SPICY DRUGS: CURES FROM CURRY? • SILICON LASER CHIPS

SCIENTIFIC AMERICAN

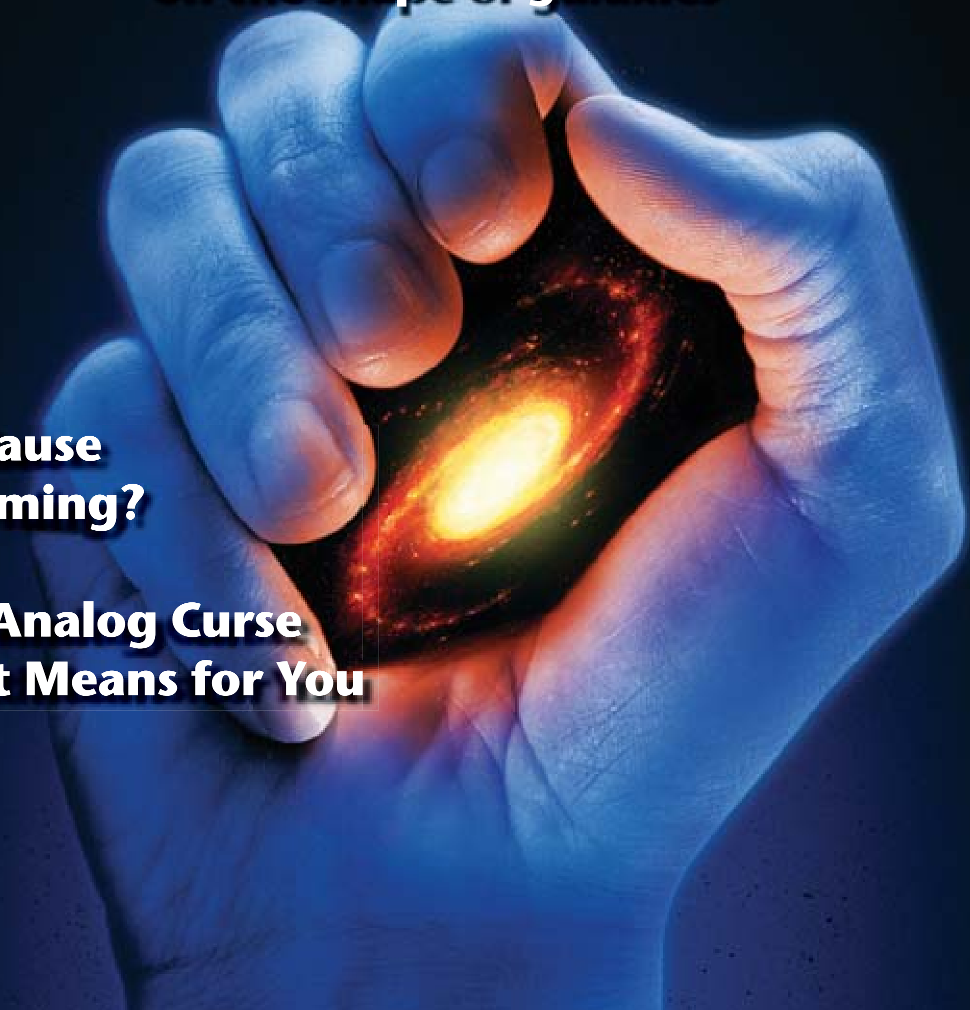
Prehistoric CSI:

Asteroids Did Not Kill These Dinosaurs

FEBRUARY 2007
WWW.SCIAM.COM

THE COSMIC GRIP OF DARK ENERGY

This mysterious force has a stranglehold on the shape of galaxies



Can Plants Cause Global Warming?

Digital TV's Analog Curse and What It Means for You

february 2007

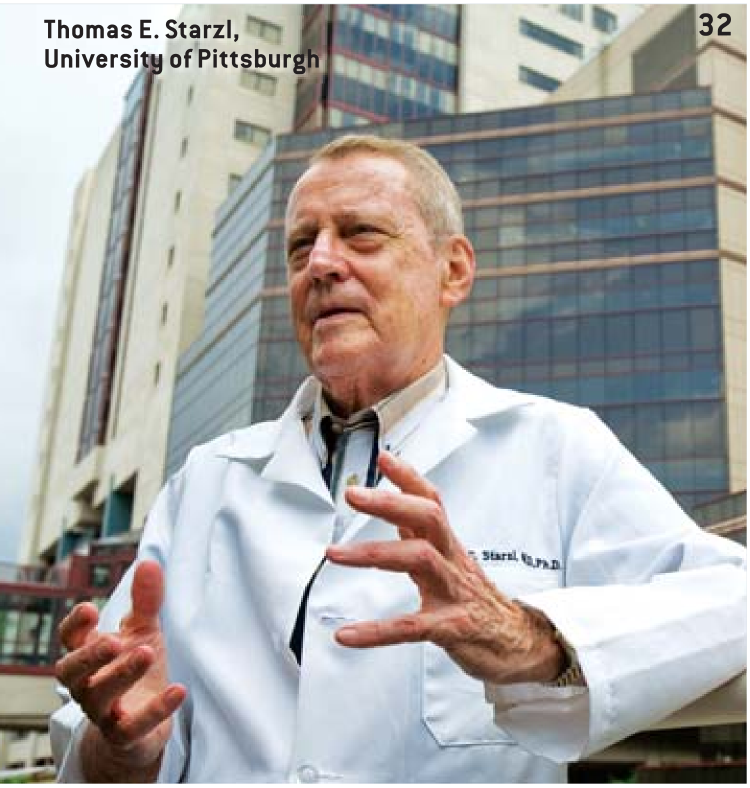
contents

SCIENTIFIC AMERICAN Volume 296 Number 2

features

34 The invisible mover of galaxies

departments



Thomas E. Starzl, University of Pittsburgh

columns

Cover image by Kenn Brown; photograph at left by Theo Anderson.

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 2007 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording for public or private use, or by any information storage or retrieval system, without the prior written permission of the publisher. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40012504. Canadian BN No. 127387652RT; QST No. Q1015332537. Publication Mail Agreement #40012504. Return undeliverable mail to Scientific American, P.O. Box 819, Stn Main, Markham, ON L3P 8A2. Subscription rates: one year $34.97, Canada $49 USD, International $55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 248-7684. Send e-mail to sacust@sciam.com** Printed in U.S.A.

Truth Time in Washington

For years, scientists worried that Republican politicians ignored science and were even downright antagonistic to it. Now that the Democrats have regained some power, we will see if they do any better. The first tests may not come on the most prominent issues, such as climate change and embryonic stem cells, but on other highly important yet less glamorous matters.

Energy. The last attempt at a comprehensive energy policy was the notorious "No Lobbyist Left Behind" act, an expensive wish list of subsidies and pork projects rather than a focused effort to clean up and secure the nation's energy supply. One obvious improvement would be to raise car mileage standards. Apart from a recent, tiny nudge, the corporate average fuel economy (CAFE) standards have been stuck since 1986. The U.S. now has the weakest automobile efficiency regulations in the industrial world; even China does better. Both the planet and U.S. technological competitiveness suffer.

WILL SCIENCE benefit from the shift in political power?

Public health. Several medical policy questions remain unresolved to the detriment of public health. The Vioxx debacle has heightened calls for better tracking of newly approved drugs for side effects. The Grassley-Dodd bill would establish a new Food and Drug Administration division that could order safety studies and take corrective actions after a drug is approved. Biodefense funding also needs reform. The Bioshield program, which is supposed to provide incentives for innovative drug development, is badly flawed, and passage of the Biodefense and Pandemic Vaccine and Drug Development Act would help to fix it. The act must, however, be amended to incorporate a viable solution to vaccine liability concerns. Fear of lawsuits is holding up the development of new vaccines, yet blanket immunity for manufacturers would fail to protect the public and would leave taxpayers liable when a vaccine caused unintended harm. The National Childhood Vaccine Injury Act of 1986, which combines liability limits with a compensation fund, offers a middle road and should be extended to all vaccines.

Paying the bills. The outgoing Congress left some financial loose ends dangling that could hurt science in the coming year. It passed only two appropriations bills for fiscal year 2007, neither of which covered science agencies, such as the National Science Foundation. Consequently, many important research and development programs remain in limbo, including those for the International Polar Year that begins next month. In the longer term, science agencies sorely need more stable funding so that they can plan effectively.

Office of Technology Assessment (OTA). To serve the country well, Congress needs not just to address specific problems but to improve how it deals with scientific and technological issues in general. From 1972 until it was canceled in late 1995, the OTA carried out high-quality analyses without getting sucked into partisan politics. Representative Rush Holt of New Jersey, a former research physicist, has repeatedly introduced legislation to restore the agency and has gained support from both sides of the aisle. Now is the time to pass that bill, instead of letting it die in committee once again.

THE EDITORS editors@sciam.com

CORBIS

How to Contact Us

EDITORIAL

For Letters to the Editors:
Letters to the Editors
Scientific American
415 Madison Ave.
New York, NY 10017-1111
or
editors@sciam.com
Please include your name and mailing address, and cite the article and the issue in which it appeared. Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

For general inquiries:
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8200
fax: 212-755-1976
or
editors@sciam.com

SUBSCRIPTIONS

For new subscriptions, renewals, gifts, payments, and changes of address:
U.S. and Canada
800-333-1199
Outside North America
515-248-7684
or
www.sciam.com
or
Scientific American
Box 3187
Harlan, IA 51537

REPRINTS

To order reprints of articles:
Reprint Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8877
fax: 212-355-0408
reprints@sciam.com

PERMISSIONS

For permission to copy or reuse material from *SA*:
www.sciam.com/permissions
or
212-451-8546 for procedures
or
Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
Please allow three to six weeks for processing.

ADVERTISING

www.sciam.com has electronic contact information for sales representatives of Scientific American in all regions of the U.S. and in other countries.

New York
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8893
fax: 212-754-1138

Los Angeles
310-234-2699
fax: 310-234-2670

San Francisco
415-503-3630
fax: 415-437-2892

Detroit
Karen Teegarden & Associates
248-642-1773
fax: 248-642-6138

Midwest
Derr Media Group
847-615-1921
fax: 847-735-1457

Southeast and Southwest
Publicitas North America, Inc.
972-386-6186
fax: 972-233-9819

Direct Response
Special Aditions Advertising, LLC
914-461-3269
fax: 914-461-3433

Australia
IMR Pty Ltd.
Tel: +612-8850-2220
Fax: +612-8850-0454

Belgium
Publicitas Media S.A.
+32-(0)2-639-8420
fax: +32-(0)2-639-8430

Canada
Derr Media Group
847-615-1921
fax: 847-735-1457

France and Switzerland
PEM-PEMA
+33-1-46-37-2117
fax: +33-1-47-38-6329

Germany
Publicitas Germany GmbH
+49-211-862-092-0
fax: +49-211-862-092-21

Hong Kong
Hutton Media Limited
+852-2528-9135
fax: +852-2528-9281

India
Convergence Media
+91-22-2414-4808
fax: +91-22-2414-5594

Japan
Pacific Business, Inc.
+813-3661-6138
fax: +813-3661-6139

Korea
Biscom, Inc.
+822-739-7840
fax: +822-732-3662

Middle East
Peter Smith Media & Marketing
+44-140-484-1321
fax: +44-140-484-1320

The Netherlands
Insight Publicitas BV
+31-35-539-5111
fax: +31-35-531-0572

Scandinavia and Finland
M&M International Media AB
+46-8-24-5401
fax: +46-8-24-5402

U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-9922

On the Web

WWW.SCIAM.COM

UPDATED EVERY WEEKDAY

Visit **www.sciam.com/ontheweb**
to find these recent additions to the site:

SPECIAL REPORT

10 Promising Treatments for the World's Biggest Health Threats

Cutting-edge pharmaceuticals now being tested could revolutionize the fight against cancer, Alzheimer's, HIV, diabetes, nicotine addiction and other devastating ailments.

NEWS

Superefficient, Cost-Effective Solar Cell Breaks Conversion Records

A tiny solar cell doubles the efficiency of common photovoltaics' conversion of sunlight to electricity by capturing the energy from a broader spectrum of light.

FACT OR FICTION?

Archimedes Coined the Term "Eureka!" in the Bath

The famed mathematician made many important scientific contributions. Was this exclamation really one of them?

PUZZLE

Sci-Doku

Try your hand at Sci-Doku, a sudoku puzzle that uses letters instead of numbers, with an added twist: a science-related clue accompanies each puzzle, and the answer is spelled out in one row or column of the solution.
Only at www.sciam.com

ASK THE EXPERTS

How much of human height is genetic, and how much is a result of nutrition?
Molecular biologist **Chao-Qiang Lai** of Tufts University answers.

New Podcasts

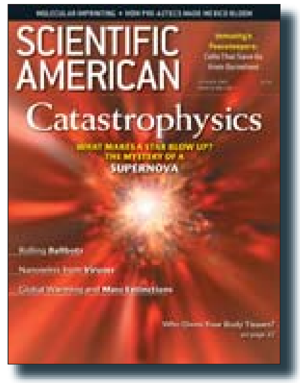
Every weekday: *60-Second Science*
Every Wednesday: *Science Talk*

Listen to the podcasts at
www.sciam.com/podcast

MATTHIAS KULKA *zefa/Corbis*

Letters

EDITORS@SCIAM.COM



THE OCTOBER 2006 issue covered the technological and theological, the massive and minuscule, the provocative and controversial. In "Impact from the Deep," Peter D. Ward took a fresh look at the earth's past with a theory that terrestrial heat and gases, not asteroids, most likely caused several mass extinctions. In "Viral Nanoelectronics," Philip E. Ross described how viruses coated with selected substances can be wrangled into self-assembling as liquid crystals, nanowires and electrodes.

Controversially, "Darwin on the Right," by Michael Shermer, and "Let There Be Light" [SA Perspectives], both about reconciling science and religion, drew the lion's share of reader reaction. Although some welcomed a peaceful coexistence, more readers' opinions were closer to that of Michael C. Brower of Andover, Mass.: "To minimize the conflict between science and religion misses the point. If scientific laws are correct, God must be remote and removed. Religious Americans believe in a God who is engaged in human affairs. Science challenges their core belief. It does not help the cause to gloss over that fact."

EARTH ATTACKS

Peter D. Ward's article, "Impact from the Deep," caused me to ask the following question. His "killer greenhouse" scenario (very high carbon dioxide and hydrogen sulfide) sounds like a stable equilibrium. So what mechanism or mechanisms would have operated to move the climate away from that state?

Dennis Jespersen
via e-mail

WARD REPLIES: Hydrogen sulfide very rapidly breaks down in our atmosphere, so it would have to be continuously generated to sustain the situation described in this mass extinction model. Remember, these events are rare and short-lived in the long span of geologic time. The earth cannot keep producing the huge volcanic phenomena known as flood basalts, which created the extraordinary heating that led to global warming, depleted oxygen in the oceans, and promoted the rise of deep-sea hydrogen sulfide sources to surface waters. As soon as flood basalt volcanism stops, the earth cools, ocean oxygenation increases and the chemocline descends—event over.

VIRAL INGENUITY

After reading "Viral Nanoelectronics," by Philip E. Ross, I wondered if Angela Belcher, her colleagues at the Massachusetts Institute of Technology or the folks at Cambrios have considered applying this technique to the creation of solar cells. Could you combine the solar cell in a "layered" configuration with the battery described in the article? If that is feasible, I would think numerous applications, ranging from cell phones to the military, would be able to take advantage of its low weight and "moldability" to fit designed space requirements.

Charles Barney
Burnsville, Minn.

BELCHER REPLIES: Indeed, we are using layering and templating for a variety of applications. Basically we change the DNA sequence that codes for the protein sequence that grows or arranges the materials for the specific device application. Solar cells are a major focus in our work at M.I.T. right now. We are trying to take advantage of the ability of biology to template different materials in a very small space and in close proximity to act as full-spectrum solar absorbers.

SCIENCE IN FAITH

As an evangelical Christian with a biology background, I appreciate and agree with most of Michael Shermer's column, "Darwin on the Right" [Skeptic], on why Christians should stop opposing evolution. He missed, however, what is in my experience the main reason so many Christians hold on so strongly to creationism. That is the belief that if we throw out the literal creation account, then we are opening the door to throwing out the very basis of Christianity, the physical and historical resurrection of Jesus Christ from the dead. As long as the two are linked, conservative Christians can never accept evolution.

Only by decoupling the two issues can Christians accept evolution. Fortunately, this separation has already happened once in Christian history, when the Protestants of the Reformation dropped the belief in the literal transformation of the Eucharist in the Mass. Once they realized that they could rationally take the Eucharist passages figuratively and still take the Resurrection literally, they followed the physical evidence and never looked back. Conservative Christians will not accept evolution until they make the same intellectual leap. How long that will take, only God knows.

Blake Adams
San Antonio, Tex.

FAITH IN SCIENCE

In "Let There Be Light" [SA Perspectives], you wrote that "scientific research … cannot begin to … offer a comprehensive moral philosophy." That is utterly absurd. Studying religions to understand why societies chose various aspects of a moral philosophy, along with working out a logical, comprehensive moral philosophy that is likely to help a society prosper, is a reasonable endeavor. I strongly suspect that this subject is well within the realm of sociologists, psychologists and political scientists.

Additionally, the implication that only the religiously observant have a moral philosophy and that we atheists who don't want intelligent design taught in school have no morals is repugnant.

Chuck Simmons
Redwood City, Calif.

You are too generous to religion. Science disproves fundamental tenets of every major religious system, from creation stories to miracles to visions of the afterlife, and after throwing these out there is nothing left beyond a nebulous "religious feeling" and some moral principles, which are largely common sense to begin with.

If this counts as "religion," then perhaps it can coexist with science, but those who profess faith both in science and in one of the established religions make a mockery of both.

Will Nelson
Tucson, Ariz.


HANS WOLF age Fotostock

SCIENCE HASN'T A PRAYER when it tries to take on the functions of religion—and vice versa.

COUNTING ON CARE

In "Contentious Calculation," by John Dudley Miller [News Scan], scientists argue over the correct number of cancer deaths attributed to Chernobyl. But one look at that heartbreaking, utterly tragic photograph of a cancer-stricken child leaning her exhausted head on her mother's arm as she awaits treatment in Kiev, Ukraine, prompts the question: Can we spend a little less time arguing how many will die and more time saving those we can? That child's face says it all.

Marie Jones
San Marcos, Calif.

MISTER T CELLS

If T-regs might be able to play a key role in curing or reducing the effects of autoimmune diseases within people, as Zoltan Fehervari and Shimon Sakaguchi suggest in "Peacekeepers of the Immune System," might it be possible that the treatment of immunodeficiency syndromes and inflammatory disorders with intravenous immunoglobulin (IVIg)—which is a blood product taken from 10,000 to 20,000 donors per batch—is somehow beneficial for the autoimmune diseases via transference of T-reg cells from the donors to the patient receiving the treatment?

Michael Trifone
Newington, Conn.

FEHERVARI AND SAKAGUCHI REPLY: High-dose intravenous immunoglobulin controls unwanted immune responses by a mechanism of action unrelated to T-regs. The fractionation of blood plasma and purification of immunoglobulin (Ig) completely removes any cellular components—including T-regs.

ERRATA In "Ballbots," by Ralph Hollis, in the box "The Problem of the Vertical," the phrase in the second sentence of the third paragraph, "which corresponds to the earth's orbital period at its surface on the equator," appeared in error. The sentence should read: "This pendulum would, in fact, have a period of about 84.4 minutes, the so-called Schuler period."

In "Uninformed Consent," by JR Minkel [News Scan], it was misstated that Washington University in St. Louis anonymized tissue samples collected by its former investigator, William Catalona. The university has not, though if asked by a donor, it would state that anonymization of samples might occur if a patient chose to withdraw from the research.

