive and technically difficult. Using electron beams to fashion structures is costly and slow. X-rays and extreme ultraviolet light can damage the equipment used in the process.

Soft Lithography

Advantages: This method allows researchers to inexpensively reproduce patterns created by electron-beam lithography or other related techniques. Soft lithography requires no special equipment and can be carried out by hand in an ordinary laboratory.

Disadvantages: The technique is not ideal for manufacturing the multilayered structures of electronic devices. Researchers are trying to overcome this drawback, but it remains to be seen whether these efforts will be successful.

Scanning Probe Methods

Advantages: The scanning tunneling microscope and the atomic force microscope can be used to move individual nanoparticles and arrange them in patterns. The instruments can build rings and wires that are only one atom wide.

Disadvantages: The methods are too slow for mass production. Applications of the microscopes will probably be limited to the fabrication of specialized devices.

Bottom-Up Methods

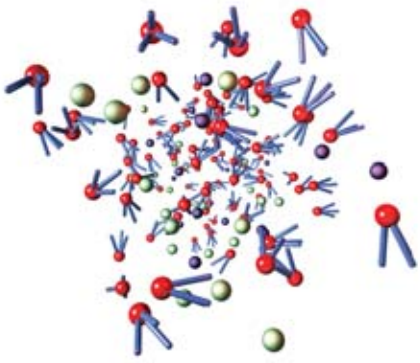
Advantages: By setting up carefully controlled chemical reactions, researchers can cheaply and easily assemble atoms and molecules into the smallest nanostructures, with dimensions between two and 10 nanometers.

Disadvantages: Because these methods cannot produce designed, interconnected patterns, they are not well suited for building electronic devices such as microchips.

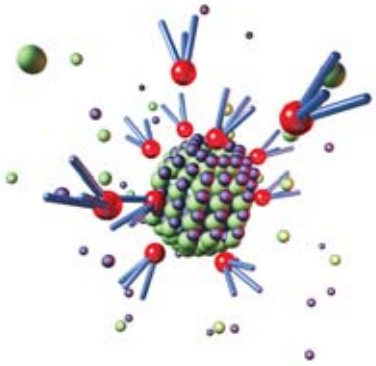
QUANTUM DOT ASSEMBLY

Crystals called quantum dots contain only a few hundred atoms and emit different wavelengths of light depending on their size. They may become useful as biological markers of cellular activity.

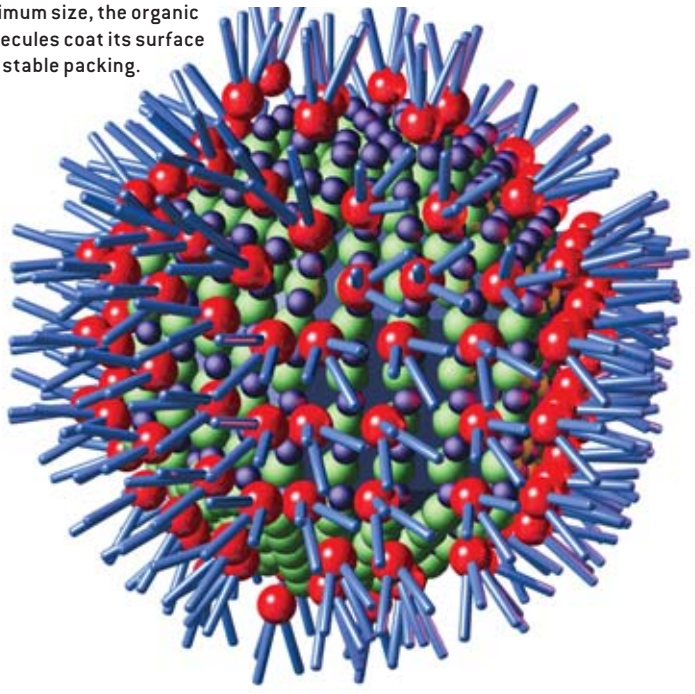
1 A chemical reaction brings together cadmium ions (*purple*), selenium ions (*green*) and organic molecules (*red spheres with blue tails*).



2 The organic molecules act as surfactants, binding to the surface of the cadmium selenide crystal as it grows.



3 When the crystal reaches its optimum size, the organic molecules coat its surface in a stable packing.



BRYAN CHRISTIE DESIGN

One procedure used to make quantum dots involves a chemical reaction between a metal ion (for example, cadmium) and a molecule that is able to donate a selenium ion. This reaction generates crystals of cadmium selenide. The trick is to prevent the small crystals from sticking together as they grow to the desired size. To insulate the growing particles from one another, researchers carry out the reaction in the presence of organic molecules that act as surfactants, coating the surface of each cadmium selenide particle as it grows. The organic molecules stop the crystals from clumping together and regulate their rate of growth. The geometry of the particles can be controlled to some extent by mixing different ratios of the organic molecules. The reaction can generate particles with a variety of shapes, including spheres, rods and tetrapods (four-armed particles similar to toy jacks).

It is important to synthesize the quantum dots with uniform size and composition, because the size of the dot determines its electronic, magnetic and optical properties. Researchers can select the size of the particles by varying the length of time for the reaction. The organic coating also helps to set the size of the particles. When the nanoparticle is small (on the scale of molecules), the organic coating is loose and allows further growth; as the particle enlarges, the organic molecules become crowded. There is an optimum size for the particles that allows the most stable packing of the organic molecules and thus provides the greatest stabilization for the surfaces of the crystals.

These cadmium selenide nanoparticles promise some of the first commercial products of nanoscience: Quantum Dot Corporation (now Invitrogen) and Evident Technologies have been developing the crystals for use as biological labels. Researchers can tag proteins and nucleic acids with quantum dots; when the sample is illuminated with ultraviolet light, the crystals will fluoresce at a specific wavelength and thus show the locations of the attached proteins. Many organic molecules also fluoresce, but quantum dots have several advantages that make them better markers. First, the color of a quantum dot's fluorescence can be tailored by changing the dot's size: the larger the particle, the more the emitted light is shifted toward the red end of the spectrum. Second, if all the dots are the same size, their fluorescence spectrum is narrow—that is, they emit a very pure color. This property is important because it allows particles of different sizes to be used as distin-

guishable labels. Third, the fluorescence of quantum dots does not fade on exposure to ultraviolet light, as does that of organic molecules. When used as dyes in biological research, the dots can be observed for conveniently long periods.

Scientists are also investigating the possibility of making structures from colloids—nanoparticles in suspension. Christopher B. Murray, now at the University of Pennsylvania, and a team at the IBM Thomas J. Watson Research Center explored the use of such colloids to create a medium for ultrahigh-density data storage. The IBM team's colloids contained magnetic nanoparticles as small as three nanometers across, each composed of about 1,000 iron and platinum atoms. When the colloid is spread on a surface and the solvent allowed to evaporate, the nanoparticles crystallize in two- or three-dimensional arrays. Studies indicate that these arrays can potentially store trillions of bits of data per square inch, giving them a capacity 10 to 100 times greater than that of present memory devices.

The Future of Nanofabrication

THE INTEREST in nanostructures is so great that every plausible fabrication technique is being examined. Although physicists and chemists are now doing most of the work, biologists may also make important contributions. The cell (whether mammalian or bacterial) is relatively large on the scale of nanostructures: the typical bacterium is approximately 1,000 nanometers long, and mammalian cells are larger. Cells are, however, filled with much smaller structures, many of which are astonishingly sophisticated. The ribosome, for example, carries out one of the most important cellular functions: the synthesis of proteins from amino acids, using messenger RNA as the template. The complexity of this molecular construction project far surpasses that of man-made techniques.

It is unclear whether "nanomachines" taken from cells will be useful. They will probably have very limited application in electronics, but they may provide valuable tools for chemical synthesis and sensing devices. Work by Carlo D. Montemagno of the University of Cincinnati showed that it is possible to engineer a primitive nanomachine with a biological engine. Montemagno extracted a rotary motor protein from a bacterial cell and connected it to a metallic nanorod—a cylinder 750 nanometers long and 150 nanometers wide that had been fabricated by lithography. The rotary motor, which was only 11 nanometers tall, was powered by adenosine triphosphate (ATP), the source of chemical energy in cells. Montemagno showed that the motor could rotate the nanorod at eight revolutions per minute. At the very least, such research stimulates efforts to fabricate functional nanostructures by demonstrating that such structures can exist.

The development of nanotechnology will depend on the availability of nanostructures. The invention of the STM and AFM has provided new tools for viewing, characterizing and manipulating these structures; the issue now is how to build them to order and how to design them to have new and useful functions. The importance of electronics applications has tended to focus attention on nanodevices that might be incorporated into future integrated circuits. And for good technological reasons, the electronics industry has emphasized fabrication methods that are extensions of those currently used to make microchips. But the explosion of interest in nanoscience has created a demand for a broad range of fabrication methods, with an emphasis on low-cost, convenient techniques.

The new approaches to nanofabrication are unconventional only because they are not derived from the microtechnology developed for electronic devices. Chemists, physicists and biologists are rapidly accepting these techniques as the most appropriate ways to build various kinds of nanostructures for research. And the methods may even supplement the conventional approaches—photolithography, electron-beam lithography and related techniques—for applications in electronics as well. The microelectronics mold is now broken. Ideas for nanofabrication are coming from many directions in a wonderful free-for-all of discovery. SA

Bottom-up methods start with atoms or molecules and build *up* to nanostructures.

MORE TO EXPLORE

Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. J. C. Love, L. A. Estroff, J. K. Kriebel, R. G. Nuzzo and G. M. Whitesides in *Chemical Reviews,* Vol. 105, No. 4, pages 1103–1170; April 2005.

New Approaches to Nanofabrication: Molding, Printing, and Other Techniques. B. D. Gates, Q. B. Xu, M. Stewart, D. Ryan, C. G. Willson and G. M. Whitesides in *Chemical Reviews,* Vol. 105, No. 4, pages 1171–1196; April 2005.

More information about nanofabrication can be found at the following Web sites:

The Bawendi group at M.I.T.: http://web.mit.edu/chemistry/nanocluster/

International SEMATECH: www.sematech.org

The Mirkin group at Northwestern University: http://chemgroups.northwestern.edu/mirkingroup

The Whitesides group at Harvard University: http://gmwgroup.harvard.edu

The Willson group at the University of Texas at Austin: http://willson.cm.utexas.edu

MOLECULAR LEGO

A modest collection of small building blocks enables the design and manufacture of nanometer-scale structures programmed to have virtually any shape desired

By Christian E. Schafmeister

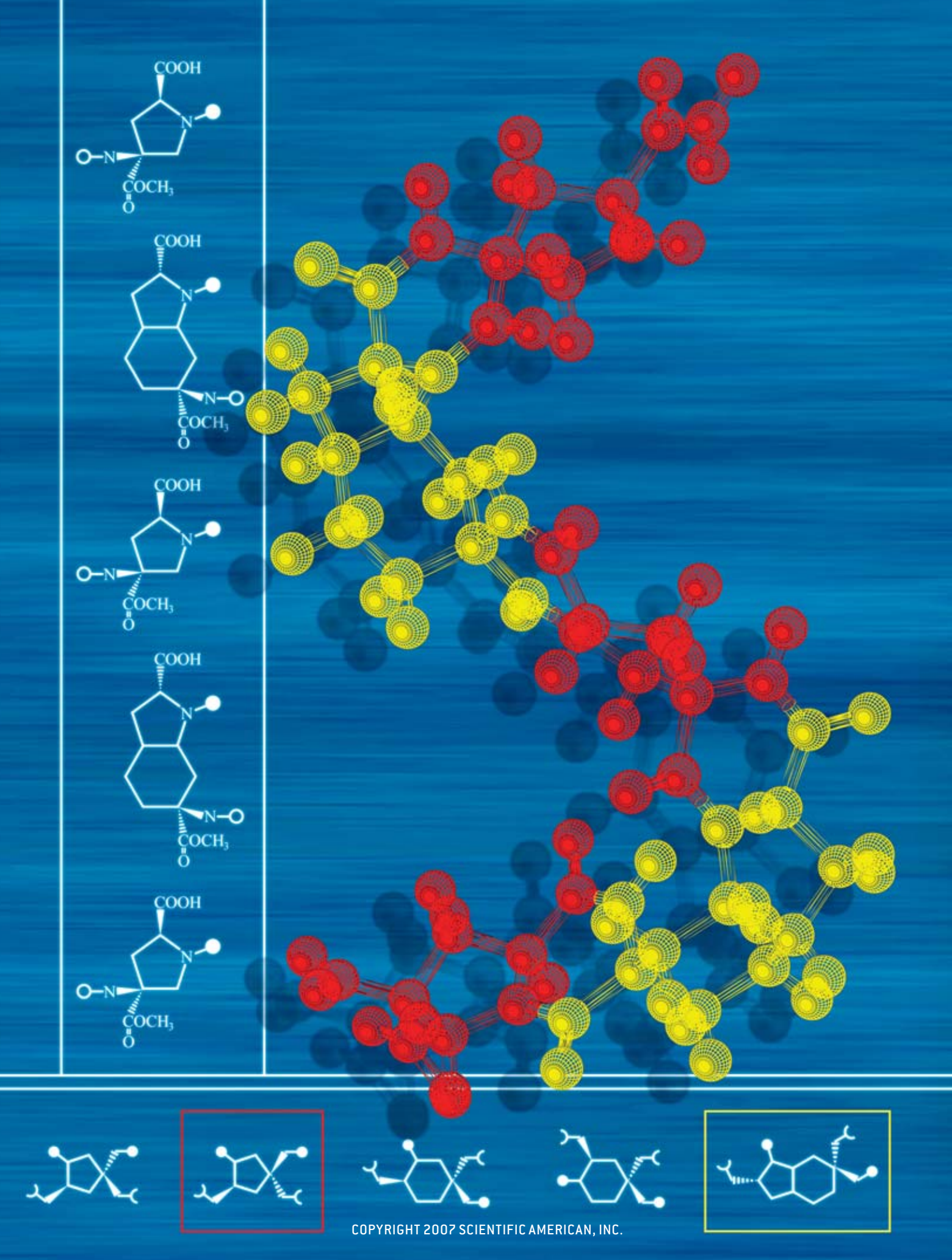
Proteins, the fundamental nanomachines of life, have provided scientists like me with many lessons in our own efforts to create nanomachinery. Proteins are large molecules containing hundreds to thousands of atoms and are typically a few nanometers (billionths of a meter) to tens of nanometers across. Our bodies contain at least 20,000 different proteins that, among other things, cause our muscles to contract, digest our food, build our bones, sense our environment and tirelessly recycle hundreds of small molecules within our cells.

As a chemistry undergraduate in 1986, I dreamed of the possibility of designing and synthesizing macromolecules (molecules containing more than 100 atoms) that could do the amazing things that proteins do and more. I have programmed computers since the first TRS-80s came out in the late 1970s, and I thought it would be wonderful if I could build complex molecular machines as easily as I could write software. I wanted to create a "programming language for matter"—a combination of software and chemistry that would enable people to describe a nanomachine's shape and would then determine the series of chemical processes that a chemist or a robot should carry out to build the nanodevice.

Unfortunately, the idea of inventing nanomachines by designing new proteins runs into a severe obstacle. Every protein generally starts as a simple, linear chain assembled from a specific sequence of amino acids drawn from a repertoire of just 20 amino acids. So far, so good, but the properties of a protein and what functions it can carry out depend on its shape. Shortly after the chain of amino acids is put together in the cell, it collapses into an intricate tangle of helices and other structures through a complex process called protein folding. The sequence of amino acids determines the final shape, but predicting what shape a particular sequence will take on is one of the most significant unsolved challenges of science and engineering (the "protein folding problem").

CREATING NEW NANOSTRUCTURES similar to proteins is made practical with a collection of building-block molecules (*bottom*) developed to join together and form rigid structures whose overall shapes are completely preplanned by the designer, like a model made of tiny, oddly shaped Lego bricks.

KEN EWARD *BioGrafx*



Some 20 years after I first entertained my vision of the future, my laboratory has at last developed a way to produce large molecules with programmable shapes and the computer software required to design them. Our approach is inspired by the modularity of natural proteins, but it does not rely on amino acid chains to collapse spontaneously into a shape—so it avoids contending with the unsolved folding problem.

We are developing this technology to create molecules that can carry out specific functions. One of our initial goals is to create sensors: large molecules that change shape and color when they bind to particular target molecules, such as glucose, toxins or chemical warfare agents. The binding event triggers the sensor molecule to swing two fluorescent groups together that alter its color, thereby signaling that the target is present in the sample. We are also using our technique to create long, hinged molecules that open and close in response to an external signal—a step toward the creation of molecular actuators, molecular valves and computer memories.

We envisage that our technique will ultimately lead to an even more advanced method of constructing nanomachines: we would use it to fashion complex nanotools such as an assembler that, like the ribosome responsible for constructing proteins inside cells, would assemble other nanomachines under external programmer control. For now, this second dream lies in the future.

Lessons from Nature

WHEN I FINISHED my undergraduate studies in 1990, I thought that the path to developing nanomachinery lay in deducing the rules of protein folding and using them to develop new proteins. I joined Robert M. Stroud and his protein crystallography group at the University of California, San Francisco. Protein crystallographers grow crystals of proteins and use x-rays to determine the exact three-dimensional arrangement of the proteins' atoms. Using this tool, I developed a deep appreciation of the complexity and beauty of protein structure. I spent four years creating 4HB1, an artificial protein of my own design. I first assembled an artificial gene and then inserted it into bacteria, which "expressed" it—that is, made the protein encoded by the gene's DNA. Next I crystallized the resultant protein and determined its x-ray crystal structure. It was thrilling to discover that 4HB1 had the conformation I had designed it to have!

Yet after all this work, 4HB1 was a molecular doorstop. It did not do anything other than exist as a well-folded artificial protein. Most disturbing was that the experience did not reveal the simple rules I needed to create other proteins of a desired shape. On the contrary, the complexity of protein folding suggested that such simple rules might not exist. While finishing my Ph.D. in 1997, I concluded that a better way to create custom-designed nanomachinery would be to construct them from a limited set of modular building blocks that did not attain their shape via the folding process of proteins.

This was not a new idea. In 1995 Brent Iverson of the University of Texas at Austin had developed building blocks that could be chained together into short polymers called oligomers. These oligomers then self-assembled into pleated structures as electron-rich donor groups pulled on electron-deficient acceptor groups in the structure.

At about the same time, Sam Gellman of the University of Wisconsin–Madison and Dieter Seebach of the Swiss Federal Institute of Technology in Zurich were developing synthetic molecules called beta-peptides, which are flexible chains of beta-amino acids—molecules that are mostly not naturally occurring and whose general structure is slightly different from that of regular amino acids (alpha-amino acids). Gellman and Seebach's short beta-peptides fold into twisted helices.

These new approaches to constructing macromolecules that held a specific shape were inspiring, but they seemed to trade one folding problem for another. The difficulty is that natural proteins and these new molecules involve chains of molecules connected by single bonds that leave the structure with a lot of freedom to bend at locations all along its length. Which way one of these molecules bends in acquiring its final shape depends on the complex interplay of attractive and repulsive forces arising when different building blocks all along the chain are brought closer together.

I had a more radical approach in mind. I wanted to eliminate the usual folding process altogether and thus gain more control over the shape of the final product. To achieve this goal, I set out to invent rigid building blocks that could be attached to one another through *pairs* of bonds to create rigid, ladderlike macromolecules. This idea had been tried before: in 1987 J. Fraser Stoddart, then at the University of Sheffield in England, introduced the concept of a "molecular Lego set" by creating molecular belts and collars from building blocks.

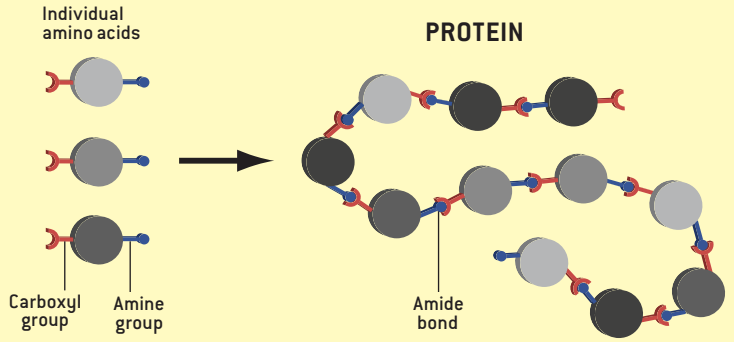
Overview/*Nano Lego*

- Proteins are nature's nanomachines, tirelessly carrying out myriad biological tasks. Because proteins are made of flexible chains of amino acids that fold up in a very complicated way, scientists cannot easily predict a new protein's shape (and hence its function).
- Now chemists have developed a library of novel molecular building blocks, called bis-amino acids, that can be strung together to form proteinlike structures that have rigid, readily predictable and designable shapes.
- Potential applications for these "bis-peptides" include medicines, enzymes for catalyzing useful reactions, chemical sensors, nanoscale valves and computer storage devices.

HOW BIS-PEPTIDES DIFFER FROM PROTEINS

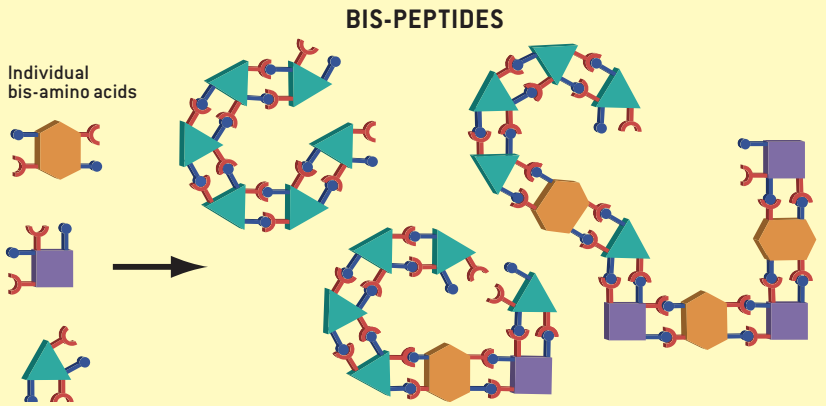
NATURAL PROTEINS

Organisms make 20 different amino acids that they string together into flexible chains that are generally called peptides when they are short and proteins when they are long. The amino acids are joined by amide bonds, which form when carboxyl and amine groups react together. The final shape of a protein depends on the complex interplay of interactions occurring among amino acids all along its length. This complexity makes it extremely hard to predict what shape a new amino acid sequence will take on. (All figures are highly schematic for clarity.)



PREDICTABLE BIS-PEPTIDES

Chemists have produced a library of building blocks called bis-amino acids that sport two pairs of carboxyls and amines. When linked together, these building blocks, or monomers, form a rigid chain called a bis-peptide that has a predictable shape directly determined by the sequence of bis-amino acids selected. Therefore, chemists can design and build precise nanostructures simply by combining bis-amino acids in a specific order.



I joined the laboratory of Gregory Verdine at Harvard University to learn synthetic organic chemistry. During two years of synthesizing unnatural amino acids and searching for a route to my larger vision, I came across a paper that described a chemical structure called a diketopiperazine. In this structure, six atoms join into a ring containing two amide bonds [*see box on next page*]. Amide bonds are the ones that link a protein's constituent amino acids together in a chain, like a line of people holding hands. A diketopiperazine arises when two amino acids come together like two people facing each other and holding both hands, their arms forming a closed ring. Chemists who synthesize proteins have developed many excellent reactions for forming amide bonds between amino acids, and they are all too familiar with the diketopiperazine structure, because it can form when it is not wanted and interfere with their efforts to synthesize proteins. I figured, though, that I could make use of diketopiperazine formation to link my building blocks.

The rest of the idea soon fell into place. In the "people" analogy, the two "arms" of an amino acid are groups of just a few atoms called the amine group and the carboxyl group. (Unlike arms, however, these groups do not actually stick out very far.) Think of one as the left arm and the other as the right, with an amide bond being a left hand holding a right hand. Each of my new building blocks, or monomers, would be like two people tied rigidly together (for example, back to back) with their arms in front of them. One monomer would connect with the next in the sequence by a person on one holding both hands of a person on the other—forming a diketopiperazine ring.

In real chemical terms, each monomer would consist of a rigid molecule of mostly carbon atoms with two amino acid groups integrated into it, and the amines and carboxyls of both amino acids would be available for bonding to other monomers. Two monomers would join by having an amino acid group on each one reacting together to form a diketopiperazine ring. We would call this kind of monomer a bis-amino acid ("bis" meaning "twice") because each one contains two amino acids. And just as chains of amino acids are called peptides, we would call our chains of bis-amino acids "bis-peptides."

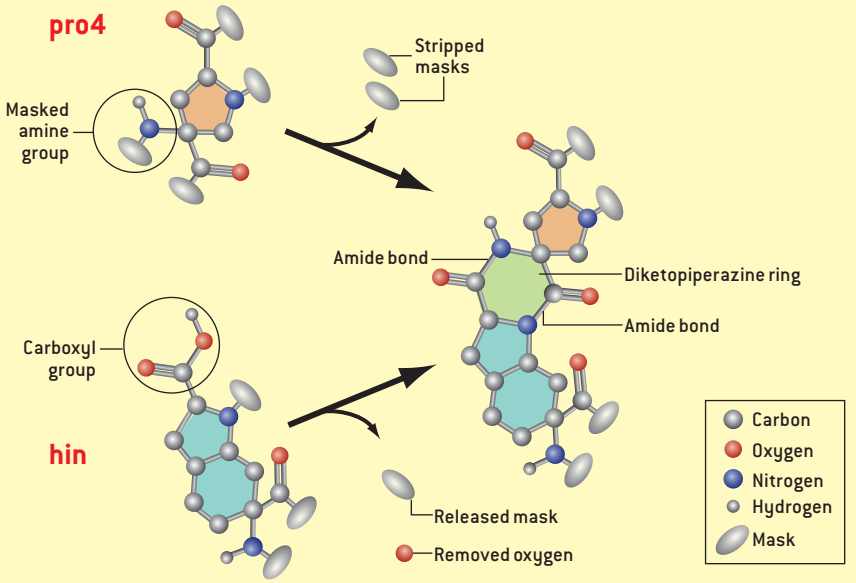
Starting from Scratch

WITH BLUEPRINTS for a collection of building blocks in hand, I launched a new lab at the University of Pittsburgh, where my students and I could develop the synthetic chemistry to make this idea work. Within two years Christopher Levins, one of my first graduate students, had synthesized our first bis-amino acids. He started with hydroxyproline, a commercially available component of collagen (the protein that makes cartilage, ligaments and tendons strong) that another group had previously used in making molecules very like our monomer design. Using a nine-step recipe that we worked out together, Levins converted hydroxyproline into four kinds of building blocks, which we named pro4(2S4S), pro4(2S4R),

MELISSA THOMAS

THE CHEMISTRY

In practice, chemists synthesize bis-amino acids with protective groups, or masks, to prevent bonds from forming among them indiscriminately. Using a series of steps (*not depicted here*), the chemists link two monomers—such as pro4 and hin, whose chemical structures are shown at the left—by inducing what is called a diketopiperazine ring (*green*) to form between them. The rigidity of this ring and of the other carbon rings within the bis-amino acids ensures the stiffness and predictable shape of the resulting chains. (Some hydrogen atoms and details of the protective groups have been omitted for clarity.)



pro4(2R4S) and pro4(2R4R). We call them "pro4" because they all resemble the amino acid proline with an additional amino acid mounted on carbon 4 (chemists identify the carbon atoms in an organic molecule by labeling them with numbers in a systematic fashion). The labels "S" and "R" indicate the orientation of the groups attached to carbon 2 and carbon 4. The completed building blocks are dry powders that are stable for months of storage at room temperature.

We construct our monomer building blocks with protective groups attached to the amines (to prevent amide bonds from forming until we want them to) and with one of the carboxyls in a modified, less reactive form called an ester. To synthesize a bis-peptide, we assemble the building blocks in the desired sequence with single bonds and then join up all the second bonds to rigidify the molecule into its final shape [*see box on opposite page*]. Levins carried out this two-part procedure to build our first short structures made of pro4 monomers.

The first part of the linking process uses a technique called solid supported synthesis. It begins with plastic beads coated with an amine group. The carboxyl group on the first building block forms an amide bond with one of the amines, fixing the building block to a bead. Using an excess of building blocks ensures that virtually all the amines on the beads have a building block attached. A quick wash with a solvent removes by-products and leftover building blocks. Then a wash with a base removes the protective group from one of the two amines on the newly added building block (the two amines have different protective groups, so only one of them is stripped). A second building block is added and attaches to the first through its carboxyl and the exposed amine group. The protection is then removed from one of its amines, a third building block is added, and so on.

This construction process goes slowly: it takes about an hour to add each successive monomer because we have to wait long enough for nearly all the exposed amines to get their building blocks. Fortunately, robots usually used for synthesizing peptides can automate the work and can easily construct many sequences in parallel.

When a chain is complete, we use strong acid to remove the beads, then strip the second amine protective group from every building block within the chain. Adding a base solution causes the newly revealed amine on every building block to attack the ester on the preceding building block and form another amide bond to it. With two amide bonds connecting each pair of adjacent building blocks, the entire molecule is now rigid and has a predictable, well-defined shape.

We soon found that bis-peptides are soluble in water and other polar organic solvents (solvents that mix readily with water). The water solubility of bis-peptides makes them easy to study and suggests that we could use them to develop new medicines, which must be able to disperse through the blood.

Programming Shapes

THE BIS-AMINO ACIDS that make up our bis-peptides join together like strangely shaped Lego bricks. In particular, each bis-amino acid behaves like a brick whose top surface of studs is tilted and twisted relative to its bot-

THE AUTHOR

CHRISTIAN E. SCHAFMEISTER is an associate professor of chemistry at Temple University, where he is developing shape-programmable molecules. He received his Ph.D. in biophysics at the University of California, San Francisco, in 1997. As a postdoctoral fellow at Harvard University, he developed a new way of making peptides more resistant to proteases, rendering them more appropriate as potential drugs. He is a member of the working group preparing the Technology Roadmap for Productive Nanosystems for the Foresight Nanotech Institute in Palo Alto, Calif.

MELISSA THOMAS

tom surface of holes. Repeatedly stacking one type of brick on top of itself allows you to make one curved shape, with the specific shape of the curve depending on which bis-amino acid is chosen. Using just two different kinds of bricks stacked in different sequences, you can make 2^N different shapes (N is the number of bricks in the stack). A bis-peptide 10 blocks long made out of our four pro4 bis-amino acids could have any one of about a million (4^{10}) shapes. The more shapes of building blocks we have, the better we will be able to control the final shape of the macromolecule. The challenge then is to design and synthesize those sequences that have useful functions.

The key to designing bis-peptides with specific shapes is knowing the precise shapes that our individual bis-amino acids take on when they are joined to one another. This information, analogous to knowing the size of each brick and the tilt and twist of its studs, would become the basis for our "programming language for matter." Having synthesized our first bis-peptides, we could then carry out measurements to determine how their pieces fit together.

We performed nuclear magnetic resonance experiments to find out which hydrogen atoms on a bis-peptide are close to one another and applied other techniques to measure the orientations of carbon-hydrogen bonds. From the results of these measurements we inferred the shape information that we needed, and we used it to create a computer-aided design program for building bis-peptides called CANDO (for *c*omputer-*a*ided *n*anostructure *d*esign and *o*ptimization).

Gregory Bird, another graduate student in my lab, used CANDO to design molecular rods and curved structures. Recently he assembled these structures, attaching a chemical group called a spin probe to each end of every sequence to verify that the results in the reaction vessel matched the design in the computer. Indeed, sequences of pro4(2S4S) and pro4(2R4R) building blocks had **C** and **S** shapes just as CANDO predicted they would.

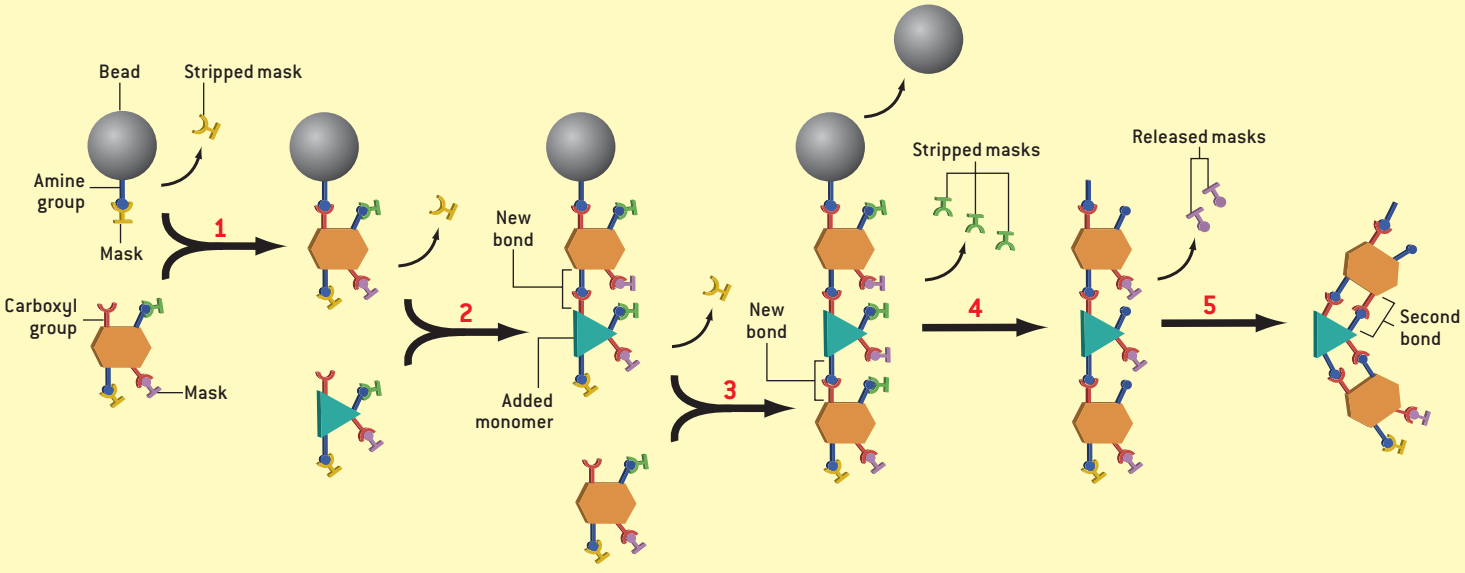
The pro4 group of bis-amino acids are like Lego bricks that have relatively small tilts, so we can use them to make rodlike and gently curving shapes, which could function like struts to hold chemical groups apart at specific distances. Many useful functions of proteins, however, come about because of cavities that can serve to bind the protein to a specific target or to hold molecules and catalyze reactions. To create compact bis-peptides that have suitable cavities, we needed to expand our repertoire of building blocks. My student Stephen Habay took the first step toward this goal by developing a bis-amino acid we call "hin" that creates a sharp turn in a bis-peptide.

Year by year our collection of monomers continues to grow, and CANDO analyses suggest that our present repertoire of 14 monomers is sufficient to create compact bis-peptides containing cavities. But as we developed new building blocks and incorporated them into bis-peptides, we ran into a problem. The reaction that forms the rigidifying second amide bond was very rapid between pro4 monomers but was

MELISSA THOMAS

HOW BIS-PEPTIDES ARE MADE

The synthesis of a bis-peptide proceeds by first assembling selected bis-amino acids in the correct sequence and then rigidifying the structure: A bead is coated with a protected amine group. The protective mask (*yellow*) is stripped away, and the first bis-amino acid latches onto it via the bis-amino acid's free carboxyl group (*1*). The process is repeated (*2* and *3*) with more bis-amino acids, producing a chain linked by single bonds. Then the bead is stripped away, as are the protective groups (*green*) on the unbonded amines (*4*). The freed amines react with the nearby masked carboxyls, releasing the masks and forming a second bond between each adjacent pair of monomers (*5*).



sluggish for all our new building blocks. Raising the reaction temperature sped things up but scrambled the resulting shapes. This problem was a huge obstacle to creating larger and more complex bis-peptides.

My student Sharad Gupta partially overcame this challenge by developing a new approach to closing the second amide bond. On each monomer he changed the ester to one that is more susceptible to the amine's attack, and, inspired by a 1970s paper, he used acetic acid as a catalyst instead of a base. The combination of heat and acid accelerated the ring-closing reaction without scrambling our bis-peptides' shapes in the way that heat and base did.

We took six months to find the combination of ester, protective group, solvent and temperature that we have settled on for now, but we will return to this problem in the future because our solution does not work well for sequences longer than about five monomers. In the meantime, we are focusing on developing some applications with the bis-peptides that we can produce efficiently—those of any length that involve only the pro4 monomers, and sequences of up to five monomers that include the others.

Developing Applications

ONE OF THE FIRST applications that we have pursued for our bis-peptides is a macromolecule that would bind tightly to the cholera toxin protein (Ctx). The protein has five identical pockets, each at the corner of a pentagon. These pockets allow Ctx to bind to the sugar GM1, which fits neatly into the pockets. The epithelial cells that line the small intestine have molecules of GM1 attached to their surface, and when Ctx binds to five of these molecules, it initiates a chain of events that leads to life-threatening diarrheal disease. Molecules that bind tightly to these pockets on Ctx could prevent the toxin from binding to human cells and stop the disease in its tracks.

Other researchers have developed small sugars that bind to these pockets individually. But those drugs do not work well, because they do not bind very tightly to Ctx and cannot compete with the five simultaneous interactions that Ctx makes with GM1 on human cells. We wondered whether we could synthesize a bis-peptide that could plug sugars into two pockets at the same time. We can attach almost anything we want at the ends of a bis-peptide, so for this application we put a small sugar on each end of rod-shaped bis-peptides that just span the distance between adjacent pockets in the Ctx protein. The experiment worked in that bis-peptides with two sugars bound to Ctx more tightly than the individual small sugars, and they bound at least as well as the natural GM1 target does.

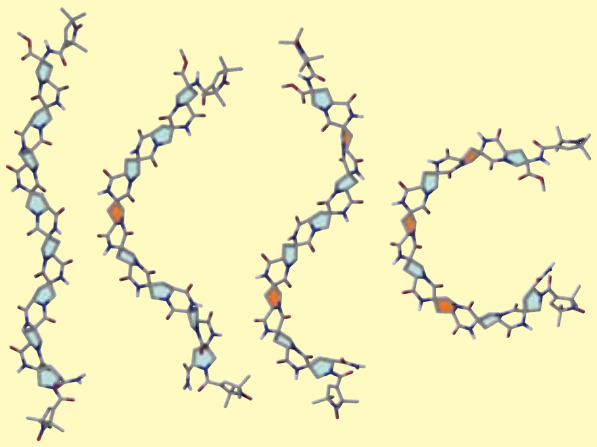
We have not, however, been able to determine whether each bis-peptide was binding two pockets of one Ctx or binding with pockets on two different Ctx molecules and thus creating a cross-linked network of Ctx molecules. Cross-linking Ctx would not be a useful way to fight cholera, because it would be effective only in a person who had a lot of Ctx (probably a lethal amount) in the body already. (If the Ctx concentration were too low, each bis-peptide might bind to one pocket on one Ctx but then have too small a chance of encountering another Ctx to create a cross-link.) But cross-linking proteins on the surfaces of viruses might be effective, and so we are now applying this approach to inhibiting viruses, including HIV and Ebola.

As well as attaching groups to the ends of a rigid rod, we have developed molecular actuators in which two rods are joined by a hinge. An actuator is a device that responds to a signal by producing motion. Our rod-hinge-rod actuators are designed to be open normally and to fold over, or close, when groups on the outer ends of the rods bind a metal or a small molecule. My student Laura Belasco made our first version of these, in which the rods are four building blocks long, the hinge is an ordinary amino acid, and a metal triggers the opening and closing. One application would be molecular valves [*see box on opposite page*]. The valve would consist of a nanoscopic hole with hinged rods attached around its rim. Outstretched, the rods would block the hole; folded, they would open it. These valves could be used to make a device that senses a patient's condition and releases the appropriate medicine in response.

Control of the opening and closing could be carried out electronically by putting groups at the end of the rods that would bind when the correct charge was present. Computer storage devices could be made out of a forest of hinged rods if they could be controlled individually in this way. Atomic force microscope tips would scan across the rows of the forest detecting which rods were standing up as the 1s and 0s, analogous to detecting the pits or no pits of

DESIGNER SHAPES

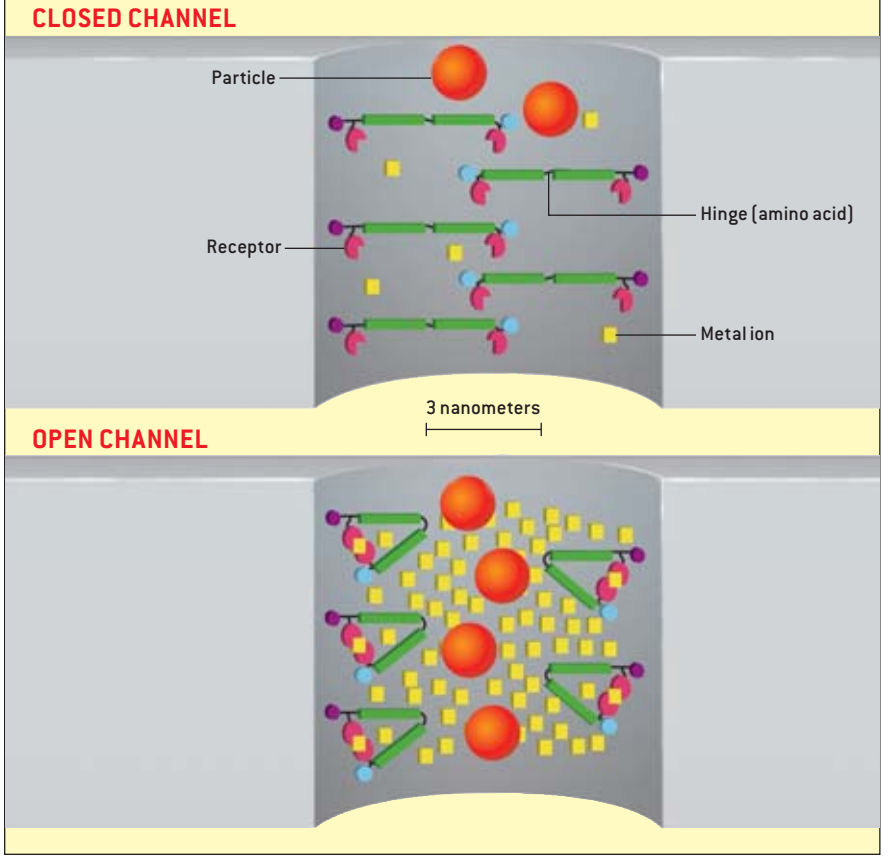
As these examples of bis-peptides synthesized by the University of Pittsburgh group show, the shapes of the molecules can vary from nearly straight rods to tight crescents with the insertion of the right monomers.



CHRISTIAN E. SCHAFMEISTER

NANOSCALE VALVES

Valves that open to a mere three-nanometer diameter using bis-peptide actuators are on the drawing board. The actuators, which have been synthesized and demonstrated, consist of two short bis-peptide rods (*green*) joined by an amino acid that serves as a hinge. In a low concentration of the triggering metal ion (*yellow*), the actuators extend from the rims of tiny holes etched through aluminum films (*gray*), blocking larger particles or molecules (*orange*) from passing (*top*). At high concentrations the ions bind to receptors (*pink*), causing the actuators to fold over, opening the channel (*bottom*).



IBM's "Millipede" drive. Erasing a pit, which is difficult for the Millipede system, would be as simple as reversing the state of the hinged rod.

The side chains of the 20 amino acids that organisms use to build their proteins are decorated with a variety of chemical groups. Proteins position these chemical groups in configurations whose shape and other properties serve to catalyze reactions, bind small molecules and carry out their many functions. Similarly, in our lab we are developing building blocks that carry an additional chemical group, which will let us create bis-peptides that display chemical groups along their ladderlike backbones. So far we have made the first such building block with a side group. If we can make macromolecules with constellations of chemical groups that mimic the active sites of enzymes—the areas where catalysis takes place—we could use them to learn how to create designer enzymes.

Twenty years from now I envision an active community of developers: dozens of groups inventing designer bis-peptide-based macromolecules and learning how to produce artificial enzymes and other useful molecular devices. Some promising anticancer drugs such as halichondrin-B and bryostatin are currently very expensive to synthesize. The rare sponges and sea creatures that produce these compounds cannot provide the quantities needed for widespread use. In 20 years we might be able to create artificial enzymes that efficiently synthesize these and other valuable compounds in an environmentally benign way. Imagine adding a drop of artificial enzymes to a barrelful of high-fructose corn syrup and a few days later harvesting gallons of bryostatin.

If we could develop artificial enzymes that break down plant cellulose into ethanol or that use light energy to combine water and carbon dioxide to create ethanol, such an undertaking would have massive benefits for society. We could even design artificial enzymes to synthesize our bis-amino acid building blocks and join them together, making it much easier to make bis-peptides.

We have developed a combination of chemistry and software for creating macromolecules with programmable shapes. Because it takes only a few days to produce bis-peptides, we can design and assemble them, test their properties and fashion the next generation on a timescale of weeks. The fascinating challenge in coming years will be to learn how to begin with a function and to design the best bis-peptide sequence for carrying it out. SA

MORE TO EXPLORE

Constructing a Molecular LEGO Set. John P. Mathias and J. Fraser Stoddart in *Chemical Society Reviews,* Vol. 21, No. 4, pages 215–225; 1992.

Foldamers: A Manifesto. S. H. Gellman in *Accounts of Chemical Research,* Vol. 31, No. 4, pages 173–180; April 1998.

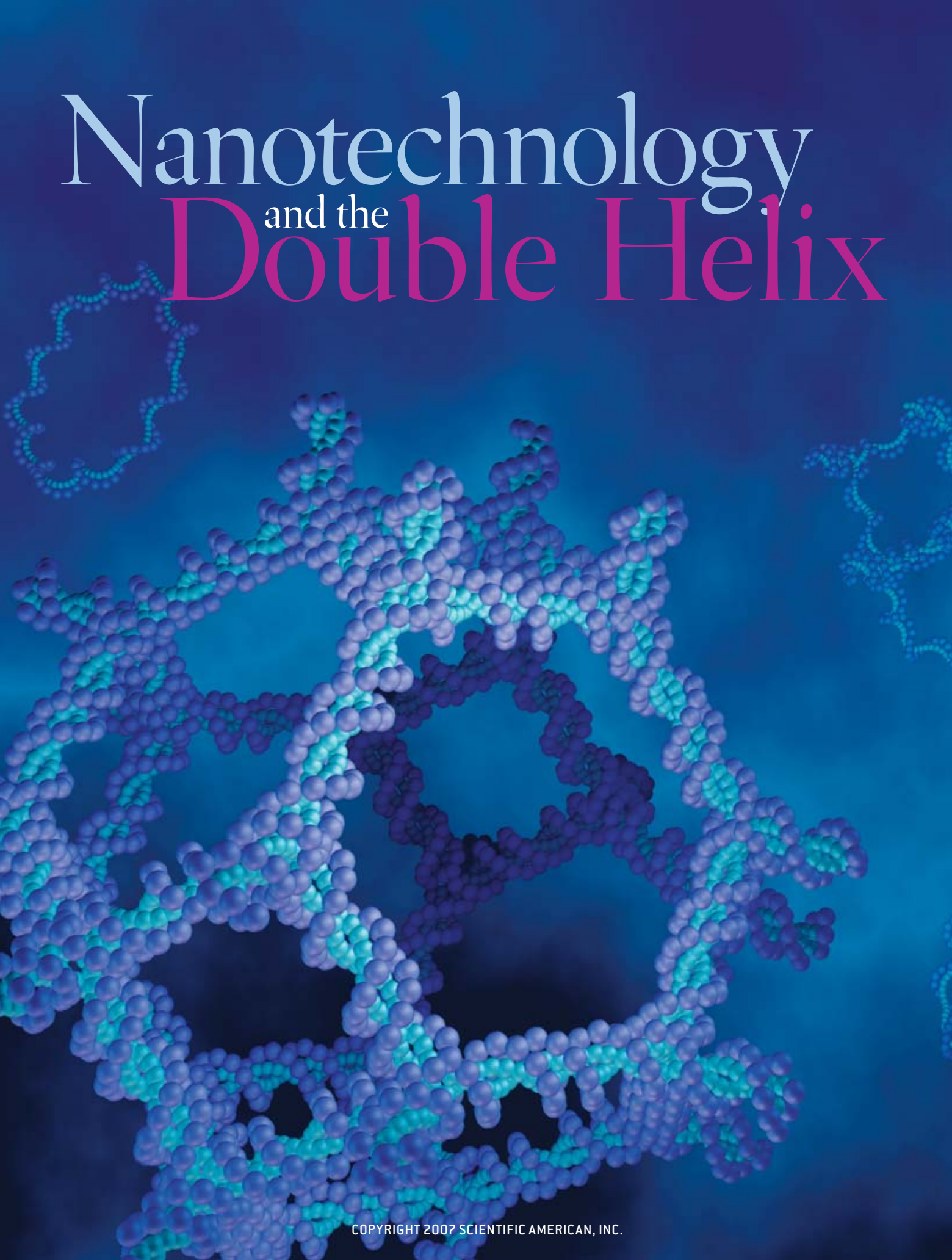
A Field Guide to Foldamers. D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes and J. S. Moore in *Chemical Reviews,* Vol. 101, No. 12, pages 3893–4011; December 2001.

Flexibility and Lengths of Bis-Peptide Nanostructures by Electron Spin Resonance. S. Pornsuwan, G. Bird, C. E. Schafmeister and S. Saxena in *Journal of the American Chemical Society,* Vol. 128, No. 12, pages 3876–3877; March 29, 2006.

Christian E. Schafmeister's laboratory site: **www.candolab.org**

MELISSA THOMAS

Nanotechnology and the Double Helix



DNA is more than just the secret of life—it is also a versatile component for making nanoscopic structures and devices

By Nadrian C. Seeman

The year 2003 witnessed the 50th anniversary of the discovery of DNA's double-helix structure by James D. Watson and Francis H. Crick. Their discovery reduced genetics to chemistry and laid the foundations for the next half a century of biology. Today thousands of researchers are hard at work deciphering the myriad ways that genes control the development and functioning of organisms. All those genes are written in the medium that is DNA.

Yet this extraordinary molecule has other uses in addition to those of biochemistry. By employing the techniques of modern biotechnology, we can make long DNA molecules with a sequence of building blocks chosen at will. That ability opens the door to new paths not taken by nature when life evolved. In 1994, for example, Leonard M. Adleman of the University of Southern California demonstrated how DNA can be used as a computational device. In this article I will discuss another nonbiological use of DNA: the building of structures and devices whose essential elements and mechanisms range from around one to 100 nanometers in size—in a word, nanotechnology.

Such structures have many potential applications. Regular lattices made of DNA could hold copies of large biological molecules in an ordered array for x-ray crystallography to determine their structure, an important step in the "rational" design of drugs. Alternatively, the lattices could serve as scaffolding for nanoelectronic components, either as a working device or as a step in the manufacture of a device. Materials could be constructed—either made of the DNA or made by it—with structures precisely designed at the molecular level. DNA machines with moving parts could be employed as nanomechanical sensors, switches and tweezers as well as for more elaborate robotic functions.

Branched DNA

THE NANOSCALE is the scale of molecules. A typical bond between two atoms is about 0.15 nanometer long. (A nanometer is a billionth of a meter.) The helix of DNA has a diameter of about two nanometers, and it twists full circle once every 3.5 nanometers or so, a distance of about 10 base pairs, which form the "rungs" of DNA's ladder [*see box at top on page 33*]. A short piece of DNA has highly specific interactions with other chemicals, depending on its sequence of base pairs. One can imagine using such pieces to recognize particular molecules or to control the composition of a material by acting as a catalyst. And for many years biologists have used DNA for its recognition properties, especially exploiting the "sticky ends" in genetic engineering. A sticky end occurs when one strand of the helix extends for several un-

DNA STRANDS SELF-ASSEMBLE into a complicated structure when their base sequences are designed to pair up with specific partners. Here a stick model of a truncated octahedron, which has six square faces and eight hexagonal faces, has formed. The edges are about 20 nanometers long. A short "hairpin" of DNA sticks out from each corner. The hairpins could be modified to link truncated octahedra together to form a regular three-dimensional scaffold.

KEN EWARD *BioGrafx* (*strands, left*); ALICE Y. CHEN (*helix, right*)

paired bases beyond the other [*see illustration at bottom on opposite page*]. The stickiness is the propensity of the overhanging piece to bond with a matching strand that has the complementary bases in the corresponding order—the base adenine on one strand pairs with thymine on the opposite strand, and cytosine binds with guanine.

At first sight, it does not appear that DNA can lead to interesting structures. Naturally occurring DNA forms a linear chain, like a long piece of twine, so that all one can envision making from it is lines or circles, perhaps snarled up or knotted in one way or another. But a linear chain is not the only form that DNA takes. During certain cellular processes, DNA exists briefly as a branched molecule. This branching occurs when DNA replicates (in preparation for cell division) and during recombination (when genetic material is swapped between matching pairs of chromosomes, as happens when sperm and eggs are produced).

The branches form when the double helix partially unravels into two strands. In replication, each strand is made into a new double helix by the addition of complementary nucleotides all along its length. (A nucleotide is the combination of a base and the corresponding section of the backbone of the helix.) More interesting is the crossover that occurs in recombination, in which two pieces of DNA break and partially unravel and the resulting four strands join up somewhat like the intersection where two highways cross.

In recombining DNA, the branch point occurs where each of the four strands switches from one partner to another. The branch point moves around because of twofold symmetry (like that of the numeral "69") in the base sequences that flank it. This symmetry means that each strand can pair up with either of two other strands. In 1979 I was working with Bruce H. Robinson, now at the University of Washington, to describe the nature of this motion when I recognized that synthetic DNA molecules lacking this symmetry could form branched molecules whose branch points do not move. To design such a junction, one would make four strands of DNA. For each strand, the sequence along half of the strand would match half of a second strand and the remaining half would match half of a third strand.

DNA's favorite structure is the conventional double helix identified by Watson and Crick. A quantity called free energy determines which structure is favored. In general, free energy determines whether a chemical reaction proceeds in the forward or reverse direction; it also determines the conformation—the folds and joins—of large molecules such as DNA, RNA and proteins. A chemical system always tends to change toward the state that has the lowest free energy. For two complementary strands of nucleotides, the free energy is minimized when they pair up to form a double helix.

The four strands of our immobile junction can come together and form the maximum amount of conventional DNA double helices only by forming a branched molecule. In general, a branch point is not favored—it increases the free energy of the molecule—but this increase is outweighed by the much greater energy saving in the four arms made of ordinary double-helix DNA. Today it is simple to synthesize such strands and implement this idea of a stable branched DNA molecule, but in 1979 it was state-of-the-art chemistry and I was a crystallographer, not an organic chemist, so mostly I just thought about the system. (It was not until 1982 that I learned how to make DNA.)

Inspiration from Escher

I FIGURED OUT that it ought to be possible to make branched DNA junctions with many arms, not just four. One day, in the fall of 1980, I went over to the campus pub to think about six-arm junctions. For some reason, I thought about Dutch artist M. C. Escher's woodcut *Depth* [*see illustration on page 34*]. I realized that the center of each fish in that picture was just like an idealized picture of the branch point of a six-arm junction. Six features extend from that center point on the fish: a head and a tail, a top fin and a bottom fin, a left fin and a right fin. The fish are organized in the same way as the molecules in a molecular crystal, with regular repeats forward and back, up and down, left and right. It struck me that if I held junctions together using sticky ends, I might be able to organize matter on the nanometer scale in the same way that Escher held his school of fish together using his imagination.

We have several good reasons for wanting to build such structures. First, we are aiming to fabricate macroscopic pieces of matter made of designed molecules joined together in a structure that is controlled with nanoscopic precision. This procedure could result in materials having novel properties or novel combinations of properties. For example, materials with designed optical properties, such as photonic crystals, could be made by constructing precisely defined arrays with specific repeat distances.

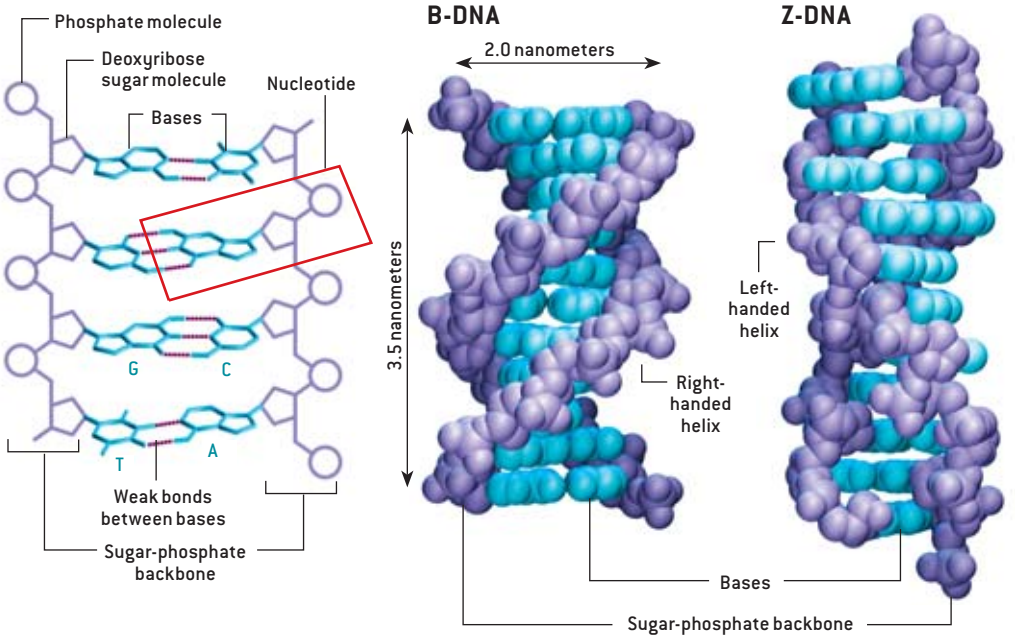
Another goal is to use DNA as scaffolding to hold other molecules in arrays, including those that do not form a regular crystalline structure on their own. In this way, one could make

Overview/*DNA Nanotech*

- DNA is an ideal molecule for building nanometer-scale structures. Strands of DNA can be programmed to self-assemble into complex arrangements by producing the strands with the appropriate combinations of complementary bases, which preferentially bond together to form stretches of double helices.
- DNA scaffolds could hold guest molecules in orderly arrays for crystallography. They could also hold molecule-size electronic devices or be used to build materials with precise molecular configurations.
- Nanometer-scale DNA machines can function by having parts of their structure change from one DNA conformation to another. These movements can be controlled by chemical means or by the use of special DNA strands.

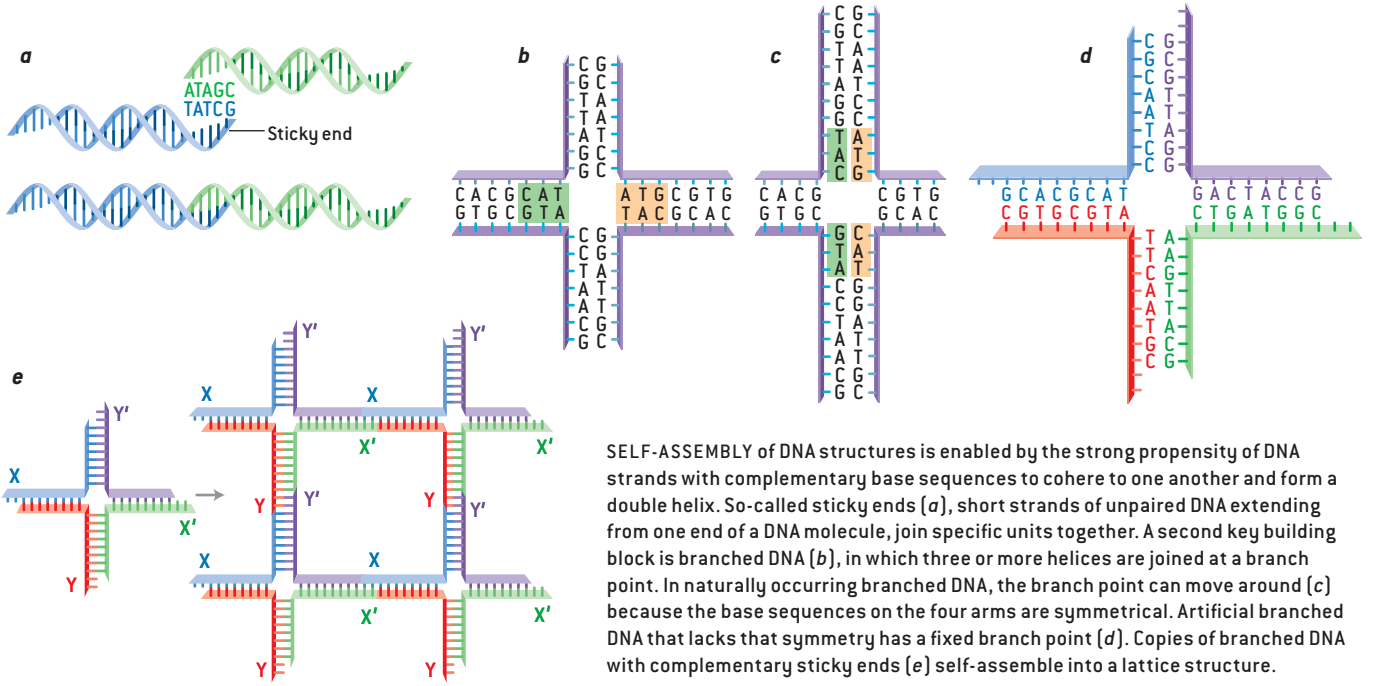
THE STRUCTURE OF DNA

DNA is a nanoscale structure, consisting of a double backbone of phosphate and sugar molecules between which complementary pairs of bases (A and T; C and G) are connected by weak bonds (*left*). DNA's most common conformation is B-DNA (*center*), which twists in a right-handed double helix about two nanometers in diameter. One full turn of the helix is about 3.5 nanometers, or 10 to 10.5 base pairs long. In special circumstances DNA can form a left-handed double helix called Z-DNA (*right*).



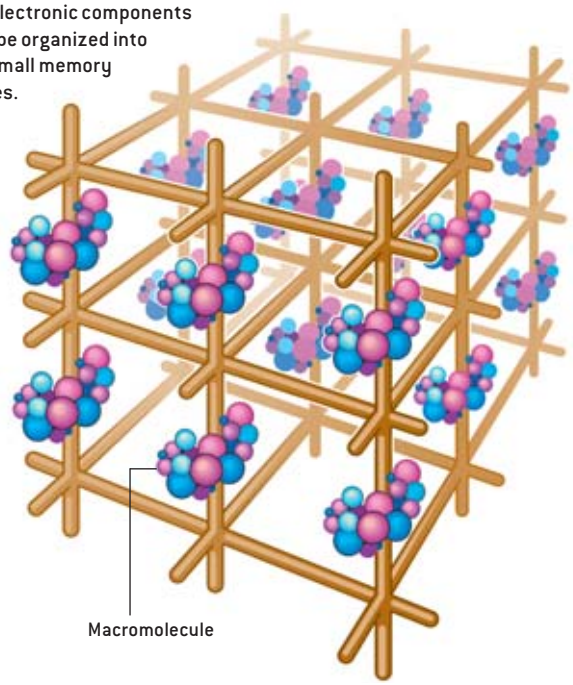
crystals for use in crystallography experiments by making DNA cages that contain large biological molecules such as proteins within them [*see right illustration on next page*]. Such cages would enable crystallographers to determine the three-dimensional structures of the enclosed molecules—a key procedure in the rational design of drugs that mesh precisely with specific parts of a targeted molecule. (This crystallographic application is the one that most strongly motivates my interest in this field.) Currently many of the receptor molecules that could be excellent drug targets do not lend themselves to conventional crystallography. In a similar fashion, one could organize nanoelectronic components into very small memory devices, as Robinson and I suggested in 1987.

Why use DNA for these purposes? The chief reason is that strands of DNA interact in the most programmable and predictable way. A sticky end that is N bases long has one of 4^N



KEN EWARD *BioGrafx* (*top*); ALICE Y. CHEN (*bottom*)

SELF-ASSEMBLY of DNA structures is enabled by the strong propensity of DNA strands with complementary base sequences to cohere to one another and form a double helix. So-called sticky ends (*a*), short strands of unpaired DNA extending from one end of a DNA molecule, join specific units together. A second key building block is branched DNA (*b*), in which three or more helices are joined at a branch point. In naturally occurring branched DNA, the branch point can move around (*c*) because the base sequences on the four arms are symmetrical. Artificial branched DNA that lacks that symmetry has a fixed branch point (*d*). Copies of branched DNA with complementary sticky ends (*e*) self-assemble into a lattice structure.



ESCHER'S WOODCUT *DEPTH* (*left*) inspired the author to consider an array of six-arm junctions connected together to form a three-dimensional molecular crystal (*below*). The center of each fish is just like the branch point of a six-arm junction. Instead of arms, six features extend from that center point: a head and a tail, a top and bottom fin, and a left and right fin. Molecular scaffolding could hold other molecules in regular arrays. For example, DNA cages containing oriented biological macromolecules as guests could be used in crystallography experiments. In a similar fashion, nanoelectronic components could be organized into very small memory devices.

possible sequences of bases. This enormous variability and the propensity of the end to bond to only a closely matching sequence provide ample scope for designing molecules that consist of a large number of DNA strands joined to one another in a completely specified manner. Furthermore, we know that two sticky ends form the classic helical DNA structure when they cohere, and these helical stretches of DNA are relatively stiff. Thus, we know not only which strands link to which other strands but also the detailed shape of the joined segments. Since the mid-1990s, it has been possible to program the shapes of DNA branched species using only their sequences. We do not have such specific information for proteins or antibodies, which are other candidates for working elements. Those also have tremendous variability, but determining what shape a protein will take and how two proteins or antibodies will join together are laborious problems that would have to be solved anew for each example.

Another reason for working with DNA is the simplicity of its synthesis with the tools of the biotechnology industry. We can manipulate DNA with many enzymes, such as restriction enzymes (which cleave DNA at particular sites) or ligases (which catalyze the joining of two molecules by covalent bonds—sturdy chemical bonds that involve the sharing of pairs of electrons between atoms). These tools can be used to make and manipulate conventional DNA, as well as exotic derivatives, in which different bases from the usual four are incorporated or in which additional molecules are attached on the outside of the DNA's backbone (the sides of the DNA ladder). Medical researchers hoping to use nucleic acids (DNA and RNA) for therapy have made many such variants. DNA is extremely well suited to making such derivatives because every nucleotide along the helix has sites where molecules can be attached.

Finally, as we will see below, DNA can be induced to form structures different from the standard double helix. We can build nanomechanical devices whose parts move—such as closing tweezers or a rotating shaft—when there is a transition from one DNA structure to another. One drawback is that DNA objects must be constructed in an aqueous solution. It is no problem, however, to dry the resulting structures (on mica, for instance) as we do to make microscopic images of our results.

THE AUTHOR

NADRIAN C. ("NED") SEEMAN trained in crystallography, but his frustrations with a macromolecular crystallization experiment led him to the idea that DNA junctions could be used in a new approach to crystallization. Ever since then, he has been trying to implement this concept and its spin-offs. For the past 19 years, Seeman has worked in the department of chemistry at New York University. When told in the mid-1980s that what he was doing was nanotechnology, his response was similar to that of M. Jourdain, the title character of Molière's *Bourgeois Gentilhomme,* who was delighted to discover that he had been speaking prose all his life.

Stick Models

THE FIRST STEP in any new scientific research program is to establish the basic feasibility of the project. In 1991 Junghuei Chen, now at the University of Delaware, and I did this by building a DNA molecule shaped like a cube formed from sticks [*see illustration below*]. Each edge of the cube is a stretch of double-helical DNA; each corner is a three-arm junction. Each corner is connected to three other corners; it is said that the cube's connectivity is three. Genetic engineers had made many linear DNA constructs, but this was the first DNA molecule with connectivity greater than two. The cube self-assembles from pieces of DNA designed to adhere to one another, but the ends of each piece do not join up. Ligases can connect these free ends, resulting in six closed loops, one for each face of the cube. Because of the helical nature of DNA, each of these loops is twisted around the loops that flank it, so the cube cannot come apart, even if all the bonds joining the base pairs together were somehow broken.

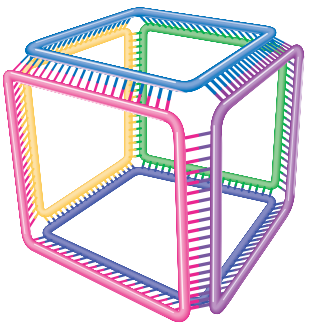
Yuwen Zhang, now at Human Genome Sciences in Rockville, Md., and I built another shape called a truncated octahedron, which is similar to but more complicated than a cube [*see illustration on page 30*]. Although three-arm junctions would have sufficed to make individual truncated octahedra, instead we built them using four-arm junctions. We intended that the extra arm sticking out at each corner could be used to connect truncated octahedra together in a larger structure, but in the end we did not continue in this direction. We had created only a very tiny quantity of truncated octahedra—enough to characterize their structure but too few to attempt to join them together—and even that minute sample had taken us to the limits of what we could do without overhauling our procedures (for example, by robotizing repetitive steps). Instead we turned to simpler components.

Another reason for changing direction was that along the way we realized that the stick polyhedra we had built were not rigid. DNA is a stiff molecule: a stretch of DNA that is two or three turns long (the lengths we use for the polyhedra edges) can wiggle around its helix's axis no more than a piece of cooked spaghetti two or three millimeters long can wiggle around its central axis. That inflexibility ensured that the edges of our stick figures were rigid, but we learned that the angles at each corner were quite variable. The polyhedra we had built were rather like structures made of toothpicks stuck into blobs of marshmallow at the corners. Such structures might have uses, but building a regular lattice is not one of them. It is much easier to self-assemble an orderly, crystallike piece of matter from bricklike components than from marshmallows.

THE POLYHEDRA we had built were rather like structures made of toothpicks stuck into BLOBS OF MARSHMALLOW at the corners.

To solve this problem, my group examined another branched motif found in biological recombination systems, the DNA double-crossover (DX) molecule. The DX molecule consists of two double helices aligned side by side, with strands crossing between the helices, yoking them together [*see box on next page*]. We characterized this molecule and established that it is stiff. We also demonstrated that a DX molecule containing another small double-helical region (called a DX + J molecule) is very stiff. This additional double-helical region creates a bump on the top of the DX molecule, which serves as a marker—a nanotech equivalent of a dab of paint.