CLARIFICATION In answer to the question "How do fast breeder reactors differ from regular nuclear power plants?" [Ask the Experts], P. Andrew Karam stated that commercial reactors typically use uranium fuel that has been enriched with up to 8 percent uranium 235. Globally, uranium 235 enrichment ranges from 2 to 8 percent, with most reactors running in the middle of that range. U.S. commercial facilities are limited to a maximum of 5 percent.

50, 100 & 150 Years Ago

FROM SCIENTIFIC AMERICAN

Viral Pattern ▪ Breathing Machines ▪ Science and Guilt

FEBRUARY 1957

THE INFLUENZA VIRUS—"The virus particle has a relatively simple structure which we may hope to understand fairly well. The normal living cell is something of which our knowledge is both enormously extensive and utterly incomplete. The infected cell presents us with a far more complex problem. It is perhaps characteristic of the growing edge of biology that when a new phenomenon like virus multiplication comes to be studied, almost all the knowledge of cellular chemistry and function gained from other types of study turns out to be irrelevant. Any attempt to picture what is happening in the infected cell must necessarily therefore be provisional and oversimplified. A virus is not an individual organism in the ordinary sense of the term but something which could almost be called a stream of biological pattern. The pattern is carried from cell to cell by the relatively inert virus particles, but it takes on a new borrowed life from its host at each infection.—Sir Frank Macfarlane Burnet" [*Editors' note: Burnet won a Nobel Prize in Physiology or Medicine in 1960.*]

FEBRUARY 1907

FRESH AIR—"Smoke helmets, smoke jackets, and self-contained breathing apparatus generally are used in mines of all kinds, fire brigades, ammonia chambers of refrigerating factories, and other industrial concerns. The curious gear is intended to supply the user with factitious but perfectly respirable air, for about four hours at a stretch. Oxygen can be supplied from a steel cylinder. The question of renewing the oxygen is often a serious one, say in the remote mining districts of South America. Some shipping companies absolutely refuse to carry compressed oxygen in steel cylinders; but now a new substance, known as 'oxylithe,' has come along. The stuff is prepared in small cakes ready for immediate use, and on coming in contact with water it gives off chemically pure oxygen."

For a variety of these apparatuses from 1907, see www.sciam.com/ontheweb

FEBRUARY 1857

MURDER—"Dr. H. Burdell was found stabbed in his own room in this city on the morning of the 29th of last month. There was bad feeling existing between him and his housekeeper, but science has removed some of what were at first strong indications of guilt. A dagger was found in her drawer faintly stained with blood; these stains are proved by chemical analysis to be rust. A very palpable bloody stain on a blue silk dress proved to be sugar or fruit preserves. A knife from the place of business of a suspected boarder, and a newspaper found in his room, showed stains which responded to the chemical tests for blood, and under the microscope showed the blood disks or red globules to be arterial."

THE PRAIRIE FARMER sowing seeds, 1857

MAELSTROM—"I have been informed by a European acquaintance that the Maelstrom, that great whirlpool on the coast of Norway, laid down in all geographies, and of which we have heard such wonderful stories, has no existence. He told me that a nautical and scientific commission, appointed by the King of Denmark, was sent to approach as near as possible to the edge of the whirlpool, sail around it, measure its circumference, observe its action, and make a report. They went out, and sailed all around where the Maelstrom was said to be, but the sea was as smooth as any other part of the German ocean. I had been instructed to believe that the Maelstrom was a fixed fact in the ocean, and that ships, and even huge whales, were sometimes dragged within its terrible liquid coils, and buried forever."

ILLINOIS FARMING—"The accompanying figure illustrates the Seed Sower of R. Hurd, of Moline, Ill. It is a perspective view of the machine at work on the broad prairie—it is drawn by an original, grave, and majestic, but, withal, a spanking team. As the machine is pulled forward, the grain or seed is deposited."

news SCAN

STEM CELLS

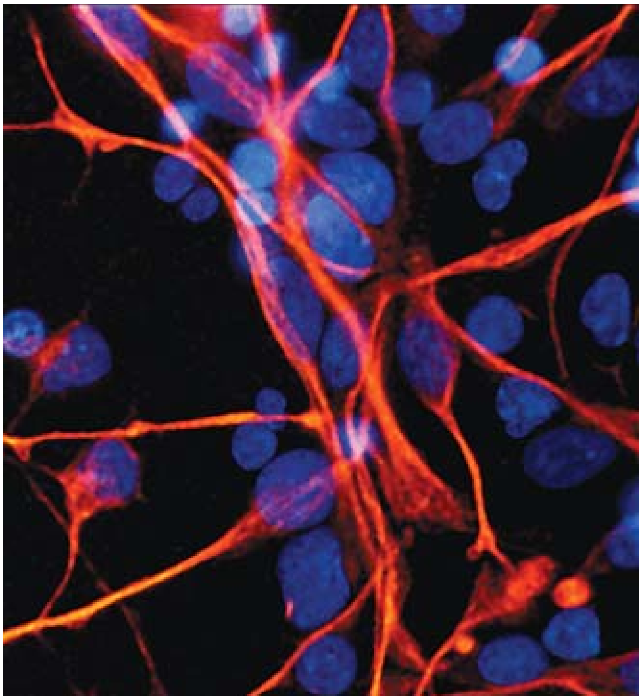
A Stroke for Stem Cells

THE BRAIN BECOMES A TARGET IN STEM CELL CLINICAL TRIALS **BY CHARLES Q. CHOI**

The first stem cell therapy targeting a major brain disorder, chronic stroke, could begin clinical trials this year if the U.S. Food and Drug Administration approves the request filed in December by stem cell firm ReNeuron in Guildford, England. This latest treatment suggests stem cell therapies are growing not only in number but in ambition.

Chronic stroke, in which patients suffer from permanent infirmity, is the leading cause of adult disability in the developed world. It afflicts 25 million people worldwide, and the number of new cases is rising 7 percent annually, mostly because of an aging population. "Right now there's next to nothing for treating chronic stroke, and what there is addresses the symptoms rather than the cause," says neurologist Justin Zivin of the University of California, San Diego.

Stem cells have the potential to regenerate body parts. In prior stroke studies on animals, stem cells injected into the brain or bloodstream migrated to sites of damage, apparently drawn by signals from damaged cells. This migration may happen because the repair pathways initiated by the damaged cells are similar to pathways triggered during embryon-



NEURAL STEM CELLS in culture (*cell nuclei in blue*) have differentiated and generated stringy protein called beta3-tubulin (*red*). Therapies with such cells are heading toward clinical trials.

RENEURON

BATTLE AGAINST BATTEN DISEASE

Several companies have initiated clinical trials to test stem cell therapies for repairing bone, easing coronary problems and preventing transplantation complications, among other uses. For therapies for the central nervous system, only one firm had a clinical trial under way in 2006: StemCells in Palo Alto, Calif. The company's phase I trial calls for injecting stem cells into six children diagnosed with a rare fatal neurodegenerative condition called Batten disease. Parents of the first treated child reported last December that their child showed subtle improvements in speech and no seizures, although a full evaluation of the treatment may take months or years.

ic development, where stem cells are key, explains ReNeuron co-founder and chief scientific officer John Sinden.

A major concern about stem cells centers on how unstable they can become when grown in the lab. ReNeuron can generate large numbers of stable cell lines by engineering cells with a modified version of the gene *c-myc*. This gene promotes cell division while activating genes that prevent chromosomal abnormalities. The scientists can switch *c-myc* on or off by introducing or withholding a synthetic compound.

ReNeuron developed cells for brain damage by splicing their modified *c-myc* into human fetal brain tissue obtained from a U.S. cell bank. They tested 120 neural stem cell lines in the lab for stability and robustness and in animals for the capacity to engraft with minimal immune rejection. Two lines showed potential: ReN001, which ReNeuron is aiming at stroke, and ReN005, which is under research for Huntington's disease.

In studies with rats that experienced stroke, ReN001 significantly improved sensory and motor function. The stem cells probably did not replace the massive number of cells lost during stroke, Sinden clarifies. Rather the cells most likely pumped out chemicals that activated repair pathways, resulting in new blood vessels and brain cells.

If their phase I clinical trial to test the safety and preliminary efficacy of this therapy gains approval, University of Pittsburgh researchers will test the therapy on 10 patients who suffer from chronic ischemic stroke—the most common form, in which clots block blood flow. Ten million to 20 million cells will be implanted directly in the brain through a small hole in the skull, and patients will be monitored over 24 months. ReNeuron has partnered with BioReliance in Glasgow, Scotland, to scale up cell production; the company has roughly one million ReN001 doses currently on hand, Sinden estimates.

Past clinical trials of stem cell therapies for chronic stroke patients used cells derived from tumors in humans and brain tissue from fetal pigs. ReNeuron's fetal cells "are closer to the neurons in [healthy] people than others used before, so they might be more effective," Zivin says. "What ReNeuron has done to create this cell line is ambitious and well thought out," adds neurologist Sean Savitz of Harvard Medical School. Savitz notes, however, that *c-myc* is associated not only with stem cells and development but also with cancer. "This is definitely not to say that it will promote tumors," he says, but the researchers "will have to continue to convince the scientific community that the cells will not divide unchecked the way they do in tumors."

Charles Q. Choi, based in New York City, is a frequent contributor.

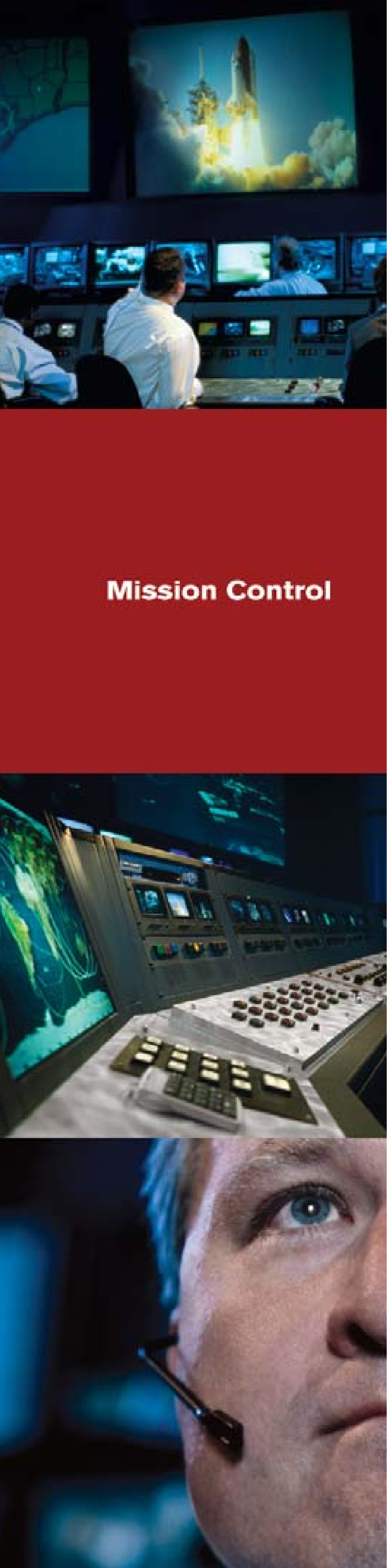
INTELLIGENCE

Unsettled Scores

HAS THE BLACK-WHITE IQ GAP NARROWED? BY MARINA KRAKOVSKY

This much is uncontested: for most of the 20th century, blacks worldwide have scored, on average, 15 points lower on most IQ tests than whites have.

What scientists cannot agree on is why. Most attribute the gap to differences in education, health and other environmental influences. Hereditarians, on the other hand, view the black-white difference as largely genetic in origin. They note, among other indirect evidence, that the disparity persists across time and around the world—



a permanence that is crucial to the debate over what explains group differences. "If black-white differences converged—if there wasn't this whopping big difference everywhere—there'd be no debate left," says J. Philippe Rushton, a psychologist at the University of Western Ontario and an outspoken hereditarian.

So it was big news when William T. Dickens and James R. Flynn published a paper in *Psychological Science* in October 2006 concluding that African-Americans have greatly reduced the racial IQ gap. Drawing on their analysis of four different IQ tests taken between 1972 and 2002, including the Stanford-Binet and the Armed Forces Qualification Test, the researchers say that blacks in the U.S. have gained four to seven IQ points on non-Hispanic whites, thus closing a quarter to half the gap.

To Dickens and Flynn, the implications are obvious. "The gap isn't fixed in the stars," says Flynn, a professor of political studies at the University of Otago in Dunedin, New Zealand, who is best known for his observations of across-the-board IQ gains over time—evidence that environmental factors strongly drive scores. Dickens, a senior fellow in economic studies at the Brookings Institution, adds that future progress toward equal opportunity should continue to reduce the disparity.

But the evidence does not justify such optimism, says Rushton, who calls himself a race realist. When Rushton peer-reviewed the study, he argued for its rejection on the grounds that the data were cherry-picked, because it ignored at least four IQ tests whose results showed either no black gains or even a slight decline. In the end, he published a forceful commentary on the study, coauthored with the eminent (and controversial) psychologist Arthur Jensen of the University of California, Berkeley.

Dickens and Flynn rebutted, stating that they had excluded samples that were not representative of the U.S. population. For example, in one sample one group had more education than the race it represented, and it was not clear which race that was. Dickens believes most IQ test publishers simply do not try to restandardize their tests using nationally representative samples, making it hard for researchers to draw definitive inferences.

Linda Gottfredson, an education professor at the University of Delaware, points to a more serious question in the findings. If black IQ is rising, and if IQ is the best predictor of academic achievement, one would expect a similar rise in standardized achievement tests. Yet, she notes, most studies show no such gain. For instance, Charles Murray, a scholar at the conservative think tank American Enterprise Institute and author of the controversial 1994 best-seller *The Bell Curve,* published a report on black achievement in the December 2006 issue of the journal *Intelligence;* he concludes that although blacks started closing the school achievement gap during the 20th century, progress stopped for people born since the mid-1970s.

Dickens explains these discrepancies in part through different ways of looking at the same data. "No one doubts that there is an achievement gap," he says. "The question is whether it has narrowed, and the answer is yes," although the gains have not been equal on all measures.

Gottfredson also points to a troublesome trend in the 30-year data that Dickens and Flynn used—namely, at each successive age, the black-white gap steadily increases. "All these data are interpreted as a slam dunk for environmental explanations, when they're at least as consistent with genetic ones," she says, adding that the effects of family environment tend to wash out with age. In the Dickens and Flynn model, on the other hand, an impoverished family environment actually creates greater disadvantages with age. For example, Dickens explains, "it's much less of a problem for a four- or five-year-old to have a high school dropout for a mother than for a 16-year-old," because poorly educated parents cannot help with advanced schoolwork.

If black progress has stalled, demographic changes may be to blame. Whereas high-IQ women of all races have lower birth rates, the birth rates of high-IQ black women in particular have "gone through the floor," says Murray—leaving a shortage of both the best genes and the best family environments for the next generation. He adds that even if it were possible to answer the genetic question once and for all, as he believes today's DNA testing can help to do, most scientists would shy away from such research for political reasons.

Because genes interact with the environment, some researchers think the question of genetic differences by race is pointless. "What does that mean?" asks Eric Turkheimer, a University of Virginia psychologist who has explored poverty's effect on realizing an individual's genetic potential. What Rushton and Jensen must be arguing, he says, "is that a gap will remain under any environmental influence. And how can you possibly know that? All you can show is it's been very frustrating to reduce."

Marina Krakovsky is based in the San Francisco Bay Area.

NEUROTECH

Chipping In

BRAIN CHIP FOR MEMORY REPAIR CLOSES IN ON LIVE TESTS **BY ANNA GRIFFITH**

Supplementing the human brain with computer power has been a staple of science fiction. But in fact, researchers have taken several steps in melding minds with machines, and this spring a team from the University of Southern California may replace damaged brain tissue in rats with a neural prosthesis.

For the past few years, researchers have demonstrated the ability to translate another creature's thoughts into action. In 2000 neurologist Miguel Nicolelis of Duke University wired a monkey with electrodes so that its thoughts could control a robotic arm. Brain-machine interfaces developed by Niels Birbaumer, a neuroscientist at the University of Tübingen in Germany, already help some paralyzed patients move a computer cursor with their brain waves to select letters for writing a message.

Theodore W. Berger and his U.S.C. colleagues have developed the first brain-machine interface to communicate back to the brain. Last January they used a silicon chip to mimic biological neurons in tissue slices of rat hippocampus, the hub for memory sorting and storage. The chip replaced a surgically removed section of the hippocampus and restored function by processing incoming neural signals into appropriate output with 90 percent accuracy.

The biomedical engineers had been on the verge of testing a chip in hippocampal slices for several years, but roadblocks slowed work. Existing electrode array technology would not function well in tissue slices, forcing the researchers to construct their own. Cutting the hippocampus slices just right to keep the neural pathway intact was also difficult.

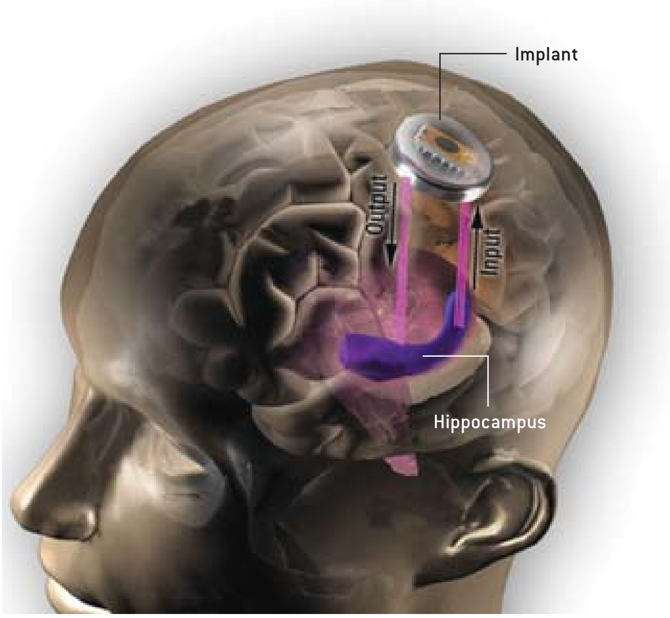
Because building the one-millimeter-square chip costs tens of thousands of dollars and takes several months, the planned spring test will actually rely on a model of that chip—specifically a larger, reprogrammable device linked to a computer called a field programmable gate array (FPGA). The FPGA will allow investigators to easily test and modify their new mathematical model of neural communication for living rats before committing it to a chip. Sam Deadwyler, a professor of physiology and pharmacology at Wake Forest University and a collaborator in the study, has demonstrated that stimulating the hippocampus of living rats with a certain pattern of activity can increase performance on a memory task, such as recalling which lever will dispense water. In a few months he will use the FPGA mathematical model to predict hippocampal activity. If the model is correct, the artificial implant should restore memory for such tasks in rats with drug-induced amnesia.

NEURON LANGUAGE

To engineer the brain chip, Theodore W. Berger of the University of Southern California first investigated the nuts and bolts of neuron communication. For more than two decades, he probed neurons in rat hippocampal tissue with electrical stimulation patterns, recorded their responses and created a comprehensive database of neuron behavior. His U.S.C. colleagues Vasilis Marmarelis and John Granacki developed mathematical equations to model neuron behavior and transformed the equations into hardware language.

If animal trials are successful, the team's first human implant will actually use a model of monkey brain activity because researchers cannot safely record activity from humans to develop a human-specific model. Though not ideal, a monkey implant might work well enough, the researchers speculate, because the programmable chip could adapt to human thought or the brain could learn to use the chip as it does with cochlear implants in associating garbled noises with words.

For more complicated animal models, U.S.C. physicist Armand Tanguay suggests a multichip module to facilitate the transition. Light beams would transmit signals between neuron units on multiple chip layers. Unlike wires, light beams pass directly through one another without interference, allowing for many more interconnections. The result: a web of light between silicon chips mimicking a dense neural network.