In collaboration with Erik Winfree of the California Institute of Technology, Furong Liu and Lisa A. Wenzler of my group at New York University used combinations of DX and DX + J molecules as tiles to make two-dimensional crystals with defined patterns. The tiles are joined together by sticky ends on each helix. One arrangement, with columns of DX

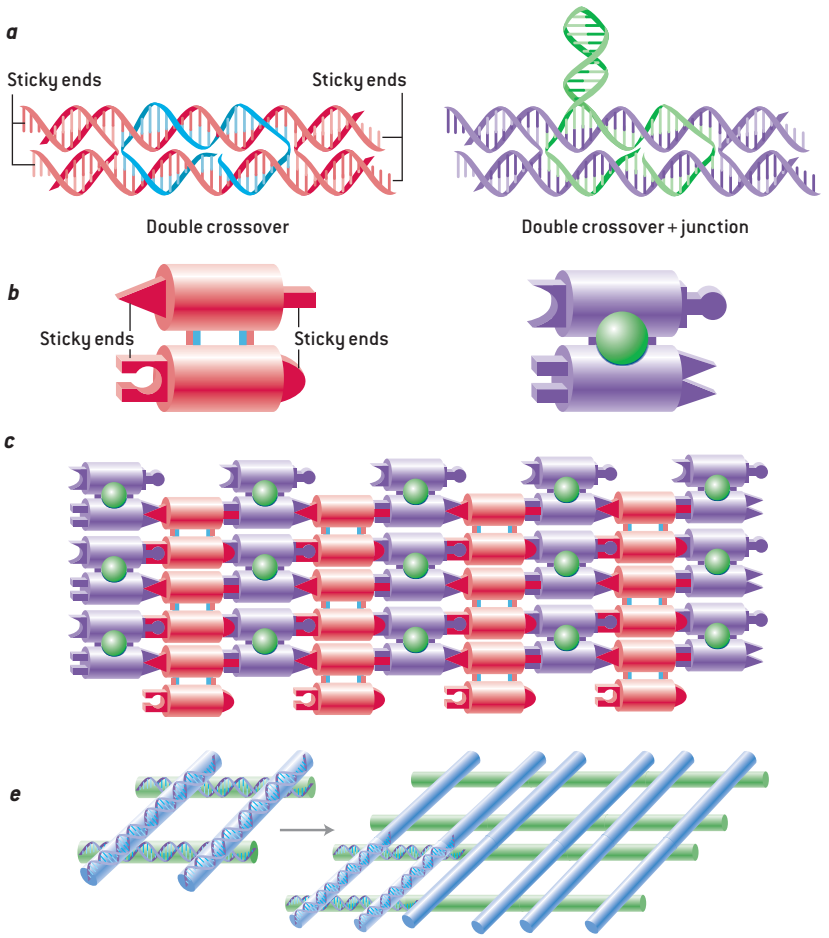


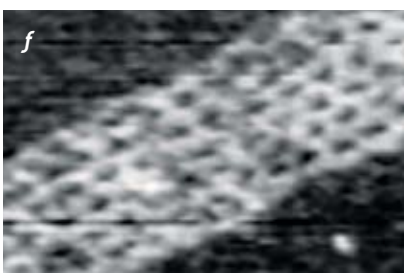
STICK CUBE (*far right*) made out of six loops of DNA demonstrated that three-dimensional structures can be built. The backbone of each DNA strand is depicted as colored spheres (a different color for each strand) and the bases as white spheres. Each edge of the cube comprises 20 nucleotide pairs, or about two complete turns of the double helix. Each corner is a three-arm junction. Simplified schematic (*near right*) depicts how the DNA strands are connected but omits the helical twists.

ALICE Y. CHEN (*left*); KEN EWARD *BioGrafx* (*right*)

STIFF DNA ARRAYS

Two-dimensional crystals can be made out of stiff bricks of DNA. The bricks (*a*) are double-crossover (DX) and double-crossover-plus-junction (DX + J) units, which cannot flop around at their joining points the way that multiarm junctions can. Each brick has four distinct sticky ends for joining bricks together. The extended green strand of the DX + J unit sticks out of the plane. Each unit is about 4 by 16 nanometers in size. For simplicity, the DX and DX + J units are shown schematically, with geometric shapes at their ends representing the sticky ends (*b*). In a solution, the sticky ends cohere and the units self-assemble in a two-dimensional pattern (*c*). The striped pattern shows up in an atomic-force microscope image of the crystal (*d*) (which is deposited onto a flat mica surface for the microscopy). The bright stripes, spaced about 32 nanometers apart, are the lines of DNA protruding from the DX + J units. Parallelograms of DNA have also been self-assembled into two-dimensional patterns (*e, f*).





tiles alternating with columns of DX + J tiles, produces a pattern of stripes separated by about 32 nanometers. We deposited the arrays on a flat mica surface and examined them with an atomic force microscope to confirm that the structure had the correct dimensions. We established that the pattern was not accidental by making a second crystal with modified tiles that link together with three DX columns for each DX + J column, to produce stripes with double the separation.

More recently, John H. Reif's group at Duke University demonstrated "DNA bar codes" made using such patterns. In these tilings, the positions of stripes were programmed to occur in a pattern representing the number "01101" (with molecules analogous to our DX and DX + J serving as 0 and 1, respectively). The pattern was programmed using an input DNA strand whose sequence encoded the 01101 pattern. The analogues of the DX and DX + J bricks self-assembled on the sections of the DNA strand corresponding to 0 and 1, respectively. Many such five-brick sequences then joined up in parallel, generating the 01101 pattern of stripes. The stripes were about 15 nanometers apart. By examining the stripes with an atomic force microscope, one is effectively using the bar code to read out the data that were encoded on the input DNA strand. This visual means of reading out the DNA sequence could greatly speed up the readout stage of DNA-based computing and might also be used for mapping mutations. In an exciting recent extension of the use of long DNA strands, Paul Rothemund of Caltech has used a viral strand of about 7,000 nucleotides to build complex patterns, including a smiley face and a map of the Western Hemisphere.

Chengde Mao, now at Purdue University, and I have made two-dimensional patterns from DNA parallelograms similar to our stick polyhedra. Copies of this unit can be joined to

ALICE Y. CHEN (*illustrations*); NADRIAN C. SEEMAN (*micrographs*)

form a crystal that extends like a waffle in two dimensions. One can tune the sizes of the cavities in the array by changing the dimensions of the parallelograms. Although individual branched junctions are floppy, arranging four of them at the corners of a parallelogram results in a well-behaved unit in a parallelogram array.

Nanomachines

CENTRAL TO NANOTECHNOLOGY are molecular-scale machines. DNA has proved to be very useful for constructing these machines. We have built several devices from DNA, but here I will focus on two that have well-defined structures. In both cases, the mechanism is based on a structural transition of DNA molecules—a change from one conformation (such as the usual double helix) to another.

Conventional DNA is a right-handed helix. Imagine walking up a spiral staircase with your left hand on the inner banister and your right hand on the outer one. Such a staircase is a right-handed helix. Conventional right-handed DNA is called B-DNA and is the most energetically favored structure in typical aqueous conditions.

Double-helical DNA can also assume a number of different structures depending on its base sequence and the chemical species present in the solution in which it is immersed. One is Z-DNA, whose structure was first characterized in 1979 by Alexander Rich and his colleagues at the Massachusetts Institute of Technology [*see box at top on page 33*]. Z-DNA is a left-handed DNA structure.

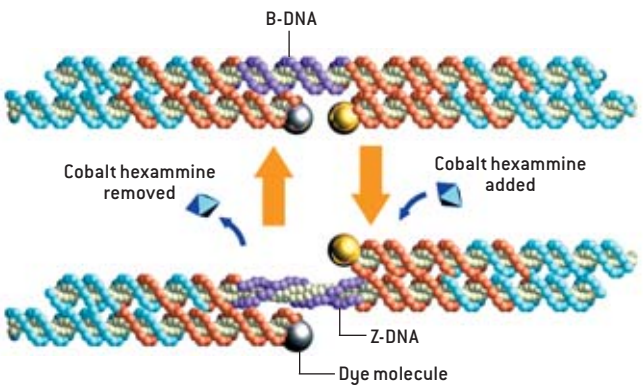
To make Z-DNA typically requires a stretch of alternating cytosine and guanine bases. The DNA backbone includes negatively charged phosphate groups, and these come close together in the Z-DNA structure. This formation is favored only if the charges of the phosphates can be screened from one another by an aqueous environment containing either a high concentration of salt or a special "effector" species, such as cobalt hexammine, $Co(NH_3)_6^{+++}$, that does the same job at a much lower concentration. The cytosine-guanine sequence requirement lets us control *where* on a DNA molecule the B-Z transition takes place (and hence *what* our machine does), and the environmental requirement lets us control *when* the transition (and hence the machine action) occurs.

My N.Y.U. colleagues Weiqiong Sun and Zhiyong Shen, Mao and I built a device consisting of two DX molecules connected by a shaft of double-helical DNA [*see illustration at right*]. In the middle of the shaft is a sequence of 20 pairs that can adopt the Z-structure in the appropriate conditions. In ordinary conditions, every part of the device will form B-DNA and the two DX molecules will both be on the same side of the shaft's axis. When cobalt hexammine is added to the solution, the central part of the shaft converts to Z-DNA and one DX molecule rotates about 3.5 turns relative to the other; the odd half-turn means that they are now on opposite sides of the shaft's axis. Removal of the cobalt hexammine reverts the device back to its original structure. We demonstrated that the motion was taking place by using spectroscopy involving two colored dyes attached to the DX molecules.

This B-Z device is quite robust, but it suffers from a flaw. Were a bunch of different B-Z devices incorporated into a larger superstructure (for example, one of the two-dimensional lattices discussed earlier), the entire structure would have only two states: every machine in the B state or every one in the Z state. To control a collection of machines individually requires devices with independent triggers. With DNA, of course, there is a natural way to do this, by using DNA strands as the triggers and having a different base sequence trigger each machine.

A CRUCIAL GOAL for nanotechnology based on DNA is to extend the successes in two dimensions TO THREE DIMENSIONS.

To implement this scheme, Hao Yan, now at Arizona State University, Xiaoping Zhang of N.Y.U., Shen and I devised a system that changes shape when different strands bind to it. The system consists of two parallel DNA double helices that each reduce to a single strand in a central crossover region. The crossover region can assume two different states according to which particular strands have been added to the solution to bind to the single-strand sections [*see box on next page*]. The two states of the device are called PX ("paranemic crossover") and JX ("juxtaposed"). When the device is in the PX state, the two helices on one side of the central junction are rotated about a half-turn from their positions in the JX state.



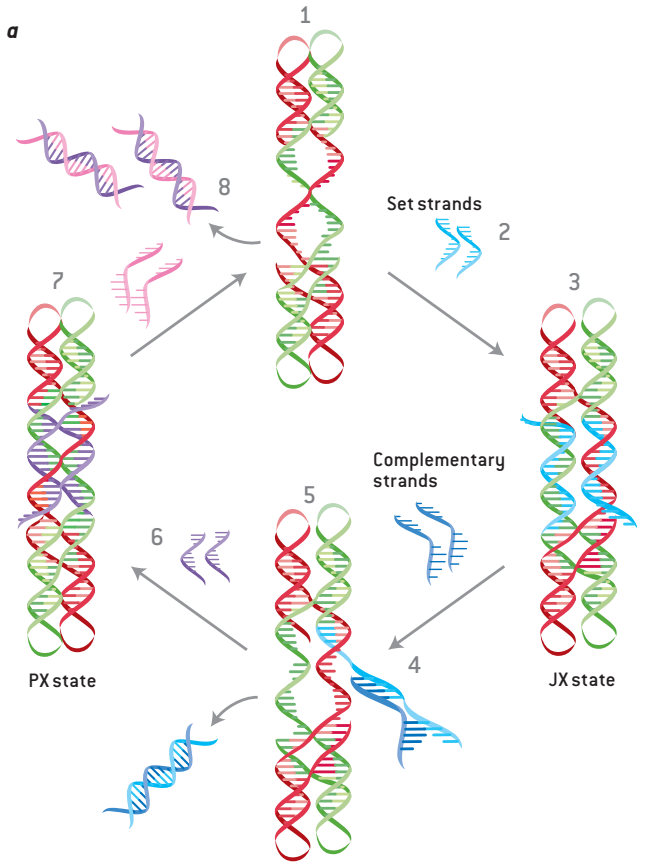
NANOMECHANICAL B-Z DEVICE that demonstrates controlled movement is made of two DX units (*blue* and *orange*) joined by a shaft of 20 base pairs (*purple*). Two colored dye molecules (*silver* and *gold spheres*) highlight the positions of the DX molecules. In the B state (*top*), the shaft is ordinary right-handed B-DNA and both DX molecules are on the same side. When cobalt hexammine is added to the solution, the shaft converts to left-handed Z-DNA [*see box at top on page 33*] and the DX units rotate through 3.5 turns relative to each other, ending up on opposite sides of the shaft.

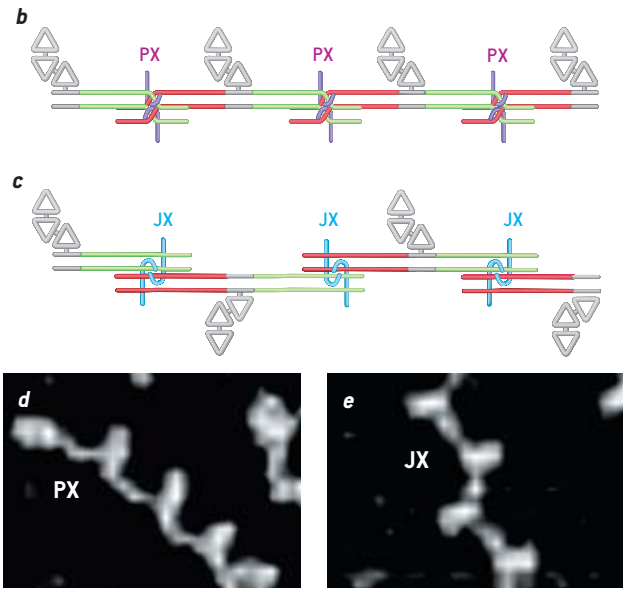
KEN EWARD *BioGrafx*

USING DNA AS A TRIGGER

Individually controllable DNA device is switched between two states (*a, steps 1–8*) by the addition and removal of specific stretches of DNA called set strands. The naked device consists of four double helices connected in the middle by two unpaired DNA strands (*1*). When the light-blue set strands are added (*2*), they bind to the unpaired strands in a way that forces the device into the "juxtaposed" (JX) state (*3*). In this state, the red and green helices are on the same side, top and bottom. The light-blue strands are stripped away when complementary strands are added (*4*), leaving the device naked again (*5*). Now the purple set strands are added (*6*), which bind in a different way, forcing the device into the so-called paranemic crossover (PX) state (*7*). This step rotates the lower part of the device, putting the red and green helices on the opposite sides. The machine's cycle can continue with the stripping away of the purple strands (*8*) and the reintroduction of the light-blue strands.

The functioning of this device was verified by connecting copies of it in a chain, with large trapezoid-shaped pieces of DNA attached as markers. When the devices are in the PX state (*b, below*), all the trapezoids are on the same side. When all the devices are in the JX state (*c*), the trapezoids alternate sides. Atomic force microscopy revealed precisely this pattern of behavior (*d, e*).





Adding a particular pair of strands (called set strands) to the solution puts the device in the JX state by binding to the central region without crossing over. To change to the PX state, we must first remove these set strands. In 2000 Bernard Yurke and his colleagues at Lucent Technologies showed that a strand can be extracted from DNA by binding the strand's full complement to it. To implement this process, our set strands have short ends that remain unpaired with the machine. When we add a full complementary strand to the solution, it begins by joining to the unpaired extension and then strips off the rest of the set strand from the device.

With the first set strands removed from the frame, we can then add different set strands, which bind to the central region and cross over there. That binding turns the two double helices and puts the device in the PX state. The process can be reversed by removing the second set strands and adding back the first ones. In this way, the double helices can be turned back and forth at will. A number of different PX-JX devices can be operated independently by adding and removing set strands designed for their individual binding regions.

We used atomic force microscopy to verify how our device moved. We made a long chain of these devices and connected a large trapezoid-shaped DNA unit to one side of each device. When all the devices are in the PX state, the trapezoids lie on the same side of the chain. When all are in the JX state, the trapezoids alternate sides, in a zigzag pattern.

In 2000 Yurke and his colleagues demonstrated nanoscopic "tweezers" made of three strands of DNA. Set strands, which Yurke calls fuel strands, opened and closed the tweezers. Other researchers have used similar methods to switch on the activity of ribozymes—enzymes made of RNA. In 1998 Michael P. Robinson and Andrew D. Ellington of the University of Texas

ALICE Y. CHEN (*illustrations*); NADRIAN C. SEEMAN (*micrographs*)

at Austin demonstrated a 10,000-fold enhancement of a ribozyme's activity by the addition of an appropriate set strand, which bound to the ribozyme, changing its conformation.

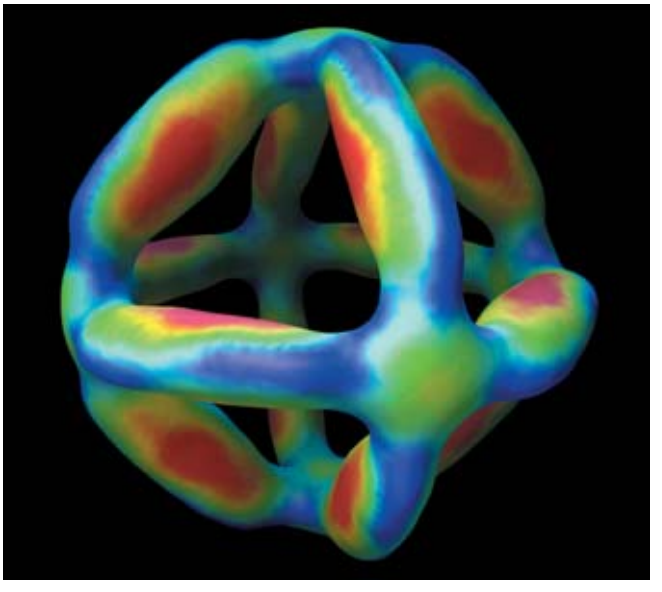
A key goal has been to incorporate DNA devices within the framework arrays. This is the first step toward DNA-based nanorobotics involving complex motions and a diversity of states. Together with Baoquan Ding, now at the Molecular Foundry at Lawrence Berkeley National Laboratory, I reported achieving this key goal in late 2006. And along with Shiping Liao, now at Barr Pharmaceuticals, I reported a multiple-state system using the PX-JX device that translates DNA signals into polymer assembly instructions. Using devices similar to the ones described here will allow us to assemble new materials with high precision. Lei Zhu, now at Florida State University, James W. Canary and Philip S. Lukeman of N.Y.U., and I recently assembled a prototype made from a small piece of nylon on a nucleic acid backbone. Someday we expect to be able to make new polymers and combinations of polymers with specific topologies (windings of their backbones) and properties.

The Future

A CRUCIAL GOAL for nanotechnology based on DNA is to extend the successes in two dimensions to three dimensions. When that has been accomplished, we will have demonstrated the ability to design solid materials by specifying a series of DNA sequences and then combining them. If the systems are highly ordered, then the crystallographic experiments involving molecules held within a regularly repeating framework mentioned earlier will be feasible.

Achieving this goal primarily entails the use of DNA as a programmable component, but neither crystallography nor nanoelectronics can rely on DNA alone. For instance, nanoelectronic components, such as metallic nanoparticles or carbon nanotubes, will have to be combined with DNA molecules in systems and liquid solutions that are compatible with both the DNA and the other components. Given the diverse chemical nature of these molecules, it has not been easy to organize metallic nanoparticles in DNA arrays. My group and those of Richard A. Kiehl of the University of Minnesota and Hao Yan, however, have been successful in this effort. In addition, even if the nanoelectronics can be constructed by DNA self-assembly, the nanomachines ultimately need to interact with the macroscopic world in a manner that is more sophisticated than the addition and removal of set strands from a solution. This challenge is likely to be formidable.

A nanotechnological dream machine is one that can replicate. Unlike linear DNA, however, branched DNA does not lend itself readily to self-replication. Yet in late 2003 William M. Shih, Joel D. Quispe and Gerald F. Joyce of the Scripps Research Institute in La Jolla, Calif., took an exciting first step toward self-replicating DNA objects. They built an octahedron from one long strand of DNA (about 1,700 bases), using five short "helper" strands to complete the assembly [*see illustration below*]. Each edge of the octahedron is made of two interlinked DNA double helices—a series of DX and PX molecules. The edges were each about 14 nanometers long, or about four turns of a double helix. A folded octahedron cannot reproduce, but in the unfolded state, the long strand is readily cloned millions of times by a standard biotechnology process called PCR (polymerase chain reaction). It is still a far cry from the replication achieved by every living organism, but by the time the Watson-Crick centenary comes around, we should have DNA-based machines that do as well. SA



DNA OCTAHEDRON shown here was built out of one long strand of DNA and five short "helper" strands. Each strut consists of two parallel, interlinked double helices. The image was reconstructed by combining data from cryo-electron microscope images of more than 600 octahedra. The colors represent relative electron density: red high and blue low.

MICHAEL PIQUE *Scripps Research Institute*

MORE TO EXPLORE

A DNA-Fuelled Molecular Machine Made of DNA. Bernard Yurke, Andrew J. Turberfield, Allen P. Mills, Jr., Friedrich C. Simmel and Jennifer L. Neumann in *Nature,* Vol. 406, pages 605–608; August 10, 2000.

Logical Computation Using Algorithmic Self-Assembly of DNA Triple Crossover Molecules. Chengde Mao, Thomas H. LaBean, John H. Reif and Nadrian C. Seeman in *Nature,* Vol. 407, pages 493–496; September 28, 2000. (*Erratum: Nature,* Vol. 408, page 750; December 7, 2000.)

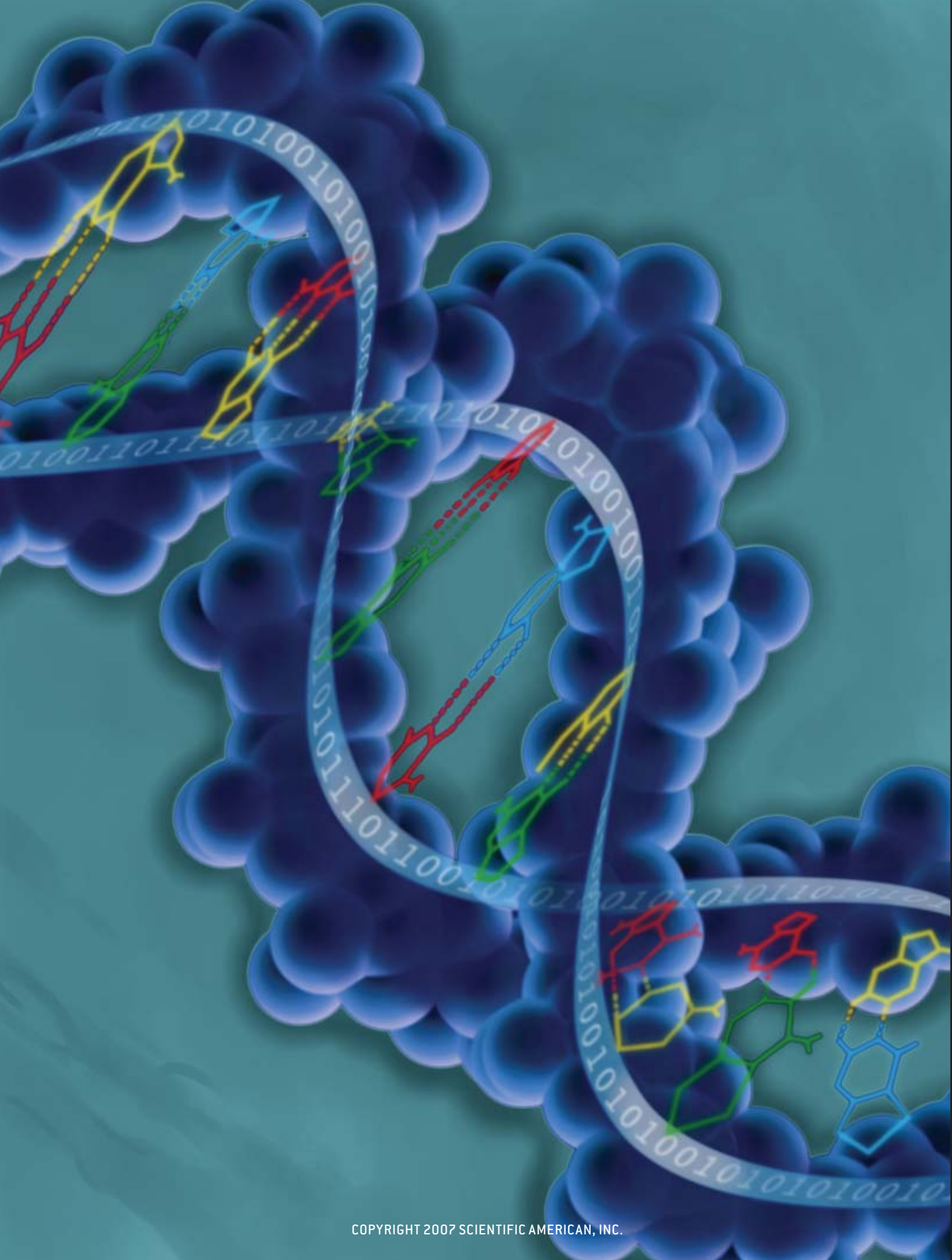
A Robust DNA Mechanical Device Controlled by Hybridization Topology. Hao Yan, Xiaoping Zhang, Zhiyong Shen and Nadrian C. Seeman in *Nature,* Vol. 415, pages 62–65; January 3, 2002.

DNA in a Material World. Nadrian C. Seeman in *Nature,* Vol. 421, pages 427–431; January 23, 2003.

DNA as an Engineering Material. Andrew Turberfield in *Physics World,* Vol. 16, No. 3, pages 43–46; March 2003.

A 1.7-Kilobase Single-Stranded DNA That Folds into a Nanoscale Octahedron. William M. Shih, Joel D. Quispe and Gerald F. Joyce in *Nature,* Vol. 427, pages 618–621; February 12, 2004.

Nadrian C. Seeman's laboratory Web site: **http://seemanlab4.chem.nyu.edu/**



BRINGING DNA COMPUTERS TO LIFE

Tapping the computing power of biological molecules gives rise to tiny machines that can speak directly to living cells

By Ehud Shapiro and Yaakov Benenson

When British mathematician Alan Turing conceived the notion of a universal programmable computing machine, the word "computer" typically referred not to an object but to a human being. It was 1936, and people with the job of computer, in modern terms, crunched numbers. Turing's design for a machine that could do such work instead—one capable of computing any computable problem—set the stage for theoretical study of computation and remains a foundation for all of computer science. But he never specified what materials should be used to build it.

Turing's purely conceptual machine had no electrical wires, transistors or logic gates. Indeed, he continued to imagine it as a person, one with an infinitely long piece of paper, a pencil and a simple instruction book. His tireless computer would read a symbol, change the symbol, then move on to the next symbol, according to its programmed rules, and would keep doing so until no further rules applied. Thus, the electronic computing machines made of metal and vacuum tubes that emerged in the 1940s and later evolved silicon parts may be the only "species" of nonhuman computer most people have ever encountered, but theirs is not the only possible form a computer can take.

Living organisms, for instance, also carry out complex physical processes under the direction of digital information. Biochemical reactions and ultimately an entire organism's operation are ruled by instructions stored in its genome, encoded in sequences of nucleic acids. When the workings of biomolecular machines inside cells that process DNA and RNA are compared to Turing's machine, striking similarities emerge: both systems process information stored in a string of symbols taken from a fixed alphabet, and both operate by moving step by step along those strings, modifying or adding symbols according to a given set of rules.

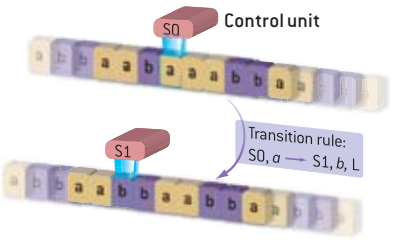
These parallels have inspired the idea that biological molecules could one day become the raw material of a new computer species. Such biological computers would not necessarily offer greater power or performance in traditional computing tasks. The speed of natural molecular machines such as the ribosome is only hundreds of operations a second, compared with billions of gate-switching operations a second in some electronic devices. But the molecules do have a unique ability: they speak the language of living cells.

The promise of computers made from biological molecules lies in their potential to operate within a biochemical environment, even within a living organism, and to interact with that environment through inputs and outputs in the form of other biological molecules. A biomolecular computer might act as

MATT COLLINS

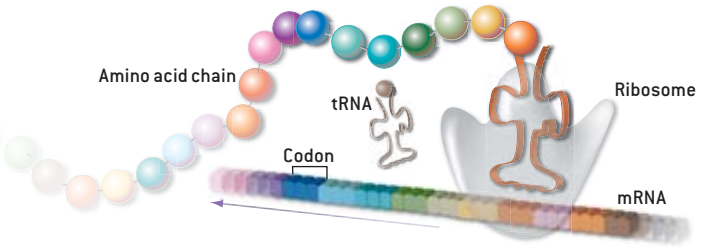
COMPUTING MACHINES: CONCEPTUAL AND NATURAL

Mathematician Alan Turing envisioned the properties of a mechanical computer in 1936, long before molecule-scale machines within cells could be seen and studied. As the workings of nature's tiny automata were later revealed, striking similarities to Turing's concept emerged: both systems store information in strings of symbols, both process these strings in stepwise fashion, and both modify or add symbols according to fixed rules.



TURING MACHINE

This hypothetical device operates on an information-encoding tape bearing symbols such as *a* and *b*. A control unit with read/write ability processes the tape, one symbol position at a time, according to instructions provided by transition rules, which note the control unit's own internal state. Thus, the transition rule in this example dictates that if the control unit's state is 0 (S0), and the symbol read is *a*, then the unit should change its state to 1 (S1), change the symbol to *b* and move one position to the left (L).



BIOLOGICAL MACHINE

An organelle found in cells, the ribosome reads information encoded in gene transcripts known as messenger RNAs (mRNAs) and translates it into amino acid sequences to form proteins. The symbolic alphabet of mRNA is made up of nucleotide trios called codons, each of which corresponds to a specific amino acid. As the ribosome processes the mRNA strand, one codon at a time, helper molecules called transfer RNAs (tRNAs) deliver the correct amino acid. The tRNA confirms the codon match, then releases the amino acid to join the growing chain.

an autonomous "doctor" within a cell, for example. It could sense signals from the environment indicating disease, process them using its preprogrammed medical knowledge, and output a signal or a therapeutic drug.

Over the past eight years we have been working toward realizing this vision. We have already succeeded in creating a biological automaton made of DNA and proteins able to diagnose in a test tube the molecular symptoms of certain cancers and "treat" the disease by releasing a therapeutic molecule. This proof of concept was exciting, both because it has potential future medical applications and because it is not at all what we originally set out to build.

Models to Molecules

ONE OF US (Shapiro) began this research with the realization that the basic operations of certain biomolecular machines within living cells—recognition of molecular building blocks, cleavage and ligation of biopolymer molecules, and movement along a polymer—could all be used, in principle, to construct a universal computer based on Turing's conceptual machine. In essence, the computational operations of such a Turing machine would translate into biomolecular terms as one "recognition," two "cleavages," two "ligations," and a move to the left or right.

Charles Bennett of IBM had already made similar observations and proposed a hypothetical molecular Turing machine in 1982. Interested in the physics of energy consumption, he speculated that molecules might one day become the basis of more energy-efficient computing devices.

The first real-world demonstration of molecules' computational power came in 1994, when Leonard M. Adleman of the University of Southern California used DNA to solve a problem that is always cumbersome for traditional computer algorithms. Known as the Hamiltonian path or the traveling salesman problem, its goal is to find the shortest path among cities connected by airline routes that passes through every city exactly once. By creating DNA molecules to symbolically represent the cities and flights and then combining trillions of these in a test tube, he took advantage of the molecules' pairing affinities to achieve an answer within a few minutes. Unfortunately, it took him considerably longer to manually fish the molecules representing the correct solution out of the mixture using the laboratory tools available to him at the time. Adleman looked forward to new technologies that would enable the creation of a more practical molecular computer.

"In the future, research in molecular biology may provide improved techniques for manipulating macromolecules," Adleman wrote in a seminal 1994 scientific paper describing the DNA experiment. "Research in chemistry may allow for

Overview/*Living Computers*

- Natural molecular machines process information in a manner similar to the Turing machine, an early conceptual computer.
- A Turing-like automaton built from DNA and enzymes can perform computations, receive input from other biological molecules and output a tangible result, such as a signal or a therapeutic drug.
- This working computer made from the molecules of life demonstrates the viability of its species and may prove a valuable medical tool.

LUCY READING-IKKANDA

the development of synthetic designer enzymes. One can imagine the eventual emergence of a general purpose computer consisting of nothing more than a single macromolecule conjugated to a ribosomelike collection of enzymes which act upon it."

Devising a concrete logical design for just such a device, one that could function as the fundamental "operational specification" for a broad class of future molecular computing machines, became Shapiro's goal. By 1999 he had a mechanical model of the design made from plastic parts. We then joined forces to translate that model into real molecules.

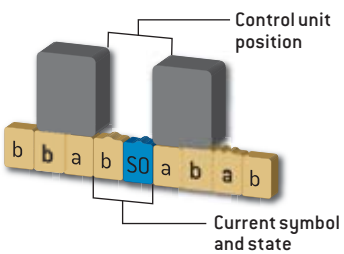
Rather than attacking the ultimate challenge of building a full-fledged molecular Turing machine head-on, however, we agreed to first attempt a very simplified Turing-like machine known as a finite automaton. Its sole job would be to determine whether a string of symbols or letters from a two-letter alphabet, such as *a* and *b,* contained an even number of *b*'s. This task can be achieved by a finite automaton with just two states and a "program" consisting of four statements called transition rules. One of us (Benenson) had the idea to use a double-stranded DNA molecule to represent the input string, four more short double-stranded DNA molecules to represent the automaton's transition rules, or "software," and two natural DNA-manipulating enzymes, *Fok*I and ligase, as "hardware."

The main logical problem we had to solve in its design was how to represent the changing intermediate states of the computation, which consist of the current internal state of the automaton and a pointer to the symbol in the input string being processed. We accomplished this with a neat trick: in each step of the computation the enzymatic hardware was actually "digesting" the input molecule, cleaving the current symbol being processed and exposing the next one. Because the symbol could be cleaved in two different locations, each resulting version of it could represent, in addition to the symbol itself, one of two possible states of the computation. Interestingly, we discovered later that this last element was similar to a

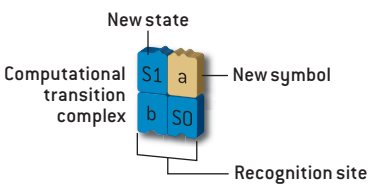
MOLECULAR TURING MACHINE MODEL

A Turing machine made of biomolecules would employ their natural ability to recognize symbols and to join molecular subunits together or cleave their bonds. A plastic model built by one of the authors (*right*) serves as a blueprint for such a system. Yellow "molecule" blocks carry the symbols. Blue software molecules indicate a machine state and define transition rules. Protrusions on the blocks physically differentiate them.