NSF-USC CENTER FOR BIOMIMETIC MICROELECTRONICS

HEADSTRONG: **Implants communicating with the hippocampus might someday restore or improve memory, if such devices succeed in rat tests.**

"Many challenges will be encountered as the researchers move from in vitro to in vivo studies in the rat," says Grace Peng, a program director at the National Institutes of Health's Division of Discovery Science and Technology. In fact, the team is not quite sure what to expect once it goes to live animals. Avoiding rejection by the immune system might mean anchoring cell adhesion molecules to the chip so that the surface of the implant looks like tissue, says U.S.C. chemist Mark Thompson. Neural plasticity, or the brain's ability to reorganize its connections, could also pose a problem by preventing the formation of stable connections between the neurons and chip. "In other application areas such as motor control or perception, plasticity and adaptability of the brain usually facilitate the effects of artificial interfaces," Peng notes optimistically.

One other possible concern, if such implants make it to human testing: Might bypassing damaged neurons in the hippocampus also bypass connections with other areas of the brain that filter what we remember? In other words, would the brain become unable to purge memories? If so, that would make the implant a truly unforgettable device.

Anna Griffith is based in Chico, Calif.

SPACE

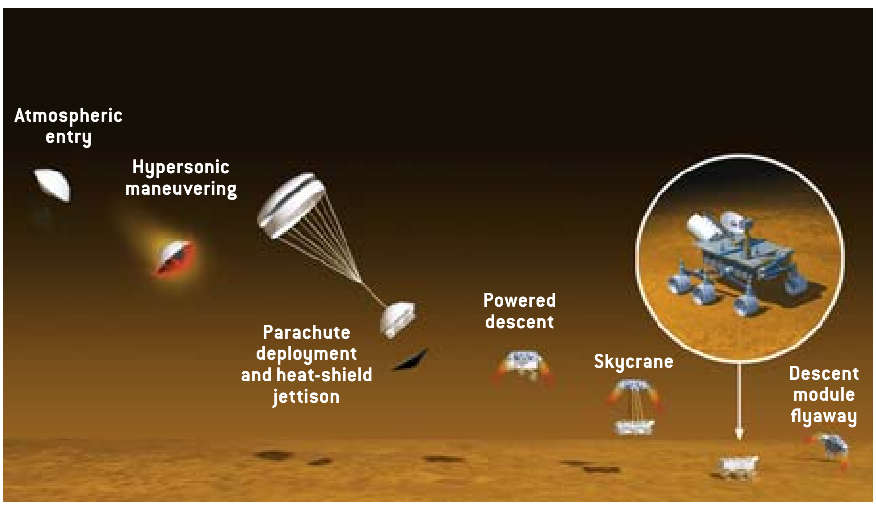
Martian Rope Trick

NASA REWRITES THE BOOK ON MARTIAN LANDINGS **BY GEORGE MUSSER**

Car commercials put their vehicles in the most improbable of places: sides of cliffs, tops of rock pillars, middles of deserts. But NASA is about to outdo them all. In 2010 it plans to land a Mini Cooper on Mars—or more precisely, a rover of about the same size and weight. And it plans to do so using a procedure never before seen in spaceflight: piloting the craft to its landing site using its entry capsule like a hypersonic flying wing, then hovering above the ground and lowering the buggy down on a long rope. "They have a very cool system," comments another rover builder, Jorge Vago of the European Space Agency.

Engineers were driven to their audacious—and not entirely uncontroversial—approach by the sheer size of the $1.5-billion Mars Science Laboratory (MSL). It outweighs by more than four times the go-cart-size Spirit and Opportunity rovers now prowling the Red Planet. Its enclosing capsule, measuring 4.5 meters across, will be bigger than even the Apollo command module was (at 3.9 meters). "We will be the biggest heat shield ever to enter an atmosphere," says Adam Steltzner of the Jet Propulsion Laboratory, who leads the team designing the entry, descent and landing sequence for the mission.

Whereas Spirit and Opportunity simply cut loose from their parachutes and bounced onto the surface, cushioned by giant air bags, MSL would require such heavy bags that they



NOVEL LANDING sequence for the Mars Science Laboratory mission includes high-speed aerial maneuvers and a "Skycrane" system to lower the rover with ropes.

DON FOLEY

would leave no room for extra payload. Touching down on landing pads would have its own issues. A legged lander is prone to tipping over, and the retrorockets would need to fire for so long that they would dig a sizable crater and stir up choking clouds of dust. The engine would have to shut down at just the right moment, which would take hair-trigger touchdown sensors. (NASA investigators suspect that the legged Mars Polar Lander crashed in 1999 because of a touchy sensor that cut the engine too soon.)

The MSL scheme seems to get around these problems. The rover's mother ship is essentially a big jet pack. As it approaches the surface, the jet pack fires up, hovers at a height of 20 meters, winches the rover down on 7.5-meter-long Kevlar cables and then descends at walking speed. If its load begins to swing, the mother ship simply moves laterally to stabilize it. No touchdown sensors are required—the mother ship registers that it has offloaded the rover when it finds it needs less rocket thrust to maintain its position. At that point, explosive charges cut the cables, and the mother ship powers off to crash-land a few hundred meters away. The name of this system, Skycrane, pays homage to the Sikorsky cargo helicopters that inspired it.

"There are some people who find this whole approach rather appalling," admits JPL's Rob Manning, chief engineer of NASA's Mars program and famed techno wizard of past rover missions. No Skycrane system has ever been tried in space, and engineers cannot conduct full-up tests on Earth because the gravity and air pressure differ so much. To be sure, they could not perform full landing tests for Spirit, Opportunity, Viking or other missions in the past, either. Then, as now, success hinged on engineers' modeling savvy.

It is not only the MSL mission that is at stake. Without a way to get large payloads to the Martian surface, even more ambitious plans, such as a sample return (let alone extraterrestrial car commercials), are stymied. "With each subsequent mission," Vago says, "you get a step closer to realizing those big missions that everyone wants."

BREATHING NEW LIFE INTO AIR BAGS

NASA's MSL is not the only weight-challenged rover in the works. In 2013 the European Space Agency plans to launch the $1-billion ExoMars rover, which carries a drilling platform to look for underground life. ExoMars is not much bigger than Spirit or Opportunity, but even so it pushes up against the limits of current air-bag technology. The project scientist, Jorge Vago, says the team is investigating a new air bag that deflates immediately on landing to absorb the shock of the impact. This approach requires fewer bags (only along the bottom of the lander). The trick will be to pop them fast enough that the rover does not bounce. In two of the five test drops done so far, some of the charges meant to blow open the bags did not go off.

INFOTECH

Dial "I" for Internet

BRITISH TELECOM BETS ALL ON INTERNET PROTOCOLS BY WENDY M. GROSSMAN

In the largest project of its kind, British Telecom will spend £10 billion (about $19 billion) between now and 2011 replacing its entire public switched telephone network—which serves 22 million subscribers—with a network based on Internet protocol (IP). The company began its first trial of what it calls 21CN (for 21st-century network) with a single exchange in Cardiff, Wales, on November 28 of last year. Rollout is scheduled to begin in 2008.

Genius or madness? The sheer expense and scale of this project mean the stakes are high. Ralph Cochrane, director of corporate development in charge of 21CN, says it has been called both. "I think we're slowly proving we're not mad," he says. Companies such as Telecom Italia, which has an IP-based long-distance backbone, have accomplished some of what British Telecom hopes to do. "But no one else is ripping out their entire network and putting in a shiny new one based on IP," Cochrane claims.

Internet protocol was designed to route data efficiently. Instead of opening up a dedicated circuit that cannot be used for anything else for the duration of a voice call—as traditional phone networks do—IP breaks up the voice stream that makes up a call into packets of data that traverse the network independently and are reassembled at their destination. The upshot is that many more calls can share the same amount of bandwidth. IP has paved the way for Skype, Vonage and other VoIP (voice over Internet protocol) providers.

British Telecom needs to make the conversion because, according to Cochrane, the proliferation of different networks is too complex and expensive to manage. In the past, each of the giant telcos commonly built a new network individually every time a novel technology (DSL, ISDN) came along or a particular sector, such as banking, wanted a special application. British Telecom's 16 national networks currently include voice, ISDN, DSL and leased lines.

British Telecom's 21CN will unify all these networks into a single network by using several hundred multiservice access nodes (MSANs)—chunks of hardware that will control all types of connections, replacing today's hodgepodge of thousands of pieces of different switching equipment. On 21CN the difference between DSL (whose available bandwidth depends on the number of users) and a leased line (whose bandwidth and service are guaranteed) will be software telling the MSAN which type of service to provide. The only thing that will remain untouched is that "first mile" of copper wiring from the exchange to individual homes. Coupled with associated changes in personnel and business practices, the new network should save British Telecom £1 billion ($1.9 billion) a year.

On the new network all telephones will be connected to MSANs. Picking up the receiver causes the nearest MSAN to send a message to one of 40 call servers, which will provide a dial tone and accept the dialed number. If the first server does not handle the number punched in, that same server will contact another one to set up the call, but the data packets that make up the voice call itself will be sent directly between MSANs.

The traditional flaw with VoIP is voice quality. Cochrane says, however, that 21CN is being built on a standard (known as multiprotocol label switching) that essentially routes packets according to labels incorporated into their headers. For a voice call or a TV program, the label might instruct the network to forward the packets without delay even if some must be dropped; for e-mail it might effectively say, "There's less hurry, but please don't lose any." This system, along with high-quality IP links that are not saturated with traffic, is intended to ensure that 21CN sounds just as good as the old network.

With only one exchange and a few hundred lines, British Telecom reported no problems during initial testing in Cardiff. A much bigger trial slated for April will network several locations—and will reveal whether the conversion is genius or madness.

Wendy M. Grossman writes about information technology from London.

PROPER PROTOCOL: British Telecom's public switched phone network will become Internet-based.

PAUL SEHEULT *Eye Ubiquitous/Corbis*

CALL FORWARDING

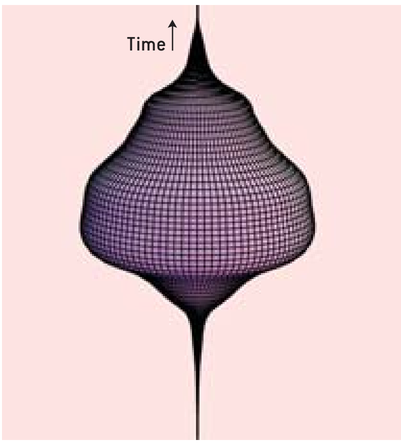
British Telecom's move to an Internet protocol system could open up opportunities for Internet service providers (ISPs)—namely, they could get into the phone business, notes Clive Feather, public policy analyst for Demon Internet (a U.K.-based ISP). "We already have a call server like BT will have," Feather says. So Demon could presumably sell voice service. With multiple phone services readily available, a customer might, for example, automatically switch among different services to obtain the lowest rates at given times or to keep personal calls separate from business activity.

PHYSICS

The Triangular Universe

INSTEAD OF STRING THEORY, FOUR-DIMENSIONAL TETRAHEDRONS BY MARK ALPERT

Imagine a landscape composed of microscopic triangular structures that constantly rearrange themselves into new patterns. Seen from afar, the landscape looks perfectly smooth, but up close it is a churning cauldron of strange geometries. This deceptively simple model is at the heart of a new theory called causal dynamical triangulation (CDT), which has emerged as a promising approach to solving the most vexing problem in physics—unifying the laws of gravity with those of quantum mechanics.



MODEL UNIVERSE generated by causal dynamical triangulation can reproduce features seen in standard cosmological theories, such as the expansion and contraction of space over time.

SOURCE: FROM "RECONSTRUCTING THE UNIVERSE," BY J. AMBJØRN, J. JURKIEWICZ AND R. LOLL, IN *PHYSICAL REVIEW D*, VOL. 72, ARTICLE NO. 064014, 2005, WITH PERMISSION FROM THE AMERICAN PHYSICAL SOCIETY

For more than 20 years, the leading contender in the quest for unification has been string theory, which posits that the fundamental particles and forces are actually minuscule strings of energy. But some scientists say this theory is misguided because it sets the strings against a fixed background; a better model, they argue, would generate not only particles and forces but also the spacetime they inhabit. In the 1980s and 1990s these researchers developed loop quantum gravity, which describes space as a network of tiny volumes only 10^{-33} centimeter across. Although this approach has achieved some notable successes, such as predicting the properties of black holes, it has yet to pass an essential test: showing that the jumble of volumes always comes together to form the familiar four-dimensional spacetime of our everyday world.

CDT is less than 10 years old, but it has already cleared this hurdle. Conceived primarily by three European theorists—Renate Loll of Utrecht University in the Netherlands, Jan Ambjørn of Copenhagen University and Jerzy Jurkiewicz of Jagiellonian University in Poland—CDT constructs spacetime geometries from simple triangular structures, in much the same way that Buckminster Fuller used triangular surfaces to create geodesic domes. The basic building block is the 4-simplex, the equivalent of a tetrahedron but in four dimensions. (Just as a tetrahedron has four triangular faces, a 4-simplex is bounded by five tetrahedrons.) Although each simplex is geometrically flat, they can be glued together in a variety of patterns to produce curved spacetimes. Because quantum theory stipulates that the structure of spacetime at very small scales must be constantly changing, the researchers determine the overall geometry by summing the probabilities of all the possible configurations of simplexes.

Previous attempts to triangulate the universe in this way ended in nonsensical results: crumpled spacetimes with an infinite number of dimensions or rolled-up geometries with just two. The key insight of CDT was to exclude the configurations that were not causal (that is, patterns that would allow an event to precede its cause). The elimination of these unrealistic tilings did the trick: in 2004 Loll, Ambjørn and Jurkiewicz used computer simulations to show that model universes constructed from hundreds of thousands of simplexes are four-dimensional. More recently, the researchers demonstrated that the large-scale shape of their universes is just like that predicted by the standard theory of cosmology.

The next big step for CDT is incorporating matter into the model to see if it can simulate the full equations of general relativity. According to Lee Smolin of the Perimeter Institute for Theoretical Physics in Waterloo, Ontario, the theory may eventually yield testable predictions, such as slight changes in the speed of high-energy photons caused by the model's nonclassical geometries at small scales. One of the pioneers of loop quantum gravity, Smolin says CDT has so far not received the attention it deserves from theoretical physicists, possibly because the approach relies so much on computer simulation. "It's not easy stuff to get into," he says. "It's hard to get at by pencil and paper."

SMALL-SCALE WEIRDNESS

One of the oddest features of the model universe generated by causal dynamical triangulation (CDT) is that its number of dimensions depends on how fine you slice it. At large scales the CDT spacetime is four-dimensional, but at the so-called Planck scale—where distances are on the order of 10^{-33} centimeter—the integration of nonclassical (and often highly curved) geometries produces a spacetime with only two dimensions. The CDT universe is analogous to a coarse sweater: a large spider can crawl across the two-dimensional surface of the cloth, but a tiny mite can only travel along the one-dimensional threads.

PSYCHOLOGY

Think of Money, Be Less Helpful

Money is an incentive to work hard, but it also promotes selfish behavior. Those conclusions may not be surprising, but psychologists at the University of Minnesota recently found that merely thinking of money makes people less likely to give help to others. Researchers subconsciously reminded some volunteers of money by showing them lucre-related words such as “salary” or by revealing a poster with currency on it. Other participants were primed with play money or neutral stimuli. All those involved in the study then performed different tasks that were unrelated to money but that assessed their behavior in social situations. When money is on the brain, people become disinclined to ask for help when faced with a difficult or even impossible puzzle. And individuals who think, even subconsciously, about money are less helpful than others, the researchers report in the November 17 *Science*. —*Ciara Curtin*

APPLIED PHYSICS

Cordless Charging

To recharge portable electronics, scientists hope to perfect a method for transmitting electrical energy wirelessly. The effect, which has not yet been demonstrated, would take advantage of induction, in which a varying magnetic field can induce electrical flow in a nearby conductor. To boost the range and power, Massachusetts Institute of Technology researchers propose introducing a short gap in a metal loop and attaching two small disks at each end. When electrified, such an object has a natural frequency that results from current flowing back and forth along the loop from one disk to the other. If a second loop has the same frequency, it should be able to receive energy from the other through the magnetic field. From a few meters away, the rate of energy transferred might reach tens of watts, or enough to power a laptop, according to simulations presented November 14 at a meeting of the American Institute of Physics. —*JR Minkel*

ANTHROPOLOGY

Deciphering Neandertal’s Faded Genes

A short, fossilized femur from a 38,000-year-old Neandertal could lead to the first full genome sequence for the closest relative of *Homo sapiens*. Working off samples from the same bone, two research groups used different methods to obtain partial DNA sequences and compared them with human DNA. The two genomes are at least 99.5 percent identical (in comparison, chimpanzees share 99 percent of the human genome). Depending on the particular genetic analysis, humans and Neandertals split from a common ancestor perhaps about 500,000 to 700,000 years ago, and the two stopped interbreeding about 370,000 years ago. The Neandertal sequence will help answer lingering questions, such as whether humans who migrated out of Africa mated with Neandertals in Europe 30,000 to 40,000 years ago. The current studies did not find any evidence of such mixing, but they did not rule out the possibility. The findings appear in the November 16 *Nature* and the November 17 *Science*. —*Nikhil Swaminathan*

NEANDERTAL DNA is at least 99.5 percent identical to that of modern humans.

CORBIS (*top*); AMERICAN MUSEUM OF NATURAL HISTORY (*bottom*); ILLUSTRATION BY MATT COLLINS

DATA POINTS: SOLAR SALE

A tiny, cost-effective solar cell nearly doubles the amount of electricity generated by sunlight as compared with standard photovoltaic cells. Developed by Spectrolab (a company that makes solar cells for spacecraft) with funding from the Department of Energy, the so-called multiple-junction solar cell consists of germanium and several layers of gallium arsenide. Each layer responds to different wavelengths and works with sunlight concentrated by lenses or mirrors, thereby maximizing the conversion efficiency.

Percent of conversion efficiency in:

Standard silicon cell: 12 to 22

Spectrolab cell: 40.7

Number of layers in Spectrolab cell: 20 to 30

Size of cell in square centimeters: 0.26685

Maximum concentrated solar input in watts per square centimeter: 24

Equivalent concentration, number of suns: 240

Electrical output in watts: 2.6

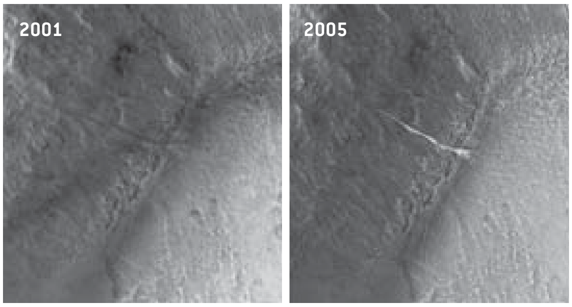
Estimated cost in cents per kilowatt-hour: 8 to 10

SOURCES: Department of Energy Announcement, December 5, 2006; Spectrolab

ASTRONOMY

Free Flow on Mars

Deposits formed in Martian gullies during the past seven years suggest that liquid water exists on Mars today. An image taken by the Mars Global Surveyor spacecraft in 2005 shows a downhill track on the wall of a crater that was not present in the previous image of the crater, taken less than four years earlier. In subsequent views of the deposit, the sun's light is coming in at different angles, but the light-colored material remains, suggesting it is not a trick of the light or the result of dry erosion. Similarly, images of another crater from February 2004 show the beginnings of a second deposit, which has grown in subsequent images, according to a report in the December 8 edition of *Science*. Finding additional examples might be tricky, given that NASA lost contact with the Mars Global Surveyor last November after nearly 10 years of operation. —*JR Minkel*



GULLY appears in a Martian crater in images taken four years apart by the now defunct Mars Global Surveyor.

GLOBAL WARMING

Methane Flatline

For 20 years, the levels of methane in the lowest layer of our atmosphere soared. But recent data indicate that the growth rate of the greenhouse gas, which is less abundant than carbon dioxide but traps 23 times more heat per molecule, began slowing in 1998. Researchers agree that the pause results from decreasing emissions of methane, but they have not pinpointed which sources have decreased and by how much. Leading hypotheses include the collapse of the Soviet Union, which resulted in a decline in energy use in the region; decreasing emissions from coal mining; and a decline in rice production. The data reported in the November 23 issue of *Geophysical Research Letters* also do not predict whether the trend will continue. Regardless, methane stabilization is good news, because it allows for more time to address the main culprit behind climate change: carbon dioxide. —*David Biello*

BRIEF POINTS

- An ancient Greek "computer," the Antikythera, could predict eclipses, count lunar cycles and probably represent the motions of the planets, among other celestial tracking tasks, according to new imaging of the corroded interior.

  Nature, November 30, 2006

- A gene from a wild cousin could boost domesticated wheat's protein, zinc and iron content by 10 to 15 percent. The gene accelerates the plant's maturity, speeding the transfer of nutrients from leaves to grain.

  Science, November 24, 2006

- Extinction "vortex": An economic model finds that placing a species on an endangered list might cause its value to ratchet up and speed its disappearance, thanks to the human thirst for rarity.

  PLoS Biology, November 28, 2006

- Hysteria in the medical sense may be real. Otherwise healthy patients who complained of limb numbness showed no appropriate brain activity when the numb appendage was stimulated, suggesting an underlying neurological defect.

  Neurology, December 12, 2006

GENETICS

Edible Cottonseeds—If You Want to Eat Them

In Joseph Heller's novel *Catch-22*, the business-minded mess officer Milo Minderbinder ends up with a surplus of Egyptian cotton, which he considers unloading by covering it with chocolate and serving it to the soldiers. Even if one did not mind cottonmouth, it would be a bad idea because cottonseeds contain the toxin gossypol. Researchers at Texas A&M University, however, have discovered a method to grow cotton plants that do not produce gossypol in their seeds, a finding that they claim could supply up to 500 million people a year with a high-protein food source. Using RNAi technology, scientists silenced the gene responsible for the production of gossypol in the cottonseed, although field tests outside the greenhouse are necessary to determine the effectiveness of the silencing. The study appears in the November 28 *Proceedings of the National Academy of Sciences USA*. —*Nikhil Swaminathan*

MMM, COTTONY: Gene silencing can render normally toxic cottonseeds safe to eat.

NASA/JPL/MALIN SPACE SCIENCE SYSTEMS (*top*); CORBIS (*bottom*)

sa Extended coverage of these News Scan briefs can be found at www.sciam.com/ontheweb

news SCAN

BY THE NUMBERS

License to Work

PROFESSIONAL LICENSING: ROAD TO SOCIAL ATHEROSCLEROSIS? BY RODGER DOYLE

Milton Friedman, the Nobel laureate economist, noted that the destruction of the medieval guilds was indispensable to the creation of the modern world. In his 1962 classic, *Capitalism and Freedom,* he explained that "there has been a retrogression, an increasing tendency for particular occupations to be restricted to individuals licensed to practice them by the state."

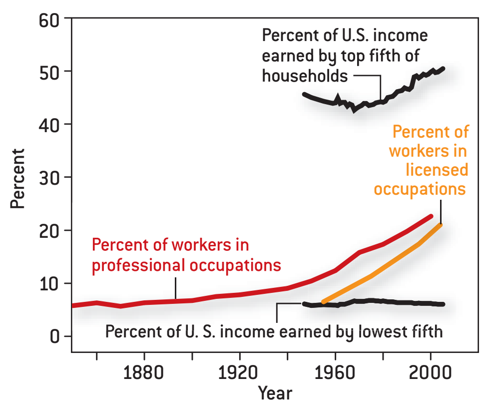
Friedman's warning came at a time when only 5 percent of jobs required licenses. Today the proportion has grown past 20 percent (*left chart*). The states, which obtained the right to grant professional licenses from the U.S. Supreme Court in 1889, now cumulatively recognize at least 800 occupations that require them. But only about 50—including medicine, dentistry, law, engineering, accounting, barbering and cosmetology—are registered in all 50 states. Currently up to half of all professionals are in occupations that require a license, but the extent of licensing varies considerably by state.

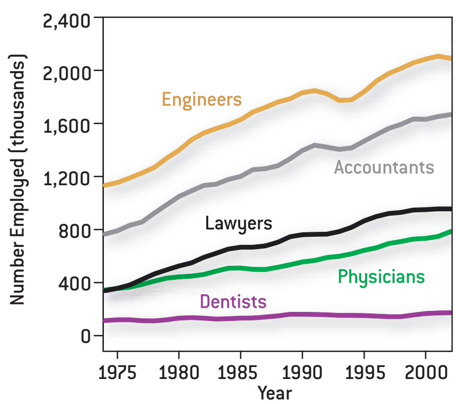
Members of licensing boards typically come from the ranks of the regulated profession, although that does not have to be the case, because the state governor appoints the members. Boards have authority to set standards for apprenticeship and administer professional entrance exams. Thus, they have a powerful means to restrict entry into the professions they regulate and to increase the income of board-certified practitioners. A study based on state-by-state data from 1960 to 1987 found that increased licensing restrictions by state dentistry boards did not lead to improved dental health—but it did boost incomes for dentists.

Additional insight into whether restrictive practices lead to higher professional incomes might come from looking at the incomes of licensed occupations and those of unlicensed occupations with comparable educational requirements. In 2005 the average dentist made 46 percent more than the average physicist, although educational requirements for the two professions are analogous. Lawyers likewise made 37 percent more than economists, whereas barbers made 51 percent more than shampooers. As shown in the left chart, restrictive practices coincide with the rising share of income going to the top fifth of the population. (Other factors, however, probably carry more weight—such as the failure to raise the minimum wage, the loss of traditional blue-collar employment since the 1970s, and government policies that favor the rich.)

Based on employment trends of the five best-paying professional occupational groups in the country (*right chart*), Friedman's implied concern that licensing inhibits job entry does not seem to be justified in most cases; engineering, accountancy, law and medicine have all grown substantially. Meanwhile, however, dentistry has not—a lag consistent with studies suggesting that it has the most restrictive licensing practices of the traditional professions.

Rodger Doyle can be reached at rodgerpdoyle@verizon.net





RODGER DOYLE; SOURCES: *HISTORICAL STATISTICS OF THE UNITED STATES: MILLENNIUM EDITION*, CAMBRIDGE UNIVERSITY PRESS, 2006, AND *CURRENT POPULATION SURVEY* (*percent of workers in professional occupations*)

LICENSE AND REGISTRATION

States with the largest percentage of workers covered by licensing:

California: 30.4
Connecticut: 30.1
Arkansas: 28.6
Michigan: 28.3
Illinois: 27.5

States with the smallest percentage:

Mississippi: 6.1
Wyoming: 11.2
Kansas: 11.9
Indiana: 12
Kentucky: 12

FURTHER READING

Licensing Occupations: Ensuring Quality or Restricting Competition? Morris M. Kleiner. W. E. Upjohn Institute for Employment Research, 2006.

Eat, Drink and Be Merry

Or why we should learn to stop worrying and love food By MICHAEL SHERMER

Among athletes who obsess about their weight, we cyclists are second to none. Training rides are filled with conversations about weight lost or gained and the latest diet regimens and food fads. Resolutions are made and broken. We all know the formula: 10 pounds of extra weight on a 5 percent grade slows your ascent by half a mile an hour. It has a ring of Newtonian finality to it. $F = MA$. The Force needed to turn the pedals equals Acceleration times that Mass on the saddle.

But most of the guys I ride with are like me: in their 40s and 50s with jobs and families, long past racing prime. We ride because it is fun, and it feels good to be fit. So why obsess over a few pounds? Because that is the cycling culture—emblematic of our society at large—and it carries its own internal calculus: the amount of guilt is directly proportional to the rise in the quantity and tastiness of the food.

"A diet that is harmful to one person may be consumed with impunity by another."

The problem is that our bodies have evolved to crave copious amounts of rich and tasty foods, because historically such foods were valuable and rare. How can we modern humans resist? We shouldn't, at least not entirely, says Barry Glassner, a University of Southern California sociologist and author of the forthcoming book *The Gospel of Food: Everything You Think You Know about Food Is Wrong* (Ecco). We have wrongly embraced what Glassner calls "the gospel of naught," the view that "the worth of a meal lies principally in what it lacks. The less sugar, salt, fat, calories, carbs, preservatives, additives, or other suspect stuff, the better the meal." The science behind this culinary religion, Glassner says, is close to naught.

When it comes to healthy absorption of nutrients, taste matters. Glassner cites a study in which "Swedish and Thai women were fed a Thai dish that the Swedes found overly spicy. The Thai women, who liked the dish, absorbed more iron from the meal. When the researchers reversed the experiment and served hamburger, potatoes, and beans, the Swedes, who like this food, absorbed more iron. Most telling was a third variation of the experiment, in which both the Swedes and the Thais were given food that was high in nutrients but consisted of a sticky, savorless paste. In this case, neither group absorbed much iron."

Speaking of iron, Atkins is out and meat is bad, right? Wrong. Glassner notes a study showing that as meat consumption and blood cholesterol levels increased in groups of Greeks, Italians and Japanese, their death rates from heart disease decreased. Of course, many other variables are involved in determining causal relations between diet and health. Glassner cites a study showing a 28 percent decrease in risk of heart attacks among nonsmokers who exercised 30 minutes a day, consumed fish, fiber and folate; they also avoided saturated fats, trans fats and glucose-spiking carbs. According to Harvard University epidemiologist Karin Michels, "It appears more important to increase the number of healthy foods regularly consumed than to reduce the number of less healthy foods regularly consumed."

It's more complicated still. Glassner reviews research showing that heart disease, cancer and other illnesses are significantly increased by "viral and bacterial infections, job stress, living in distressed neighborhoods, early deficits such as malnutrition, low birth weight, lack of parental support, and chronic sleep loss during adolescence and adulthood." Another study found that such diseases "are higher in states where participation in civic life is low, racial prejudice is high, or a large gap exists between the incomes of the rich and poor and of women and men."

To clarify this cornucopia of data, Glassner quotes the former editor of the *New England Journal of Medicine,* Marcia Angell: "Although we would all like to believe that changes in diet or lifestyle can greatly improve our health, the likelihood is that, with a few exceptions such as smoking cessation, many if not most such changes will produce only small effects. And the effects may not be consistent. A diet that is harmful to one person may be consumed with impunity by another."

As the preacher said in Ecclesiastes 8:15: "Then I commended mirth, because a man hath no better thing under the sun, than to eat, and to drink, and to be merry." SA

Michael Shermer is publisher of Skeptic *(www.skeptic.com). His new book is* Why Darwin Matters.

BRAD HINES

Moving beyond Kyoto

To seriously address the issue of global climate change, policymakers need to establish a framework that extends through the end of the century   By JEFFREY D. SACHS

Late in 2006 several events moved the U.S. and other countries closer to serious global negotiations to control greenhouse gas (GHG) emissions. It is therefore timely to ask what a meaningful global agreement would entail. A solid starting point is the 1992 U.N. Framework Convention on Climate Change, the international treaty that binds countries to act on the problem and under which specific measures, such as the Kyoto Protocol, are adopted. The signatories to the Framework Convention, including the U.S. and almost all other countries, declared the objective to be the "stabilization of greenhouse gas concentrations in the atmosphere at a level which would prevent dangerous anthropogenic [man-made] interference with the climate system." The Kyoto Protocol, adopted in 1997, did not implement this idea very well: it took a short-term view of a long-term objective and as a result lost clarity, credibility and support along the way. The key now is to move beyond it.

The costs of emissions control will be far lower than the costs of inaction.

The Kyoto Protocol calls on the high-income countries and the postcommunist nations of eastern Europe and the former Soviet Union to reduce their GHG emissions as of 2012 by around 6 percent compared with the 1990 level. This commitment is far better than nothing (a fair description of the Bush administration's nonpolicy), but it has two major flaws. First, it leaves out the developing countries, which soon will emit more than half of the world's GHGs. Without the active participation of China, India and other developing countries, stabilization of emissions is simply impossible. Second, the Kyoto Protocol takes the long-term objective of stabilization of GHG concentrations and transforms it into a short-term target on emissions reductions, with no clear link between the two. The main actions for stabilization will have to be long-term changes in technology, which exceed the 2012 horizon of the Kyoto Protocol.

This time around, it is better to start with a long-term view. "Dangerous anthropogenic interference" will most likely kick in when carbon concentrations in the atmosphere are at 450 to 550 parts per million (ppm). The world's current trajectory of energy use, deforestation and industrial growth could easily take us to twice that range by the end of the century. *The Stern Review,* an excellent new report by the U.K. Treasury, makes clear that the consequences could be catastrophic: melting of ice sheets, with a huge rise of ocean levels; massive crop failures; increased transmission of diseases; and potentially calamitous effects on ecosystem services.

The world should therefore agree to stabilize GHG concentrations in the 450 to 550 range (my esteemed colleague Jim Hansen urges the lower end of the range, others the higher end). A midcentury goal, perhaps 50 ppm lower, would provide a 40-year target consistent with the end-century target. As new scientific evidence arises, the goals would be periodically adjusted. With the two long-term anchors set, the world's governments could then agree on strategies for reaching them. These strategies would include market incentives to reduce emissions; greatly expanded research on sustainable energy use, land use and industrial development; and technology transfers from rich to poor countries.

The Stern Review makes clear that the costs of such control will be far lower than the costs of inaction. Low-cost, high-benefit efforts look promising in at least three major areas: improved energy efficiency, energy technologies that reduce greenhouse gas emissions, and sustainable land use. Smart technologies can probably keep the long-term annual costs of GHG stabilization at below 1 percent of global GDP. Rich countries can help poor countries to adopt the needed technologies.

It is time, therefore, to aim for a sensible long-term framework in which all countries will participate. The economics are right. The U.S. Congress is set to back such a course. The White House will as well, soon after 2008 and, with some luck, even before. SA

Jeffrey D. Sachs is director of the Earth Institute at Columbia University.

An expanded version of this essay is available at ***www.sciam.com/ontheweb***

BRUCE GILBERT/EARTH INSTITUTE

Graft and Host, Together Forever

Thomas E. Starzl pioneered organ transplantation with antirejection drugs—an approach he hopes to end through a phenomenon called microchimerism By MARGUERITE HOLLOWAY

The dogs bound into the office, two of them. Not quite Hounds of the Baskervilles, but large enough. It is a dramatic lead. A minute or so later Thomas E. Starzl follows them in, and, while his pets nap and his assistant provides this or that document or letter, he settles in to recount a story, a compelling narrative of a field that has come full circle. All the classical elements are there: a missed turn, bad timing, a paradigm shift, some overturned dogma and a satisfying, hopeful conclusion: organ transplantation without a lifetime of antirejection drugs.

THOMAS E. STARZL: RETHINKING REJECTION

- **Advocates immune tolerance as a transplant strategy. One method: infusing cells from organ donors into recipients weeks before surgery.**
- **Immunosuppressive drugs have boosted survival rates, but they also increase the risk of infections, cancer, hypertension and other illnesses.**
- **"He is a visionary, a tremendously large figure. You need people like that in emerging fields." —*Nicholas L. Tilney, Harvard Medical School***

The 80-year-old Starzl, a transplant surgeon and researcher at the University of Pittsburgh, where he has an institute named after him, is legendary for his groundbreaking work over the past five decades. He was the first person to perform human liver transplants. He developed new techniques for transplant surgery, helped to make kidney transplantation viable and was one of the first researchers to try xenografts—in the 1960s he placed baboon kidneys in six patients. (None of the transplants lasted long.)

Crucially, he experimented with, combined and developed drugs to suppress the immune system, thereby preventing organ rejection. He advocated widespread use of these immunosuppressants, and because of these drugs, the number of transplants has grown every year for the past several decades; in 2005 surgeons performed 28,107 transplants of the kidney, liver, pancreas, heart, lung and intestine, according to the United Network for Organ Sharing. But although the drugs permit transplants and save lives, they also have debilitating and sometimes deadly side effects, because the weakened immune system cannot fight viruses or cancers. Transplant specialists have considered the chemicals to be a necessary evil: freed from their dampening influence, the patient's immune system would rebound and reject the foreign organ.

In 1992 Starzl observed something that convinced him to rethink the way immunosuppressants are used. He had brought together many of his former patients, including some he had operated on in the early 1960s. He learned that some of them had stopped taking their drugs long ago and were doing just fine. Starzl tested these patients, hoping to see something consistent; he observed donor cells in various tissues and blood.

THEO ANDERSON

The phenomenon is called microchimerism, a condition in which a small number of cells from two individuals coexist in one body. Twins can be microchimeric, having traded cells in utero; mothers and their children can be microchimeric as well, for the same reason. (Often this coexistence is peaceable; there is some evidence, however, that microchimerism could play a role in autoimmune disorders.)

For Starzl, these shared cells are the key to tolerance—acceptance of the graft by the host. His hypothesis, essentially, is that the body comes to terms with "other" by dealing with it in an incremental way, by coming to see some circulating donor cells as "self" and paving the way for acceptance on a larger scale. The presence of large numbers of donor cells in recipients has long been observed in bone marrow transplantation, a discipline Starzl believes has advanced basic science more than organ transplantation has to date. He argues that evidence of microchimerism in his patients finally unites the science of organ transplantation with that of bone marrow, allowing his field to move beyond being "totally drug-related, which is kind of humiliating."

In microchimerism, a small number of cells from two individuals coexist in one body—a potential key to transplantation without immunosuppressants.

Considering his long advocacy for antirejection drugs, some observers have characterized Starzl's new approach as a reversal. But Starzl insists it is not. He points to a paper he wrote in 1963 about tolerance in kidney transplant patients as anticipating his mind-set today—it was just that the timing was not right; the science was not advanced enough to make sense of what he saw, of the paradigm shift in the wings. "That is the foundation of this idea that I have been pursuing ever since," he declares.

At Pittsburgh, Starzl's colleagues have incorporated his thinking about microchimerism into their procedures. They treat with a powerful immunosuppressant three weeks before surgery, then inject donor cells into the patient and follow with low doses of drugs until the operation. "You're treating the recipient's cells, so they won't cause a host-versus-graft response, but also you're treating the donor's cells, some of which are immunocompetent, so they don't try to reject the recipient," Starzl explains. After surgery, physicians prescribe lower-than-normal doses, trying to wean patients to very low levels or off the drugs completely. According to Ron Shapiro, a colleague of Starzl's at Pittsburgh, various transplant centers are following a similar protocol, including one in Galveston, Tex., and several in Europe. Researchers at Stanford and Harvard universities, among other places, also are working to understand the underlying science of microchimerism and tolerance.

Still, microchimerism as the mechanism explaining tolerance is not convincing to many in Starzl's field, observes Fritz H. Bach of Harvard Medical School. "The microchimerism idea and data have never gained traction; I think most individuals do not believe the concept as explaining tolerance," he sums up.

"His hypothesis is in dispute—the evidence isn't there," concurs David E. R. Sutherland of the University of Minnesota, who believes microchimerism is a consequence of not rejecting, rather than a cause. "Why some people don't reject when they come off immunosuppression is not understood. We don't understand why, and there is no reproducible protocol."