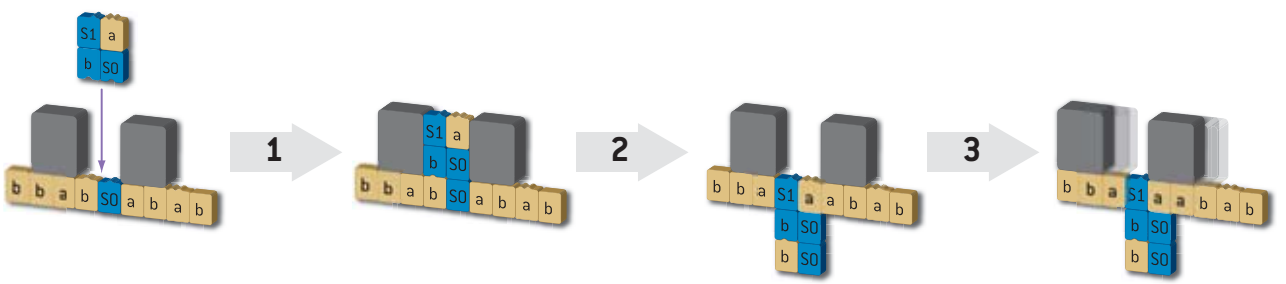
HOW IT WORKS



The machine operates on a string of symbol molecules. In its control unit position at the center, both a symbol and the machine's current state are defined.



One "computational transition" is represented by a molecule complex containing a new state and symbol for the machine and a recognition site to detect the current state and symbol. The example shown represents a transition rule: "If current state is S0 and current symbol is *b,* change state to S1 and symbol to *a,* then move one step to the left."



A free-floating computational transition complex slides into the machine's control unit (*1*). The molecule complex binds to and then displaces the current symbol and state (*2*). The control unit can move one position to the left to accommodate another transition complex (*3*). The process repeats indefinitely with new states and symbols as long as transition rules apply.

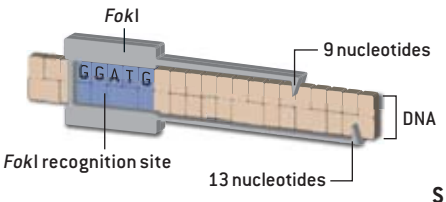
LUCY READING-IKKANDA (*illustrations*); WEIZMANN INSTITUTE OF SCIENCE (*photograph*)

BUILDING A MOLECULAR AUTOMATON

Because living organisms process information, their materials and mechanisms lend themselves readily to computing. The DNA molecule exists to store data, written in an alphabet of nucleotides. Cellular machinery reads and modifies that information using enzymes and other molecules. Central to this operating system are chemical affinities among molecules allowing them to recognize and bind with one another. Making molecules into a Turing-like device, therefore, means translating his concepts into their language.

A simple conceptual computer called a finite automaton can move in only one direction and can read a series of symbols, changing its internal state according to transition rules. A two-state automaton could thus answer a yes-no question by alternating between states designated 1 and 0. Its state at the end of the calculation represents the result.

Raw materials for a molecular automaton include DNA strands in assorted configurations to serve as both input and software and the DNA-cleaving enzyme *Fok*I as hardware. Nucleotides, whose names are abbreviated A, C, G and T, here encode both symbols and the machine's internal state.



HARDWARE

The *Fok*I enzyme (*gray*) always recognizes the nucleotide sequence GGATG (*blue*) and snips a double DNA strand at positions 9 and 13 nucleotides downstream of that recognition site.

SOFTWARE

Transition rules are encoded in eight short double-stranded DNA molecules containing the *Fok*I recognition site (*blue*), followed by spacer nucleotides (*green*) and a single-stranded sticky end (*yellow*) that will join to its complementary sequence on an input molecule.

SYMBOL AND STATE

Combinations of symbols *a*, *b* or *terminator* (*t*) and machine states 1 or 0 are represented by four-nucleotide sequences. Depending on how the five-nucleotide sequence TGGCT is cleaved into four nucleotides, for example, it will denote symbol *a* and a state of either 1 or 0.

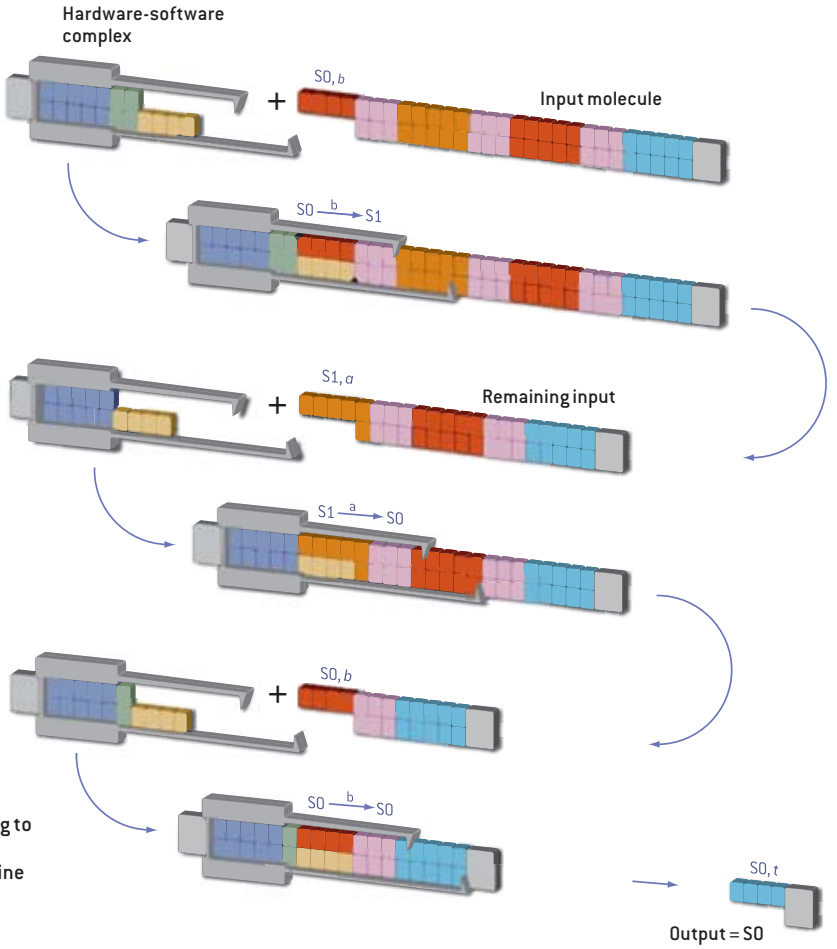
AUTONOMOUS COMPUTATION

A hardware-software complex recognizes its complementary state/symbol combination on the input molecule. The molecules join to form a hardware-software-input complex, then *Fok*I cleaves the input molecule to expose the next symbol.

A new hardware-software complex, representing a different transition rule, recognizes the next state and symbol on what remains of the input molecule.

Reactions continue until no rule applies or the terminator symbol is revealed.

In this example, computational cleavages leading to the final output (*far right*) have produced a four-nucleotide terminator symbol indicating a machine state of 0, the calculation's result.



LUCY READING-IKKANDA

design that Paul Rothemund, a former student of Adleman, once proposed for a molecular Turing machine.

Remarkably, the resulting computer that our team announced in 2001 was autonomous: once the input, software and hardware molecules were placed in an appropriate buffer solution in a test tube, computation commenced and proceeded iteratively to completion without any human intervention.

As we tested this system, we also realized that it not only solved the original problem for which we had intended it—determining whether a symbol occurs an even number of times in a string—it could do more. A two-state, two-symbol finite automaton has eight possible symbol-state-rule combinations (2^3), and because our design was modular, all eight possible transition rules could be readily implemented using eight different transition molecules. The automaton could therefore be made to perform different tasks by choosing a different "program"—that is, a different mix of transition molecules.

In trying a variety of programs with our simple molecular automaton, we also found a way of further improving its performance. Among our tests was an omission experiment, in which the automaton's operation was evaluated with one molecular component removed at a time. When we took away ligase, which seals the software molecule to the input molecule to enable its recognition and cleavage by the other enzyme, *Fok*I, the computation seemed to make some progress nonetheless. We had discovered a previously unknown ability of *Fok*I to recognize and cleave certain DNA sequences, whether or not the molecule's two strands were sealed together.

The prospect of removing ligase from our molecular computer design made us quite happy because it would immediately reduce the required enzymatic hardware by 50 percent. More important, ligation was the only energy-consuming operation in the computation, so sidestepping it would allow the computer to operate without an external source of fuel. Finally, eliminating the ligation step would mean that software molecules were no longer being consumed during the computation and could instead be recycled.

The ligase-free system took our group months of painstaking effort and data analysis to perfect. It was extremely inefficient at first, stalling out after only one or two computational steps. But we were driven by both the computational and biochemical challenges, and with help and advice from Rivka Adar and other colleagues, Benenson finally found a solution. By making small but crucial alterations to the DNA sequences used in our automaton, we were able to take advantage of *Fok*I's hitherto unknown capability and achieve a quantum leap in the computer's performance. By 2003 we had an autonomous, programmable computer that could use its input molecule as its sole source of fuel [*see box on opposite page*]. In principle, it could therefore process any input molecule, of any length, using a fixed number of hardware and software molecules without ever running out of energy.

And yet from a computational standpoint, our automaton still seemed like a self-propelled scooter compared with the Rolls-Royce of computers on which we had set our sights: the biomolecular Turing machine.

Diagnosing Disease

BECAUSE THE TWO-STATE finite automaton was too simple to be of any use in solving complex computational problems, we considered it nothing more than an interesting demonstration of the concept of programmable, autonomous biomolecular computers, and we decided to move on. Focusing our efforts for a while on trying to build more complicated automata, however, we soon ran up against the problem recognized by Adleman: the "designer enzymes" he had yearned for a decade earlier still did not exist.

No known naturally occurring enzyme or enzyme complex can perform the specific recognitions, cleavages and ligations, in sequence and in tandem, with the flexibility needed to realize the Turing machine design. Natural enzymes will have to be customized or entirely new synthetic enzymes engineered. Because science does not yet have this ability, we found ourselves with a logical design specification for a biomolecular Turing machine but forced to wait until the parts needed to build it are invented.

That is why we returned to our two-state automaton to see if we could at least find something useful for it to accomplish. With medical applications already in mind, we wondered if the device might be able to perform some kind of simple diagnosis, such as determining whether a set of conditions representing a particular disease is present.

For this task, just two states suffice: we called one state *yes* and the other *no*. The automaton would begin the computation in the *yes* state and check one condition at a time. If a condition on its checklist were present, the *yes* state would hold, but if any condition were not present, the automaton would change to the *no* state and remain that way for the rest of the computational process. Thus, the computation would end in *yes* only if all the disease conditions held, but if one condition were not met the "diagnosis" would be negative.

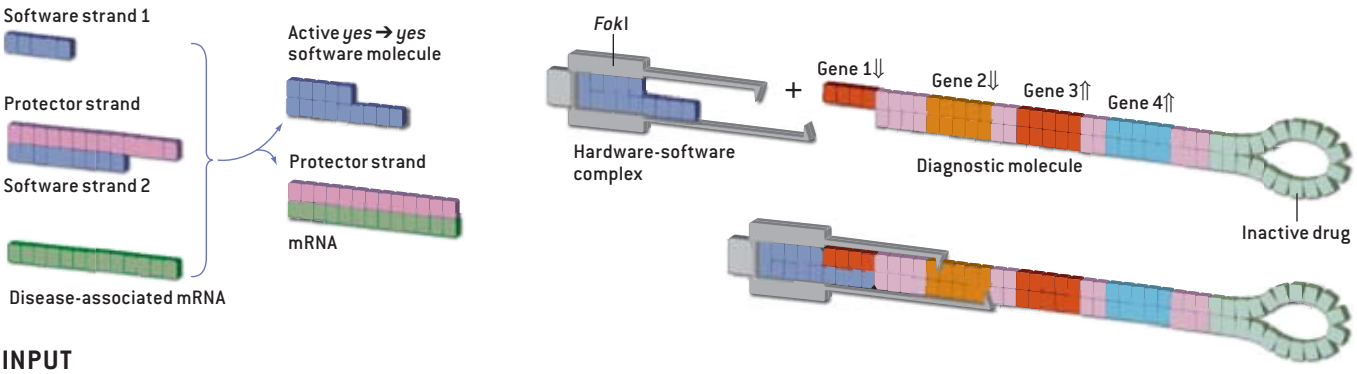
To make this logical scheme work, we had to find a way to connect the molecular automaton to its biochemical environ-

THE AUTHORS

EHUD SHAPIRO and *YAAKOV BENENSON* began collaborating to build molecular automata in 1999. Shapiro is a professor in the departments of computer science and biological chemistry at the Weizmann Institute of Science in Rehovot, Israel, where he holds the Harry Weinrebe Professorial Chair. He was already an accomplished computer scientist and software pioneer with a growing interest in biology in 1998 when he first designed a model for a molecular Turing machine. Benenson, just completing a master's degree in biochemistry at the Technion in Haifa, became Shapiro's Ph.D. student the following year. Now a Bauer Fellow at Harvard University's Faculty of Arts and Sciences Center for Systems Biology, Benenson is working on new molecular tools to probe and affect live cells.

DNA DOCTOR

Having shown that an automaton made from DNA and enzymes can perform abstract yes-or-no computations, the authors sought to give the device a practical query that it might face inside a living cell: Are indicators of a disease present? If the answer is yes, the automaton can output an active drug treatment. The basic computational concept is unchanged from the earlier design: complexes of "software" transition molecules and enzymatic "hardware" process symbols within a diagnostic molecule, cleaving it repeatedly to expose subsequent symbols. In addition, the new task requires a means for disease indicators to create input for the computation and mechanisms to confirm the diagnosis and deliver treatment.



INPUT

Gene transcripts called messenger RNAs (mRNAs) serve as disease indicators. By interacting with software molecules, mRNAs influence which of them is ultimately used in the computation. In this example, the two strands of a *yes* → *yes* transition molecule start out separated, with one bound to a single protector strand. The protector has a strong affinity for the disease-associated mRNA, however. If that mRNA is present, the protector will abandon its software strand to bind to the mRNA. The single software strands will then bind to one another, forming an active *yes* → *yes* transition molecule.

COMPUTATION

Complexes of transition-molecule software and *Fok*I enzymatic hardware process a series of symbols within the diagnostic molecule that represent underactivity (⇓) or overactivity (⇑) by specific genes. The automaton starts the computation in a *yes* state, and if all disease indicators are present, it produces a positive diagnosis. If any symptom is missing, the automaton transitions to *no* and remains in that state.

ment so that it could sense whether specific disease conditions were present. The general idea that the environment could affect the relative concentrations of competing transition molecules—and thus affect the computation—had already been suggested in the blueprint for the molecular Turing machine. To apply this principle to sense disease symptoms, we had to make the presence or absence of a disease indicator a determinant of the concentration of software molecules that testify for the symptom.

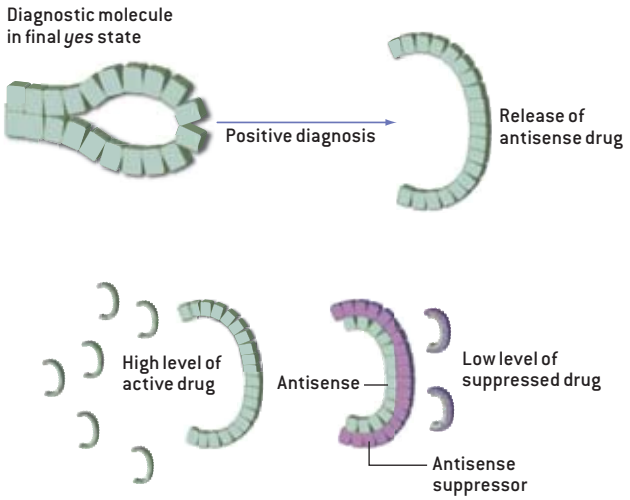
Many cancers, for example, are characterized by abnormal levels of certain proteins in the cell as a result of specific genes either overexpressing or underexpressing their encoded protein. When a gene is expressed, enzymes in the cell's nucleus copy its sequence into an RNA version. This molecular transcript of the gene, known as messenger RNA (mRNA), is then read by a structure called the ribosome that translates the RNA sequence into a string of amino acids that will form the protein. Thus, higher- or lower-than-normal levels of specific mRNA transcripts can reflect gene activity.

Benenson devised a system wherein some transition molecules would preferentially interact with these mRNA sequences. The interaction, in turn, would affect the transition molecules' ability to participate in the computation. A high level of mRNA representing a disease condition would cause the *yes* → *yes* transition molecules to predominate, increasing the probability that the computer would find the symptom to be present [*see box above*]. In practice, this system could be applied to any disease associated with abnormal levels of proteins resulting from gene activity, and it could also be adapted to detect harmful mutations in mRNA sequences.

Once we had both an input mechanism that could sense disease symptoms and the logical apparatus to perform the diagnosis, the next question became, What should the computer do when a disease is diagnosed? At first, we considered having it produce a visible diagnostic signal. In the molecular world, however, producing a signal and actually taking the next logical step of administering a drug are not that far apart. Binyamin Gil, a graduate student on our team, proposed and implemented a mechanism that enables the computer to release a drug molecule on positive diagnosis.

Still, our plan was not complete. Perhaps the central question in computer hardware design is how to build a reliable system from unreliable components. This problem is not unique to biological computers—it is an inherent property of complex systems; even mechanical devices become more unreliable as scale diminishes and the number of components increases. In our case, given the overall probabilistic nature of the computation and the imprecise behavior of biomolecular elements, some computations would inevitably end with a positive diagnosis even if several or all of the disease symptoms were absent, and vice versa. Fortunately, this probabilistic behavior is measurable and repeatable, so we could compensate for it with a system of checks and balances.

LUCY READING-IKKANDA



OUTPUT

After a positive diagnosis, final cleavage of the diagnostic molecule releases the treatment, a single-stranded so-called antisense DNA molecule (*top*). To compensate for diagnostic errors, the authors also created negative versions of the diagnostic molecules to perform parallel computations. When disease indicators are absent, these automata release a drug suppressor. With thousands of both types of diagnostic molecules computing simultaneously, the majority will make the correct diagnosis, and either the antisense molecule will outnumber its suppressors (*bottom*), or vice versa.

We created two types of computation molecules: one designed to release a drug when the computation terminates in the *yes* state, the other to release a suppressor of that same drug when the computation terminates in *no*. By changing the relative concentrations of the two types of molecules, we could have fine control over the threshold of diagnostic certainty that would trigger administration of an active drug.

Human physicians make this kind of decision whenever they weigh the risk to a patient of a possible disease against the toxicity of the treatment and the certainty of the diagnosis. In the future, if our molecular automaton is sent on a medical mission, it can be programmed to exercise similar judgment.

Dawn of a New Species

AS IT TURNED OUT, our simple scooter carried us much further than we had believed it could and in a somewhat different direction than we had first imagined. So far our biomolecular computer has been demonstrated only in a test tube. Its biological environment was simulated by adding different concentrations of RNA and DNA molecules and then placing all the automaton components in the same tube. Now our goals are to make it work inside a living cell, to see it compute inside the cell and to make it communicate with its environment.

Just delivering the automaton into the cell is challenging because most molecular delivery systems are tailored for either DNA or protein. Our computer contains both, so we are trying to find ways to administer these molecules in tandem.

Another hurdle is finding a means of watching all aspects of the computation as they occur within a cell, to confirm that the automaton can work without the cell's activities disrupting computational steps or the computer's components affecting cellular behavior in unintended ways. And finally, we are exploring alternative means of linking the automaton to its environment. Very recent cancer research suggests that microRNAs, small molecules with important regulatory functions inside cells, are better indicators of the disease, so we are redesigning our computer to "talk" to microRNA instead of mRNA.

Although we are still far from applying our device inside living cells, let alone in living organisms, we already have the important proof of concept. By linking biochemical disease symptoms directly with the basic computational steps of a molecular computer, our test-tube demonstration affirmed that an autonomous molecular computer can communicate with biological systems and perform biologically meaningful assessments. Its input mechanism can sense the environment in which it operates; its computation mechanism can analyze that environment; and its output mechanism can affect the environment in an intelligent way based on the result of its analysis.

Thus, our automaton has delivered on the promise of biomolecular computers to enable direct interaction with the biochemical world. It also brings computational science full circle back to Turing's original vision. The first computing machines had to deviate from his concept to accommodate the properties of electronic parts. Only decades later, when molecular biologists began revealing the operations of tiny machines inside living cells, did computer scientists recognize a working system similar to Turing's abstract idea of computation.

This is not to suggest that molecules are likely to replace electronic machines for all computational tasks. The two computer species have different strengths and can easily coexist. Because biomolecules can directly access data encoded in other biomolecules, however, they are intrinsically compatible with living systems in a way that electronic computers will never be. And so we believe our experiments suggest that this new computer species is of fundamental importance and will prove itself valuable for a wide range of applications. The biomolecular computer has come to life. SA

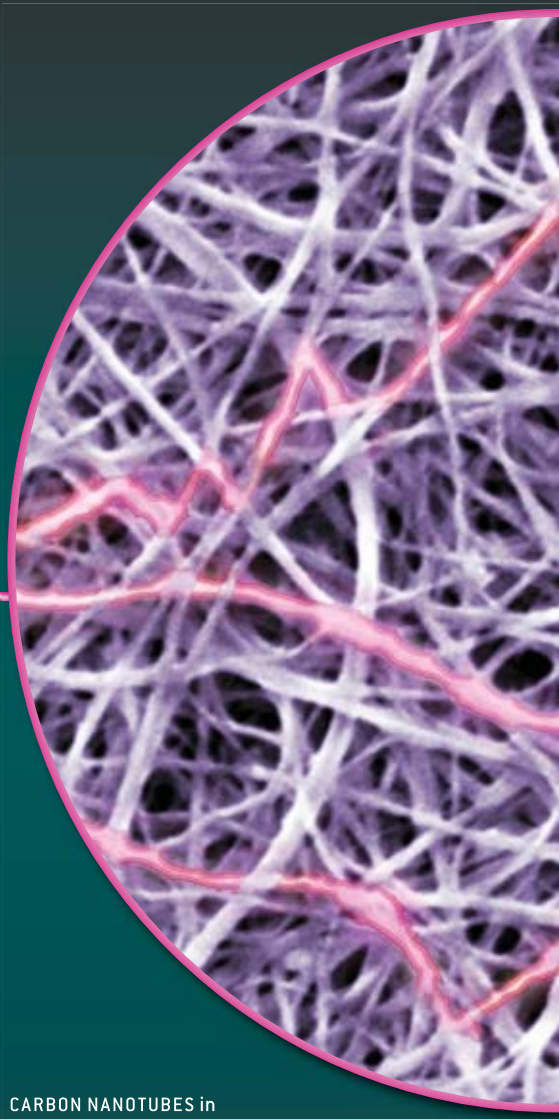
MORE TO EXPLORE

A Mechanical Turing Machine: Blueprint for a Biomolecular Computer. Presented by Ehud Shapiro at the 5th International Meeting on DNA Based Computers, Massachusetts Institute of Technology, June 14–15, 1999. **www.weizmann.ac.il/udi/press**

Programmable and Autonomous Computing Machine Made of Biomolecules. Y. Benenson, T. Paz-Elizur, R. Adar, E. Keinan, Z. Livneh and E. Shapiro in *Nature,* Vol. 414, pages 430–434; November 22, 2001.

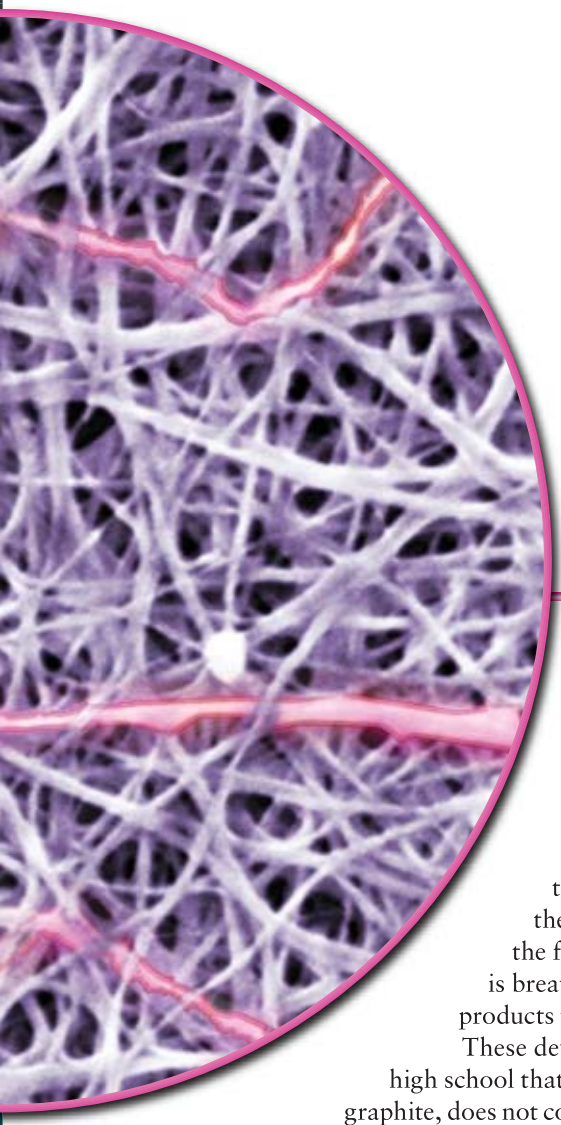
An Autonomous Molecular Computer for Logical Control of Gene Expression. Y. Benenson, B. Gil, U. Ben-Dor, R. Adar and E. Shapiro in *Nature,* Vol. 429, pages 423–429; May 27, 2004.

CARBON NANONETS SPARK NEW ELECTRONICS



CARBON NANOTUBES in a weblike mesh ensure multiple alternative pathways for electrons (*pink highlights*), providing surefire electrical conduction. The entire field of view is 0.7 micron in diameter.

JOHN YORSTON *Carl Zeiss SMT* AND ANDREW PAUL LEONARD *APL Microscopic*; COLORIZED BY JEAN-FRANCOIS PODEVIN



Random networks of tiny carbon tubes could make possible low-cost, flexible devices such as "electronic paper" and printable solar cells

By George Gruner

In many classic science-fiction stories, alien life is based on silicon—the substance at the core of modern electronics technology—rather than on carbon, the fundamental building block of earthly biology. Scientists have even speculated that they might someday create silicon life-forms. Instead the opposite is starting to happen: carbon is serving as the foundation for electronic devices—and in the process is breathing new life into the quest for inexpensive, flexible products that offer a broad range of capabilities.

These developments may surprise those of us who learned in high school that carbon, in its familiar incarnations of diamond and graphite, does not conduct electricity well. During the past 15 years, however, researchers have discovered new forms of carbon: very small structures comprising a few hundred to 1,000 atoms, through which electrons travel with ease. Of particular interest is the carbon nanotube, a molecule that resembles rolled-up chicken wire, only the "wire" is a sheet of carbon atoms that is 100 million times as small as the version used for chicken coops.

Investigators have found that random networks of carbon nanotubes—called nanonets—can perform a variety of basic electronic functions. Using novel chemistry, researchers can make such networks mimic the conductive properties of metals such as copper or the less conductive characteristics of semiconductors such as silicon. These innovations have paved the way for this single material to assume different roles in electronic devices.

Further, engineers can construct such carbon-based devices by employing simple fabrication methods. Researchers can dissolve the tubes in a liquid and spray the resulting solution to form thin layers on, say, flexible plastic sheets. They can also lay or print these materials on other layers that have various electronic functions, for example, substances that emit light when a voltage is applied.

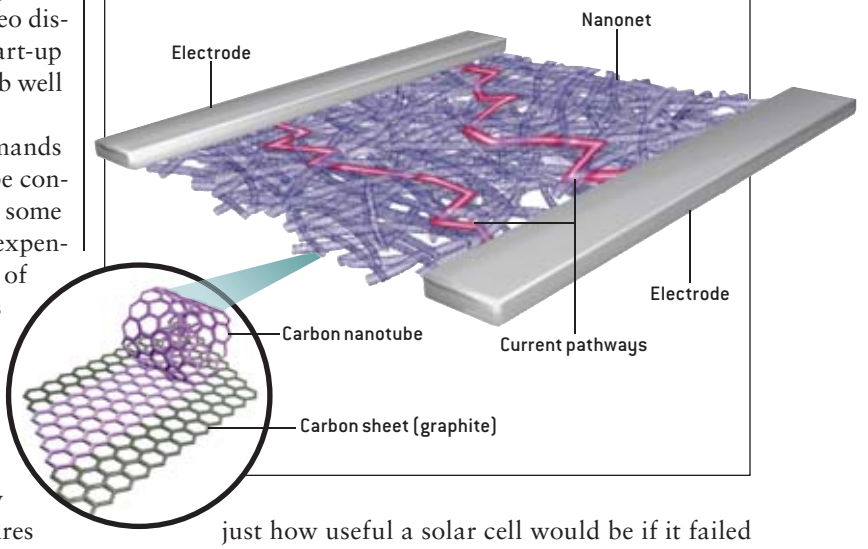
It takes little imagination to see how this kind of straightforward system could form the basis of many extremely cheap but handy products: "electronic paper" that could display information on sheets that roll up like conventional newsprint; chemical sensors; wearable electronic devices; solar cells that could be printed onto rooftop tiles; or scads of simple radio-frequency identification (RFID) sensors for monitoring warehouse or store inventories. For such applications, the expensive, lightning-fast processing power of integrated chips like Intel's Pentium processors or Samsung's video displays is not needed; rather R&D laboratories and start-up firms are racing to find technologies that can do the job well enough at low cost [*see table on page 55*].

Such exciting applications would make severe demands on today's electronic materials: they would need to be conductive, flexible, lightweight, transparent (at least for some applications, such as solar cells and displays) and inexpensive. But most conductors are metals, the majority of which are not transparent, whereas, as a rule, thin films of materials that are transparent, such as diamond, are insulators (they do not transmit electricity). Light can pass through one special class of metals, called metal oxides, however. The best known, indium tin oxide, is frequently used where engineers need see-through electrodes. But metal oxides are costly. They are also heavy and brittle, and their manufacture requires high processing temperatures and multibillion-dollar fabrication facilities.

Another alternative is an unusual category of plastics known as conducting polymers. Although common plastic substances are insulators, chemists have in recent decades managed to convert some polymers into semiconductors and even full-fledged conductors. Polymers can be produced using room-temperature techniques. Lightweight and flexible, they can easily take on multiple forms and are, of course, dirt cheap. On the downside, weak bonds hold together the atoms in most plastics. The bonds can break rather readily, which leads most polymers to degrade over time. Consider just how useful a solar cell would be if it failed after only a few warm, sunny days.

FROM NANOTUBES TO NANONETS

Thin networks of carbon nanotubes, or nanonets, can serve as electronic devices. Each nanotube is a one-atom-thick sheet of carbon, called graphite, rolled into a cylinder with a diameter of about a nanometer (*inset*)—approximately 50,000 times as small as the width of a human hair. Electric current passes through the interconnected tubes from one electrode to another.



A Better Wire

ENTER CARBON-BASED NANOWIRES. Carbon nanotubes were first discovered several decades ago, but no one realized their value at the time. Then, in 1991, Japanese chemist Sumio Ijima of NEC Corporation rediscovered them. These tiny tubes of carbon have a diameter of around one nanometer—about the same as a strand of a DNA molecule [*see box above*]. The electrical conductivity of the tubes is comparable to that of copper and surpasses that of any polymer by several orders of magnitude. They can also carry more than 100 times more current than the best metals. Carbon nanotubes are, in addition, physically robust: they can be bent easily, they do not react with most chemicals and they resist damage from day-to-day use.

Manufacturers make nanotubes by reducing coal into its component atoms using the heat of an electric arc or a laser, which creates a so-called carbon plume. They then add catalysts to the plume, which promotes the formation of various types of carbon molecules. This relatively straightforward procedure produces what is essentially soot—carbon molecules in many forms, including spherical structures called buckyballs as well as other "fullerenes" and carbon nanotubes. Fabricators must then laboriously separate the nanotubes out of the mixture. Techniques focus on separating out only the long, nearly perfect specimens that have a single "chicken wire" wall (rather than multiple, concentric walls). Suitable nanotubes are thus currently quite pricey, but mak-

JEAN-FRANCOIS PODEVIN

Overview/*Nanonet Electronics*

- Carbon nanotubes—minuscule cylinders of rolled-up carbon sheets—conduct electricity well, which could make them useful for many exciting electronic applications. But manufacturing products that use single tubes is expensive and suffers from significant reliability problems.
- Random networks of many carbon nanotubes, called nanonets, enable numerous basic electronic functions at low cost. The durable nature of nanonets also makes them suitable for portable devices.
- Carbon nanonets should find use in sensors, solar cells, electronic paper, and flexible touch screens and displays within a few years.

ers are confident that costs will drop significantly if market demand rises enough to justify high-volume manufacturing facilities.

When a single nanotube is employed to build a transistor, the voltage-activated switch that is the workhorse of modern electronics, the resulting device can easily outperform the transistors on the silicon chips in today's computers. But single carbon nanotubes will not replace silicon and copper in the foreseeable future. The main obstacle lies in their manufacturability, which is one of the most vexing problems afflicting nanotechnology's commercialization. Current devices based on a single nanotube can take days to make, because they often must be assembled by hand. Another difficulty is performance variability. Nanotubes come in slightly different shapes and forms, which affects their electrical attributes.

From One Wire to a Network

ALTHOUGH INDIVIDUAL TUBES differ from one to the next, researchers realized that this variation could be averaged out by using many tubes together—any shortcomings present in some of the tubes could be compensated for by better-performing counterparts. The simplest example is a random network of nanotubes [*see box on opposite page*]. Just as an interstate highway system can offer alternative routes when you encounter a traffic jam on one roadway, so, too, can a random assembly of electrically conducting nanotubes—a nanonet—speed the transmission of electrons by providing alternative pathways. Investigators soon established that these nearly two-dimensional random networks offered interesting properties in their own right.

First, the nanonet's many pathways and connections guarantee good electrical conductance between one electrode and another, despite possible manufacturing flaws. A good analogy is the freeway system that serves the Los Angeles metropolitan area. No one would want to attempt to traverse the City of Angels by hiking cross-country or driving the slow, stoplight-strewn surface roads; instead travelers take the freeway. The same concept applies to the nanonet, which allows electrons to jump on the tubes and move around on what is essentially a nanoscale freeway system. The multiple avenues provided by these networks also afford a considerable resistance to failure, or fault tolerance; if one route breaks from use, others are there to take up the slack.

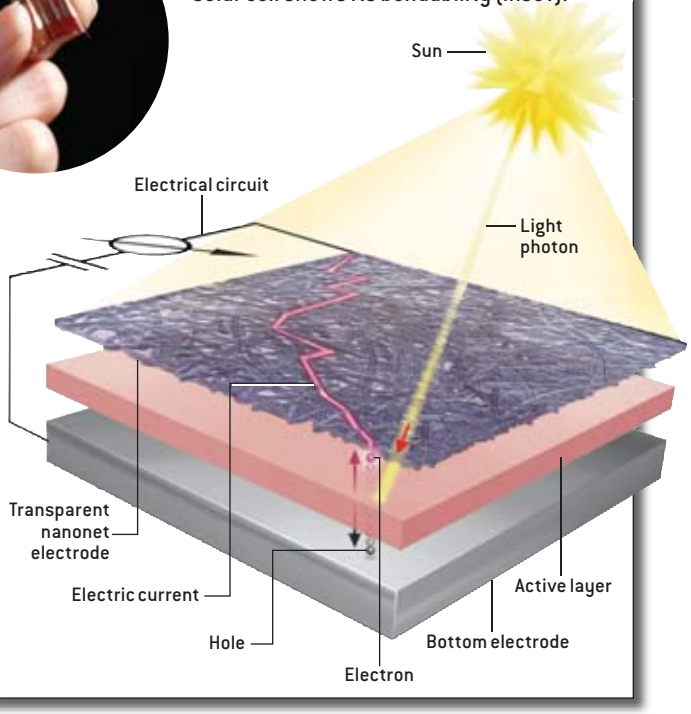
A conductive nanonet is in fact a simple example of the concept of percolation, which describes how objects, materials or electric currents move through a random medium. Imagine dropping pickup sticks on a tabletop one at a time. With only a few sticks, the chances of finding a connected pathway (by going from one stick to the next) from one end of the table to the other are slim. In fact, below a certain critical density of sticks, the odds drop to zero. But as the number of sticks increases, the pile will eventually surpass that critical density, the so-called percolation transition, where at first one and then more and more pathways form. If the pickup-stick approach were applied to copper wires on the tabletop, at some point the network would achieve electrical conduction across the table as well—with the current dependent on the density of the copper wires. Theorists studied this concept some time ago, and my group at the University of California, Los Angeles, was able to map out such a transition in networks of nanotubes.