In addition, Starzl is legendary for his campaigns, for approaches he has advocated despite their controversy—or, as some of his colleagues describe, despite their lack of supporting evidence. His xenografts in the 1960s were one such passion, and their ultimate failure earned Starzl criticism. In the 1980s he backed a drug called FK506, or tacrolimus, that had proved highly toxic in some animal studies. "Tom persisted in using tacrolimus despite initial problems with toxicity that discouraged his peers, and he turned out to be correct," describes Nicholas L. Tilney of Harvard Medical School, current president of the Transplantation Society and author of the 2003 book *Transplant: From Myth to Reality* (Yale University Press). The Food and Drug Administration approved tacrolimus in 1994, and today it is used in roughly 80 percent of kidney and 90 percent of liver and pancreas transplants.

"He really did push the envelope," Tilney says. "He did bring things up over the years, many of which turned out to be right and many of which didn't. He likes to stir the pot."

Whether microchimerism proves to be the mechanism engendering tolerance, no one disputes the influence Starzl has and continues to have on his field—and his gift for communicating his vision in a compelling narrative. Sutherland compares Starzl with Dostoyevsky. "He has that sense of how to bring things together, and he is doing that now," he says. "He just does things. He darts through, makes end runs, and he moves the field forward."

Shapiro agrees. "The interesting thing is that everyone thinks he is crazy, but when he moves on to the next thing, the first thing he says becomes conventional wisdom," he says. "Everyone knows he is very smart. But he is actually a little smarter even than that." Future organ recipients certainly hope that assessment of Starzl is right. SA

Page Intentionally Blank

The Universe's INVISIBLE HAND

Dark energy does more than hurry along the expansion of the universe. It also has a stranglehold on the shape and spacing of galaxies

BY CHRISTOPHER J. CONSELICE

WHAT TOOK US SO LONG? Only in 1998 did astronomers discover we had been missing nearly three quarters of the contents of the universe, the so-called dark energy—an unknown form of energy that surrounds each of us, tugging at us ever so slightly, holding the fate of the cosmos in its grip, but to which we are almost totally blind. Some researchers, to be sure, had anticipated that such energy existed, but even they will tell you that its detection ranks among the most revolutionary discoveries in 20th-century cosmology. Not only does dark energy appear to make up the bulk of the universe, but its existence, if it stands the test of time, will probably require the development of new theories of physics.

Scientists are just starting the long process of figuring out what dark energy is and what its implications are. One realization has already sunk in: although dark energy betrayed its existence through its effect on the universe as a whole, it may also shape the evolution of the universe's inhabitants—stars, galaxies, galaxy clusters. Astronomers may have been staring at its handiwork for decades without realizing it.

Ironically, the very pervasiveness of dark energy is what made it so hard to recognize. Dark energy, unlike matter, does not clump in some places more than others; by its very nature, it is spread smoothly everywhere. Whatever the location—be it in your kitchen or in intergalactic space—it has the same density, about 10^{-26} kilogram per cubic meter, equivalent to a handful of hydrogen atoms. All the dark energy in our solar system amounts to the mass of a small asteroid, making it an utterly inconsequential player in the dance of the planets. Its effects stand out only when viewed over vast distances and spans of time.

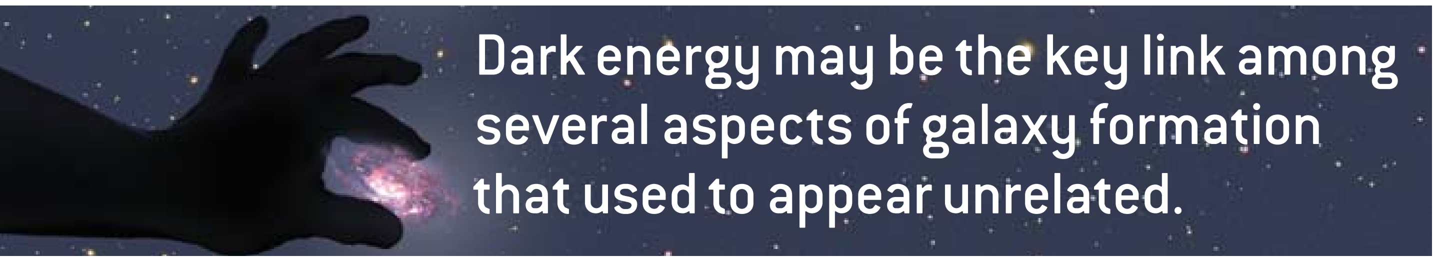
Since the days of American astronomer Edwin Hubble, observers have known that all but the nearest galaxies are moving away from us at a rapid rate. This rate is proportional to distance: the more distant a galaxy is, the faster its recession. Such a pattern implied that galaxies are not moving through space in the conventional sense but are being carried along as the fabric of space itself stretches [see "Misconceptions about the Big Bang," by Charles H. Lineweaver and Tamara M. Davis; SCI-

JEAN-FRANCOIS PODEVIN

entific American, March 2005]. For decades, astronomers struggled to answer the obvious follow-up question: How does the expansion rate change over time? They reasoned that it should be slowing down, as the inward gravitational attraction exerted by galaxies on one another should have counteracted the outward expansion.

The first clear observational evidence for changes in the expansion rate involved distant supernovae, massive exploding stars that can be used as markers of cosmic expansion, just as watching driftwood lets you measure the speed of a river. These observations made clear that the expansion was slower in the past than today and is therefore accelerating. More specifically, it had been slowing down but at some point underwent a transition and began speeding up [see "Surveying Space-time with Supernovae," by Craig J. Hogan, Robert P. Kirshner and Nicholas B. Suntzeff; Scientific American, January 1999, and "From Slowdown to Speedup," by Adam G. Riess and Michael S. Turner; Scientific American, February 2004]. This striking result has since been cross-checked by independent studies of the cosmic microwave background radiation by, for example, the Wilkinson Microwave Anisotropy Probe (WMAP).

One possible conclusion is that different laws of gravity apply on supergalactic scales than on lesser ones, so that galaxies' gravity does not, in fact, resist expansion. But the more generally accepted hypothesis is that the laws of gravity are universal and that some form of energy, previously unknown to science, opposes and overwhelms galaxies' mutual attraction, pushing them apart ever faster. Although dark energy is inconsequential within our galaxy (let alone your kitchen), it adds up to the most powerful force in the cosmos.



Cosmic Sculptor

As astronomers have explored this new phenomenon, they have found that, in addition to determining the overall expansion rate of the universe, dark energy has long-term consequences for smaller scales. As you zoom in from the entire observable universe, the first thing you notice is that matter on cosmic scales is distributed in a cobweblike pattern—a filigree of filaments, several tens of millions of light-years long, interspersed with voids of similar size. Simulations show that both matter and dark energy are needed to explain the pattern.

That finding is not terribly surprising, though. The filaments and voids are not coherent bodies like, say, a planet. They have not detached from the overall cosmic expansion and established their own internal equilibrium of forces. Rather they are features shaped by the competition between cosmic expansion (and any phenomenon affecting it) and their own gravity. In our universe, neither player in this tug-of-war is overwhelmingly dominant. If dark energy were stronger, expansion would have won and matter would be spread out rather than concentrated in filaments. If dark energy were weaker, matter would be even more concentrated than it is [*see box on page 40*].

The situation gets more complicated as you continue to zoom in and reach the scale of galaxies and galaxy clusters. Galaxies, including our own Milky Way, do not expand with time. Their size is controlled by an equilibrium between gravity and the angular momentum of the stars, gas and other material that make them up; they grow only by accreting new material from intergalactic space or by merging with other galaxies. Cosmic expansion has an insignificant effect on them. Thus, it is not at all obvious that dark energy should have had any say whatsoever in how galaxies formed. The same is true of galaxy clusters, the largest coherent bodies in the universe—assemblages of thousands of galaxies embedded in a vast cloud of hot gas and bound together by gravity.

Yet it now appears that dark energy may be the key link among several aspects of galaxy and cluster formation that not long ago appeared unrelated. The reason is that the formation and evolution of these systems is partially driven by interac-

Overview/*Dark Energy*

- Dark energy is best known as the putative agent of cosmic acceleration—an unidentified substance that exerts a kind of antigravity force on the universe as a whole.
- Less well known is that dark energy also has secondary effects on material within the universe. It helped to imprint the characteristic filigree pattern of matter on large scales. On a smaller scale, it appears to have choked off the growth of galaxy clusters some six billion years ago.
- On a still smaller scale, dark energy has reduced the rate at which galaxies yank on, bang into and merge with one another. Such interactions shape galaxies. Had dark energy been weaker or stronger, the Milky Way might have had a lower star formation rate, so the heavy elements that constitute our planet might never have been synthesized.

Supernova 1994D in galaxy NGC 4526

NASA, ESA, THE HUBBLE KEY PROJECT TEAM AND THE HIGH-Z SUPERNOVA SEARCH TEAM

tions and mergers between galaxies, which in turn may have been driven strongly by dark energy.

To understand the influence of dark energy on the formation of galaxies, first consider how astronomers think galaxies form. Current theories are based on the idea that matter comes in two basic kinds. First, there is ordinary matter, whose particles readily interact with one another and, if electrically charged, with electromagnetic radiation. Astronomers call this type of matter "baryonic" in reference to its main constituent, baryons, such as protons and neutrons. Second, there is dark matter (which is distinct from dark energy), which makes up 85 percent of all matter and whose salient property is that it comprises particles that do not react with radiation. Gravitationally, dark matter behaves just like ordinary matter.

According to models, dark matter began to clump immediately after the big bang, forming spherical blobs that astronomers refer to as "halos." The baryons, in contrast, were initially kept from clumping by their interactions with one another and with radiation. They remained in a hot, gaseous phase. As the universe expanded, this gas cooled and the baryons were able to pack themselves together. The first stars and galaxies coalesced out of this cooled gas a few hundred million years after the big bang. They did not materialize in random locations but in the centers of the dark matter halos that had already taken shape.

Since the 1980s a number of theorists have done detailed computer simulations of this process, including groups led by Simon D. M. White of the Max Planck Institute for Astrophysics in Garching, Germany, and Carlos S. Frenk of Durham University in England. They have shown that most of the first structures were small, low-mass dark matter halos. Because the early universe was so dense, these low-mass halos (and the galaxies they contained) merged with one another to form larger-mass systems. In this way, galaxy construction was a bottom-up process, like building a dollhouse out of Lego bricks. (The alternative would have been a top-down process, in which you start with the dollhouse and smash it to make bricks.) My colleagues and I have sought to test these models by looking at distant galaxies and how they have merged over cosmic time.

Galaxy Formation Peters Out

DETAILED STUDIES INDICATE that a galaxy gets bent out of shape when it merges with another galaxy. The earliest galaxies we can see existed when the universe was about a billion years old, and many of these indeed appear to be merging. As time went on, though, the fusion of massive galaxies became less common. Between two billion and six billion years after the big bang—that is, over the first half of cosmic history—the fraction of massive galaxies undergoing a merger dropped from half to nearly nothing at all. Since then, the distribution of galaxy shapes has been frozen, an indication that smashups and mergers have become relatively uncommon.

In fact, fully 98 percent of massive galaxies in today's universe are either elliptical or spiral, with shapes that would be disrupted by a merger. These galaxies are stable and comprise mostly old stars, which tells us that they must have formed early and have remained in a regular morphological form for quite some time. A few galaxies are merging in the present day, but they are typically of low mass.

The virtual cessation of mergers is not the only way the universe has run out of steam since it was half its current age. Star formation, too, has been waning. Most of the stars that exist today were born in the first half of cosmic history, as first convincingly shown by several teams in the 1990s, including

EVIDENCE FOR DARK ENERGY

SUPERNOVA EXPLOSIONS
In an expanding universe, galaxies move apart at a speed that depends on the distance between them. Supernovae offer a way to measure this effect: their spectral redshift reveals the speed of their host galaxies, and their brightness reveals distance. It turns out that galaxies billions of years ago were moving slower than a simple extrapolation from the current rate of expansion would imply. The expansion rate must have increased over that time—the hallmark of dark energy.

COSMIC MICROWAVE BACKGROUND RADIATION
Images of the background radiation contain spots whose apparent size reflects the overall geometry of space and therefore the density of the universe. This quantity exceeds the amount of matter (both ordinary and exotic), so a missing component such as dark energy must make up the difference. In addition, the background radiation has been slightly reworked by the gravitational fields of cosmic structures. The amount of reworking depends on how the expansion rate has changed over time and matches what dark energy would do.

GALAXY CONFIGURATION
Galaxies are not sprinkled randomly through the heavens. Instead they are arranged in patterns, one of which resembles the spots in the microwave background. It can be used to measure the total mass of the universe and confirm the need for dark energy.

GRAVITATIONAL LENSING
A lump of mass can serve as a lens; its gravity bends light. Such a lens can produce multiple images, like a fun-house mirror, if the light source is directly behind it—an alignment that becomes more probable the bigger the universe is, which in turn depends on the amount of dark energy. A weaker lens can still bend light by a small angle that depends on its mass. Studies of this process have revealed how clumps of matter have grown over time and found the imprint of dark energy.

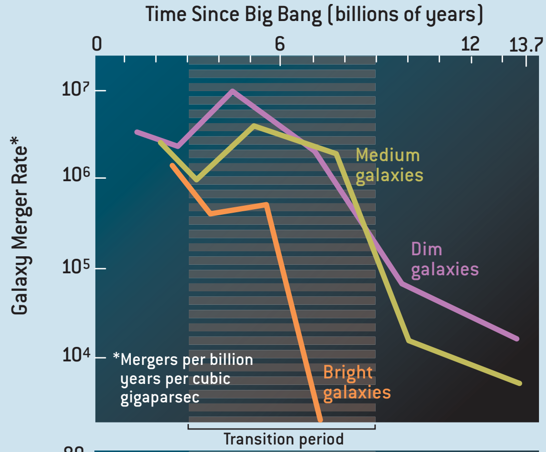
GALAXY CLUSTERS
X-ray observations trace the evolution of the mass of galaxy clusters. Dark energy is required to explain when and how they formed.

Dark Energy Takes Charge

Despite getting a quick start after the big bang, the construction of the universe soon petered out. Initially galaxies merged together, changed shape and formed stars at a brisk pace, but this activity began to wane during the period when dark energy became comparable in strength to matter (*brown area in graphs*). Coincidence?

GALAXIES STOPPED MERGING

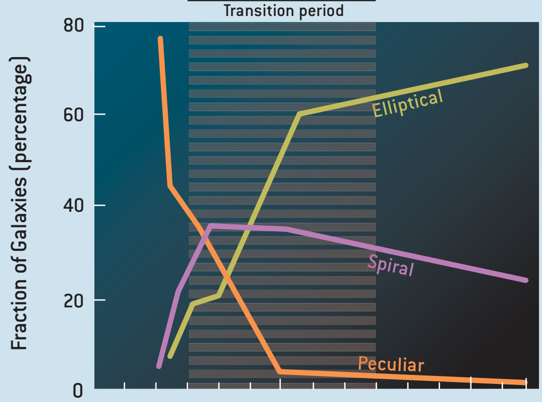
The brightest galaxies stopped colliding and amalgamating at a cosmic age of about six billion years. Less luminous galaxies can still merge but have become much less likely to do so.



Merging galaxies NGC 4676

GALAXIES SETTLED INTO REGULAR SHAPES

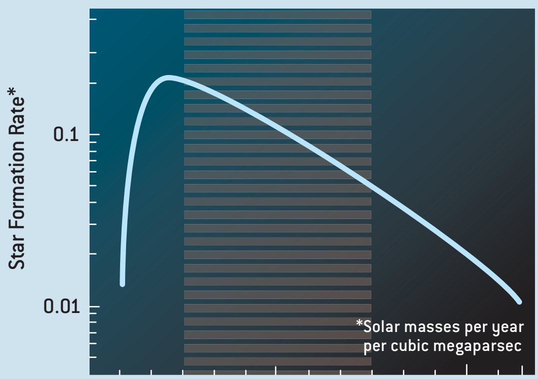
Early on, most galaxies looked peculiar—a sign they were merging with one another. As mergers became less frequent, spiral and elliptical shapes become prevalent.



Hickson Compact Galaxy Group 87

STAR FORMATION WANED

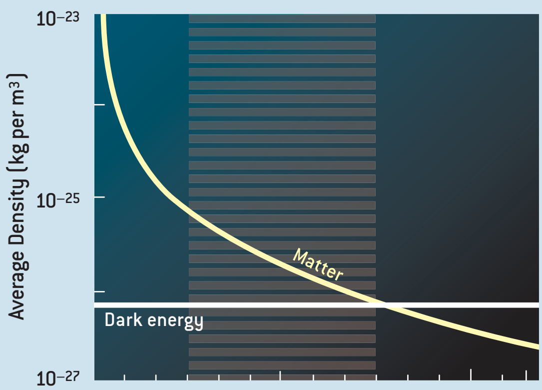
The early universe was a cauldron of star formation, but the formation rate soon peaked and began to drop. It is now lower than it has ever been.

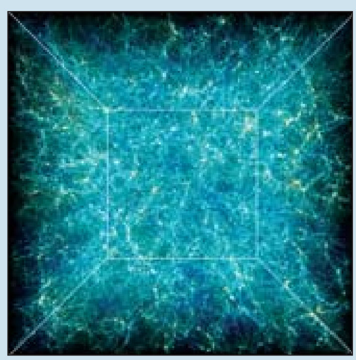


New Stars in Trapezium Cluster

DARK ENERGY BECAME A PLAYER

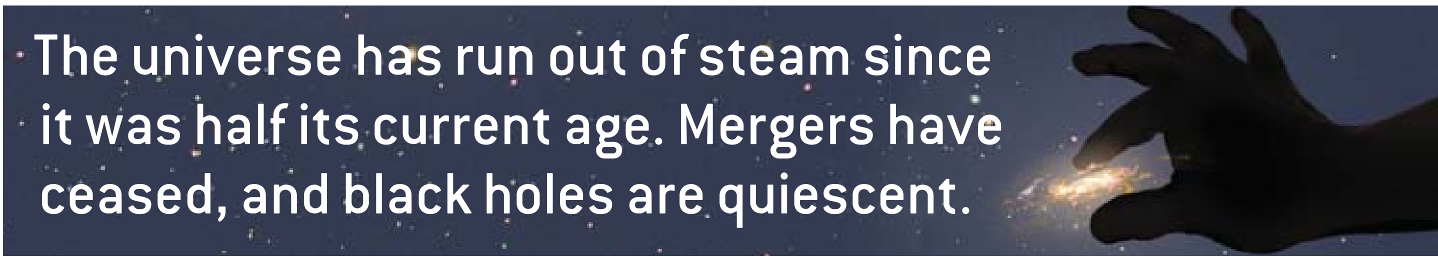
All these observed trends can be related to one simple fact: as the universe expanded, matter was spread thinner, and as its density approached that of dark energy (whose density is constant in the simplest model), the rate of expansion began to switch from decelerating to accelerating. Galaxies were pulled apart faster and became less likely to bump into one another or sweep up gas to fuel star formation.





Simulation of Matter Distribution

LUCY READING-IKKANDA (*graphs*); ANDREW HOPKINS *University of Sydney* (*source of star formation rate graph*); NASA/ACS SCIENCE & ENGINEERING TEAM (*merging galaxies*); NASA/ESA/STScI/HUBBLE HERITAGE TEAM *Photo Researchers, Inc.* (*Hickson Group*); NASA/K. L. LUHMAN *Harvard-Smithsonian Center for Astrophysics* AND G. SCHNEIDER, E. YOUNG, G. RIEKE, A. COTERA, H. CHEN, M. RIEKE AND R. THOMPSON *Steward Observatory, University of Arizona* (*Trapezium cluster*); IMAGE GENERATION BY THE VISUALIZATION INSTITUTE AT THE UNIVERSITY OF STUTTGART, DATA FROM THE VIRGO SUPERCOMPUTING CONSORTIUM OF THE MAX PLANCK SOCIETY IN GARCHING AND THE INSTITUTE FOR COMPUTATIONAL COSMOLOGY (*simulation of matter distribution*) ones led by Simon J. Lilly, then at the University of Toronto, Piero Madau, then at the Space Telescope Science Institute, and Charles C. Steidel of the California Institute of Technology. More recently, researchers have learned how this trend occurred. It turns out that star formation in massive galaxies shut down early. Since the universe was half its current age, only lightweight systems have continued to create stars at a significant rate. This shift in the venue of star formation is called galaxy downsizing [see "The Midlife Crisis of the Cosmos," by Amy J. Barger; Scientific American, January 2005]. It seems paradoxical. Galaxy formation theory predicts that small galaxies take shape first and, as they amalgamate, massive ones arise. Yet the history of star formation shows the reverse: massive galaxies are initially the main stellar birthing grounds, then smaller ones take over.