Nanonets can be, in addition, highly transparent—an advantage in applications that require light transmission. Just as the freeway pavement covers only a small fraction of the natural terrain, a web of long and skinny wires allows passage of most of the incident light—a fraction that approaches 100 percent for what can be considered one-dimensional nanowires.

Finally, much like a spiderweb, a network of nanowires typically is more robust than the same material in undifferentiated bulk, which often tends to break when bent. These characteristics make the nanonet architecture eminently suited for applications in which resistance to day-to-day use and misuse is important. Think about how many times you have dropped your cell phone or iPod.

CHEAPER SOLAR CELLS

Transparent and flexible carbon nanonet electrodes (*shown in exploded view*) would be significantly cheaper than the rigid indium tin oxide electrodes used in today's solar cells. The carbon nanonets would conduct electric charges (negative electrons and positive "holes") dislodged by light photons from an active (charge-creating) layer of a semiconductor to an external circuit. A sample printed nanonet solar cell shows its bendability (*inset*).



JEAN-FRANCOIS PODEVIN; BRAD SWONETZ *Redux* (*solar cell, Stanford University*)

Weaving a Nanonet

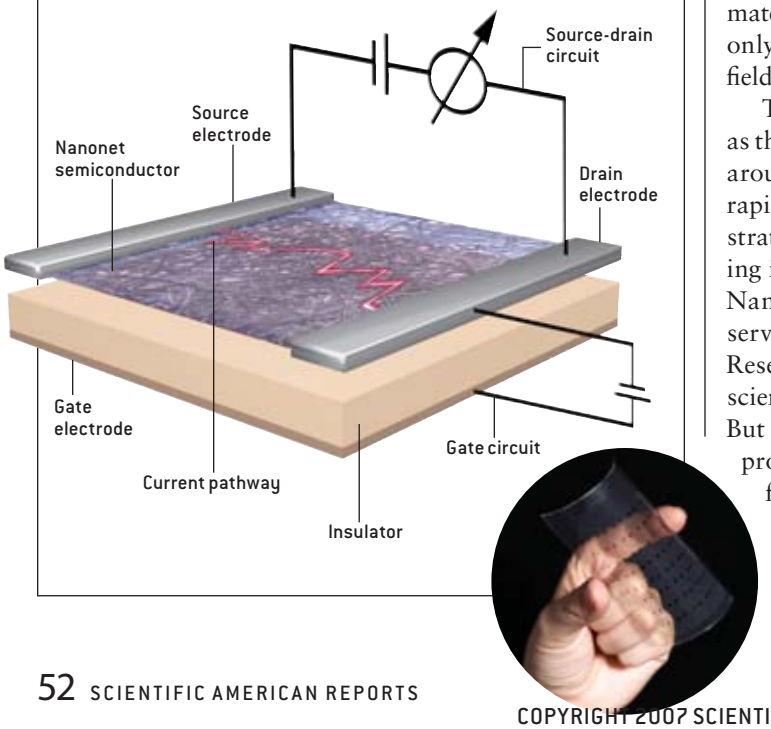
THESE PERFORMANCE BENEFITS augur well for the technology's potential in real-world applications, but any new replacement material must, of course, be more than competitive in terms of function and cost with current materials. Nanotube films initially made a couple of years ago—by my team, by a group led by physicist Siegmar Roth at the Max Planck Institute for Solid State Research in Stuttgart, Germany, and by one at the University of Texas at Austin—were not up to the task. Finding the optimal processing routes and the most advantageous way to deposit the tubes onto surfaces was not a trivial problem.

Clearly, one cannot fabricate thin films of such networks by merely throwing down a tube at a time like playing pickup sticks; another strategy is needed. One might, for instance, dissolve the tubes in a solvent (water, alcohol, organic liquids) and then spray the resulting fluid onto a surface, but that is not as easy as it sounds. When mixed in a liquid, the tubes tend to bundle together, requiring a chemical additive to keep them apart. Some agents, called surfactants (soaps), accomplish this job by completely surrounding the tubes. But surfactants, if they remain on the tubes after they are sprayed onto a surface, impede the flow of electrons between tubes (blocking the freeway ramps, so to speak). Through steady trial-and-error efforts with innumerable solvents, surfactants and processing procedures, however, researchers have created simple (room-temperature) avenues to make such thin films of nanotube networks. A method pioneered by my team and a group led by chemist Andrew Rinzler at the University of Florida yields films that have the lowest electrical resistance and thus the best operating performance to date among nanonet-based devices.

As researchers experimented with the conductivity of the tubes, they learned that the material could be transparent, a property that is important for applications such as displays and solar cells. The discovery that carbon nanonets are transparent to light came about as a by-product of research on their conductivity. The first indication that nanonets could be clear arose in 2001, when my former postdoctoral associate Leonardo Degiorgi and his group at the Swiss Federal Institute of Technology in Zurich, along with physicist David Tanner and his co-workers at the University of Florida, studied their optical characteristics. To measure nanonet conductivity precisely, they fabricated thick films: these were too deep to transmit light, but the data led both research groups to conclude that a thinner film would be both transparent and a good conductor. After these groups made this determination, Rinzler's team (collaborating with Katalin Kamarás and her colleagues at the Central Institute of Physics in Budapest) and mine at U.C.L.A. followed up with direct measurements of a nanonet film's optical transparency. Today scientists can fabricate tailor-made films with different levels of transparency and electrical conductivity by changing the thickness of the films.

TRANSPARENT TRANSISTORS

Carbon nanonet films, tailored to perform like semiconductors, can serve as the basis for field-effect transistors—the building blocks of computers, cellular phones and other digital devices. This switchlike mechanism (*shown in exploded view*) uses a small electric voltage provided by the gate electrode to greatly boost the current in the source-drain circuit. In the inset, a technician bends a plastic sheet onto which an array of see-through nanonet transistors has been printed.



Nanonet Transistors

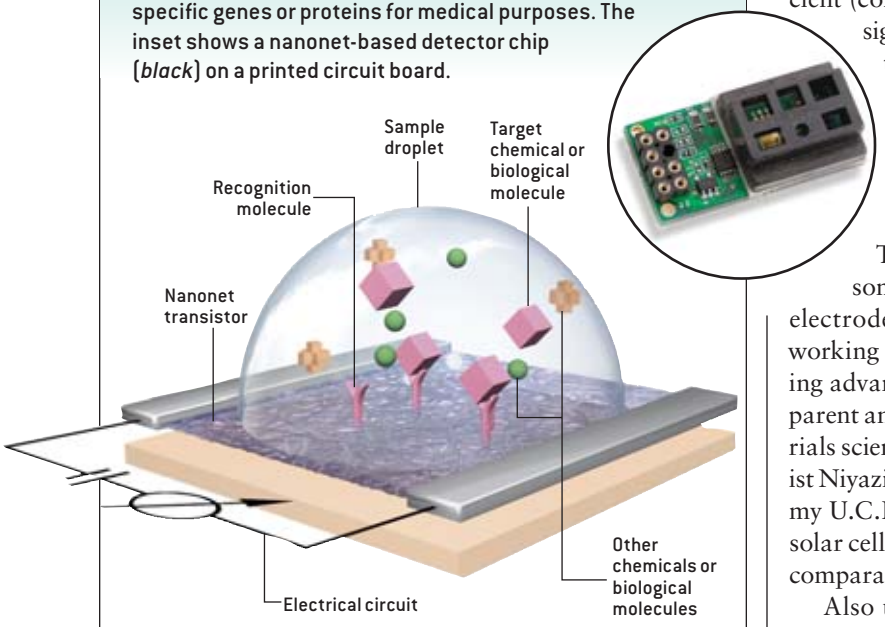
RESEARCHERS SOON TURNED their attention from making nanonet conductors to nanonet semiconductors that could serve as the basis for transistors. A transistor requires materials whose conductivity changes greatly in response to only small incremental inputs, such as altering an electric field [*see box at left*].

The notion that carbon nanotube networks could serve as the backbone of thin-film field-effect transistors emerged around seven years ago. Thereafter, progress was relatively rapid—with advances in creating nanonets on flexible substrates, demonstrating the transparency of the devices, coming in short order. Working in parallel, my R&D group at Nanomix, a start-up firm in Emeryville, Calif., where I served as chief scientist, and a research team at the Naval Research Laboratory in Washington, D.C., led by materials scientist Eric Snow, produced nanonet transistors in 2003. But these devices were formed on rigid glass substrates at processing temperatures of 900 degrees Celsius—too hot for use with flexible plastic substrates that melt at 120 degrees C. Nanomix researchers Keith Bradley and Jean-Christophe Gabriel, in collaboration with my U.C.L.A. team, manufactured the first flexible nanotube network transistors on plastic in 2003. Soon af-

JEAN-FRANCOIS PODEVIN; BRAD SWONETZ *Redux* (*transistor, Unidym*)

LOW-COST SENSORS

A nanonet device can become a sensor with the addition of "recognition molecules" that react with a target chemical or biological molecule. When the target binds to a recognition molecule, it alters the sensor's electrical output. Such devices can detect many chemicals, including a blood-borne cancer marker called prostate-specific antigen (PSA) and, soon, microorganisms such as the anthrax bacterium, a potential bioweapon. Arrays of nanonet sensors, each with different recognition molecules, could cheaply detect specific genes or proteins for medical purposes. The inset shows a nanonet-based detector chip (*black*) on a printed circuit board.



terward, my colleagues and I at U.C.L.A., working with Roth's group at the Max Planck Institute, managed to fabricate devices that were also transparent, making them suitable for applications such as portable visual displays. Physicist John Rogers and his colleagues at the University of Illinois achieved similar success only a few months later. Although these field-effect transistors operated at fast rates—the key metric for such devices—other necessary characteristics, such as low-voltage function, were lacking. The goal was to run the devices at voltages less than those standard batteries can provide to save power, but this feat was attained only recently by Rogers and by chemist Tobin Marks of Northwestern University, who employed specially made polymers to insulate the devices' conductive parts.

Nanonets in Action

CARBON NANONETS CAN OFFER distinct advantages in many portable products, a conclusion that becomes more obvious when one compares them with some of the current contenders for these applications, including films composed of organic or polymeric metals and some semiconductors. For these uses, electronic materials must exhibit good electrical conductance (otherwise, applied current heats them up, resulting in power losses) and high optical transparency (because the viewer of a display, for example, needs to see the layers that lie underneath).

Such substances will enable the development of what people variously call printed, plastic, disposable or macroelectronic products. One example is the photovoltaic cell. Typical solar cells made of single-crystal silicon have excellent performance (they convert as much as 18 percent of incoming light into electricity) but are bulky, heavy and costly to manufacture. Instead imagine a razor-thin solar cell that, though less efficient (converting only 5 or 6 percent of incoming light), is significantly cheaper to fabricate and offers the potential for easy mass production of large-area systems, both of which can compensate for the material's lower performance levels [*see box on page 51*].

In a solar cell, incoming sunlight dislodges electrons and their positively charged counterparts, called holes, in the middle layer of the device. The electrons then migrate to one electrode, power some electrical load and return to the holes via another electrode to complete the circuit. Several companies are working to perfect a cell's active (charge-creating) layers using advanced polymers and other substances that are transparent and flexible. Together with Michael McGehee's materials science group at Stanford University and physical chemist Niyazi Serdar Sariciftci of the University of Linz in Austria, my U.C.L.A. team has produced flexible, proof-of-concept solar cells with nanonet electrodes that exhibit performance comparable to that of indium tin oxide electrodes.

Also under consideration are nanonet-based films that would lie at the heart of an inexpensive, flexible and lightweight touch screen or visual display. A touch screen, for instance, consists of two sheets of electrodes separated by insulating spacers. When a finger touches the top sheet at some point, the electrodes there meet, completing an electrical circuit specific for that location that is formed by a pattern of smooth, thin layers of conductive materials that have been imprinted on the bottom sheet. In collaboration with Richard Kaner's group at U.C.L.A., my team has fabricated and tested proof-of-concept devices based on nanonets.

Nanonets also work in light-emitting diodes, which resemble photovoltaics that run in reverse so that they create light when electricity passes between the electrodes. In collaboration with Marks's group at Northwestern, my team has recently demonstrated proof-of-concept light-emitting

THE AUTHOR

GEORGE GRUNER, a Distinguished Professor of Physics at the University of California, Los Angeles, investigates fundamental issues relating to nanoscale phenomena. He has served as the chief technology officer at Nanomix in Emeryville, Calif., and he is founder of Unidym in Menlo Park, Calif.; both companies concentrate on developing novel applications of nanotechnology. Gruner's interests include skiing, classical music and technology assessment—not necessarily in that order. His research is supported by the National Science Foundation.

JEAN-FRANCOIS PODEVIN; NANOMIX "SENSATION" (*sensor*)

diodes with excellent performance (sufficient to meet the requirements for use in televisions, for example), as has a research group at the University of Montreal led by Richard Martel.

Transistors made from nanonets will, in addition, find use in printed electronics. Tests indicate that the operating speed of carbon nanonets lags somewhat behind that of crystalline silicon, from which most integrated chips are fabricated, but their conductivity and durability advantages over polymers make them attractive to device manufacturers. Although nanotube films cannot yet work in laptop computers or television sets, they are competitive in many other products—especially those that require a material that is cheap, flexible, lightweight, environmentally friendly and resistant to abuse. The first such application is expected to be large-area visual displays, called active-matrix displays. The transistors in a display must run rapidly so that the images can be readily refreshed.

Of course, the kind of portable devices that will use these displays will need power sources as well—cheap, lightweight, razor-thin and disposable batteries and supercapacitors. Nanonets could also play an important role in such power devices, serving not only as electrodes but as high-surface-area components for collecting electric charge to store it for later discharge.

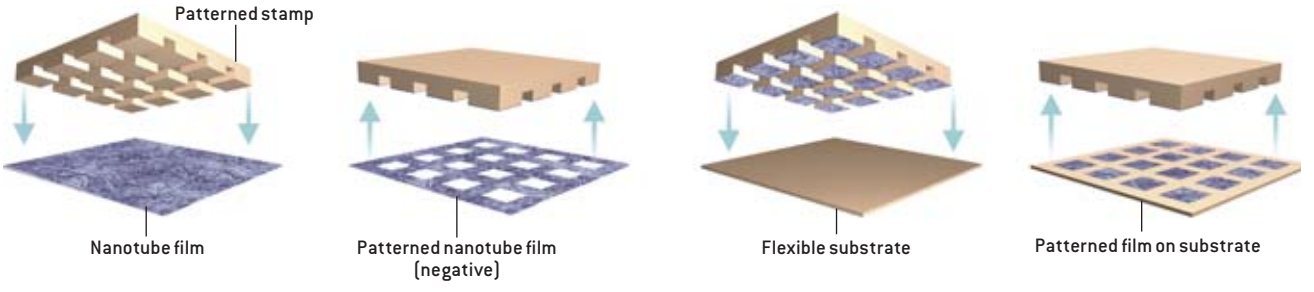
Many Pathways to Take

THE NASCENT CARBON NANONET industry has only just begun to perfect this fledgling technology. There is little doubt that the recent feasibility studies that I have described will soon be followed by working prototypes and eventually products based on those new devices. Today the industry is at the stage where the silicon chip business was half a century ago. The nanotubes are improving steadily, and researchers are successfully sorting those that conduct electricity as well as metals do from those that are semiconduc-

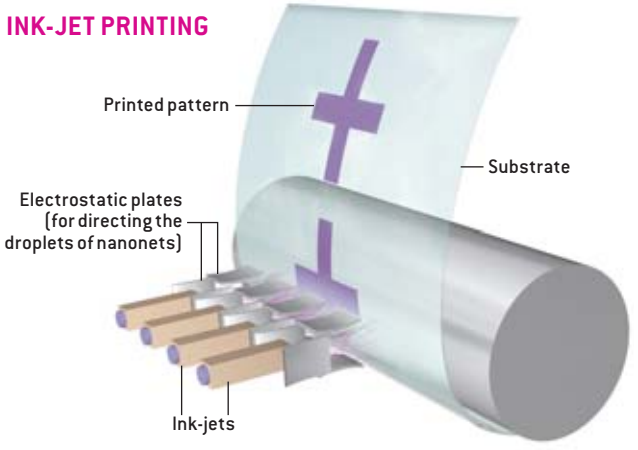
PRINTING A NANONET

Prospective makers of products based on carbon nanonets are developing several inexpensive ways to "print" an engineered pattern of the material onto a flexible polymer surface to produce, for example, an electronic circuit. The simplest method resembles using an ink pad and a stamp (*top*). A patterned stamp comes into contact with a nanonet layer, parts of which stick to the stamp's bottommost surface. The primed stamp presses down on the surface of a substrate, printing the nanonet pattern onto it. Manufacturers are also working on two mass-production techniques, including the use of standard ink-jets (*bottom left*) to spray a liquid containing dispersed nanotubes onto substrates, and a variant of offset printing, in which a nanonet solution substitutes for ink (*bottom right*).

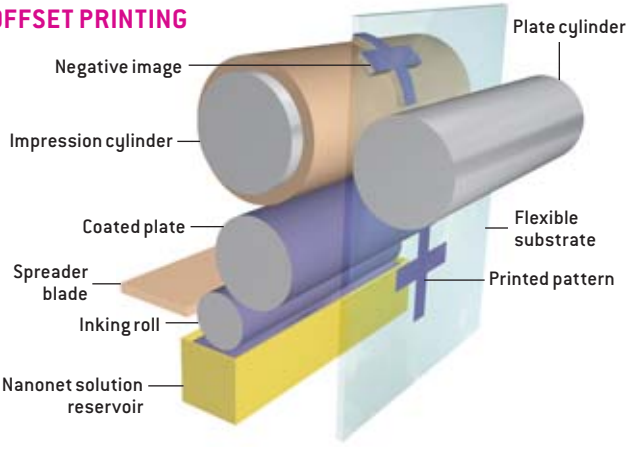
STAMP PRINTING



INK-JET PRINTING



OFFSET PRINTING



JEAN-FRANCOIS PODEVIN

Carbon Nanonet Research and Product Development

Hundreds of start-up firms and more established companies are producing or working to develop carbon nanotube materials, carbon nanonet films and the electronic devices that incorporate them. A new technology typically passes through the following sequence of developmental stages: concept, R&D, proof of concept, prototype, product development and production.

ORGANIZATION	PRODUCT FOCUS	STATUS
HIGH-GRADE MATERIALS FOR ELECTRONICS		
CarboLex, Lexington, Ky. (www.carbolex.com)	Electric arc– and chemical vapor deposition (CVD)–based fabrication	Production
Carbon Solutions, Riverside, Calif. (www.carbonsolution.com)	Electric arc–based fabrication	Production
SouthWest NanoTechnologies, Norman, Okla. (www.swnano.com)	CVD-produced specialty nanotubes	Production
Thomas Swan, Consett, England (www.thomas-swan.co.uk)	High-volume CVD-based fabrication	Production
Unidym, Menlo Park, Calif. (www.unidym.com)	CVD- and carbon monoxide–based fabrication	Production
TRANSPARENT FILMS		
Battelle Memorial Institute, Columbus, Ohio (www.battelle.org)	Transparent coatings	R&D
Eastman Kodak, Rochester, N.Y. (www.kodak.com)	Transparent optical coatings	R&D, prototype
Eikos, Franklin, Mass. (www.eikos.com)	Conducting ink	Product development
Unidym (*see above*)	Films for touch screens, solar cells, light-emitting diodes	Product development
DEVICES		
DuPont, Wilmington, Del. (www.dupont.com)	Transparent electronics	R&D
IBM, Armonk, N.Y. (www.ibm.com)	Computer-compatible transistors and interconnects	R&D
Intel, Santa Clara, Calif. (www.intel.com)	Interconnects	R&D
Motorola, Schaumburg, Ill. (www.motorola.com)	Biological and chemical sensors	Prototype
Nanomix, Emeryville, Calif. (www.nano.com)	Chemical and biological sensors	Product development, R&D
Nantero, Woburn, Mass. (www.nantero.com)	Novel memory technology	Prototype
Samsung, Seoul, South Korea (www.samsung.com)	Displays	R&D
Unidym (*see above*)	Printed electronics for displays	Proof of concept

tors, which will further better device performance. Meanwhile investigators have made progress on a process that resembles silicon doping, in which special molecules are attached to the tubes to finely alter their electrical properties. Many observers believe that it is only a matter of time before such films exceed the performance of traditional metals and start to make inroads into silicon-based digital electronics technology.

Carbon nanonets have just recently left the realm of science fiction and entered that of practical reality. Like silicon, this budding technology is highly unlikely to lead to artificial life anytime soon, but it has every chance of enabling innovative products that in the not too distant future will improve our everyday lives. SA

MORE TO EXPLORE

Carbon Nanotubes. Mildred Dresselhaus, Gene Dresselhaus, Peter Eklund and Riichiro Saito in *Physics World,* Vol. 11, pages 33–38; January 1998.

Nanotubes for Electronics. Philip G. Collins and Phaedon Avouris in *Scientific American,* Vol. 283, No. 6, pages 62–69; December 2000.

Carbon Nanotubes and Related Structures. Peter J. F. Harris. Cambridge University Press, 2001.

Carbon Nanotubes—The Route Toward Applications. R. H. Baughman, A. A. Zakhidov and W. A. de Heer in *Science,* Vol. 297, pages 787–792; August 2, 2002.

Carbon Nanotube Films for Transparent and Plastic Electronics. George Gruner in *Journal of Materials Chemistry,* Vol. 16, No. 35, pages 3533–3539; 2006.

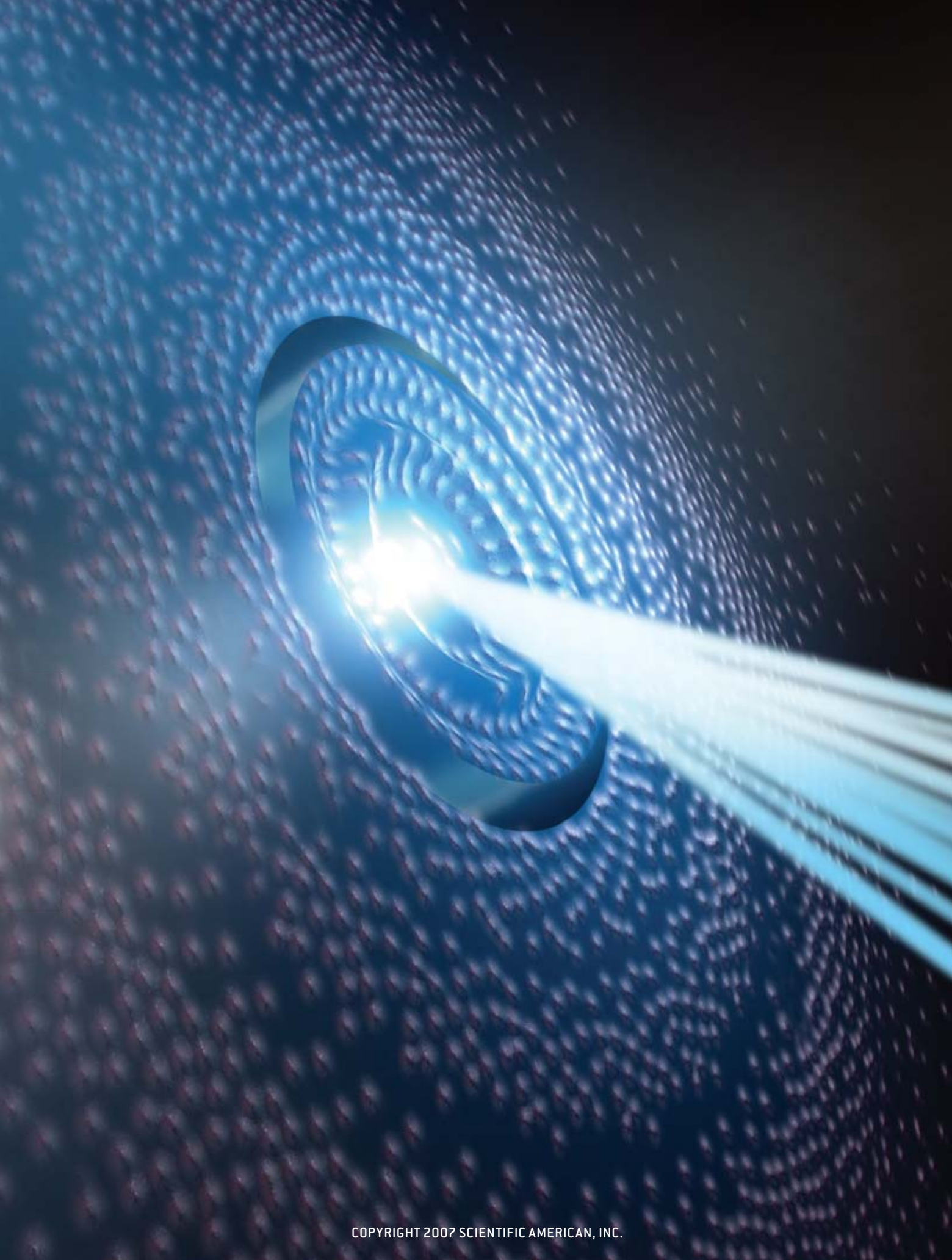
Carbon Nanotube Transistors for Biosensing Applications. George Gruner in *Analytical and Bioanalytical Chemistry,* Vol. 384, pages 322–335; 2006.

A technology that squeezes electromagnetic waves into minuscule structures may yield a new generation of superfast computer chips and ultrasensitive molecular detectors

The Promise of PLASMONICS

By Harry A. Atwater

LIGHT BEAM striking a metal surface can generate plasmons, electron density waves that can carry huge amounts of data. If focused on a surface etched with a circular groove, as in this artist's rendering, the beam produces concentric waves, organizing the electrons into high- and low-density rings.



Light is a wonderful medium for carrying information.

Optical fibers now span the globe, guiding light signals that convey voluminous streams of voice communications and vast amounts of data. This gargantuan capacity has led some researchers to prophesy that photonic devices—which channel and manipulate visible light and other electromagnetic waves—could someday replace electronic circuits in microprocessors and other computer chips. Unfortunately, the size and performance of photonic devices are constrained by the diffraction limit; because of interference between closely spaced light waves, the width of an optical fiber carrying them must be at least half the light's wavelength inside the material. For chip-based optical signals, which will most likely employ near-infrared wavelengths of about 1,500 nanometers (billionths of a meter), the minimum width is much larger than the smallest electronic devices currently in use; some transistors in silicon integrated circuits, for instance, have features smaller than 100 nanometers.

Recently, however, scientists have been working on a new technique for transmitting optical signals through minuscule nanoscale structures. In the 1980s researchers experimentally confirmed that directing light waves at the interface between a metal and a dielectric (a nonconductive material such as air or glass) can, under the right circumstances, induce a resonant interaction between the waves and the mobile electrons at the surface of the metal. (In a conductive metal, the electrons are not strongly attached to individual atoms or molecules.) In other words, the oscillations of electrons at the surface match those of the electromagnetic field outside the metal. The result is the generation of surface plasmons—density waves of electrons that propagate along the interface like the ripples that spread across the surface of a pond after you throw a stone into the water.

Overview/*Plasmonics*

- Researchers have discovered that they can squeeze optical signals into minuscule wires by using light to produce electron density waves called plasmons.
- Plasmonic circuits could help the designers of computer chips build fast interconnects that could move large amounts of data across a chip. Plasmonic components might also improve the resolution of microscopes, the efficiency of light-emitting diodes, and the sensitivity of chemical and biological detectors.
- Some scientists have even speculated that plasmonic materials could alter the electromagnetic field around an object to such an extent that it would become invisible.

Over the past decade investigators have found that by creatively designing the metal-dielectric interface they can generate surface plasmons with the same frequency as the outside electromagnetic waves but with a much shorter wavelength. This phenomenon could allow the plasmons to travel along nanoscale wires called interconnects, carrying information from one part of a microprocessor to another. Plasmonic interconnects would be a great boon for chip designers, who have been able to develop ever smaller and faster transistors but have had a harder time building minute electronic circuits that can move data quickly across the chip.

In 2000 my group at the California Institute of Technology gave the name "plasmonics" to this emerging discipline, sensing that research in this area could lead to an entirely new class of devices. Ultimately it may be possible to employ plasmonic components in a wide variety of instruments, using them to improve the resolution of microscopes, the efficiency of light-emitting diodes (LEDs), and the sensitivity of chemical and biological detectors. Scientists are also considering medical applications, designing tiny particles that could use plasmon resonance absorption to kill cancerous tissues, for example. And some researchers have even theorized that certain plasmonic materials could alter the electromagnetic field around an object to such an extent that it would become invisible. Although not all these potential applications may prove feasible, investigators are eagerly studying plasmonics because the new field promises to literally shine a light on the mysteries of the nanoworld.

Shrinking Wavelengths

FOR MILLENNIA, alchemists and glassmakers have unwittingly taken advantage of plasmonic effects when they created stained-glass windows and colorful goblets that incorporated small metallic particles in the glass. The most notable example is the Lycurgus cup, a Roman goblet dating from the fourth century A.D. and now held in the British Museum [*see illustration on page 62*]. Because of plasmonic excitation of electrons in the metallic particles suspended within the glass matrix, the cup absorbs and scatters blue and green light—the relatively short wavelengths of the visible spectrum. When viewed in reflected light, the plasmonic scattering gives the cup a greenish hue, but if a white light source is placed within the goblet, the glass appears red because it transmits only the longer wavelengths and absorbs the shorter ones.

Research into surface plasmons began in earnest in the 1980s, as chemists studied the phenomenon using Raman spectroscopy, which involves observing the scattering of laser light off a sample to determine its structure from molecular vibrations. In 1989 Thomas Ebbesen, then at the NEC Re-

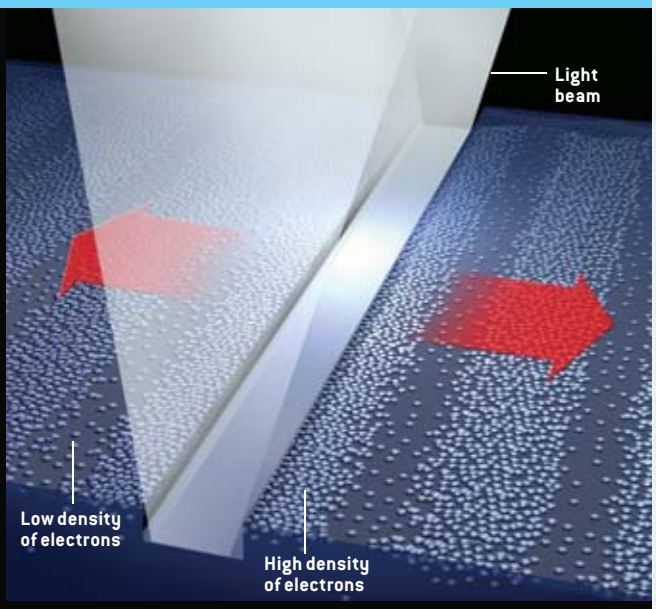
PHIL SAUNDERS *Space Channel Ltd. (preceding pages)*

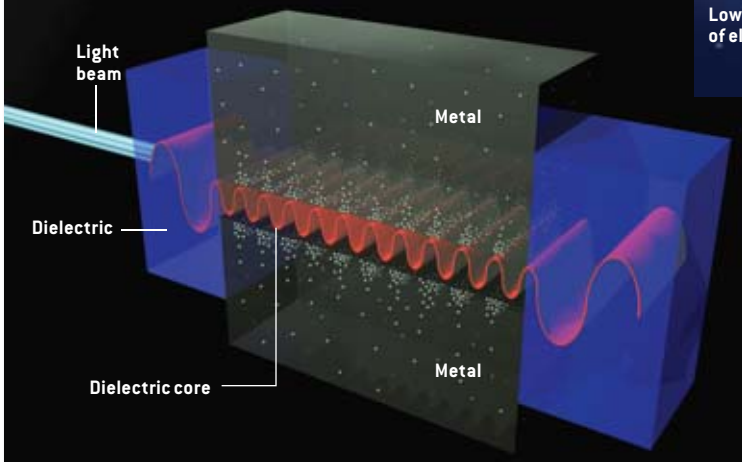
FUNNELING LIGHT INTO TINY WIRES

The study of plasmonics is relatively new, but researchers have already developed prototype devices that demonstrate the promise of the technology.

PLANAR WAVEGUIDE

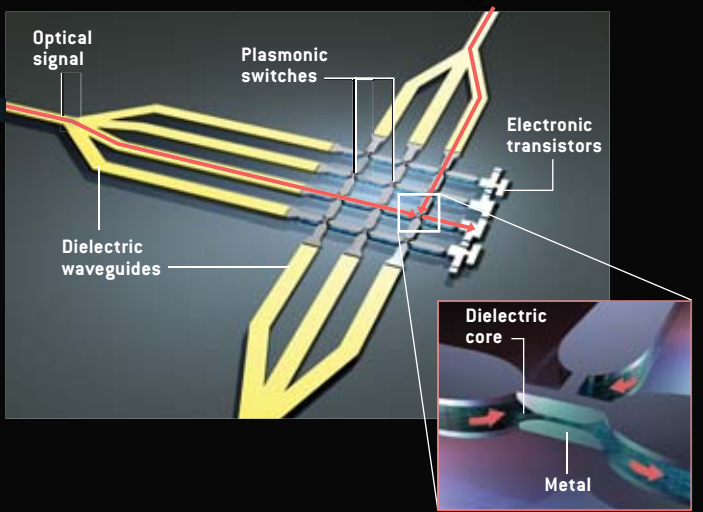
Plasmons always flow along the boundary between a metal and a dielectric (a nonconductive material such as air or glass). For example, light focused on a straight groove in a metal will generate plasmons that propagate in the thin plane at the metal's surface (the boundary between the metal and air). A plasmon could travel as far as several centimeters in this planar waveguide—far enough to convey a signal from one part of a chip to another—but the relatively large wave would interfere with other signals in the nanoscale innards of a processor.





PLASMON SLOT WAVEGUIDE

Scientists have built much smaller plasmonic circuits by putting the dielectric at the core and surrounding it with metal. The plasmon slot waveguide squeezes the optical signal, shrinking its wavelength by a factor of 10 or more. Researchers have constructed slot waveguides with widths as small as 50 nanometers—about the same size as the smallest electronic circuits. The plasmonic structure can carry much more data than an electronic wire, but it cannot transmit a signal farther than 100 microns.



A FASTER CHIP

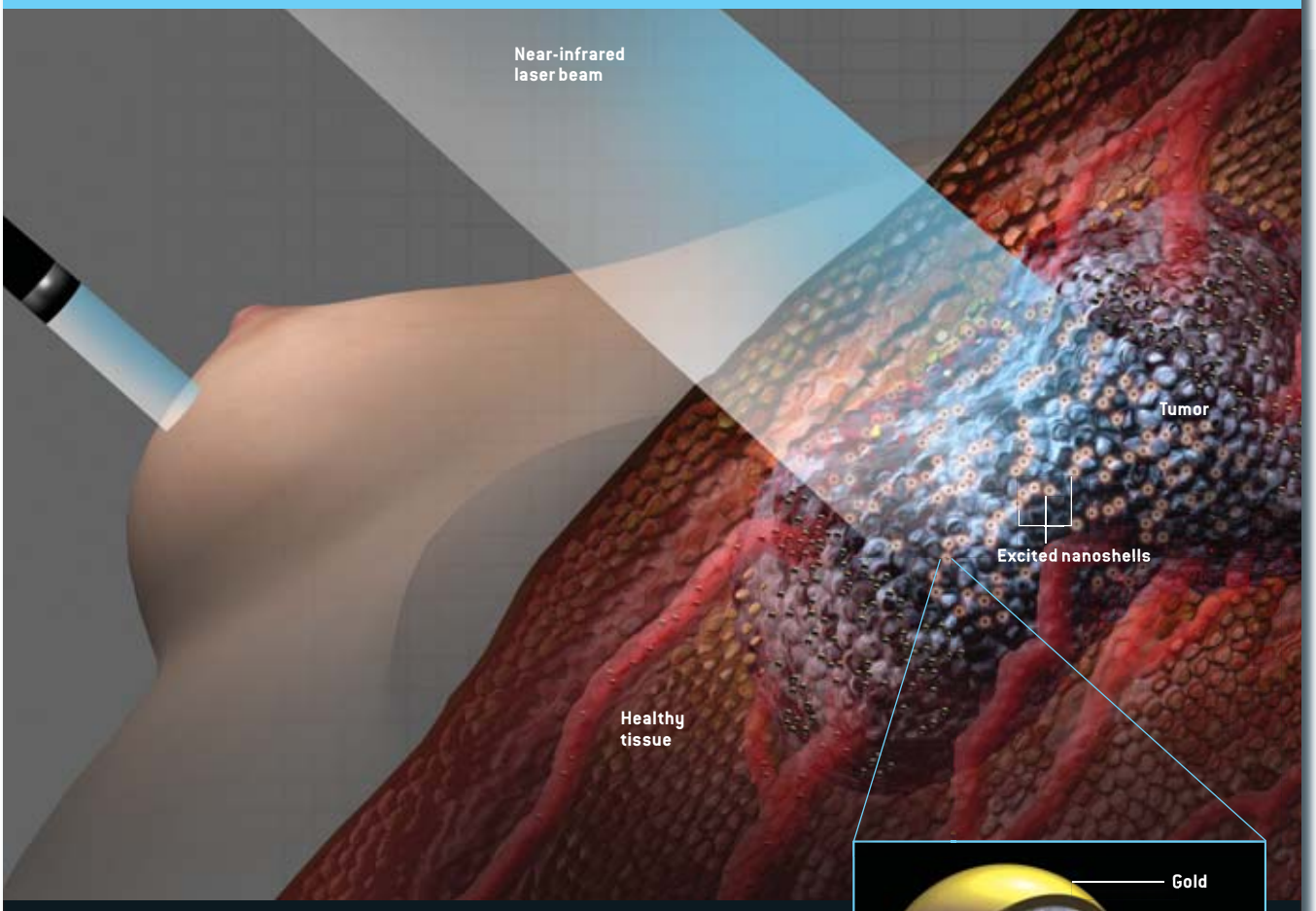
Slot waveguides could significantly boost the speed of computer chips by rapidly funneling large amounts of data to the circuits that perform logical operations. In the rendering at the left, relatively large dielectric waveguides deliver optical signals to an array of plasmonic switches (dubbed "plasmonstors"), which in turn distribute the signals to electronic transistors. The plasmonstors are composed of slot waveguides that measure 100 nanometers across at their broadest points and only 20 nanometers across at the intersections (*inset*).

PHIL SAUNDERS *Space Channel Ltd.*

search Institute in Japan, found that when he illuminated a thin gold film imprinted with millions of microscopic holes, the foil somehow transmitted more light than was expected from the number and size of the holes. Nine years later Ebbesen and his colleagues concluded that surface plasmons on the film were intensifying the transmission of electromagnetic energy.

The field of plasmonics received another boost with the discovery of novel "metamaterials"—materials in which electron oscillations can result in astounding optical properties. Two new classes of tools have also accelerated progress in plasmonics: recent increases in computational power have enabled investigators to accurately simulate the complex electromagnetic fields generated by plasmonic effects, and novel methods for constructing nanoscale structures have made it possible to

PLASMONIC THERAPY FOR CANCER



A proposed cancer treatment would employ plasmonic effects to destroy tumors. Doctors would inject nanoshells—100-nanometer-wide silica particles with an outer layer of gold (*inset*)—into the bloodstream. The nanoshells would embed themselves in a fast-growing tumor. If near-infrared laser light were pointed at the area, it would travel through the skin and induce resonant electron oscillations in the nanoshells, heating and killing tumor cells without harming the surrounding healthy tissue.

build and test ultrasmall plasmonic devices and circuits.

At first glance, the use of metallic structures to transmit light signals seems impractical, because metals are known for high optical losses. The electrons oscillating in the electromagnetic field collide with the surrounding lattice of atoms, rapidly dissipating the field's energy. But the plasmon losses are lower at the interface between a thin metal film and a dielectric than inside the bulk of a metal because the field spreads into the nonconductive material, where there are no free electrons to oscillate and hence no energy-dissipating collisions. This property naturally confines plasmons to the metallic surface abutting the dielectric; in a sandwich with dielectric and metal layers, for example, the surface plasmons propagate only in the thin plane at the interface [*see top illustration in box on preceding page*].

Because these planar plasmonic structures act as waveguides, shepherding the electromagnetic waves along the metal-dielectric boundary, they could be useful in routing signals on a chip. Although an optical signal suffers more loss in a metal than in a dielectric such as glass, a plasmon can travel in a thin-film metal waveguide for several centimeters before dying out. The propagation length can be maximized if the waveguide employs an asymmetric mode, which pushes a greater portion of the electromagnetic energy away from the guiding metal film and into the surrounding dielectric, thereby lowering loss. Because the electromagnetic fields at the top and bottom surfaces of the metal film interact with each other, the frequencies and wavelengths of the plasmons can be adjusted by changing the thickness of the film. In the late 1990s research groups led by Sergey Bozhevolnyi of

PHIL SAUNDERS *Space Channel Ltd.*

Aalborg University in Denmark and Pierre Berini of the University of Ottawa developed planar plasmonic components, operating at the telecommunications wavelength of 1,500 nanometers, that could perform many of the same functions—such as splitting guided waves—usually done by all-dielectric devices. These structures could prove useful in transmitting data from one part of a chip to another, but the electromagnetic fields accompanying the plasmons are too large to convey signals through the nanoscale innards of a processor.

Plasmons propagate like the ripples that spread across the surface of a pond after you throw a stone in the water.

To generate plasmons that can propagate through nanoscale wires, researchers have explored more complex waveguide geometries that can shrink the wavelength of the signal by squeezing it into a narrow space. In the late 1990s my lab group and a team led by Franz Aussenegg and later Joachim Krenn of the University of Graz in Austria launched parallel efforts to produce these "subwavelength" surface-plasmon waveguides. Working with me at Caltech, Stefan Maier built a structure consisting of linear chains of gold dots, each less than 100 nanometers across. A visible beam with a wavelength of 570 nanometers triggered resonant oscillations in the dots, generating surface plasmons that moved along the chains, confined to a flattened path only 75 nanometers high. The Graz group achieved similar results and imaged the patterns of the plasmons carried along the chains. The absorption losses of these nanowires were relatively high, however, causing the signal to die out after it traveled a few hundred nanometers to a few microns (millionths of a meter). Thus, these waveguides would be suitable only for very short-range interconnections.

Fortunately, the absorption losses can be minimized by turning the plasmonic waveguides inside out, putting the dielectric at the core and surrounding it with metal [*see middle illustration in box on page 59*]. In this device, called a plasmon slot waveguide, adjusting the thickness of the dielectric core changes the wavelength of the plasmons. With support from the Air Force Office of Scientific Research, my lab at Caltech and Mark Brongersma's Stanford University group have shown that plasmon slot waveguides are capable of transmitting a signal as far as tens of microns. Hideki Miyazaki of the National Institute for Materials Science in Japan obtained a striking result by squeezing red light (with a wavelength of 651 nanometers in free space) into a plasmon slot waveguide that was only three nanometers thick and 55 nanometers wide. The researchers found that the wavelength of the surface plasmon propagating through the device was 51 nanometers, or about 8 percent of the free-space wavelength.

Plasmonics can thus generate signals in the soft x-ray range of wavelengths (between 10 and 100 nanometers) by exciting materials with visible light. The wavelength can be reduced by more than a factor of 10 relative to its free-space value, and yet the frequency of the signal remains the same. (The fundamental relation between the two—frequency times wavelength equals the speed of light—is preserved because the electromagnetic waves slow as they travel along the metal-dielectric interface.) This striking ability to shrink the wavelength opens the path to nanoscale plasmonic structures that could replace purely electronic circuits containing wires and transistors.

Just as lithography is now used to imprint circuit patterns on silicon chips, a similar process could mass-produce minuscule plasmonic devices with arrays of narrow dielectric stripes and gaps. These arrays would guide the waves of positive and negative charge on the metal surface; the alternating charge densities would be very much akin to the alternating current traveling along an ordinary wire. But because the frequency of an optical signal is so much higher than that of an electrical one—more than 400,000 gigahertz versus 60 hertz—the plasmonic circuit would be able to carry much more data. Moreover, because electrical charge does not travel from one end of a plasmonic circuit to another—the electrons bunch together and spread apart rather than streaming in a single direction—the device is not subject to resistance and capacitance effects that limit the data-carrying capacity of integrated circuits with electrical interconnects.

Plasmonic circuits would be even faster and more useful if researchers could devise a "plasmonstor" switch—a three-terminal plasmonic device with transistorlike properties. My lab at Caltech and other research groups have recently developed low-power versions of such a switch. If scientists can produce plasmonstors with better performance, the devices could serve as the core of an ultrafast signal-processing system, an advance that could revolutionize computing 10 to 20 years from now.

THE AUTHOR

HARRY A. ATWATER is Howard Hughes Professor and Professor of Applied Physics and Materials Science at the California Institute of Technology. His research interests center on subwavelength-scale photonic devices for computing, imaging and renewable energy applications. In addition to devising plasmonic nanostructures, his group is actively exploring the use of new materials for solar power generation (photovoltaics), as well as the solar-driven generation of chemical fuels.

Nanoshells and Invisibility Cloaks

THE POTENTIAL USES of plasmonic devices go far beyond computing, however. Naomi Halas and Peter Nordlander of Rice University have developed structures called nanoshells that consist of a thin layer of gold—typically about 10 nanometers thick—deposited around the entire surface of a silica particle about 100 nanometers across. Exposure to electromagnetic waves generates electron oscillations in the gold shell; because of the coupling interaction between the fields on the shell's inner and outer surfaces, varying the size of the particle and the thickness of the gold layer changes the wavelength at which the particle resonantly absorbs energy. In this way, investigators can design the nanoshells to selectively absorb wavelengths as short as a few hundred nanometers (the blue end of the visible spectrum) or as long as nearly 10 microns (the near infrared).

This phenomenon has turned nanoshells into a promising tool for cancer treatment. In 2004 Halas, working with her Rice colleague Jennifer West, injected plasmonic nanoshells into the bloodstream of mice with cancerous tumors and found that the particles were nontoxic. What is more, the nanoshells tended to embed themselves in the rodents' cancerous tissues rather than the healthy ones because more blood was circulated to the fast-growing tumors. (The nanoshells can also be attached to antibodies to ensure that they target cancers.)

Fortunately, human and animal tissues are transparent to radiation at certain infrared wavelengths. When the researchers directed near-infrared laser light through the mice's skin and at the tumors, the resonant absorption of energy in the embedded nanoshells raised the temperature of the cancerous tissues from about 37 degrees Celsius to about 45 degrees C.

The photothermal heating killed the cancer cells while leaving the surrounding healthy tissue unharmed. In the mice treated with nanoshells, all signs of cancer disappeared within 10 days; in the control groups, the tumors continued to grow rapidly. Houston-based Nanospectra Biosciences is currently seeking permission from the Food and Drug Administration to conduct clinical trials of nanoshell therapy in patients with head and neck cancer.

Plasmonic materials may also revolutionize the lighting industry by making LEDs bright enough to compete with incandescent bulbs. Beginning in the 1980s, researchers recognized that the plasmonic enhancement of the electric field at the metal-dielectric boundary could increase the emission rate of luminescent dyes placed near the metal's surface. More recently, it has become evident that this type of field enhancement can also dramatically raise the emission rates of quantum dots and quantum wells—tiny semiconductor structures that absorb and emit light—thus increasing the efficiency and brightness of solid-state LEDs. In 2004 my Caltech colleague Axel Scherer, together with co-workers at Nichia Corporation in Japan, demonstrated that coating the surface of a gallium nitride LED with dense arrays of plasmonic nanoparticles (made of silver, gold or aluminum) could increase the intensity of the emitted light 14-fold.

Furthermore, plasmonic nanoparticles may enable researchers to develop LEDs made of silicon. Such devices, which would be much cheaper than conventional LEDs composed of gallium nitride or gallium arsenide, are currently held back by their low rates of light emission. My group at Caltech, working with a team led by Albert Polman of the FOM Institute for Atomic and Molecular Physics in the Netherlands, has shown that coupling silver or gold plasmonic nanostructures to silicon quantum-dot arrays could boost their light emission by about 10 times. Moreover, it is possible to tune the frequency of the enhanced emissions by adjusting the dimensions of the nanoparticles. Our calculations indicate that careful tuning of the plasmonic resonance frequency and precise control of the separation between the metallic particles and the semiconductor materials may enable us to increase radiative rates more than 100-fold, allowing silicon LEDs to shine just as brightly as traditional devices.

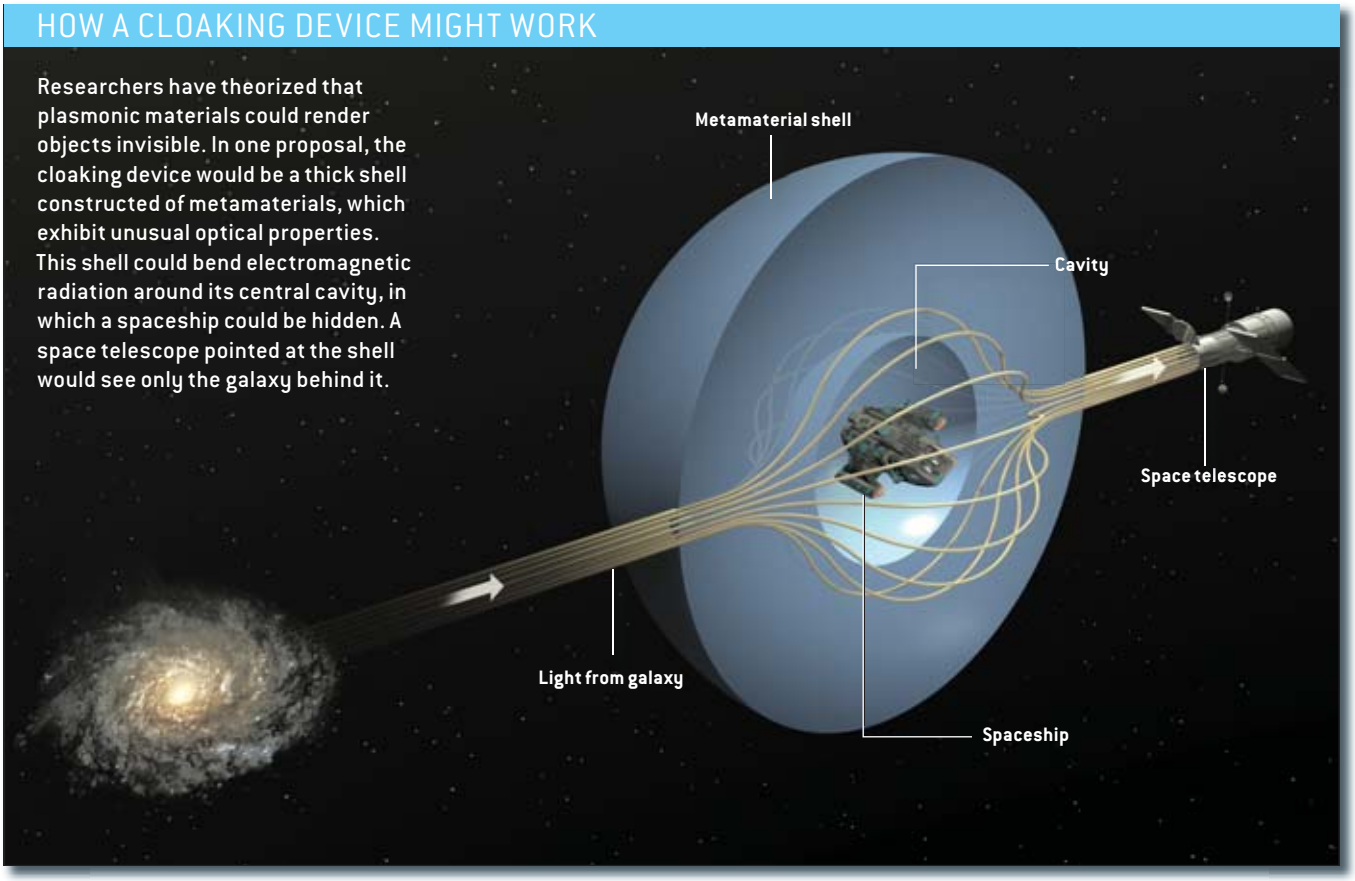
Scientists are even working on a plasmonic analogue to a laser. Mark Stockman of Georgia State University and David Bergman of Tel Aviv University have described the physics of such a device, which they called a SPASER (for *s*urface *p*lasmon *a*mplification of *s*timulated *e*mission of *r*adiation). Although the SPASER exists only in theory so far, the researchers have suggested routes to fabricating it using semiconductor quantum dots and metal particles. Radiative energy from the quantum dots would be transformed into plasmons, which

LYCURGUS CUP, a Roman goblet dating from the fourth century A.D., changes color because of the plasmonic excitation of metallic particles within the glass matrix. When a light source is placed inside the normally greenish goblet, it looks red.

BRITISH MUSEUM, LONDON/BRIDGEMAN ART LIBRARY

HOW A CLOAKING DEVICE MIGHT WORK

Researchers have theorized that plasmonic materials could render objects invisible. In one proposal, the cloaking device would be a thick shell constructed of metamaterials, which exhibit unusual optical properties. This shell could bend electromagnetic radiation around its central cavity, in which a spaceship could be hidden. A space telescope pointed at the shell would see only the galaxy behind it.



PHIL SAUNDERS *Space Channel Ltd.*

would then be amplified in a plasmonic resonator. Because the plasmons generated by a SPASER would be much more tightly localized than a conventional laser beam, the device could operate at very low power and selectively excite very small objects. As a result, SPASERs could make spectroscopy more sensitive and pave the way for hazardous-materials detectors that could identify minute amounts of chemicals or viruses.

Perhaps the most fascinating potential application of plasmonics would be the invention of an invisibility cloak. In 1897 H. G. Wells published *The Invisible Man,* a tale of a young scientist who discovers how to make his own body's refractive index equal to that of air, rendering him invisible. (A material's refractive index is the ratio of the speed of light in a vacuum to the speed of light in the material.) Exciting a plasmonic structure with radiation that is close to the structure's resonant frequency can make its refractive index equal to air's, meaning that it would neither bend nor reflect light. The structure would absorb light, but if it were laminated with a material that produces optical gain—amplifying the transmitted signal just as the resonator in a SPASER would—the increase in intensity would offset the absorption losses. The structure would become invisible, at least to radiation in a selected range of frequencies.

A true invisibility cloak, however, must be able to hide anything within the structure and work for all frequencies of visible light. The creation of such a device would be more difficult, but some physicists say it is possible. In 2006 John B. Pendry of Imperial College London and his colleagues showed that a shell of metamaterials could, in theory, reroute the electromagnetic waves traveling through it, diverting them around a spherical region within [*see box above*].

Although Wells's invisible man may never become a reality, such ideas illustrate the rich array of optical properties that inspire researchers in the plasmonics field. By studying the elaborate interplay between electromagnetic waves and free electrons, investigators have identified new possibilities for transmitting data in our integrated circuits, illuminating our homes and fighting cancer. Further exploration of these intriguing plasmonic phenomena may yield even more exciting discoveries and inventions. SA

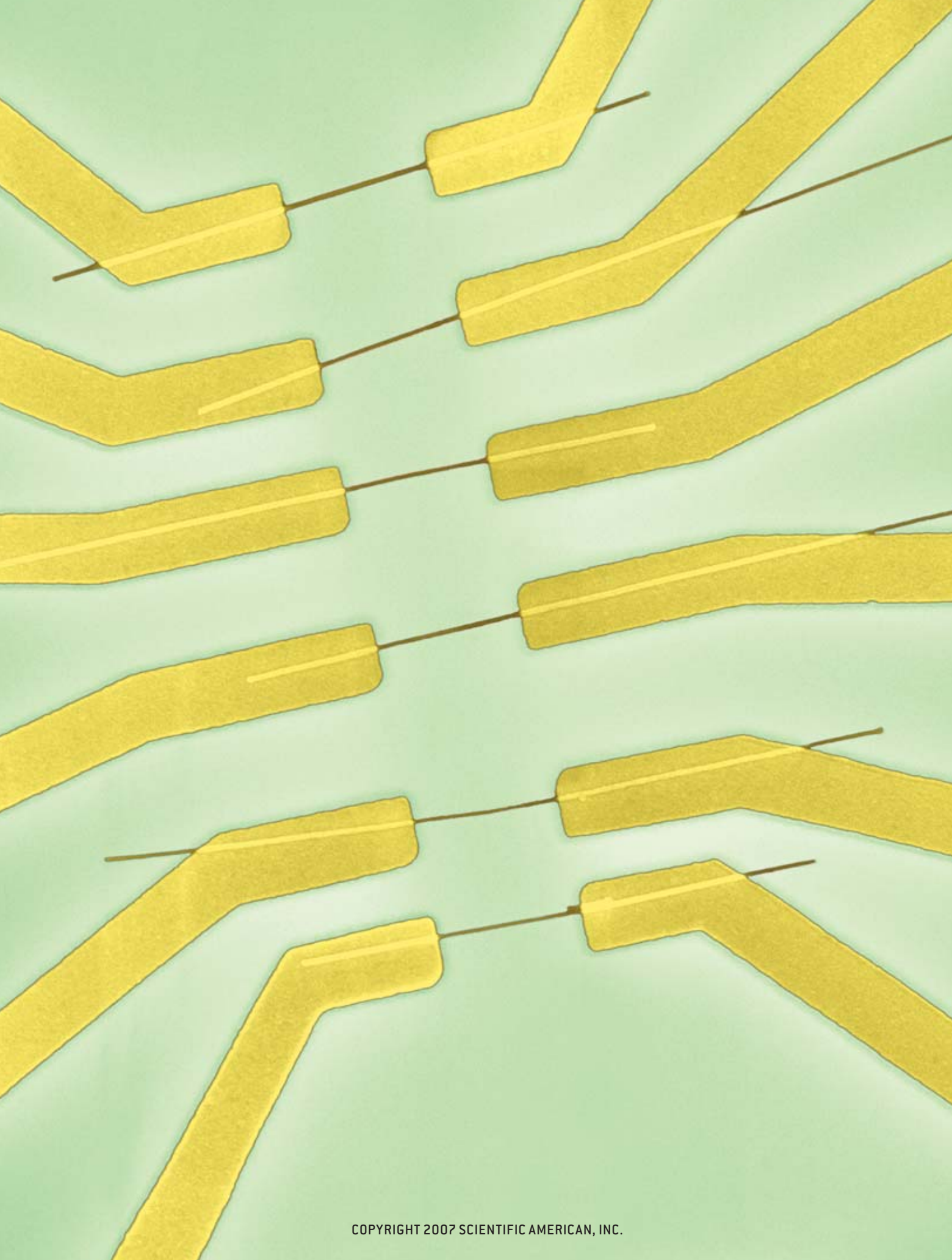
MORE TO EXPLORE

Surface Plasmon Subwavelength Optics. William L. Barnes, Alain Dereux and Thomas W. Ebbesen in *Nature,* Vol. 424, pages 824–830; August 14, 2003.

Plasmonics: Localization and Guiding of Electromagnetic Energy in Metal/Dielectric Structures. Stefan A. Maier and Harry A. Atwater in *Journal of Applied Physics,* Vol. 98, No. 1, Article No. 011101, 10 pages; July 2005.

Plasmonics: Merging Photonics and Electronics at Nanoscale Dimensions. Ekmel Ozbay in *Science,* Vol. 311, pages 189–193; January 13, 2006.

Plasmonics: Fundamentals and Applications. Stefan A. Maier. Springer Verlag, 2007.



The Incredible Shrinking Circuit

RESEARCHERS HAVE BUILT NANOTRANSISTORS AND NANOWIRES. NOW THEY JUST NEED TO FIND A WAY TO PUT THEM ALL TOGETHER

BY CHARLES M. LIEBER

NANOWIRES, each about five to 10 nanometers in diameter, may represent the future of electronics. They are the brown lines, made of indium phosphide, connecting the gold electrodes in this micrograph. These wires have been put to truly diverse uses—as memory storage and logic gates and as arrays of light-emitting diodes.

Do we really need to keep on making circuits smaller? The miniaturization of silicon microelectronics seems so inexorable that the question seldom comes up—except maybe when we buy a new computer, only to find that it becomes obsolete by the time we leave the store. A state-of-the-art microprocessor today has on the order of 500 million transistors; by 2015 it could have nearly five billion. Yet within the next two decades this dramatic march forward will run up against scientific, technical and economic limits. A first reaction might be, So what? Aren't five billion transistors enough already?

Yet when actually confronted with those limits, people will no doubt want to go beyond them. Those of us who work to keep computer power growing are motivated in part by the sheer challenge of discovering and conquering unknown territory. But we also see the potential for a revolution in medicine and so many other fields, as extreme miniaturization and new ways of building electronics enable people and machines to interact in ways that are not possible with existing technology.

Overview/*Nanoelectronics*

- Silicon chips, circuit boards, soldering irons: these are the icons of modern electronics. But the electronics of the future may look more like a chemistry set. Conventional techniques can shrink circuits only so far; engineers may soon need to shift to a whole new way of organizing and assembling electronics. One day your computer may be built in a beaker.
- Researchers have created nanometer-scale electronic components—transistors, diodes, relays, logic gates—from organic molecules, carbon nanotubes and semiconductor nanowires. Now the challenge is to wire these tiny components together.
- Unlike conventional circuit design, which proceeds from blueprint to photographic pattern to chip, nanocircuit design will probably begin with the chip—a haphazard jumble of as many as 10^{24} components and wires, not all of which will even work—and gradually sculpt it into a useful device.
- The first nanocircuits may be added on top of conventional circuits to create hybrid systems that extend the life of microelectronics and lead smoothly to the introduction of nanoelectronics.

As the word suggests, microelectronics involves components that measure roughly one micron on a side (although lately the components have shrunk to a size of about 50 nanometers). Going beyond microelectronics means more than simply shrinking components by a factor of 10 to 1,000. It also involves a paradigm shift for how we think about putting everything together.

Microelectronics and nanoelectronics both entail three levels of organization. The basic building block is usually the transistor or its nanoequivalent—a switch that can turn an electric current on or off as well as amplify signals. In microelectronics, transistors are made out of chunks of semiconductor—a material, such as impure silicon, that can be manipulated to flip between conducting and nonconducting states. In nanoelectronics, transistors might be organic molecules or nanoscale inorganic structures.

The use of molecules for electronic devices remained theoretical until a recent confluence of advances in chemistry, physics and engineering.

The next level of organization is the interconnection—the wires that link transistors together in order to perform arithmetic or logical operations. In microelectronics, wires are metal lines typically hundreds of nanometers to tens of microns in width deposited onto the silicon; in nanoelectronics, they are nanotubes or other wires as narrow as one nanometer.

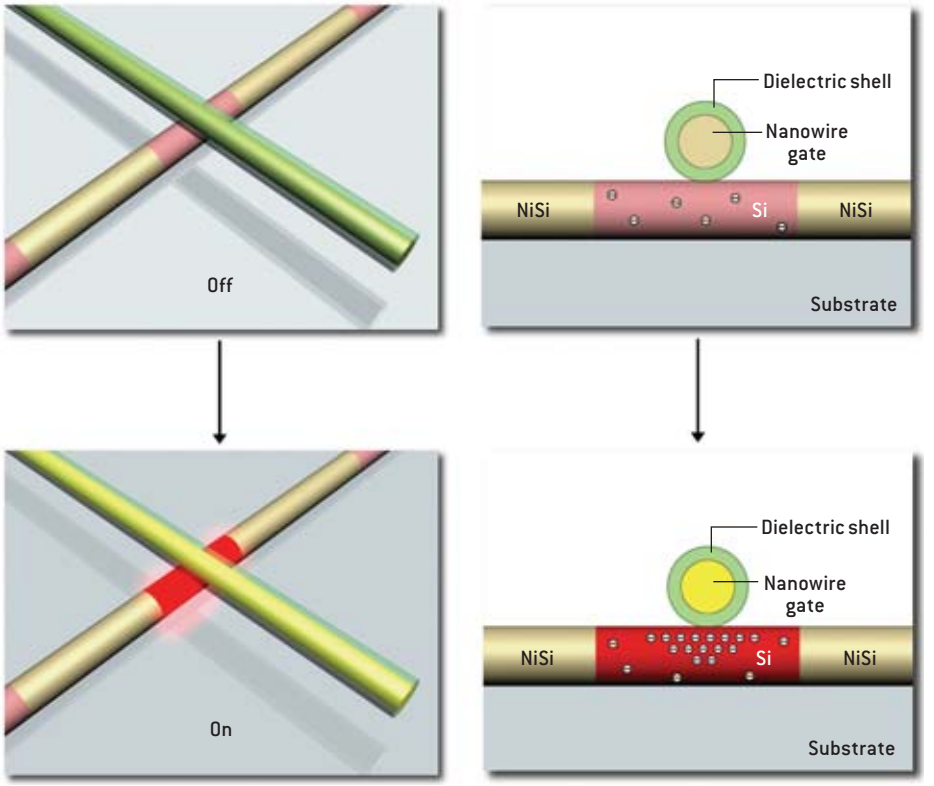
At the top level is what engineers call architecture—the overall way the transistors are interconnected, so that the circuit can plug into a computer or other system and operate independently of the lower-level details. Nanoelectronics researchers have not quite gotten to the point of testing different architectures, but we do know what abilities they will be able to exploit and what weaknesses they will need to compensate for.

In other ways, however, microelectronics and nanoelectronics could not be more different. To go from one to the other, many believe, will require a shift from top-down manufacturing to a bottom-up approach. To build a silicon chip today, fabrication plants start with a silicon crystal, lay down a pattern using a photographic technique known as lithography, and etch away the unwanted material using acid or plasma. That procedure simply does not have the precision for devices that are mere nanometers in width. Instead researchers use the methods of synthetic chemistry to produce building blocks by the mole (6×10^{23} pieces) and assemble a portion of them into progressively larger structures. Thus far the progress has been impressive. But if this research is a climb up Mount Everest, we have barely reached the base camp.

IMAGE BY FELICE FRANKEL, WITH TECHNICAL HELP FROM YU HUANG *Harvard University* (*page 64*)

NANOTRANSISTORS

NANOWIRE TRANSISTORS could be a key building block of electronics on the nanometer scale. The lower element of the crossed structure conducts electricity like a tiny wire once a voltage applied to the upper wire creates charge carriers (*electrons shown*) and switches it on. In the diagram, the lower nanowire consists of an active or switchable silicon (Si) region (*red*) sandwiched between metallic nickel silicide (NiSi) regions (*tan*) in a structure that can repeat to give multiple switching elements. The upper nanowire is isolated from direct contact by an insulating dielectric shell (*green*), and the color change (*to yellow*) highlights the voltage change applied to this control nanowire.



FANG QIAN *Lieber Group, Harvard University*

Smallifying Machines

THE USE OF MOLECULES for electronic devices was suggested more than three decades ago in a seminal paper by Arieh Aviram of IBM and Mark A. Ratner, now at Northwestern University. By tailoring the atomic structures of organic molecules, they proposed, it should be possible to concoct a transistorlike device. But their ideas remained largely theoretical until a recent confluence of advances in chemistry, physics and engineering.

Several groups have worked to evaluate Aviram and Ratner's ideas, including teams at the University of California, Berkeley, the California Institute of Technology, Hewlett-Packard, Yale University and Rice University. They have demonstrated that sandwich structures containing thousands of molecules clustered together can carry electrons from one metal electrode to another. Each molecule is about 0.5 nanometer wide and one or more nanometers long. The research groups have shown that the clusters can behave as on/off switches and might thus be usable in computer memory; once on, they will stay on for 10 minutes or so. That may not sound like a long time, but computer memory typically loses its information instantly when the power is turned off; even when the power is on, the stored information leaks away and must be "refreshed" every 0.1 second or so.

The switching mechanism for the molecules has been the subject of considerable debate. Some researchers believe it involves oxidation reduction to induce conduction, whereas others have presented strong evidence for conduction through metal filaments that form reversibly between metal contacts separated by the molecules. This latter idea is a well-known phenomenon being investigated for nonvolatile memory in conventional microelectronics.

In the on position, the clusters of molecules may conduct electricity as much as 1,000 times better than in the off position. That ratio is actually rather low compared with that of typical semiconductor transistors, whose conductivity varies a millionfold. Researchers are now working to understand the switching process itself in order to improve the observed characteristics.

My own research group at Harvard University is one of

THE AUTHOR

CHARLES M. LIEBER spent much of his childhood building—and breaking—stereos, cars and model airplanes. He is now the Mark Hyman Professor of Chemistry at Harvard University, where he directs a group of 25 undergraduate, graduate and postdoctoral researchers who focus on nanoscale science and technology. Lieber founded NanoSys, Inc., with Larry Bock of CW Ventures and Hongkun Park of Harvard in 2001. Work from Lieber's laboratory also helped to form another nanotech company, Nantero, focused on nonvolatile memory. Lieber is currently working to start a new company focused on nanoelectronics for personalized medicine.

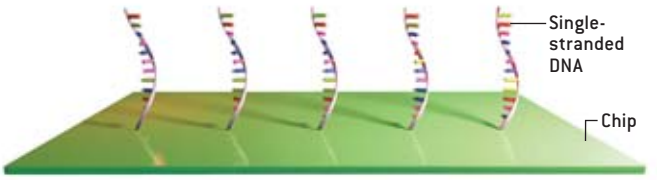
DNA Computing

Why limit ourselves to electronics? Most efforts to shrink computers assume that these machines will continue to operate much as they do today, using electrons to carry information and transistors to process it. Yet a nanoscale computer could operate by completely different means. One of the most exciting possibilities is to exploit the carrier of genetic information in living organisms, DNA.

The molecule of life can store vast quantities of data in its sequence of four bases (adenine, thymine, guanine and cytosine), and natural enzymes can manipulate this information in a highly parallel manner. The power of this approach was first brought to light by computer scientist Leonard M. Adleman of the University of Southern California in 1994. He showed that a DNA-based computer could solve a type of problem that is particularly difficult for ordinary computers—the Hamiltonian path problem, which is related to the infamous traveling salesman problem.