Another oddity is that the buildup of supermassive black holes, found at the centers of galaxies, seems to have slowed down considerably. Such holes power quasars and other types of active galaxies, which are rare in the modern universe; the black holes in our galaxy and others are quiescent. Are any of these trends in galaxy evolution related? Is it really possible that dark energy is the root cause?

The Steady Grip of Dark Energy

SOME ASTRONOMERS HAVE PROPOSED that internal processes in galaxies, such as energy released by black holes and supernovae, turned off galaxy and star formation. But dark energy has emerged as possibly a more fundamental culprit, the one that can link everything together. The central piece of evidence is the rough coincidence in timing between the end of most galaxy and cluster formation and the onset of the domination of dark energy. Both happened when the universe was about half its present age.

The idea is that up to that point in cosmic history, the density of matter was so high that gravitational forces among galaxies dominated over the effects of dark energy. Galaxies rubbed shoulders, interacted with one another, and frequently merged. New stars formed as gas clouds within galaxies collided, and black holes grew when gas was driven toward the centers of these systems. As time progressed and space expanded, matter thinned out and its gravity weakened, whereas the strength of dark energy remained constant (or nearly so). The inexorable shift in the balance between the two eventually caused the expansion rate to switch from deceleration to acceleration. The structures in which galaxies reside were then pulled apart, with a gradual decrease in the galaxy merger rate as a result. Likewise, intergalactic gas was less able to fall into galaxies. Deprived of fuel, black holes became more quiescent.

This sequence could perhaps account for the downsizing of the galaxy population. The most massive dark matter halos, as well as their embedded galaxies, are also the most clustered; they reside in close proximity to other massive halos. Thus, they are likely to knock into their neighbors earlier than are lower-mass systems. When they do, they experience a burst of star formation. The newly formed stars light up and then blow up, heating the gas and preventing it from collapsing into new stars. In this way, star formation chokes itself off: stars heat the gas from which they emerged, preventing new ones from forming. The black hole at the center of such a galaxy acts as another damper on star formation. A galaxy merger feeds gas into the black hole, causing it to fire out jets that heat up gas in the system and prevent it from cooling to form new stars.

Apparently, once star formation in massive galaxies shuts down, it does not start up again—most likely because the gas in these systems becomes depleted or becomes so hot that it cannot cool down quickly enough. These massive galaxies can still merge with one another, but few new stars emerge for want of cold gas. As the massive galaxies stagnate, smaller galaxies continue to merge and form stars. The result is that massive galaxies take shape before smaller ones, as is observed. Dark energy perhaps modulated this process by determining the degree of galaxy clustering and the rate of merging.

Dark energy would also explain the evolution of galaxy clusters. Ancient clusters, found when the universe was less than half its present age, were already as massive as today's clusters. That is, galaxy clusters have not grown by a significant amount in the past six billion to eight billion years. This lack of growth is an indication that the infall of galaxies into clusters has been curtailed since the universe was about half its current age—a direct sign that dark energy is influencing the way galaxies are interacting on large scales. Astronomers knew as early as the mid-1990s that galaxy clusters had not grown much in the past eight billion years, and they attributed

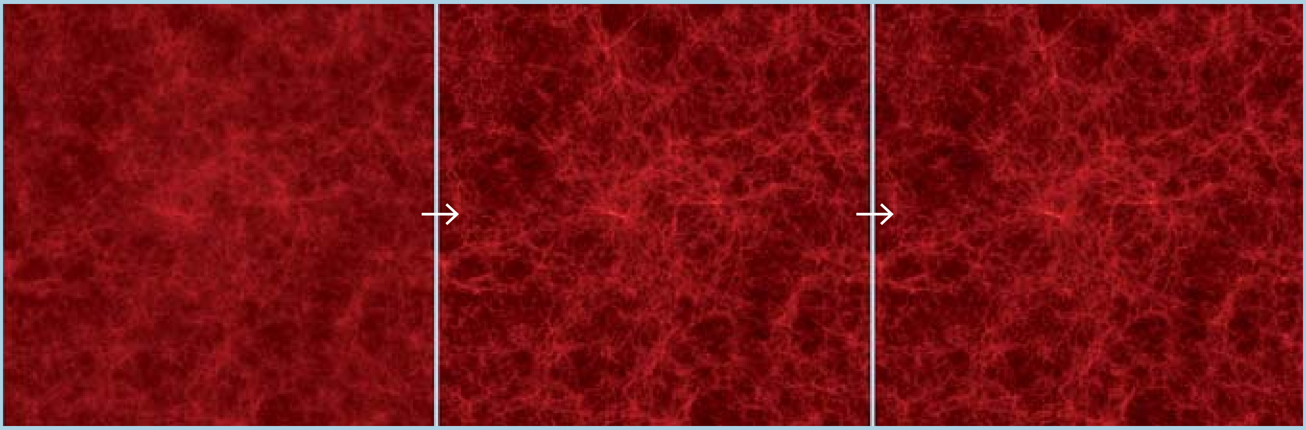
THE AUTHOR

CHRISTOPHER J. CONSELICE is an astronomer and Lecturer at the University of Nottingham in England, where he recently moved from the California Institute of Technology. He specializes in the formation of galaxies and leads several observational programs in infrared and visible light with telescopes both on the ground and in space. A lover of both the heavens and the earth, he comes from a family of Pennsylvanian farmers and spends his free time boating, fishing, biking and caving.

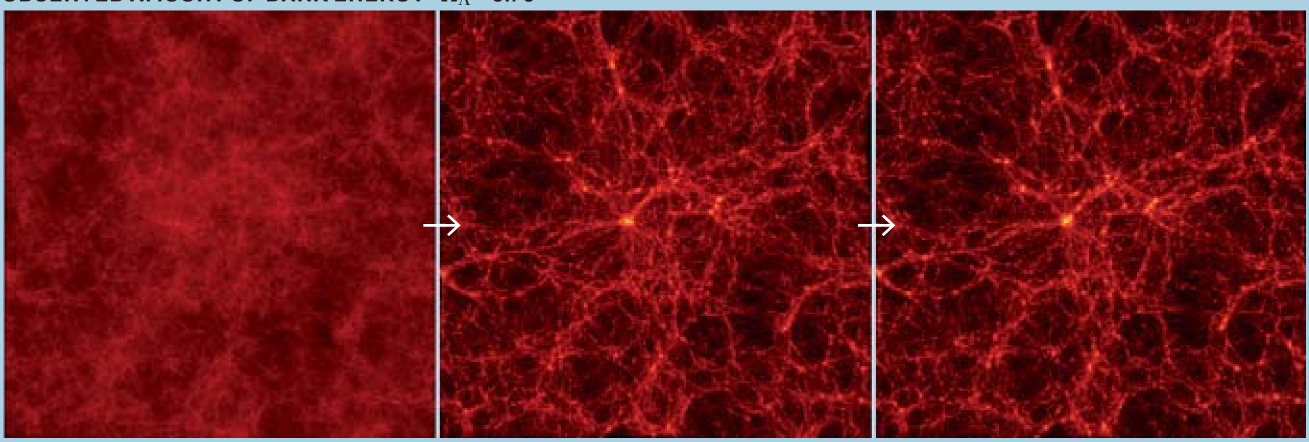
What-If Scenarios

If the universe had more dark energy in it, it would look radically different. Cosmic acceleration would have started sooner, pulled material apart faster and nipped the formation of large structures in the bud. The converse would happen if the universe had less dark energy. Each box below shows a region that is now one billion light-years across and contains 27 million particles, each representing a galaxy. These simulations assume that the dark energy density is constant in space and time. The quantity Ω_Λ is the governing cosmological parameter; it represents the density of dark energy today.

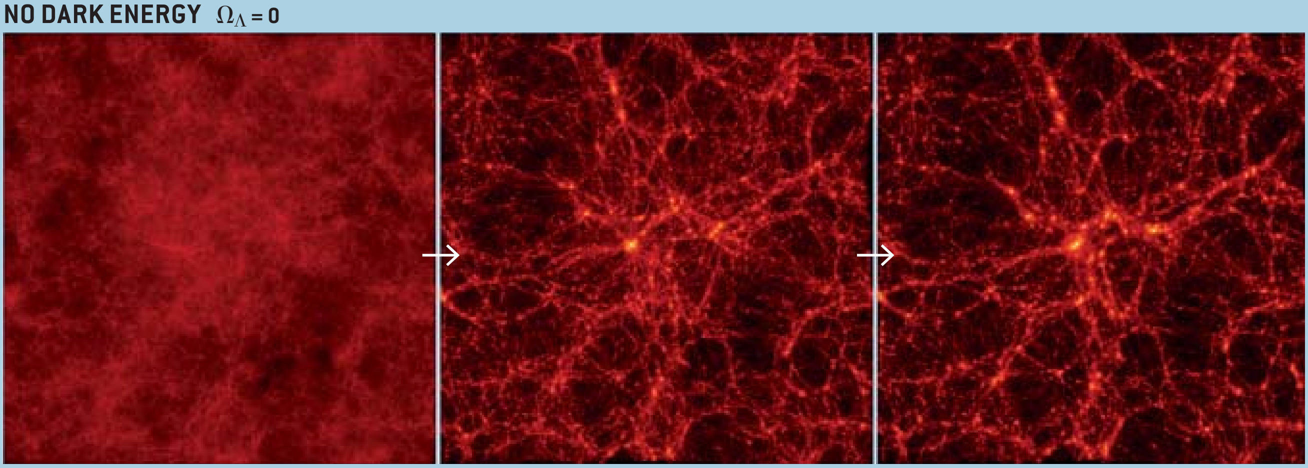
MORE DARK ENERGY $\Omega_\Lambda = 0.99$



OBSERVED AMOUNT OF DARK ENERGY $\Omega_\Lambda = 0.75$



NO DARK ENERGY $\Omega_\Lambda = 0$



EARLY UNIVERSE: When the universe is a sixth of its current size, matter is evenly distributed in all three scenarios. Dark energy has not yet exerted its influence.

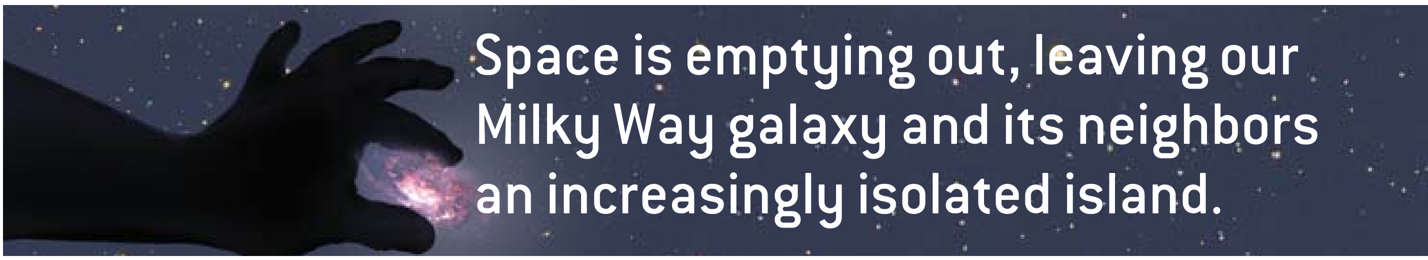
TRANSITION PERIOD: When the universe is 75 percent of its current size, the effects of dark energy are stark. In the high dark energy scenario (*top*), the universe looks amorphous. In the other two scenarios, structure formation still continues, producing a cobweb pattern.

TODAY: In a universe with the observed amount of dark energy (*middle*), large-scale structure formation has ended, leaving the cobweb frozen in place. In a zero dark energy scenario (*bottom*), the cobweb continues to develop.

LIANG GAO, CARLOS FRENK AND ADRIAN JENKINS *Institute for Computational Cosmology, Durham University* this to a lower matter density than theoretical arguments had predicted. The discovery of dark energy resolved the tension between observation and theory.

An example of how dark energy alters the history of galaxy clusters is the fate of the galaxies in our immediate vicinity, known as the Local Group. Just a few years ago astronomers thought that the Milky Way and Andromeda, its closest large neighbor, along with their retinue of satellites, would fall into the nearby Virgo cluster. But it now appears that we shall escape that fate and never become part of a large cluster of galaxies. Dark energy will cause the distance between us and Virgo to expand faster than the Local Group can cross it.

By throttling cluster development, dark energy also controls the makeup of galaxies within clusters. The cluster environment facilitates the formation of a zoo of galaxies such as the so-called lenticulars, giant ellipticals and dwarf ellipticals. By regulating the ability of galaxies to join clusters, dark energy dictates the relative abundance of these galaxy types.



This is a good story, but is it true? Galaxy mergers, black hole activity and star formation all decline with time, and very likely they are related in some way. But astronomers have yet to follow the full sequence of events. Ongoing surveys with the Hubble Space Telescope, the Chandra X-ray Observatory and sensitive ground-based imaging and spectroscopy will scrutinize these links in coming years. One way to do this is to obtain a good census of distant active galaxies and to determine the time when those galaxies last underwent a merger. The analysis will require the development of new theoretical tools but should be within our grasp in the next few years.

Striking a Balance

AN ACCELERATING UNIVERSE dominated by dark energy is a natural way to produce all the observed changes in the galaxy population—namely, the cessation of mergers and its many corollaries, such as loss of vigorous star formation and the end of galactic metamorphosis. If dark energy did not exist, galaxy mergers would probably have continued for longer than they did, and today the universe would contain many more massive galaxies with old stellar populations. Likewise, it would have fewer lower-mass systems, and spiral galaxies such as our Milky Way would be rare (given that spirals cannot survive the merger process). Large-scale structures of galaxies would have been more tightly bound, and more mergers of structures and accretion would have occurred.

Conversely, if dark energy were even stronger than it is, the universe would have had fewer mergers and thus fewer massive galaxies and galaxy clusters. Spiral and low-mass dwarf irregular galaxies would be more common, because fewer galaxy mergers would have occurred throughout time, and galaxy clusters would be much less massive or perhaps not exist at all. It is also likely that fewer stars would have formed, and a higher fraction of our universe's baryonic mass would still be in a gaseous state.

Although these processes may seem distant, the way galaxies form has an influence on our own existence. Stars are needed to produce elements heavier than lithium, which are used to build terrestrial planets and life. If lower star formation rates meant that these elements did not form in great abundance, the universe would not have many planets, and life itself might never have arisen. In this way, dark energy could have had a profound effect on many different and seemingly unrelated aspects of the universe, and perhaps even on the detailed history of our own planet.

Dark energy is by no means finished with its work. It may appear to benefit life: the acceleration will prevent the eventual collapse that was a worry of astronomers not so long ago. But dark energy brings other risks. At the very least, it pulls apart distant galaxies, making them recede so fast that we lose sight of them for good. Space is emptying out, leaving our galaxy and its immediate neighbors an increasingly isolated island. Galaxy clusters, galaxies and even stars drifting through intergalactic space will eventually have a limited sphere of gravitational influence not much larger than their own individual sizes.

Worse, dark energy might be evolving. Some models predict that if dark energy becomes ever more dominant over time, it will rip apart gravitationally bound objects, such as galaxy clusters and galaxies. Ultimately, planet Earth will be stripped from the sun and shredded, along with all objects on it. Even atoms will be destroyed. Dark energy, once cast in the shadows of matter, will have exacted its final revenge.

MORE TO EXPLORE

A Direct Measurement of Major Galaxy Mergers at $z \lesssim 3$. Christopher J. Conselice, Matthew A. Bershady, Mark Dickinson and Casey Papovich in *Astronomical Journal,* Vol. 126, No. 3, pages 1183–1207; September 2003. **www.arxiv.org/abs/astro-ph/0306106**

Dark Energy. Robert R. Caldwell in *Physics World,* Vol. 17, No. 5, pages 37–42; May 2004. **http://physicsweb.org/articles/world/17/5/7**

The Extravagant Universe: Exploding Stars, Dark Energy, and the Accelerating Cosmos. Robert P. Kirshner. Princeton University Press, 2004.

The Infinite Cosmos: Questions from the Frontiers of Cosmology. Joseph Silk. Oxford University Press, 2006.

TRACKING AN

The case was cold—the bones in the mass grave were 70 million years old. But critical clues pointed to the killer's identity

By Raymond R. Rogers and David W. Krause

RAYMOND R. ROGERS

LARGE MEAT-EATING DINOSAUR *Majungatholus atopus* (*above*) met an untimely end some 70 million years ago in what is now northwestern Madagascar (*opposite page, top*). Members of the authors' team carefully excavated the remains, including a jaw with serrated teeth used to slice through flesh (*right*), and packed them in plaster for transport to the U.S. (*opposite page, bottom*), where the researchers studied the fossils in detail for clues to the cause of death.

ANCIENT KILLER

One body rests on its left side, head and neck pulled back toward the pelvis—a classic death pose. The arms and legs are still in their anatomically correct positions, but closer inspection reveals that bones of the hands and feet are dislocated, although most parts are present and accounted for. The skull, too, is somewhat disjointed, and here again the component pieces lie near one another. Curiously, the tip of the tail is missing altogether. Nearby rest more corpses in markedly different states of preservation and disarray. Some are still largely intact, others represented by only a skull, a shoulder blade, a single limb bone. Did the unfortunate creatures die here, or were they brought together after their demise? Did they all perish at the same instant, or did their deaths transpire over time? And what killed them?

Our team of Malagasy and American paleontologists and geologists started asking such questions as soon as we discovered the mass grave in the summer of 2005 in the ancient sediments of northwestern Madagascar, an island whose Venetian-red soils

COLD CASE

inspired its nickname, the Great Red Island. We turned up some intriguing information as we searched for the answers, but how we went about the task is perhaps as interesting as what we found.

Before doing anything, we named the site, designating it MAD05-42 to indicate the year it was found and its sequence in the discovery of fossil localities in this area. The second task was identifying the dead, and based on our discoveries elsewhere in the region, we quickly discerned that most of the remains were dinosaurs of various species.

This dinosaur burial ground is not unique for northwestern Madagascar. It matches a pattern we have seen repeatedly over a decade of geologic investigation in the semiarid grasslands near the remote village of Berivotra. There we have uncovered layer on layer of mass death, with the remains of animals big and small, young and old, entombed together in spectacular bonebeds. And so as we worked to uncover what killed the animals in MAD05-42, we also could not help but wonder why we find so many bonebeds here and why they are so beautifully preserved.

Opening a Very Cold Case

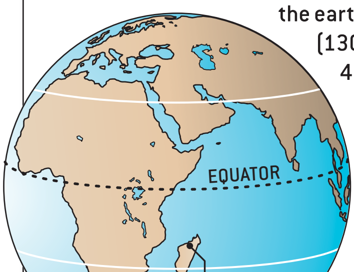
WE WERE MILLIONS OF YEARS too late to use most of the tools of modern-day coroners. To tease out the clues hidden in the bones and rock, we had to turn to geologic dating techniques and to the field of inquiry known as taphonomy, which explores the fate of organic remains as they cross from the living world to the dead.