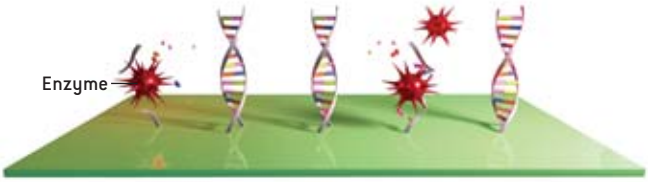
Adleman started by creating a chemical solution of DNA. The individual DNA molecules encoded every possible pathway between two points. By going through a series of separation and amplification steps, Adleman weeded out the wrong paths—those, for example, that contained points they were not supposed to contain—until he had isolated the right one. More recently, Lloyd M. Smith's group at the University of Wisconsin–Madison implemented a similar algorithm using gene chips, which may lend themselves better to practical computing (*diagram*).

Despite the advantages of DNA computing for otherwise intractable problems, many challenges remain, including the high incidence of errors caused by base-pair mismatches and the huge number of DNA nanoelements needed for even a modest computation. DNA computing may ultimately merge with other types of nanoelectronics, taking advantage of the integration and sensing made possible by nanowires and nanotubes. *—C.M.L.*

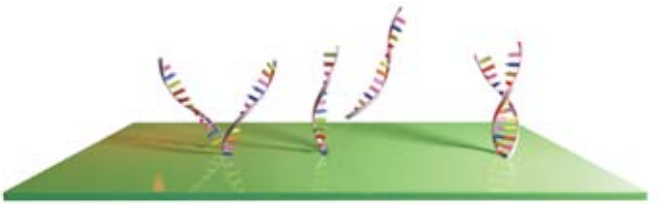


1 Single DNA strands are attached to a silicon chip. They encode all possible values of the variables in an equation that the researchers want to solve.

2 Copies of a complementary strand—which encodes the first clause of the equation—are poured onto the chip. These copies attach themselves to any strand that represents a valid solution of the clause. Any invalid solutions remain a single strand.



3 An enzyme removes all the single strands.



4 Other processes melt away the added complementary strands. These steps are repeated with all the clauses of the equation.

5 The DNA strand that survives this entire process represents the solution to the whole equation.

BRYAN CHRISTIE DESIGN

To overcome the unreliability of nanodevices, we may rely on sheer numbers—the gizmos are so cheap that plenty of spares are available.

several that have focused not on organic molecules but on long, thin, inorganic wires. One example is the carbon nanotube, which is typically about 1.4 nanometers in diameter. Not only can these nanoscale wires carry much more current, atom for atom, than ordinary metal wires, they also can act as tiny transistors. By functioning both as interconnections and as components, nanowires kill two birds with one stone. Another advantage is that they can exploit the same basic physics as standard silicon microelectronics, which makes them easier to understand and manipulate.

In 1997 Cees Dekker's group at the Delft University of Technology in the Netherlands and Paul L. McEuen's group, then at the University of California, Berkeley, independently reported highly sensitive transistors made from metallic carbon nanotubes. These devices could be turned on and off by a single electron but required very low temperatures to operate.

More recent efforts have focused on semiconductor carbon nanotubes, which can function as field-effect transistors, as first shown by Dekker and his co-workers. In addition, Hongjie Dai of Stanford University, Ali Javey, now at U.C. Berkeley, and Phaedon Avouris of the IBM Thomas J. Watson Research Center have shown that nanotube transistors can exhibit extremely high performance—exceeding that of conventional transistors—and can be configured into basic circuits such as logic gates and ring oscillators. Finally, my group has demonstrated a very different type of switch, a nanoscale electromechanical relay made from carbon nanotubes.

BRIAN TIMKO *Lieber Group, Harvard University*

CHEMISTRY MEETS BIOLOGY: This micrograph of an array, in which nanowires connect with a rat cortical neuron (*center*), shows the potential for creating functional interfaces between nanodevices and live cells or other biological systems.

Hot Wire

A MAJOR PROBLEM with nanotubes is that they are difficult to make uniform. Because a slight variation in diameter or helicity can spell the difference between a conductor and a semiconductor, a large batch of nanotubes may contain only a few working devices. In April 2001 Avouris and his co-workers started with a mixture of conducting and semiconducting nanotubes and, by either applying a current between metal electrodes or reacting with gaseous etchant, selectively removed the conducting nanotubes until just semiconducting ones were left. The solution is only partial, however, because it leaves behind open space (thus reducing device density) where metallic nanotubes once were.

My group has also been working on a different type of nanoscale wire, which we term the semiconductor nanowire. It is about the same size as a carbon nanotube, but its composition is easier to control precisely. To synthesize these wires, we start

with a metal catalyst, which defines the diameter of the growing wire and serves as the site where molecules of the desired material tend to collect. As the nanowires grow, we incorporate chemical dopants (impurities that add or remove electrons), thereby controlling whether the nanowires are *n*-type (having extra electrons) or *p*-type (having a shortage of electrons or, equivalently, a surfeit of positively charged "holes").

The availability of *n*- and *p*-type materials, which are the essential ingredients of transistors, diodes and other electronic devices, has opened up a new world for us. We have assembled a wide range of devices, including both major types of transistors (field-effect and bipolar); inverters, which transform a 0 signal to a 1; and light-emitting diodes, which pave the way for optical interconnections. Our bipolar transistors were the first molecular-scale devices ever to amplify a current. Xiangfeng Duan, a former member of my lab, was the first to assemble memory from crisscrossing *n*- and *p*-type nanowires. The memory can store information for 10 minutes or longer by trapping charge at the interface between the crossing nanowires [*see box below*].

Breaking the Logjam

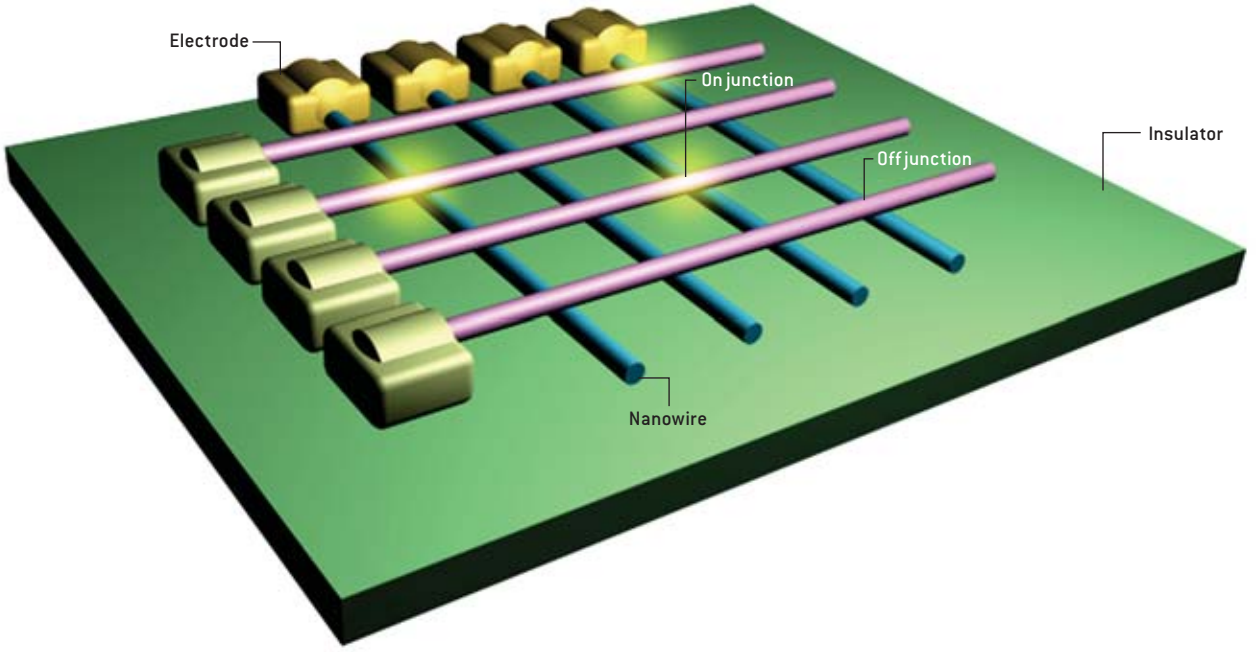
BUILDING UP AN ARSENAL of molecular and nanoscale devices is just the first step. Interconnecting and integrating these devices is perhaps the much greater challenge. First, the nanodevices must be connected to molecular-scale wires. To date, organic-molecule devices have been hooked up to conventional metal wires created by lithography. It will not be easy to substitute nanowires, because we do not know how to make a good electrical connection without ruining these tiny wires in the process. Using nanowires and nanotubes both for the devices and for the interconnections has, however, been shown to solve that problem.

Second, once the components are attached to nanowires, the wires themselves must be organized into, for example, a two-dimensional array. Duan and another member of my team, Yu Huang, made the first significant breakthrough: they assembled nanocircuits by means of fluid flows. Just as sticks and logs can flow down a river, nanoscale wires can be drawn into parallel lines using fluids.

More recent work by members of my lab has expanded these basic ideas in several very significant directions that bode well for large-scale integration and manufacturing. First, Song Jin and Dongmok Whang showed that the Langmuir-Blodgett technique could be used to organize nanowires en masse on the surface of water and then transfer them at controlled density and orientation to centimeter-scale substrates. Since then, Javey and SungWoo Nam have shown that nanowires can be

NANOWIRE ARRAY

Crisscrossing nanowires neatly solves a major problem in molecular-scale electronics: How do you connect wires to components such as transistors or diodes? The wires do double duty, serving both as wires and as components. Each junction is a component, in this case a transistor or diode switch depending on the compositions and structures of the two distinct types of nanowires. To flip a switch on or off, a certain voltage is applied to the two nanowires. Crisscrossed semiconductor nanowires have been employed to create switches that are turned on and off electrically and can form memory and logic arrays—key steps toward the assembly of a nanocomputer.



FANG QIAN *Lieber Group, Harvard University*

Soon nanodevices may have useful applications—for example, as ultrasensitive detectors of single virus particles and pieces of DNA.

directly printed onto moderate-scale wafers with controlled orientation and density, and Guihua Yu in my lab and Anyuan Cao of the University of Hawaii at Manoa have blown "polymer bubble" nanowire films that can be transferred to commercial-scale wafers and even very large flexible substrates.

These processes create interconnections in the direction of alignment, thus yielding parallel nanowire arrays. To add wires in other directions, we repeat the process, building up additional layers of nanowires. For instance, to produce a right-angle grid, we first lay down a series of parallel nanowires, then rotate the direction by 90 degrees and lay down another series. By using wires of different compositions for each layer, we can rapidly assemble an array of functional nanodevices using equipment not much more sophisticated than that in a high school chemistry lab. A grid of diodes, for example, consists of a layer of conducting nanotubes above a layer of semiconducting nanotubes, or a layer of *n*-type nanowires atop a layer of *p*-type nanowires. In both cases, each junction serves as a diode.

Intimately linked to all these efforts is the development of architectures that best exploit the unique features of nanoscale devices and the capabilities of bottom-up assembly. Although we can make unfathomable numbers of dirt-cheap nanostructures, the devices are much less reliable than their microelectronic counterparts, and our capacity for assembly and organization is still quite primitive.

In collaboration with André DeHon of the University of Pennsylvania, my group has been working on highly regular architectures based on crossed-nanowire arrays that can be generalized for universal computing machines. For memory, the architecture starts with a two-dimensional array of crossed nanowires or suspended electromechanical switches in which one can store information at each cross point. The same basic architecture is being pursued by researchers at Caltech and Hewlett-Packard, and it resembles the magnetic-core memory that was common in computers of the 1950s and 1960s.

Law of Large Numbers

TO OVERCOME the unreliability of individual nanodevices, we may rely on sheer numbers—the gizmos are so cheap that plenty of spares are always available. Researchers who work on defect tolerance have shown that computing is possible even if many of the components fail, although identifying and mapping the defects can be slow and time-consuming. Ultimately we hope to partition the enormous arrays into subarrays whose reliability can be easily monitored. The optimum size of these subarrays will depend on the defect levels typically present in molecular and nanoscale devices.

Another significant hurdle faced by nanoelectronics is "bootstrapping." How do engineers get the circuit to do what they want it to? In microelectronics, circuit designers work like architects: they prepare a blueprint of a circuit, and a fabrication plant builds it. In nanoelectronics, designers will have to work like computer programmers. A fabrication plant will create a raw nanocircuit—billions on billions of devices and wires whose functioning is rather limited. From the outside, it will look like a lump of material with a handful of wires sticking out. Using those few wires, engineers will somehow have to configure those billions of devices. Such challenges are what keeps me tremendously excited about the field as a whole.

Even before we solve these problems, nanodevices may have useful applications. For example, Gengfeng Zheng in my group has used semiconductor nanowires as ultrasensitive detectors. This technology has even been used to detect single virus particles and single pieces of DNA and, with the assembly of many sensors, could sequence the entire human genome on a single chip. The technology could also serve in minimally invasive medical devices and, as Fernando Patolsky, now at Tel Aviv University, and Brian Timko in my group have shown, could be used to build artificial synapses or two-way interfaces to live neurons.

Although substantial work remains before nanoelectronics makes its way into computers, this goal now seems less hazy than it was even a year ago. As we gain confidence, we will learn not just to shrink digital microelectronics but also to go where no digital circuit has gone before. We might assemble and interconnect multiple layers of unique functional building blocks to enable truly 3-D computational engines and nanoelectronic systems. Nanoscale devices that exhibit quantum phenomena, for example, could be exploited in quantum encryption and quantum computing. And building nanoelectronic devices on biocompatible polymers could usher in a totally new form of smart tissue or hybrid bionanoelectronic brains. The richness of the nanoworld will change the macroworld. SA

MORE TO EXPLORE

The author's Web site: http://cmliris.harvard.edu

The Avouris group: www.research.ibm.com/nanoscience

The Caltech/U.C.L.A./Hewlett-Packard team: www.its.caltech.edu/~heathgrp/ and www.hpl.hp.com/research/qsr/

The Dai group: www.stanford.edu/dept/chemistry/faculty/dai/group

The DeHon group: www.seas.upenn.edu/~andre

The Dekker group: www.mb.tn.tudelft.nl/user/dekker

The Rice / Yale team: www.jmtour.com and www.eng.yale.edu/reedlab

The Smith group: www.chem.wisc.edu/Smith/home.php

Less is more in Medicine

Sophisticated forms of nanotechnology will find some of their first real-world applications in biomedical research, disease diagnosis and, possibly, therapy

BY A. PAUL ALIVISATOS

The 1966 film *Fantastic Voyage* treated moviegoers to a bold vision of nanotechnology applied to medicine: through mysterious means, an intrepid team of doctors and their high-tech submarine were shrunk to minute size so that they could travel through the bloodstream of an injured patient and remove a life-threatening blood clot in his brain. In the past 40 years, great strides have been made in fabricating complex devices at ever smaller scales, leading some people to believe that such forms of medical intervention are possible and that tiny robots will soon be coursing through everyone's veins. Indeed, in some circles the idea is taken so seriously that worries have emerged about the dark side of such technology: Could self-replicating nanometer-scale automatons run amok and destroy the entire biological world?

In my view, shared by most investigators, such thoughts belong squarely in the realm of science fiction. Still, nanotechnology can potentially enhance biomedical research tools—for example, by providing new kinds of labels for experiments done to discover drugs or to reveal which sets of genes are active in cells under various conditions. Nanoscale devices could, moreover, play a part in quick diagnostic screens and in genetic tests, such as those meant to determine a person's susceptibility to different disorders or to reveal which specific genes are mutated in a patient's cancer. Investigators are also studying them as improved contrast agents for noninvasive imaging and as drug-delivery vehicles. The emerging technologies may not be as photogenic as a platelet-size Raquel Welch blasting away at a clot with a laser beam, but they are every bit as dramatic because, in contrast, the benefits they offer to patients and researchers are real.

How exactly can nanotechnology do all these things? The answer hinges on one's definitions. All of biology is arguably a form of nanotechnology. After all, even the most complicated creature is made up of tiny cells, which themselves are constructed of nanoscale building blocks: proteins, lipids, nu-

HIGHLY MAGNIFIED VIALS contain solutions of quantum dots—semiconductor nanocrystals—of specific sizes. The precise size of a quantum dot determines the color it emits after exposure to light. By attaching different sizes of dots to different biological molecules, investigators can track the activities of many tagged molecules simultaneously.

The technologies may not be as photogenic as a minute Raquel Welch blasting away at a clot, but they are just as dramatic because their benefits are real.

cleic acids and other complex biological molecules. But by convention the term "nanotechnology" is usually restricted to artificial constructions made, say, from semiconductors, metals, plastic or glass. A few inorganic structures of nanometer scale—minute crystals, for instance—have already been commercialized, notably as contrast agents.

Magnetic Attraction

NATURE ITSELF provides a beautiful illustration of the usefulness of such inorganic crystals in a biological context: humble magnetotactic (magnetic-sensing) bacteria. Such organisms, which live in bodies of water and their muddy bottoms, thrive only at one depth in the water or sediment. Above this position, oxygen is too abundant for their liking; below, too scarce. A bacterium that drifts away from the right level must swim back, and so, like many of its cousins, the microbe wields a whiplike tail for propulsion. But how does the buoyant cell tell up from down when gravity has essentially no effect on it?

The answer is that this bacterium has fixed within it a chain of about 20 magnetic crystals that are each between 35 and 120 nanometers in diameter. Together these crystals constitute a miniature compass. Because the magnetic field of the earth is inclined in most places (it points not only north but also downward in the Northern Hemisphere and upward in the Southern), a magnetotactic bacterium can follow a magnetic field line up or down to its desired destination.

This compass is a marvel of natural nanoscale engineering. For one, it is made of the perfect material—either magnetite or greigite, both highly magnetic iron minerals. The use of multiple crystals is also no accident. At very small scales, the larger a magnetic particle is, the longer it stays magnetized. But if the particle becomes too large, it will spontaneously form two separate magnetic domains with oppositely directed magnetizations. Such a crystal has little overall magnetization and does not make for a very effective compass needle. By building its compass out of crystals that are of just the right size to exist as stable, single magnetic domains, the bacterium makes the best use of every bit of iron it lays down. Interestingly, when people design media for hard-disk storage, they follow exactly the same strategy, using magnetic nanocrystals that are of the proper size to be both stable and strong.

Artificial magnetic crystals of similar dimension might soon serve biomedical research in a novel way. Two groups, one in Germany and the other at my institution, the University of California, Berkeley, are exploring the use of magnetic nanoparticles to detect particular biological entities, such as microorganisms that cause disease.

Their method, like many of the techniques applied today, requires suitable antibodies, which bind to specific targets. The magnetic particles are affixed, as labels, to selected antibody molecules, which are then applied to the sample under study. To detect whether the antibodies have latched onto their target, an investigator applies a strong magnetic field (which temporarily magnetizes the particles) and then examines the specimen with a sensitive instrument capable of detecting the weak magnetic fields emanating from the probes. Labeled antibodies that have not docked to the sample tumble about so rapidly in solution that they give off no magnetic signal. Bound antibodies, however, are unable to rotate, and together their magnetic tags generate a readily detectable magnetic field.

Because the unbound probes produce no signal, this approach does away with the time-consuming washing steps usually required of such assays. The sensitivity demonstrated with this experimental technique is already better than with standard methods, and anticipated improvements in the apparatus should soon boost sensitivity by a factor of several hundred.

Despite these advantages, the magnetic method probably will not completely replace the widespread practice of labeling probes with a fluorescent tag, typically an organic molecule that glows with a characteristic hue when it is energized by light of a particular color. Colors are very useful in various diagnostic and research procedures, such as when more than one probe needs to be tracked.

The world of modern electronics is

IMAGE BY FELICE FRANKEL, WITH TECHNICAL HELP FROM K. F. JENSEN, M. G. BAWENDI, C. MURRAY, C. KAGAN, B. DABBOUSI AND J. RODRIGUEZ-VIEGO *M.I.T.* (*page 72*)

Overview/*Nanomedicine*

- Nanometer-scale objects made of inorganic materials can serve in biomedical research, disease diagnosis and even therapy.
- Biological tests measuring the presence or activity of selected substances become quicker, more sensitive and more flexible when certain nanoscale particles are put to work as tags or labels.
- Nanoparticles could be used to deliver drugs just where they are needed, avoiding the harmful side effects that so often result from potent medicines.
- Artificial nanoscale building blocks may one day be used to help repair such tissues as skin, cartilage and bone—and they may even help patients regenerate organs.

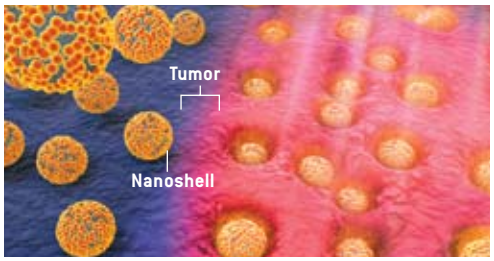
A GRAND PLAN FOR MEDICINE

The National Nanotechnology Initiative includes among its goals, or "grand challenges," a host of futuristic improvements in the detection, diagnosis and treatment of disease. Some are depicted here. The goals, many of which are far from being realized, also feature new aids for vision and hearing, rapid tests for detecting disease susceptibility and responses to drugs, and tiny devices able to find problems—such as incipient tumors, infections or heart problems—and to relay the information to an external receiver or fix them on the spot.

1 GOAL: Improved Imaging

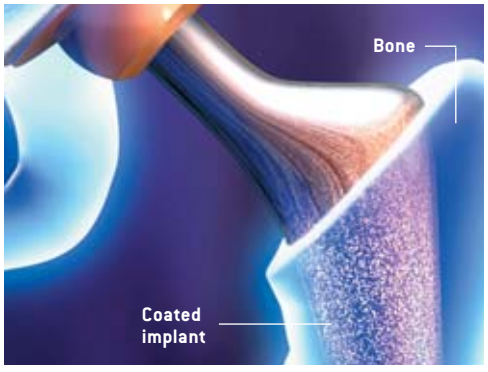
Improved or new contrast agents would detect problems at earlier, more treatable stages. They might, for instance, reveal tumors (*red*) only a few cells in size.

2 GOAL: New Ways to Treat Disease



Nanoparticles would deliver treatments to specifically targeted sites, including places that standard drugs do not reach easily. For example, gold nanoshells (*spheres*) that were targeted to tumors might, when hit by infrared light, heat up enough to destroy the growths.

3 GOAL: Superior Implants



Nanometer-scale modifications of implant surfaces would improve implant durability and biocompatibility. For instance, an artificial hip coated with nanoparticles might bond to the surrounding bone more tightly than usual.

JEFF JOHNSON *Hybrid Medical Animation*

also full of light-emitting materials. Every CD player, for instance, reads the disc with light from a solid-state laser diode, which is made of an inorganic semiconductor. Imagine carving out a vanishingly small piece of that material, a scoop the size of a protein molecule. The result is a semiconductor nanocrystal, or, in the talk of the trade, a "quantum dot." Like nanoscale magnetic crystals, these minuscule dots have much to offer biomedical researchers.

As the name suggests, quantum dots owe their special properties to the weird rules of quantum mechanics, the same rules that restrict the electrons in atoms to certain discrete energy levels. An organic dye molecule absorbs only photons of light with just the right energy to lift its electrons from their quiescent state to one of the higher levels available to them. That is, the incident light must be exactly the right wavelength, or color, to do the job. The molecule will subsequently emit a photon when the electron falls back to a lower energy level. This phenomenon is quite different from what happens in bulk semiconductors, which allow electrons to occupy two broad

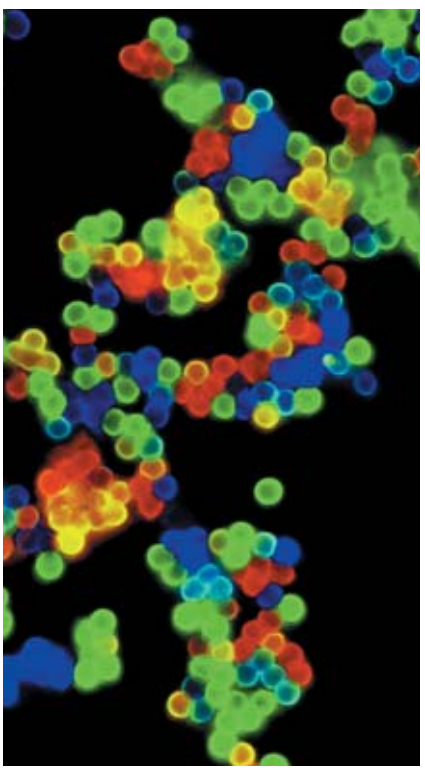
THE AUTHOR

A. PAUL ALIVISATOS is a professor in the department of chemistry at the University of California, Berkeley, where he received his doctorate in physical chemistry in 1986. A Fellow of both the American Association for the Advancement of Science and the American Physical Society, Alivisatos has received many awards for his research on the physical properties of nanocrystals. He is the scientific founder of Quantum Dot Corporation, now a subsidiary of Invitrogen, which is working to commercialize the use of semiconductor nanocrystals as fluorescent labels in biomedical tests.

For therapy, one might encapsulate drugs within nanometer-scale packages that control the medicines' release in sophisticated ways.

bands of energy. Such materials can absorb photons in a broad range of colors (all those that have enough energy to bridge the gap between these two bands), but they emit light only at one specific wavelength, corresponding to the band-gap energy. Quantum dots are an intermediate case. Like bulk semiconductors, they absorb photons of all energies above the threshold of the band gap. But the wavelength of light a quantum dot emits—its color—depends very strongly on the dot's size. Hence, a single type of semiconducting material can yield an entire family of distinctly colored labels.

Physicists first studied quantum dots in the 1970s, thinking that they might one day fashion new electronic or optical devices. Few of the pioneering investigators had any idea that these objects could help diagnose disease or discover new drugs. And none of them would have dreamed that the first real-world applications of quantum dots would be in biology and medicine. Making quantum dots that would function properly in biological systems did indeed require years of research, but they are now a reality.



LATEX BEADS filled with quantum dots of single colors glow at nearly the same wavelengths as the dots themselves. Researchers have also loaded selections of different dots into single beads. Their aim is to create a huge variety of distinct labels for biological tests [*see "Nano Bar Codes" in box on opposite page*].

The Rainbow Coalition

THE FIRST COMMERCIALLY available quantum dots for biological imaging were created by combining techniques developed in my laboratories and those of Moungi Bawendi at the Massachusetts Institute of Technology. Today multiple-color dots are extensively used by biomedical researchers the world over and by clinical pathologists in hospitals.

Semiconductor nanocrystals have several advantages over conventional dye molecules. Small inorganic crystals can withstand significantly more cycles of excitation and light emission than can typical organic molecules, which soon decompose. And this stability allows investigators to track the goings-on in cells and tissues for longer intervals than can now be achieved. But the greatest benefit semiconductor nanocrystals offer is less subtle—they come in more colors.

Biological systems are very complex, and frequently several components must be observed simultaneously. Such tracking is difficult to achieve, because each organic dye must be excited with a different wavelength of light. But quantum dots make it possible to tag a variety of biological molecules, each with a crystal of a different size (and hence color). And because all these crystals can be energized with a single light source, they can all be monitored at once.

This approach is being pursued actively, but quantum dots offer even more interesting possibilities. Imagine a small latex bead filled with a combination of quantum dots. The bead could, for instance, contain five different sizes of dots, or five colors, in a variety of concentrations. After the bead is illuminated, it will give off light, which when spread out by a prism will produce five distinct spectral lines with prescribed intensities—a spectral bar code, if you like. Such beads allow for an enormous number of distinct labels (billions, potentially), each of which could be attached, say, to DNA molecules composed of different sequences of genetic building blocks.

With these kinds of beads, technicians could easily compare the genetic material in a sample against a library of known DNA sequences, as might be done if an investigator wanted to find out which genes were active in certain cells or tissues. They would simply expose the sample to the full beaded library and read the spectral bar codes of the library DNAs that bind to sequences in the sample. Because binding takes place only when genetic sequences match closely (or, more precisely, when one sequence complements the other), the results would immediately reveal the nature of the genetic material in the sample.

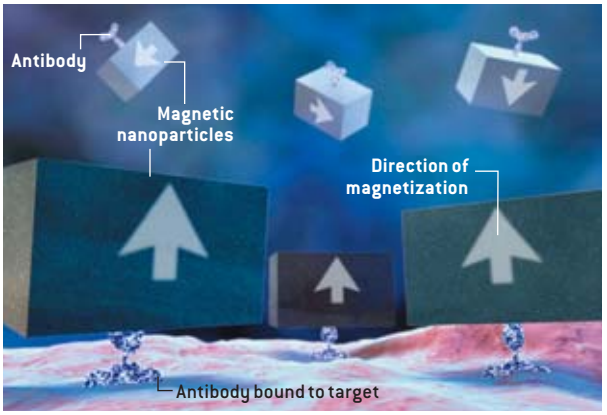
Semiconductor quantum dots should soon serve biomedical researchers in this way, but they are not the only nanostructures useful for optically sensing the genetic composition of biological specimens. Another example emerges from the work of Chad A. Mirkin and Robert L. Letsinger of Northwestern University, who developed an ingenious method to test for the presence of a specific genetic sequence in solution. Their scheme employs 13-nanometer gold particles studded with DNA.

The trick here is to use two sets of gold particles. The first set carries DNA

SHUMING NIE *Emory University and Georgia Institute of Technology*

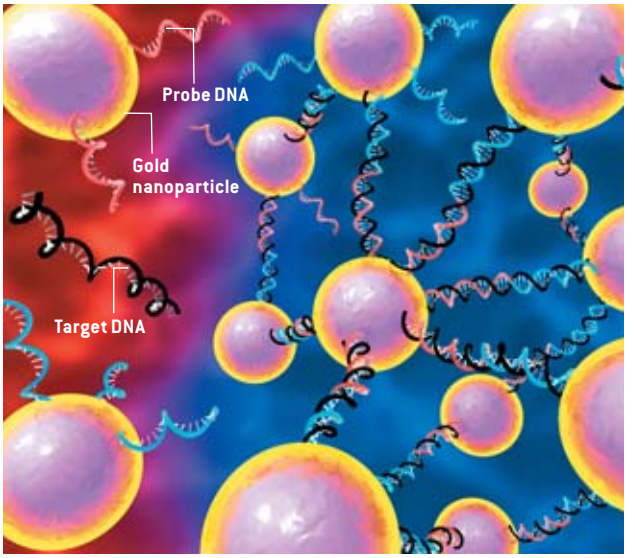
BIONANOTECH IN ACTION

The items here could one day enhance the speed and power of biomedical tests, such as those used to screen small samples of material for the presence of particular genetic sequences. For clarity, the images have not been drawn to scale.



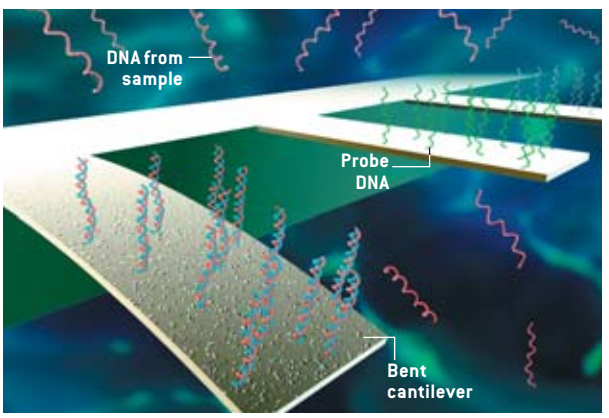
MAGNETIC TAGS

Many tests reveal the presence of a molecule or disease-causing organism by detecting the binding of an antibody to that target. When antibodies labeled with magnetic nanoparticles bind to their target on a surface (*foreground*), brief exposure to a magnetic field causes these probes collectively to give off a strong magnetic signal. Meanwhile unbound antibodies tumble about in all directions, producing no net signal. This last property makes it possible to read the results without first washing away any probes that fail to find their target.



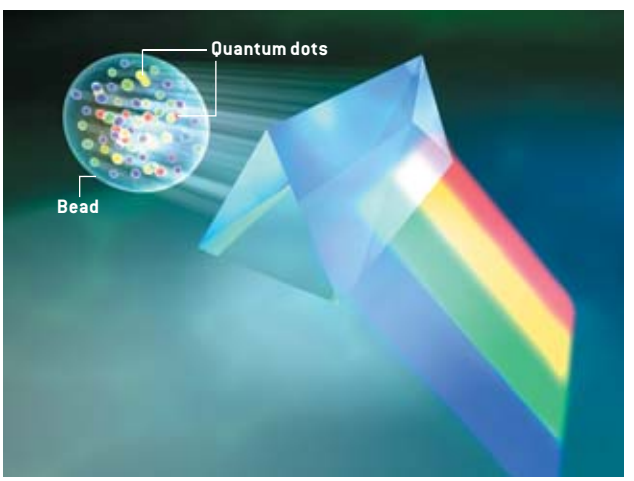
GOLD PARTICLES

Gold nanoparticles studded with short segments of DNA could form the basis of an easy-to-read test for the presence of a genetic sequence (*black*) in a sample under study. DNA complementary to half of such a sequence (*red*) is attached to one set of particles in solution, and DNA complementary to the other half (*blue*) is attached to a second set of particles. If the sequence of interest is present in the sample, it will bind to the DNA tentacles on both sets of spheres, trapping the balls in a dense web. This agglomeration will cause the solution to change color (*from red to blue*).



CLEVER CANTILEVERS

Biological samples can be screened for the presence of particular genetic sequences using small beams (cantilevers) of the type employed in atomic force microscopes. The surface of each cantilever is coated with DNA able to bind to one particular target sequence. A sample is then applied to the beams. Binding induces a surface stress, which bends the affected beams by nanometers—not much, but enough to reveal that the bent beams found their specific targets in a sample.



NANO BAR CODES

Latex beads filled with several colors of nanoscale semiconductors known as quantum dots can potentially serve as unique labels for any number of different probes. In response to light, the beads would identify themselves (and, thus, their linked probes) by emitting light that separates into a distinctive spectrum of colors and intensities—a kind of spectral bar code.

JEFF JOHNSON *Hybrid Medical Animation*

that binds to one half of the target sequence; the second set carries DNA that binds to the other half. DNA with the complete target sequence readily attaches to both types of particles, linking them together. Because each particle has multiple DNA tentacles, bits of genetic material carrying the target sequence can glue many particles together. And when these gold specks aggregate, their optical properties shift markedly, changing the test solution from red to blue. Because the outcome of the test is easy to see without any instrumentation at all, such a system might be particularly useful for home DNA testing.

Feeling the Force

NO DISCUSSION of bionanotechnology would be complete without at least a brief mention of one of the hottest instruments in science today—the atomic force microscope. Such devices probe materials in the same way an old-fashioned phonograph reads the grooves in a record: by dragging a sharp point over the surface and detecting the resulting deflections. The tip of an atomic force microscope is, however, much finer than a phonograph needle, so it can sense far smaller structures. Regrettably, fabricating tips that are both fine and sturdy for these microscopes has proved to be quite difficult.

The solution appeared in 1996, when workers at Rice University affixed a slender carbon nanotube to the tip of an atomic force microscope, making it possible to probe samples just a few nanometers in size. In 1998 Charles M. Lieber and his co-workers at Harvard University applied this approach to probing biomolecules, providing a very high resolution means to explore complex biological molecules and their interactions at the most basic level.

But atomic force microscopy may soon be applied to more than just making fundamental scientific measurements. In 2000 James K. Gimzewski, then at the IBM Zurich Research Laboratory, showed with collaborators at IBM and the University of Basel that an array of micron-scale arms, or cantilevers, much like the ones employed in atomic force microscopes, could be used to screen samples for the presence of certain genetic sequences. They attached short strands of DNA to the tops of the cantilevers. When genetic material carrying a complementary sequence binds to the anchored strands, it induces a surface stress, which bends the cantilevers subtly—by just nanometers—but enough to be detected. By fabricating devices with many cantilevers and coating each with a different type of DNA, researchers should be able to test a biological sample rapidly for the presence of specific genetic sequences (as is now done routinely with gene chips) by nanomechanical means without the need for labeling.

This example, like the others described earlier, illustrates that the connections between nanotechnology and the practice of medicine are often indirect, in that much of the new work promises only better research tools or aids to diagnosis. But in some cases, nano-objects being developed may themselves prove useful for therapy. One might, for instance, encapsulate drugs within nanometer-scale packages that control the medicines' release in sophisticated ways.

Consider a class of artificial molecules called organic dendrimers. More than two decades ago Donald A. Tomalia of the Michigan Molecular Institute in Midland fashioned the first of these intriguing structures. A dendrimer molecule branches successively from inside to outside. Its shape resembles what one would get by taking many sprigs from a tree and poking them into a foam ball so that they shot out in every direction. Dendrimers are globular molecules about the size of a typical protein, but they do not come apart or unfold as easily as proteins do, because they are held together with stronger chemical bonds.

Like the lush canopies of mature trees, dendrimers contain voids. That is, they have an enormous amount of internal surface area. Interestingly, they can be tailored to have a range of different cavity sizes—spaces that are just perfect for holding therapeutic agents. Dendrimers can also be engineered to transport DNA into cells for gene therapy, and they might work more safely than

Petite Plumbing Jobs

Microfluidics enhances biomedical research

Most of the nanotechnologies now being developed for biomedical use take the form of minute objects immersed in comparatively large quantities of fluid, be it water, blood or a complex experimental concoction. But investigators are also building devices to manipulate tiny amounts of such liquids. These so-called microfluidic systems pump solutions through narrow channels, controlling the flow with diminutive valves and intense electric fields.

The ability to handle vanishingly small quantities of a solution in this way allows biomedical researchers to carry out many different experiments on what might be only a modest amount of sample—and to do so in an efficient manner, with hundreds of tests being performed, say, on the surface of a single glass slide. Microfluidic devices also offer researchers the means to carry out experiments that could not otherwise be done; for example, to deliver test solutions of specific compositions to different parts of a cell under study.

Although many of the components being created for these systems are considerably larger than a micron, some experimental devices include nanoscale dimensions. Notably, Harold G. Craighead's team at Cornell University has devised methods for sorting different sizes of DNA fragments in water according to how fast the fragments traverse passages measuring 100 nanometers across or travel through microchannels that repeatedly narrow to a depth of 75 to 100 nanometers. These or other nanofluidic devices could potentially increase the speed and reduce the costs of separating DNA molecules for sequencing and could in theory be adapted for separating proteins or other molecules. —*A.P.A.*

the other leading method: genetically modified viruses.

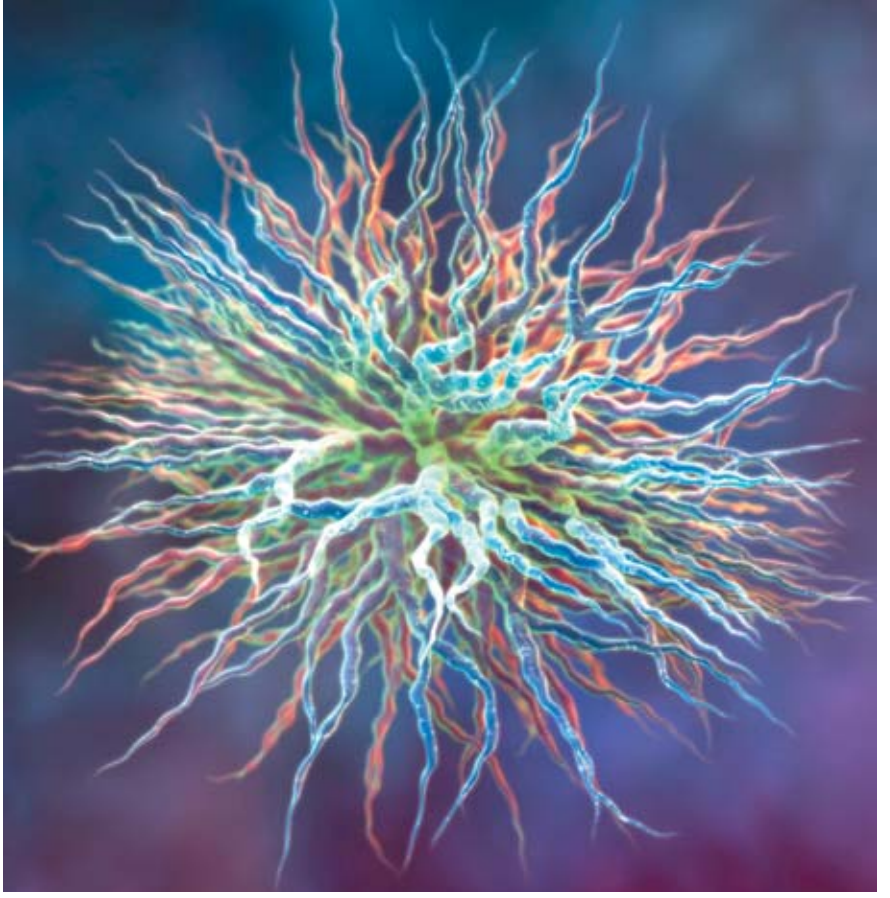
Other types of nanostructures possess high surface area, and these, too, may prove useful for delivering drugs where they are needed. But dendrimers offer the greatest degree of control and flexibility. It may be possible to design dendrimers that spontaneously swell and liberate their contents only when the appropriate trigger molecules are present. This ability would allow a custom-made dendrimer to release its load of drugs in just the tissues or organs needing treatment.

Other drug-delivery vehicles on the horizon include hollow polymer capsules under study by Helmuth Möhwald of the Max Planck Institute of Colloids and Interfaces in Golm, Germany. In response to certain signals, these capsules swell or compress to release drugs. Also intriguing are so-called nanoshells, invented by Naomi Halas and her co-workers at Rice.

Nanoshells are extremely small beads of glass coated with gold. They can be fashioned to absorb light of almost any wavelength, but nanoshells that capture energy in the near-infrared are of most interest because these wavelengths easily penetrate several centimeters of tissue. Nanoshells injected into the body can therefore be heated from the outside using a strong infrared source. Such a nanoshell could be made to deliver drug molecules at specific times by attaching it to a capsule made of a heat-sensitive polymer. The capsule would release its contents only when gentle heating of the attached nanoshell caused it to deform.

A more dramatic application envisioned for nanoshells is in cancer therapy. The idea is to link the gold-plated spheres to antibodies that bind specifically to tumor cells. Heating the nanoshells sufficiently would in theory destroy the cancerous cells, while leaving nearby tissue unharmed. FDA approval for clinical trials of nanoshells is pending.

It is, of course, difficult to know for certain whether nanoshells will ultimately fulfill their promise. The same can be said for the myriad other minus-



JEFF JOHNSON *Hybrid Medical Animation*

ORGANIC DENDRIMER, shown in an artist's conception, could be roughly the size of a protein molecule. Dendrimers harbor many internal cavities and are being eyed as drug-delivery vehicles.

cule devices being developed for medical use—among them, one-nanometer buckyballs made from just a few dozen carbon atoms. Yet it seems likely that some of the objects being investigated today will be serving doctors in the near future. Even more exciting is the prospect that physicians will make use of nanoscale building blocks to form larger structures, thereby mimicking the natural processes of biology. Such materials might eventually serve to repair damaged tissues. Research on these bold strategies is just beginning, but at least one enterprise already shows that the notion has merit: building scaffoldings on which to grow bone. Samuel I. Stupp of Northwestern is pioneering this approach using synthetic molecules that combine into fibers to which bone cells have a strong tendency to adhere.

What other marvels might the future hold? Although the means to achieve them are far from clear, sober nanotechnologists have stated some truly ambitious goals. One of the "grand challenges" of the National Nanotechnology Initiative is to find ways to detect cancerous tumors that are a mere few cells in size. Researchers also hope eventually to develop ways to regenerate not just bone or cartilage or skin but also more complex organs, using artificial scaffoldings that can guide the activity of seeded cells and can even direct the growth of a variety of cell types. Replacing hearts or kidneys or livers in this way might not match the fictional technology of *Fantastic Voyage*, but the thought that such medical therapies might actually become available in the not so distant future is still fantastically exciting. SA

MORE TO EXPLORE

Ultrasensitive Magnetic Biosensor for Homogeneous Immunoassay. Y. R. Chemla, H. L. Grossman, Y. Poon, R. McDermott, R. Stevens, M. D. Alper and J. Clarke in *Proceedings of the National Academy of Sciences USA,* Vol. 97, No. 27, pages 14268–14272; December 19, 2000.

The author's Web site: **www.cchem.berkeley.edu/~pagrp/**

Naomi Halas Nanophotonics Group: **www.ece.rice.edu/~halas/**

Information on using nanotechnology to combat cancer is available at **http://nano.cancer.gov**

Information about quantum dots and their use in biomedicine is available at **www.qdots.com**

FRED GAMBINO

Shamans of Small

Like interstellar travel, time machines and cyberspace, nanotechnology has become one of the core plot devices on which science-fiction writers draw

By Graham P. Collins

"Reading Drexler's *Engines of Creation* in 1990 went into the making of the world in *Queen City Jazz,* my first novel, though the book drew from many other sources: Shakers, ragtime, jazz, American literature, even Krazy Kat." So writes Kathleen Ann Goonan in the Summer 2001 *SFWA Bulletin,* the quarterly of the Science Fiction and Fantasy Writers of America, in a brief essay about her award-nominated novel *Crescent City Rhapsody,* the third book of her musically structured Nanotech Quartet.

Unsurprisingly, Goonan is far from the only science-fiction writer to take inspiration from K. Eric Drexler's vision of molecular nanotechnology, for it is a vision that connects to numerous preexisting themes of science fiction and offers writers an extraordinarily broad palette of capabilities, all imbued with the appearance of scientific plausibility. Touted by its proponents to be upon us within a decade or few, nanotechnology also gives science-fiction writers a chance to engage in the art of predicting and warning about possible futures. This role as an ad hoc think tank is one that innumerable science-fiction writers and fans take on enthusiastically, not only in their fiction but also in endless earnest panel discussions at conventions, in online newsgroups and discussion boards, and in articles labeled (sometimes optimistically) nonfiction. It is the culture of the intensely technophilic—even those who write of techno-dystopias and apocalypses are enrapt in a love-hate relationship with science and technology. The borders between four dominions—those of scientists, writers, readers and science-fiction fans—are hopelessly blurred, with countless individuals holding joint citizenships.

MICROSCOPIC CRAFT, built using nanotech, battle foreign invaders the way an immune system does. Even the implant-enhanced eye sees only a fog sparkling with laser light.

NANOTITLES DISCUSSED

QUEEN CITY JAZZ. Kathleen Ann Goonan. Tor, 1994.
CRESCENT CITY RHAPSODY. Kathleen Ann Goonan. Avon Eos, 2000.
FANTASTIC VOYAGE: MICROCOSM. Kevin J. Anderson. Onyx Books, 2001.
MICROCOSMIC GOD: THE COMPLETE STORIES OF THEODORE STURGEON, Vol. 2. Theodore Sturgeon. Edited by Paul Williams. North Atlantic Books, 1995.
BLOOD MUSIC. Greg Bear. Arbor House, 1985.
DUNE. Frank Herbert. Chilton Books, 1965.
NANOWARE TIME. Ian Watson. Tor, 1991.
/ [SLANT]. Greg Bear. Tor, 1997.
THE DIAMOND AGE. Neal Stephenson. Bantam, 1995.
QUEEN OF ANGELS. Greg Bear. Warner Books, 1990.
YEAR'S BEST SF 12. Edited by David G. Hartwell and Kathryn Cramer. Eos, 2007.

STORY EXCERPT: EVERYDAY NANO-NOVELTIES

We Were Out of Our Minds with Joy, by David Marusek (from *Asimov's Science Fiction,* November 1995)

I held patents for package applications in many fields, from emergency blankets and temporary skin, to military camouflage and video paint. But my own favorites, and probably the public's as well, were my novelty gift wraps. My first was a video wrapping paper that displayed the faces of loved ones (or celebrities if you had no loved ones) singing "Happy Birthday" to the music of the New York Pops. That dated back to 2025 when I was a molecular engineering student.

My first professional design was the old box-in-a-box routine, only my boxes didn't get smaller as you opened them, but larger, and in fact could fill the whole room until you chanced upon one of the secret commands, which were any variation of "stop" (whoa, enough, cut it out, etc.) or "help" (save me, I'm suffocating, get this thing off me, etc.).

Next came wrapping paper that screamed when you tore or cut it. That led to paper that resembled human skin. It molded itself perfectly and seamlessly (except for a belly button) around the gift and had a shelf life of fourteen days. You had to cut it to open the gift, and of course it bled. We sold mountains of that stuff.

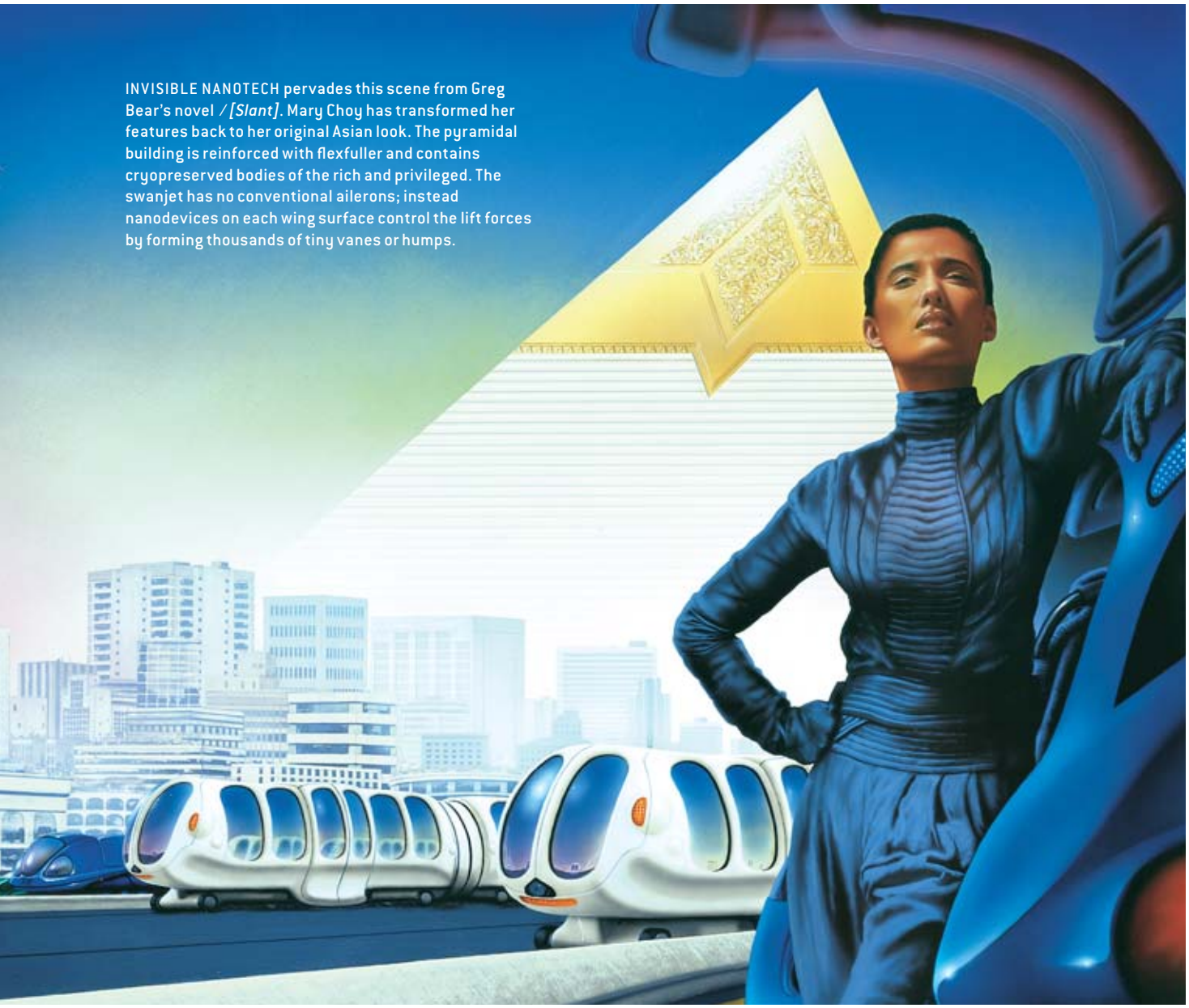
But it would be a mistake to think that science fiction's central role is one of serious prognostication. The question "What if . . . ?" lies at the heart of science fiction, but what comes after the ellipsis and the answers that stories give are ultimately not science but literature—that strange mix of entertainment and meaningful enrichment of life. The art of any fiction writer is the art of the storyteller. As Kathryn Cramer (writer, anthologist and daughter of physicist and fiction writer John G. Cramer) writes in *The Ascent of Wonder:*

> The majority of science fiction stories are not plausible extrapolations upon our current situation, using available information; rather they are Escheresque impossible objects which use the principles of science in much the same way that Escher used rules of geometric symmetry—the rules give form to the impossible imaginative content.

Antediluvian Nanotech

MANY OF DREXLER'S imaginings have antecedents in science fiction and feed into old, potent themes of the genre. Science fiction has long been fascinated with machines in general, such as in the stories of Jules Verne. The absolute control of matter promised by nanomachines is a variant of the dream that *Homo sapiens* can achieve complete mastery over nature and has utter freedom to shape its own destiny. The dark vision of nanobots running amok is a new wrinkle on the old golem / Frankenstein myth, the dangers of meddling with godlike powers or bringing too much hubris to science. The bright vision of the world, and indeed the nature of humanity, being transformed into something transcendent and new is another science-fiction standby.

Nanotech burst into the collective consciousness of technology aficionados at a good time to interface neatly with the mid-1980s wave of cyberpunk stories, in which characters experience the completely programmable virtual realities of cyberspace. With full-scale



INVISIBLE NANOTECH pervades this scene from Greg Bear's novel */ [Slant]*. Mary Choy has transformed her features back to her original Asian look. The pyramidal building is reinforced with flexfuller and contains cryopreserved bodies of the rich and privileged. The swanjet has no conventional ailerons; instead nanodevices on each wing surface control the lift forces by forming thousands of tiny vanes or humps.

JIM BURNS

The prospect of absolute control of matter using nanomachines

is a variant of the dream that *Homo sapiens* can achieve complete mastery over nature.

molecular nanotech it is not just *virtual* reality that is programmable. The intelligent agents and viruses of cyberspace become free to roam about in the air that we breathe and within our bodies—a curious inversion of people loading their consciousnesses into machines.

Elements suggestive of now common nanotech themes appeared in science fiction well before the advent of the term "nanotech." The concept of microscopic surgery appeared in the 1966 movie *Fantastic Voyage*, novelized by Isaac Asimov. Of course, instead of using molecular machinery built of real atoms, a large scientific-looking contraption magically reduces people and machinery to microscopic scale by shrinking their very atoms, in violation of numerous physical principles. Interestingly, the 2001 novel *Fantastic Voyage: Microcosm*, by Kevin J. Anderson, uses the miniaturization technology of *Fantastic Voyage* to explore the dormant body of an alien from a shot-down UFO. Lo and behold, the Lilliputian explorers find themselves confronting alien nanotechnology.

MINIATURIZED PEOPLE in the 1966 movie *Fantastic Voyage* swim like nanobot surgeons in a patient's bloodstream.

Progenitors of nanotech fiction extend back even further. In the classic 1941 story by Theodore Sturgeon, "Microcosmic God," a scientist creates a society of miniature creatures ("Neoterics") that evolve at a rapid pace and produce technological wonders. This theme is picked up in a modern way in *Blood Music,* by Greg Bear, first published as a novelette in the magazine *Analog* in 1983 and expanded into a novel in 1985. Despite predating the popularization of nanotech, *Blood Music* is frequently cited as a seminal nanotech story and included in nanotech anthologies. In the story, a researcher creates intelligent cells, "noocytes," that escape from confinement and spread like an epidemic through humanity, destroying it but also seemingly bringing about a transcendental change to a new form of existence. It's the end of the world as we know it, but we'll all feel fine afterward.

Television has also picked up on Sturgeon's concept. In the 1996 Halloween episode of *The Simpsons,* Lisa accidentally creates a microscopic society (ingredients: a tooth, Coca-Cola and an electric shock delivered by Bart) that rapidly advances from the Stone Age through the Renaissance and then far beyond our own technology. The theme is combined more soberly with the modern concept of nanotech in the episode "Evolution" of *Star Trek: The Next Generation,* which aired in 1989, just three years after Drexler's *Engines of Creation* hit the bookstores. Boy wonder Wesley Crusher accidentally releases some "nanites," tiny robots designed to work in living cells, which proceed to evolve into a highly intelligent society that invisibly infests the systems of the starship *Enterprise* and starts wreaking havoc. Fortunately, in classic *Next Generation* style, at the last minute, contact is made with the evolved nanites, and a mutually acceptable peaceful outcome is negotiated: they are placed on a convenient planet where they'll have more room to live and grow. If only every conflict, plague or technological disaster in the real world were solvable with such ease and rationality.

The central feature of molecular nanotechnology, precise manipulation of atoms, crops up in a classic science fantasy of 1965:

> The stuff was dancing particles within her[. . . .] She began recognizing familiar structures, atomic linkages: a carbon atom here, helical wavering . . . a glucose molecule. An entire chain of molecules confronted her and she recognized a protein . . . a methyl-protein configuration.
>
> [. . .] she moved into it, shifted an oxygen mote, allowed another carbon mote to link, reattached a linkage of oxygen . . . hydrogen.

Anyone can obtain free food from public nanotech matter compilers,

but they're not up to the cordon bleu standards of *Star Trek*'s replicators.

> The change spread … faster and faster as the catalysed reaction opened its surface of contact.

The novel? *Dune,* by Frank Herbert, in which computers are banned throughout the empire and spaceship pilots navigate through hyperspace by means of drug-induced precognition. While Paul Atreides is on his way to becoming the messianic ruler of the known universe, his mother, Lady Jessica, takes part in a ritual involving "the water of life"—a deadly poison related to the "melange," or "spice," that feeds supernatural powers of intuition and prescience. To survive the ritual, Jessica's consciousness dives down into inner space and slows time to a crawl to analyze the chemical composition of the poison on her tongue and, using a "psychokinesthetic extension of herself," transform it into a catalyst that rapidly detoxifies all the rest of the poison, turning it into a potent but not deadly narcotic.

Yet except for the use of psychokinesis in place of a technological framework, the entire process sounds like a nanotech engineer working at a virtual-reality station to design a molecule. Or like a scenario from *Unbounding the Future: The Nanotechnology Revolution,* by Drexler, Chris Peterson and Gayle Pergamit, in which a tourist experiences a museum exhibit that simulates the molecular-scale world, complete with scaling (slowing down) of time. In essence, nanotechnology offers to make *Dune*'s fantasy of complete human control over the self and the rest of the universe into a reality but in a mass-produced industrial fashion rather than through intensive individual training and drug-enhanced psychic powers.

Nanotech is also prefigured in *Dune* through the Ixians, traders from a rare high-tech corner of the universe who are "supreme in machine culture. Noted for miniaturization." Indeed, devices from *Dune* such as the "hunter-seeker," a tiny poison-tipped flying needle, would be completely at home in nanotech stories.

Hot Stuff

NANOTECHNOLOGY'S use in science fiction takes many forms, classifiable by a number of measures. The role of nanotech ranges from a central part of the plot to a relatively incidental part of the fictional world. The nanotech may be developed by humans, or it may be a gift from aliens, or it may *be* the aliens. The technology may work according to well-defined rules, or it may be arbitrary magic, scantily clad in trappings of science. The rules may be expressly mentioned in the text (perhaps even laboriously described), or the work may rely on the reader to tap into a science-fictional consensus reality, acquired from reading earlier stories, of what generic Acme nanotech can and can't do.

Both of the latter alternatives relate to two approaches to presenting technology in science fiction. At one extreme, the text practically contains a research paper on the author's hyperspace theory and blueprints of the first starship (rather like some of the letters *Scientific American* receives). At the other, famously pioneered by Robert A. Heinlein, the technology is dropped in without explanation: "The door dilated." In three words we know we are in a future with strange new technologies, and we are really *there* because in the future commonplace devices such as dilating doors need no more explanation than cellular phones do today.

When you read a large number of nanotechnology stories in a short space of time, some amusing recurring themes appear. On the one hand, nanotechnology often becomes a means to accomplish anything within the realm of the imagination, while conveniently ignoring the constraints of physical laws. Curiously, on the other hand, these stories reveal some of the actual technical challenges that molecular nanotechnologists might confront if they ever were to execute their designs for real-world nanobots. For example, it seems that most everyone writing nanotech fiction is aware that highly active nanocritters will generate heat, a problem of some concern when said nanocritters are functioning inside your body.

In one of the early stories, the 1989 novella *Nanoware Time,* by Ian Watson, the nanoware has been brought to humankind by seemingly benevolent aliens that look like giant golden centipedes. When the nanoware is injected into a person, it takes root in the subject's brain, supplying him or her with the power to ... (can you guess?) harness "demons" from a parallel dimension. These demons have no will of their own but possess extraordinary powers that the nanowired person can then use, for instance, to shield himself from the vacuum of space, propel a starship across the galaxy, fire bolts of energy for good or ill, and so on. The nanoware is really just a technological cloak for supernatural magic. In another era the alien device would have been some other mind-enhancing black box or injectable drug. Yet because this is nanoware—nanobots that rewire the hardware and software of your "wetware" (your brain)—one does have to worry about the heat generated by its functioning. A few of the early human volunteers fried their brains before the right parameters for humans were worked out:

> Heat was a byproduct of all the rapid molecular activity in the skull while the busy little nanomachines built the nanoware. Thus some brains got cooked.
>
> Vance was among the survivors. [His] brain damage was repaired by other nanos; sort of repaired. He'd been rehabilitated, retrained as a waste recycler.

STORY EXCERPT: DYSTOPIAN NANO-FUTURE

No Love in All of Dwingeloo, by Tony Daniel
(from *Asimov's Science Fiction,* November 1995)

[We lived] on Kokopelli Station. There were more people living up-cable, thousands and thousands more. Still, this was just a trickle of refugees compared with the billions who had died below. Earth was horribly worse, and it was then that I realized I was seeing the *future.*

The squabbles and wars of the present had played themselves out and we'd done it, we'd ruined the planet. The seas were biohazard cauldrons, seething with an ecology of war viruses. The land was haunted by nanoplasms, the primal form that life had taken, been reduced to. Sea and land were at war—over nothing, any longer—just a meaningless perpetual struggle between viral life and nano algorithms caught in a perpetual loop. Those who crafted the weaponry were dead.

A few million humans survived on the coasts, in the land between the warring elements. They were temporarily immune to the nano, but none could say for how long, since the nano evolved, its sole purpose finding ways to beat back the living, zombie sea—and, incidentally, to remake whatever people remained into a substance that could not wield a gun and could not think to use one.

Yet I found myself completely, unshakably content. Kokopelli was safe. We had severed all but one cablelift, and created defenses that kept the muck below at bay.

In one of the many plot threads in the 1997 novel */ [Slant],* by Greg Bear, a sequel to the landmark *Queen of Angels* (1990), four people are found dead in an illicit body-modification clinic: they were cooked, literally, when the body-modifying nanotech ran amok. The cause is promptly uncovered by investigators when they examine the jars of pastelike "nano" on the shelf:

> Mary picks up a bottle, reverses it to read the label.[…] The label confirms her suspicions.[…]
>
> "This isn't medical grade," Mary says. "It's for gardens.[…] Any real expert could reprogram it. Apparently they didn't have a real expert."

Presumably the victims were broiled because a bug in the badly reprogrammed garden-grade nano made it run wild, generating far too much heat in the process.

Later in the book a group of criminals who have infiltrated a huge tetrahedral building make use of some illicitly obtained MGN—military-grade nano. Sprayed from a canister like fire-extinguisher foam, the nano deconstructs objects present in the building's garage and rebuilds the atoms into intelligent robotic weaponry. During this process the garage heats up like an oven, but not too hot, because "at about four hundred degrees, nano cooks itself."

A spectacular case of spontaneous human nanocombustion occurs in one of the most surreal sections of Neal Stephenson's tour de force *The Diamond Age.* A secluded cult known as the Drummers is infected with millions of nanoprocessors. When two processors meet in someone's bloodstream, they compare notes, perform a computation and then go on their way: a kind of distributed, Internet-like supercomputer. The computation proceeds mostly at a steady pace, but occasionally it advances in a spurt of activity when myriad parallel threads of the computation are brought together for synthesis by an orgy (exchange of bodily fluids is the key means of transferring these processors and their data between people). The orgy culminates when the nanoprocessors are loaded into one unlucky woman who is promptly incinerated by the heat of the nano-orgy that ensues in her bloodstream.

To access the computation's result, the other Drummers mix her ashes into a soup—highly reminiscent of the Martian process of "grokking" the dead (in essence, ritual cannibalism to honor and fully appreciate the deceased) in Robert A. Heinlein's 1961 novel *Stranger in a Strange Land,* but with the patina of a scientific rationale.

Stephenson's *The Diamond Age* and Bear's *Queen of Angels* are comprehensive depictions of societies completely changed by nanotechnology. In *Queen of Angels,* a large proportion of the populace has been "therapied," in which injected nanotech devices infiltrate a person's brain to correct psychological imbalances and weaknesses. Many people undergo extensive nano-enabled body modification, ranging from practical enhancements for their occupation to beautification and the addition of exotic features. A complex tension runs through the society because of prejudices and attitudes about "transforms," "high naturals," and "therapied" and "simple untherapied" individuals.

In Bear's world, nano comes in jars like paste. In Stephenson's *The Diamond Age,* the key to nano is "the feed," a type of nanopipeline that runs into every household, supplying atoms as needed by matter compilers, which are as common as microwave ovens are today. Anyone can obtain free food from public matter compilers, but they're not up to the cordon bleu standards of *Star Trek*'s replicators; rice they can do, but green vegetables come out as a paste. Airborne nanotech is ubiquitous, ranging from almond-size surveillance monitors to microscopic attack and defense craft engaged in a constant struggle, like an immune system battling invaders. On bad days in the poor section of town, this ongoing contest looks like a fog shot through with firefly sparkles of laser light. A sootlike coating composed of casualties from this conflict settles on everything and everyone. Wealthier enclaves, such as that of the Vickys (neo-Victorians), are protected from such troubles by a deep defensive perimeter of airborne nanobots.

Nanofiction is not without humor.

Nano crops up incessantly,
although most works that mention it do not have it on center stage.

The Diamond Age, particularly early in the book, is told with abundant wit and drollery. The story opens with the exploits of a spectacularly stupid lowlife named Bud. A parody of the Walkman generation, Bud gets around on in-line skates capable of a top speed of more than 100 kilometers an hour, and his music system is "a phased acoustical array splayed across both eardrums like the seeds on a strawberry." He's sometimes a little "hinky" on the skates: implanted nanosites incessantly twitch his muscle fibers to maximize their bulk. Together with a testosterone pump in his forearm, "it was like working out at a gym night and day, except you didn't have to actually do anything and you never got sweaty."

And isn't that, in the end, what much of nanotech is about? A quasi-scientific way to get what you want, effortlessly and at minimal cost.

There can be no doubt that "nano" is a permanent addition to the tools of the sci-fi trade. Consider the 2007 edition of the anthology series *Year's Best SF 12*, edited by David G. Hartwell and Kathryn Cramer. "Nano" is right there in large type, the first word on page 1, because Nancy Kress's "Nano Comes to Clifford Falls" leads off the volume. The piece is a relatively uncomplicated look at how ordinary people, represented by the residents of a small town "far out on the plains," respond to the cornucopia that is four nanomachines that have been given to the town. Three little ones are stationed inside the town hall, and "the Big Gray" stands out front. The mayor is in charge of the machines, and the town folk get to place orders for food, clothing and anything else that appears in the nanocatalogue. It is as if everyone in town—and the rest of the country—has won the lottery. But what happens when nobody has to work to make ends meet?

The other story in the volume to feature nano right in the title is Rudy Rucker's "Chu and the Nants." Nants are patented "bio-mimetic self-reproducing nanomachines" that work together in swarms. Micron-size (upward of 1,000 nanometers), each one somehow has a gigabyte of memory and a processor that runs at about a billion updates per second. The U.S. sends an egg case of nants to Mars, where they proliferate, consuming the entire planet over the course of two years and forming themselves into a Dyson sphere—a shell enclosing the entire inner solar system. Why does the U.S. do this? Because otherwise China would do it first.

The nant-sphere forms a computer with 10^{48} bytes of memory, and president Dick Dibbs and his advisers expect it to be a "strategic military planning tool":

> "That's why they could short-circuit all the environmental review processes." Ond gave a wry chuckle and shook his head. "But it's not going to work out like they expect. A transcendently intelligent nant-sphere is supposed to obey an imbecile like Dick Dibbs? Please."

For the short time that it remains under Earthly control, the sphere also serves as a screen on which advertisements are displayed across the entire sky: "plugs for automobiles, fast food chains and credit cards," along with promos of Dibbs (who, having undergone a "life-extending DNA-modification that made him legally a different person," is now eligible for a third and fourth term in office). A similar swarm of nants is to be released on Earth to convert everyone and everything into a virtual reality under Dibbs's control. Rucker's forthcoming novel, *Postsingular*, will pick up where "Chu and the Nants" leaves off.

These two short stories, singled out in a fairly arbitrary fashion, nonetheless provide a snapshot of today's science-fictional nanotechnology. Nano crops up incessantly, although most works that mention it do not have it on center stage. Themes range from those driven by examination of social issues, as in the Kress piece, to those in which technological innovations run rampant, as in Rucker's. Of course, any one-dimensional scale does no justice to the degree of variety in the genre; science fiction spans more dimensions than string theory.

The ongoing state of *real* nanotechnology research, for the most part, has little to do with how the details used in fiction evolve over time. The science of science fiction lives in a parallel world to our reality, and what is known to be true or possible in our reality is mirrored only fitfully. And we wouldn't want it any other way. SA

Graham P. Collins is a staff editor and writer. He is also an occasional science-fiction writer and the Webmaster of the Science Fiction and Fantasy Writers of America.

MORE TO EXPLORE

The Encyclopedia of Science Fiction. Second edition. Edited by John Clute and Peter Nicholls. Palgrave, 1993. Also available from Grolier on CD-ROM as *The Multimedia Encyclopedia of Science Fiction.*

The Ascent of Wonder: The Evolution of Hard Science Fiction. Edited by David G. Hartwell and Kathryn Cramer. Tor Books, 1997. The introduction is available online at **ebbs.english.vt.edu/exper/kcramer/aow.html**

Analog Science Fiction and Fact: **www.analogsf.com**

Asimov's Science Fiction: **www.asimovs.com**

Locus magazine's online index of science fiction: **www.locusmag.com/index/**

Magazine of Fantasy & Science Fiction: **www.sfsite.com/fsf/**

Science Fiction and Fantasy Writers of America: **www.sfwa.org**