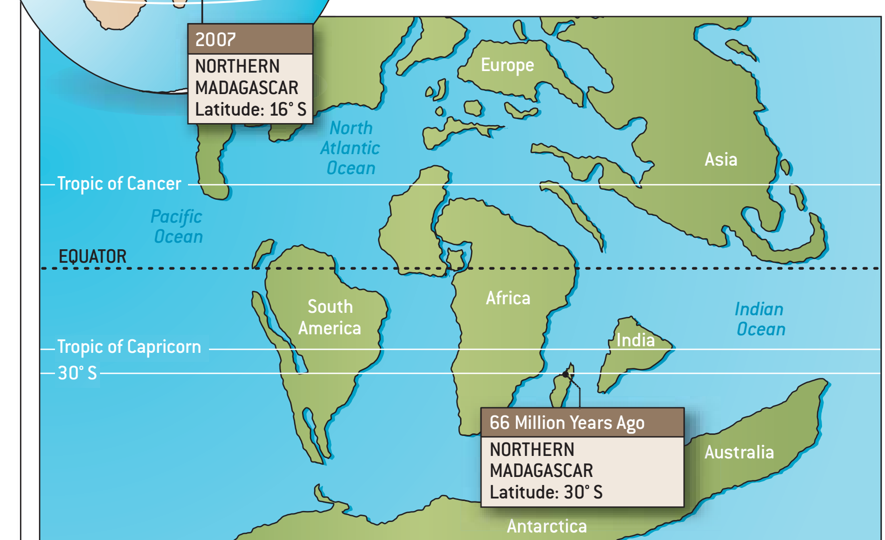
After naming the site, we disinterred the bones from the rocks in which they were embedded. We started with shovels and rock hammers to remove the overlying sediment, then moved on to dental picks and fine brushes to expose the bones themselves. We took great care not to damage the fragile bone surfaces. Once we fully exposed the skeletal remains, we mapped and photographed them exactly where we found them to record any significant spatial relationships. We next soaked the delicate bones with consolidating glues and carefully jacketed them in protective coats of burlap and plaster. After the plaster set, we catalogued the bones and packed them for the long journey to our laboratories in the U.S., where later we painstakingly removed any remaining sediment and studied the bones in detail, looking in particular for any marks on the surfaces that might reveal the killer's identity.

At the site we determined that the dead were preserved in a distinctive body of sedimentary rock known as the Maevarano Formation, situated a few tens of meters below rocks laid down at the Cretaceous/Tertiary boundary—the time, 65 million years ago, that all dinosaurs (apart from birds) and many other creatures suffered extinction on a global scale [see "Repeated Blows," by Luann Becker; SCIENTIFIC AMERICAN, March 2002, and "The Day the World Burned," by David A. Kring and Daniel D. Durda; SCIENTIFIC AMERICAN, December 2003]. The deathbed lay 44.5 meters beneath the mass extinction horizon and 14.5 meters beneath the local top of the Maevarano Formation. Measuring the radioactive decay of minerals from volcanic rocks in the layers below the formation yielded ages of approximately 88 million years. Marine sediments above and interbedded with the formation, laid down by seas that ebbed and flowed along the western shores of the island, contained seashells and tiny skeletons of single-celled microorganisms dated from other sites to near the end, but not the very end, of the Cretaceous period. All the temporal evidence thus indicates that the deaths occurred approximately 70 million years ago. Whatever killed the dinosaurs in quarry MAD05-42 was unrelated to the great global extinction that took place several million years later.

GEOLOGIC EVIDENCE

The history of the Madagascar landmass provides clues about what killed so many animals there 70 million years ago. At the beginning of the Mesozoic era (250 million years ago), Madagascar rested in the heart of Gondwana. Tectonic activity rearranged the earth's lithospheric plates, and by the early Cretaceous (130 million years ago), Madagascar was positioned about 400 kilometers away from the African mainland; it then rejoined the African plate and began moving northward. When the deaths under investigation occurred (*map below*), northern Madagascar lay near 30 degrees south latitude, where climatic conditions alternated between prolonged droughts and intense rainy periods.





LUCY READING-IKKANDA

More Mass Death

One of the first sites discovered, MAD93-18, reveals the recurrent nature of mass mortality in ancient Madagascar more dramatically than any other site. MAD93-18 has three discrete deathbeds stacked one on top of the other. Excavations there produced nearly complete skeletons of the large sauropod *Rapetosaurus* (*bottom left* and *bottom right*) as well as the skeletal remains of many other animals new to science, including a primitive bird, *Rahonavis ostromi* (*right*), which had small, fragile bones.

CATHERINE FORSTER *Stony Brook University (top)*; DESIRE RANDRIANARISTA *Centre ValBio (International Training Center for the Study of Biodiversity) (bottom left)*; RAYMOND R. ROGERS *(bottom right)*

Taphonomy advanced our inquiry as well. Taphonomic study examines bone modification (were the bones burned, broken or bitten?), carcass disturbance (dismemberment and selective removal of body parts by scavengers or predators), and burial history (how the bodies were buried and what happened to them after burial). The study of fossilization processes—essentially, what turns bone into stone—also falls within the realm of this science.

When we considered the dead in quarry MAD05-42 from a taphonomic perspective, we could tell that they perished over a prolonged period, perhaps weeks or months, because their corpses revealed variable postmortem histories. For example, some carcasses were largely intact and others were dismembered and scattered widely, which would not have happened instantaneously. In addition, some bones were in exquisite condition, whereas others showed evidence of advanced weathering and surface degradation. When the animals in an ancient bonebed have died at different times, we describe the site as "time-averaged" and use taphonomic clues to assess the amount of time between the first death and the last. Although we cannot determine exactly how much time transpired in the formation of this particular deathbed, we do know that death did not come at the same instant for the animals entombed here.

Scene of the Crime

THE GEOLOGIC HISTORY of the Madagascar landmass also provided important clues about what killed these dinosaurs. At the onset of the Mesozoic era (250 million years ago), Madagascar rested in the heart of Gondwana (the southern half of the supercontinent Pangaea), sandwiched between Africa and India, with Antarctica near its southern tip. Tectonic activity soon rearranged lithospheric plates on a massive scale, and by the late Jurassic (160 million years ago), Madagascar had rifted away from Africa and was moving southward, with India in tow. By the late Cretaceous (88 million years ago), Madagascar had resutured to the African plate, albeit

ANCIENT SOILS and rocks present convincing evidence of a semiarid climate during the late Cretaceous. Vertically oriented root casts (*lower right*) indicate that plants had adapted to dry conditions by seeking deep sources of moisture. Many of the relic roots also reveal clumps of calcium carbonate, which tend to form in arid regions. Other sediments, laid down in sandy rivers (*top photographs*), indicate that water flow varied dramatically. When the rivers were flowing, currents drove dunes downstream, resulting in stacked sets of inclined layering (*top right*). Dinosaurs and many other animals visited the shallow rivers in search of nourishment—in fact, the bones of quarry MAD05-42 lie in one of these ancient channels.

RAYMOND R. ROGERS

about 400 kilometers distant from the mainland, and the Indian subcontinent and Antarctica had rifted away, leaving the world's fourth largest island isolated in the Indian Ocean.

After reuniting with the African plate, Madagascar shifted northward toward its present location in the Southern Hemisphere tropics. But when the deaths under investigation transpired, some 70 million years ago, northern Madagascar was positioned near 30 degrees south latitude—still far from the Tropic of Capricorn (which now passes through *southern* Madagascar) and arguably well within the influence of subtropical weather patterns. Today major deserts and semideserts occur in swaths between 15 and 35 degrees north and south of the equator. These arid belts reflect large-scale atmospheric circulation patterns (known as the Hadley cell) that drive masses of hot, dry air down to earth after they have relinquished their moisture near the equator. The high-pressure zones that result from the descending air cells tend to keep rainfall at bay most of the time, but when rain does come, it can be intense.

Rocks of the Maevarano Formation present convincing evidence of a semiarid and seasonal climate during the late Cretaceous. Most telling are the red oxidized paleosols (ancient soils), which contain beautifully preserved, vertically oriented casts of roots. Vertical roots are common today where plants have adapted to dry conditions by seeking ever deeper sources of moisture and nutrients. Furthermore, many of the relic roots of the Maevarano Formation are encrusted with calcium carbonate or interspersed with irregular clumps of this mineral that are called carbonate nodules. In the modern world, oxidized soils enriched in calcium carbonate tend to occur in semiarid to arid regions where evaporation and transpiration limit the effects of precipitation.

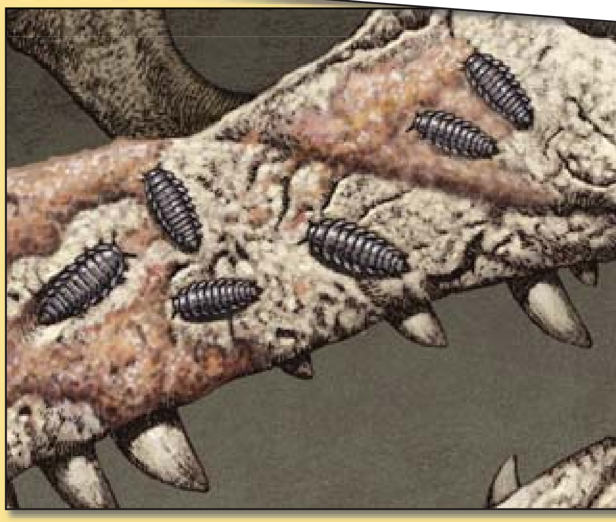
Other sediments of this ancient terrain were laid down in shallow, sandy rivers. These, too, provide telling evidence of a subtropical history, with clear indication that the flow of water fluctuated dramatically and probably seasonally. When the rivers were flowing, currents drove ripples and dunes downstream, resulting in stacked sets of inclined layering that geologists refer to as cross-stratification.

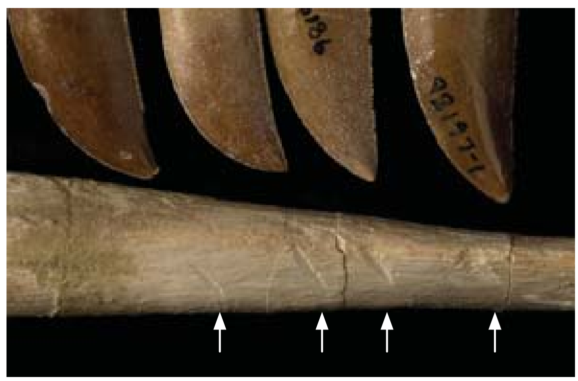
Dinosaurs and many other animals

THE AUTHORS

RAYMOND R. ROGERS and *DAVID W. KRAUSE* have excavated and investigated the spectacular bonebeds of Madagascar since 1996. Rogers is an associate professor and chair of the geology department at Macalester College and a research associate at the Field Museum in Chicago and the Science Museum of Minnesota. He earned a Ph.D. in geology from the University of Chicago in 1995. Krause is Distinguished Service Professor in the department of anatomical sciences at Stony Brook University and a research associate at the Field Museum. He received a doctorate in geology from the University of Michigan at Ann Arbor in 1982. Krause's work in Madagascar, the world's fourth poorest country, led him to establish the Madagascar Ankizy Fund (www.ankizy.org), a nonprofit organization that builds schools and provides temporary clinics for children in remote areas of the island.

OVAL PITS in the dinosaur bones (*photograph*) are signs that carrion beetles infested the corpses, fed on the flesh and then laid their eggs (*illustration below*).





TELLTALE MARKS on the bones, and the finer grooves within them (*arrows*), match exactly the size and spacing of the teeth, and the flesh-slicing serrations on the front and back edges, of the dinosaur *Majungatholus*.

Dining on the Dead

Widespread death in Madagascar during the late Cretaceous provided a smorgasbord for those that fed on the dead. Necrophagy, the act of consuming corpses or carrion, is a niche that must be filled if biological recycling is to proceed efficiently, and modern practitioners range from bacteria to large-bodied vertebrates. Along with our colleague Eric M. Roberts, who accompanied us to Madagascar a decade ago as an undergraduate (and is now a lecturer at the University of Witwatersrand in Johannesburg), we have found traces of necrophagous insect activity in the dinosaur bones of Madagascar: centimeter-long oval pits, usually in what had been spongy tissue inside the bones. These pits are signs that adult beetles infested the corpses, fed on the carrion and then laid their eggs nearby. After hatching, the larvae fed as well and used their robust mandibles to excavate the pits, which served as pupation chambers.

Insects were not the only creatures feeding on the dead. Analysis of bite marks has yielded irrefutable evidence that dinosaurs also dined here. Working in collaboration with Kristina Curry Rogers of the Science Museum of Minnesota, we have documented tooth marks of the seven-meter-long theropod *Majungatholus atopus* on a selection of bones from at least three separate bonebeds. Comparing the shape and size of the tooth marks with the jaws and teeth of various carnivores, we have been able to rule out all the other sharp-toothed suspects.

A few of the bitten bones in our sample belong to *Rapetosaurus,* a previously unknown long-necked sauropod dinosaur that Curry Rogers described as part of her dissertation work at Stony Brook University. Yet the vast majority of tooth-marked bones, primarily ribs and vertebrae, belong to *Majungatholus* itself. Cannibalism as an ecologic strategy is not at all uncommon among living animals, and it certainly should not be unexpected among dinosaurs. Unearthing the evidence to prove it, however, has been another matter, and in the bonebeds of Madagascar we have documented the only well-substantiated case of cannibalism among dinosaurs. Unfortunately, the bite-mark evidence does not definitively disclose whether *Majungatholus* actually killed the individuals it dined on—and thus practiced predation on its own species—or simply opportunistically scavenged their remains. *—R.R.R. and D.W.K.*

RAYMOND R. ROGERS (*top*); RAÚL MARTIN (*middle*); MARYLOU STEWART *Stony Brook University* (*bottom*)

no doubt frequented these rivers in search of water, nourishment or refuge. Indeed, the bones in quarry MAD05-42 are scattered in one of these ancient river channels. Part of the time the rivers were dry; at other times they roiled with thick milk shake–like slurries of mud and sand. We will return to these slurries—they play an important role in our story.

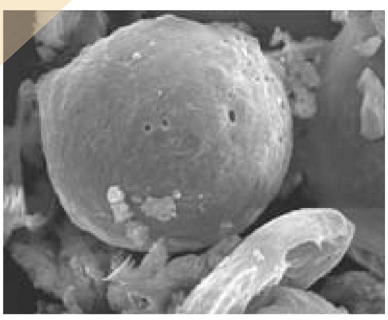
Identifying a Killer

A LONE ANIMAL can meet its end in many ways—perhaps too many, if the goal is to positively identify a killer in the fossil record. But the options diminish significantly in cases of mass death, such as those in the Maevarano Formation. For help in narrowing the possibilities down to one, we looked again to taphonomy. The bonebeds of Madagascar generally preserve the remains of more than one type of animal, be it multiple species of dinosaurs, as in MAD05-42, or a more diverse array, as in quarry MAD93-18 [*see box on page 45*], which has yielded the skeletal remains of fishes, turtles, snakes, crocodiles, three different types of nonavian dinosaurs, birds and mammals. This killer was indiscriminate, paying no attention to size,

DROUGHT WAS THE KILLER. Animals congregated in the parched riverbeds, where they perished as food and water disappeared. The large dinosaur *Majungatholus* fed on *Rapetosaurus* (*foreground*) as well as on its own kind (*far left*). Birds (*Rahonavis*) also scavenged the remains. The immediate cause of death could have been dehydration, heat stress, malnutrition, even poisoning, as the stagnant water turned toxic. Other carnivorous animals and insects feasted on the carcasses until they, too, met their end when torrential rains triggered mudflows that encased the dead and dying, preserving their remains for 70 million years.

age, taxonomy or habitat—a fact that tends to exclude a predator, such as a meat-eating dinosaur or crocodile, because modern predators generally show at least some degree of prey selection.

Nor is there any support for a disease-based scenario (it is difficult, though, to test for disease with fossilized bones). Because the creatures died at different times, we do not suspect dramatic instantaneous events such as earthquakes, floods or fire. Whatever killed the animals acted over time and struck victims individually after they had arrived at the river under their own volition. We also have irrefutable evidence that the killer struck repeatedly in different locales but with the same basic modus operandi. These animals were not dropped in their tracks during one bad day in the late Cretaceous—there were many bad days.



ALGAL SPORES found in the rocks encasing the bones suggest that the shrinking pools of water may have been toxic.

When all the evidence is brought to bear, we can confidently pinpoint one killer: drought. The opportunity certainly existed; this was a subtropical ecosystem with clear indications of aridity and seasonality. Moreover, we can see that animals congregated in the desiccating riverbeds, probably around remaining pools of water, where again and again they perished as good drinking water and nourishment disappeared. Today lethal droughts, especially in parts of Africa and interior Australia, prompt animals to gather around remaining resources. During an extended drought, thousands of animals can succumb at the site of their last hope for a drink, and their bodies may accrue in localized "dead zones" over several years.

Studies of modern drought-related mortality indicate that the unlucky animals preserved in the Maevarano bonebeds could have ultimately died from any number of causes: dehydration, heat stress, malnutrition, perhaps even poisoning as their dwindling water supplies turned foul and noxious. In fact, we have some tantalizing evidence that algal blooms occurred in the stagnant pools of water that drew the animals together. Michael Zavada, a Cretaceous pollen expert from East Tennessee State

RAÚL MARTIN (*top*); MICHAEL ZAVADA *East Tennessee State University* (*bottom*)

University, has isolated tiny algal spores in the rocks associated with the bones [*see bottom illustration on opposite page*]; whether these spores represent telltale clues of toxic algal blooms, however, has yet to be confirmed.

But how were the animals' bodies preserved, many of them so exquisitely? Biological remains tend to fare poorly at the ground surface, where scavengers hold sway and the sun slowly but inexorably bleaches even the largest of bones until they splinter and eventually turn to dust. When long-term preservation in the fossil record is at stake, burial should occur as soon as possible after death. Indeed, it could be argued that from a fossil's perspective, rapid burial is the single most critical key to immortality.

Fortunately for those of us who study these fossils, a very efficient undertaker was operating in conjunction with the killer weather. The drought conditions that periodically spelled disaster in the parched riverbeds eventually had to come to an end, and when the rains returned, as they did with a vengeance, they triggered debris flows. Viscous slurries of green mud and sand mobilized by rain-induced erosion poured over the bones and encased them. The sedimentary characteristics of the burial beds reflect a special category of fluid flow in which turbulence is suppressed, and both water and sediment move en masse in an essentially plastic fashion. This type of mass flow, often known as a mudflow, is not uncommon today. The lethal mudslides in Guatemala in 2005, unleashed by the torrential rains of Hurricane Stan, are a recent example.

Again and again after the fatal droughts had taken their toll, thick beds of mud and sand flowed over the bodies and the scattered bones, whether they belonged to animals that had died minutes or months before, and effectively packaged them all together in a protective and permanent sedimentary tomb. It would be another 70 million years until the tombs were cracked open and the amazing stories within revealed. SA

MORE TO EXPLORE

Monsters of Madagascar. John Flynn and David Krause in *National Geographic,* Vol. 198, No. 2, pages 44–57; August 2000.

Cannibalism in the Madagascan Dinosaur *Majungatholus atopus.* Raymond R. Rogers, David W. Krause and Kristina Curry Rogers in *Nature,* Vol. 422, pages 515–518; April 3, 2003.

The Natural History of Madagascar. Edited and translated by Steven M. Goodman and Jonathan P. Benstead. University of Chicago Press, 2004.

METHANE, Plants and

CLIMATE CHANGE

The surprising recent finding that living plants produce methane does not throw doubt on the cause of global warming. Human activities—not plants—are the source of the surge in this and other greenhouse gases

By Frank Keppler and Thomas Röckmann

What do you do as a scientist when you discover something that clearly contradicts the textbooks? The two of us faced this problem head-on when experiments we were running in 2005 showed that living vegetation produces the greenhouse gas methane. The established view held that only microbes that thrive without oxygen (anaerobic bacteria) can manufacture this gas. But our tests unexpectedly revealed that green plants also make methane—and quite a lot of it.

The first thing we did was look for errors in our experimental design and for every conceivable scenario that could have led us astray. Once we satisfied ourselves that our results were valid, though, we realized we had come across something very special, and we began to think about the consequences of our findings and how to present them to other researchers. Difficult as this discovery had been for us to accept, trying to convince our scientific peers and the public was almost impossible—in large part because we had to explain how such an important source of methane could have been overlooked for decades by the many able investigators studying methane and puzzling over climate change.

Natural Gas

MOST PEOPLE KNOW methane (often written as the chemical formula CH_4) as natural gas. Found in oil fields and coal beds as well as in natural gas fields, it has become an important source of energy and will most likely remain so given the limited reserves of oil on the planet. Approximately 600 million metric tons of it—both anthropogenic (from human activities) and natural—rise into the atmosphere every year. Most of these emissions have been thought to come from the decay of nonfossil organic material as a result of activity by anaerobic bacteria. Wetlands such as swamps, marshes and rice paddies provide the greatest share. Cattle, sheep and termites also make methane, as a by-product of anaerobic microbial digestion in their gut. Forest and savanna fires release methane, as does the combustion of fossil fuels [*see box on page 55*]. Over the years, researchers have gained considerable knowledge about the global methane cycle, and the consensus of the Intergovernmental Panel on Climate Change (IPCC) in 2001 was that the major sources had probably been identified (although the proportion each source contributes was still uncertain).

Nevertheless, some observations were difficult to explain. For instance, large fluctuations of atmospheric methane during the ice ages and warm ages, which have been reconstructed from air bubbles trapped in ice cores, remained a mystery. But no scientist in 2001 would have factored in direct emis-

KENN BROWN

sions of methane by plants, because no one suspected that biological production of methane by anything other than microbial anaerobic processes was possible.

Knowing the sources of methane and how much they emit is important because methane is an extremely efficient greenhouse gas. Much more carbon dioxide is spewed into the atmosphere every year, but one kilogram of methane warms the earth 23 times more than a kilogram of carbon dioxide does. As a result of human activities, the concentration of methane in the atmosphere has almost tripled over the past 150 years. Will it continue to increase into the 21st century? Can emissions be reduced? Climate scientists need to answer such questions, and to do so we must know the origin and fate of this important gas.

Startling Findings

THE IDEA OF INVESTIGATING plants as methane emitters grew out of research we had been conducting on chloromethane, a chlorinated gas that destroys ozone and was thought to come mainly from the oceans and forest fires. A few years ago, while working at the Department of Agriculture and Food Science in Northern Ireland, we discovered that aging plants provide most of the chloromethane found in the atmosphere. Because methane, like chloromethane, is released during the burning of biomass, we wondered whether intact plants might also release methane.

To satisfy our curiosity, we collected 30 different kinds of tree leaves and grasses from tropical and temperate regions and placed them in small chambers with typical concentrations of atmospheric oxygen. To our amazement, all of the various kinds of leaves and plant litter produced methane. Usually a gram of dried plant material releases between 0.2 and three nanograms (one billionth of a gram) of methane an hour. These relatively tiny amounts were difficult to monitor, even using our highly sensitive state-of-the-art equipment.

We collected 30 different kinds of tree leaves and grasses. To our AMAZEMENT all of the various kinds of leaves and plant litter produced METHANE.

The task was made still more challenging because we had to differentiate between methane produced by plant tissue and the high background levels normally present in ambient air. We believe this difficulty is what prevented biologists from observing the phenomenon earlier. The secret to our discovery was that we removed the interfering effect of the natural methane background by flushing the chambers with methane-free air before the start of each experiment. We were then able to measure the methane released by plant tissue.

Our curiosity fueled, we undertook similar experiments with living plants [*see box on page 56*], and we found that the rates of methane production increased dramatically, jumping to 10 to 100 times those of leaves detached from plants. By running a series of experiments, we excluded the possibility that bacteria that thrive without oxygen produced the methane. Finally, we were absolutely convinced that living plants release methane in significant quantities. We could provide no immediate answers about the mechanism of how they did this, although we suspect that pectin, a substance in the walls of the plant cells, is involved. We decided that this question would have to await further research, which is currently under way. Because of methane's role in climate change, however, we realized it was crucial to begin to take into account the quantity of gas released into the atmosphere by this newly discovered source.

How much might plants be contributing to the planet's methane totals? It was immediately obvious to us that even though a single leaf or plant made only tiny amounts of methane, these small bits would add up quickly because plants cover a substantial part of the globe. We were nonetheless astounded by the figure generated by our calculations: between 60 million and 240 million metric tons of methane come from plants every year—this constitutes 10 to 40 percent of annual global emissions. Most of it, about two thirds, originates in the vegetation-rich tropics. We knew, of course, that extrapolating global estimates from a limited sample of laboratory measurements was open to error. Still, the final number seemed extremely large—and if it surprised us, it would be heresy to many of our scientific peers.

Fortunately for us, support for our work soon came from an unexpected source. A group of environmental physicists in Heidelberg, Germany, was observing the earth's atmosphere from space. In 2005 the scientists' satellite measurements revealed "clouds" of methane over tropical forests [*see illustration on page 57*]. They reported that their observa-

Overview/*Nature's Surprise*

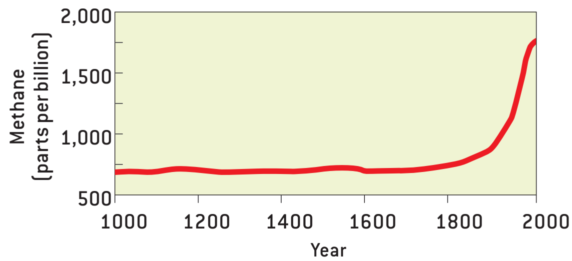
- The established view has been that methane (natural gas) is produced by microbes that thrive without oxygen, but experiments by the authors' team unexpectedly revealed that living plants also manufacture this potent greenhouse gas.
- Although this startling finding can explain many previously puzzling observations, a number of scientists are still skeptical, in particular about the amount of methane that plants generate. Knowing the sources of methane and how much they emit is important because of methane's role in trapping heat.
- An early misinterpretation of the finding suggested that forests might actually be contributing to global warming, but the authors emphasize that plants do not contribute to the recent increase in methane and global warming.

THE TEXTBOOK VIEW

In the past 150 years, methane emissions into the atmosphere have roughly tripled (*graph*), and today some 600 million metric tons are sent into the air annually. That rise is a concern because methane, like carbon dioxide, traps heat in the earth's atmosphere and therefore contributes to global warming.

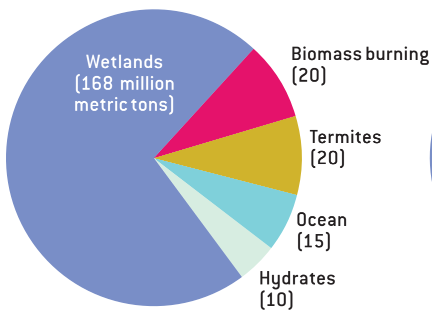
Until the authors and their colleagues published their recent discoveries, traditional thinking held that all natural releases of methane resulted from the activity of bacteria that thrive in wet, oxygen-poor environments. Such environments include swamps and rice paddies as well as the digestive systems of termites and ruminants. And analyses of the sources of the gas in the environment (*pie charts*) indicated that the dramatic rise in methane concentrations since the mid-1800s has stemmed from human industrial activities (such as the use of fossil fuels for energy) and increased rice cultivation and breeding of ruminants (because of population growth). The authors' work casts no doubt on the explanation for why methane concentrations in the atmosphere have increased, but estimates of the relative contributions to methane levels from natural sources will have to be revised.

Methane Concentration in the Atmosphere

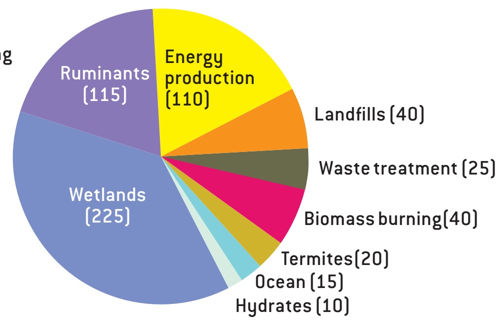


Methane Emissions

Preindustrial times: 233 million metric tons a year



Today: 600 million metric tons a year



LAURIE GRACE AND JANA BRENNING

tions could not be explained by simply using the current understanding of the global methane budget. In light of our findings, however, their work made sense: green vegetation was the source of the methane clouds.

Recently further support has come from Paul J. Crutzen, a 1995 Nobel Prize winner, and his colleagues. After our findings were published in January 2006, they reanalyzed measurements made in 1988 of air samples from the Venezuelan savanna and concluded that 30 million to 60 million metric tons of methane could be released from vegetation in these regions. Crutzen said that "looking back to 1988, we could have made the discovery, but accepting the general wisdom that methane can only be produced under anaerobic conditions, we missed the boat."

Despite this support for our work, many scientists are still skeptical about methane emissions from plants, especially about our estimate of how much methane comes from vegetation. A number of our scientific colleagues are therefore recalculating the budget for the plant source, using different methods from ours but applying our emission rates. Of course, we keenly await an independent verification of our laboratory findings.

THE AUTHORS

FRANK KEPPLER and *THOMAS RÖCKMANN* first discovered methane emissions from plants when they were working together at the Max Planck Institute for Nuclear Physics in Heidelberg, Germany. Keppler earned a Ph.D. in environmental geochemistry from the University of Heidelberg in 2000. He recently received a European Young Investigator Award (EURYI) to build his own research group at the Max Planck Institute for Chemistry in Mainz. Röckmann received his Ph.D. from the University of Heidelberg. In 2005 he was appointed full professor at the Institute for Marine and Atmospheric Research Utrecht in the Netherlands.

Solving an Old Puzzle

OUR FINDINGS WOULD EXPLAIN a trend that has puzzled climate scientists for years: fluctuations in methane levels in parallel with changes in global temperatures. Ice cores serve as natural archives that store information about atmospheric composition and climate variability going back almost a million years. Tiny bubbles of air trapped in the ice reveal the relative concentrations of atmospheric gases in the past [*see box on next page*]. We see in the ice cores, for example, that variations of past carbon dioxide levels are closely linked to changes in global temperatures. During ice ages, carbon dioxide concentrations are low; during warm spells, levels increase.

In general, methane concentrations follow the same trend as carbon dioxide, but the reason has been unclear. Scientists have tried to use models of wetlands (the only major natural source of methane previously believed to exist) to reconstruct the curious variations of past methane levels. Yet they found it difficult to reproduce the reported differences in atmospheric methane levels between glacial and interglacial periods.

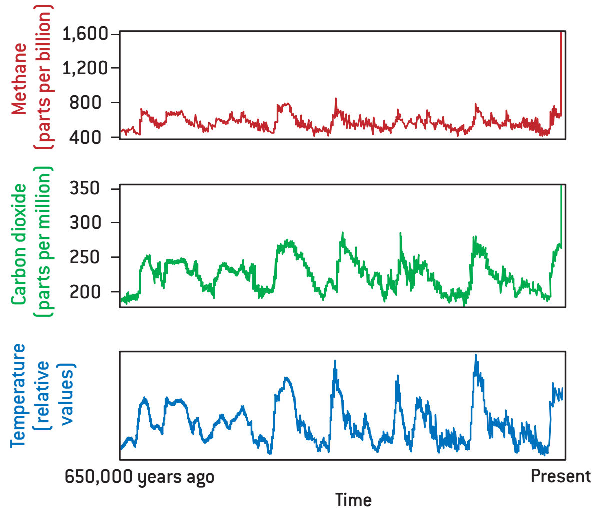
THE NEW VIEW

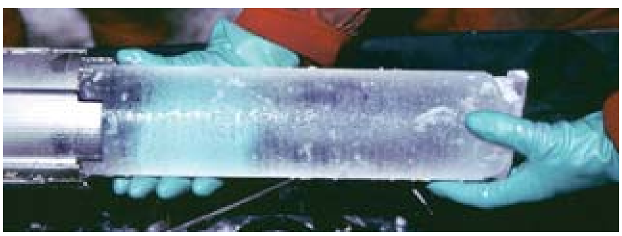
The authors' team scrutinized the gases emitted by plant debris and by living plants. To their surprise, the scientists found that both plant debris and growing vegetation produce methane. This important source of emissions had been overlooked until the team performed experiments in chambers that had been flushed of methane, which allowed the researchers to measure the minute amounts of the gas that plants give off.

The new view could explain puzzling fluctuations in methane levels that mirror changes in levels of carbon dioxide and in global temperatures (*graphs*). Scientists have tracked these changes by studying ice cores, in which trapped bubbles preserve information about the composition of the atmosphere going back almost a million years; concentrations of deuterium in the ice provide information about temperature. High atmospheric carbon dioxide concentrations and rising temperatures most likely led to a large increase in vegetation, which could have been accompanied by correspondingly large releases of methane.

AUTHORS' EXPERIMENT detected minuscule quantities of methane produced by living vegetation (*rye grass in photograph*).

Parallel Patterns



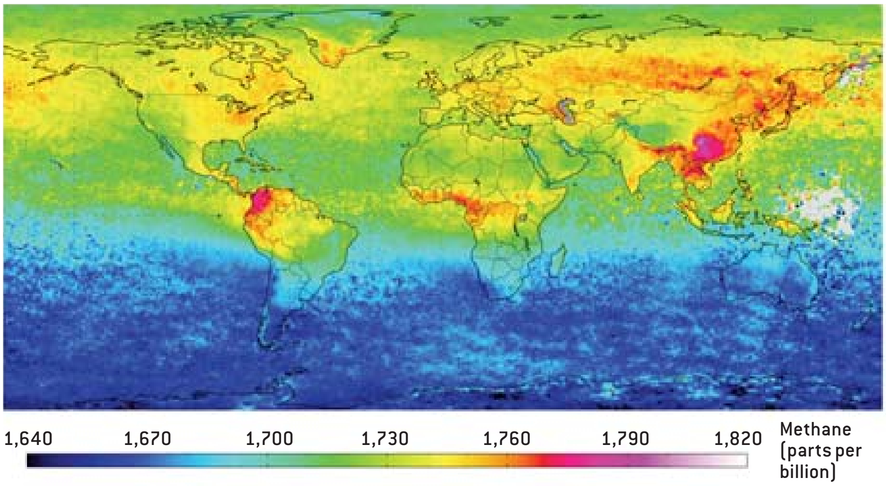


ICE CORE (*far left*) contains bubbles that reveal the composition of the ancient atmosphere. The gas bubbles in the micrograph of a thin cut (*left*) are dark in color and one to three millimeters across.

Another explanation that has been suggested involves the gas in a form known as methane hydrates [see "Flammable Ice," by Erwin Suess, Gerhard Bohrmann, Jens Greinert and Erwin Lausch; Scientific American, November 1999]. These develop at high pressure, such as that found on the ocean floor. An unknown but possibly very large quantity of methane is trapped in this form in ocean sediments. The sudden release of large volumes of methane from these sediments into the atmosphere has been suggested as a possible cause for rapid global warming events in the earth's distant past. Yet recent results from polar ice core studies show that marine methane hydrates were stable at least over the past 40,000 years, indicating that they were not involved in the abrupt increases of atmospheric methane during the last glacial cycle.

We know that terrestrial vegetation is very sensitive to environmental changes, and thus the total amount of vegetation on the planet varies as the climate cools down and warms up during glacial cycles. In light of our findings, such variations should now be seriously considered as a possible cause of declines in methane levels during glacial periods and rises during the interglacials. During the last glacial maximum—around 21,000 years ago—the plant growth of the Amazon forests was only half as extensive as today, and tropical vegetation might thus have released much less methane. Since that time, global surface temperature and carbon dioxide concentrations have risen, leading to enhanced plant growth and, we would expect, to more and more methane released from vegetation.

Similar climate scenarios may have occurred during other periods of the earth's history, particularly at mass extinction events, such as the Permian-Triassic boundary (250 million years ago) and the Triassic-Jurassic boundary (200 million years ago). Extremely high atmospheric carbon dioxide concentrations as well as rising temperatures could have resulted in a dramatic increase in vegetation biomass. Such global warming periods could have been accompanied by a massive release of methane from vegetation and by more heating. Though speculative, the assumption that emissions may have been as much as 10 times higher than at present is not totally unreasonable. If this is so, methane emissions from vegetation, in addition to emissions of the gas from wetlands and perhaps from the seafloor, could be envisaged as a driving force in historic climate change.



SATELLITE IMAGES of the earth's atmosphere provided support for the authors' controversial finding. In 2005 environmental physicists observed clouds of methane over tropical forests. Although the standard model of methane production cannot explain this observation, the authors' discovery made sense of the curious clouds: the abundant green vegetation of the tropics was emitting the methane.

FROM "SATELLITE CHARTOGRAPHY OF ATMOSPHERIC METHANE FROM SCIAMACHY ON BOARD ENVISAT: ANALYSIS OF THE YEARS 2003 AND 2004," BY C. FRANKENBERG ET AL., IN *JOURNAL OF GEOPHYSICAL RESEARCH*, VOL. 111; 2006

Media Misinterpretations

WHEN YOU SEE A REPORT on your scientific work on the BBC World News immediately following news about bird flu and the situation in Iraq, on the very day your work has first been published, you realize that you have found something with great societal relevance. This realization was reinforced the next day as our research appeared in newspapers around the world, often in front-page headlines.

Unfortunately, extensive media coverage can lead to exaggerations, and in our case it resulted in the misinterpretation of our results. In particular, many reports claimed that plants may be responsible for global warming; in one instance, we saw the headline "Global Warming—Blame the Forests" on the front page of a reputable newspaper.

When you then receive many e-mails and phone calls from individuals asking whether they should cut down all the trees in their garden to fight global warming, you realize that something has gone badly wrong in the communication to the public. We felt compelled to issue another press release to address the misinterpretations.

In our second press release we emphasized that if our finding is true, plants have been emitting methane into the atmosphere for hundreds of millions of years. Those emissions have contributed to the natural greenhouse effect, without which life as we know it would not be possible. Plants are not responsible, however, for the dramatic increase in methane concentrations since the start of industrialization. This surge was brought about by human activities.

Our discovery also led to intense speculation that methane emissions by plants could diminish or even outweigh the carbon storage effect of reforestation programs. If that were correct, it would have important implications for countries attempting to implement the Kyoto Protocol to minimize global carbon emissions, because, under the protocol, tree-planting programs can be used in national carbon dioxide mitigation strategies. But our calculations show that the climatic benefits gained by establishing new forests to absorb carbon dioxide would far exceed the relatively small negative effect of adding more methane to the atmosphere (which may reduce the overall carbon uptake of the trees by 4 percent at most). The potential for reducing global warming by planting trees is most definitely positive.

In the heat of this debate, people forgot a crucial fact: plants are the green lung of our planet—they provide the oxygen that makes life as we know it possible. They perform many other beneficial tasks as well. As just two crucial examples, they provide a natural environment that fosters biodiversity, and they control the tropical water cycle. The problem is not the plants; it is the global large-scale burning of fossil fuels.

A more legitimate concern is whether the methane produced by vegetation can have an impact on climate in the near future. Although plants are not responsible for the massive increase of methane in the atmosphere since preindustrial times, they do tend to grow faster. As we can expect methane emissions from vegetation to increase with temperature, this would lead to even more warming. This vicious cycle would be a natural phenomenon except for its speed, which is accelerated mainly by anthropogenic activities such as burning fossil fuels. The large plant feedback to global climate change that most likely happened in the past, however, is probably unlikely today because so many forests have been cut down.

Although it is too early to say exactly how our revelation might influence predictions for climate change in the more distant future, it is clear that all new assessments should consider emissions of methane by plants. SA

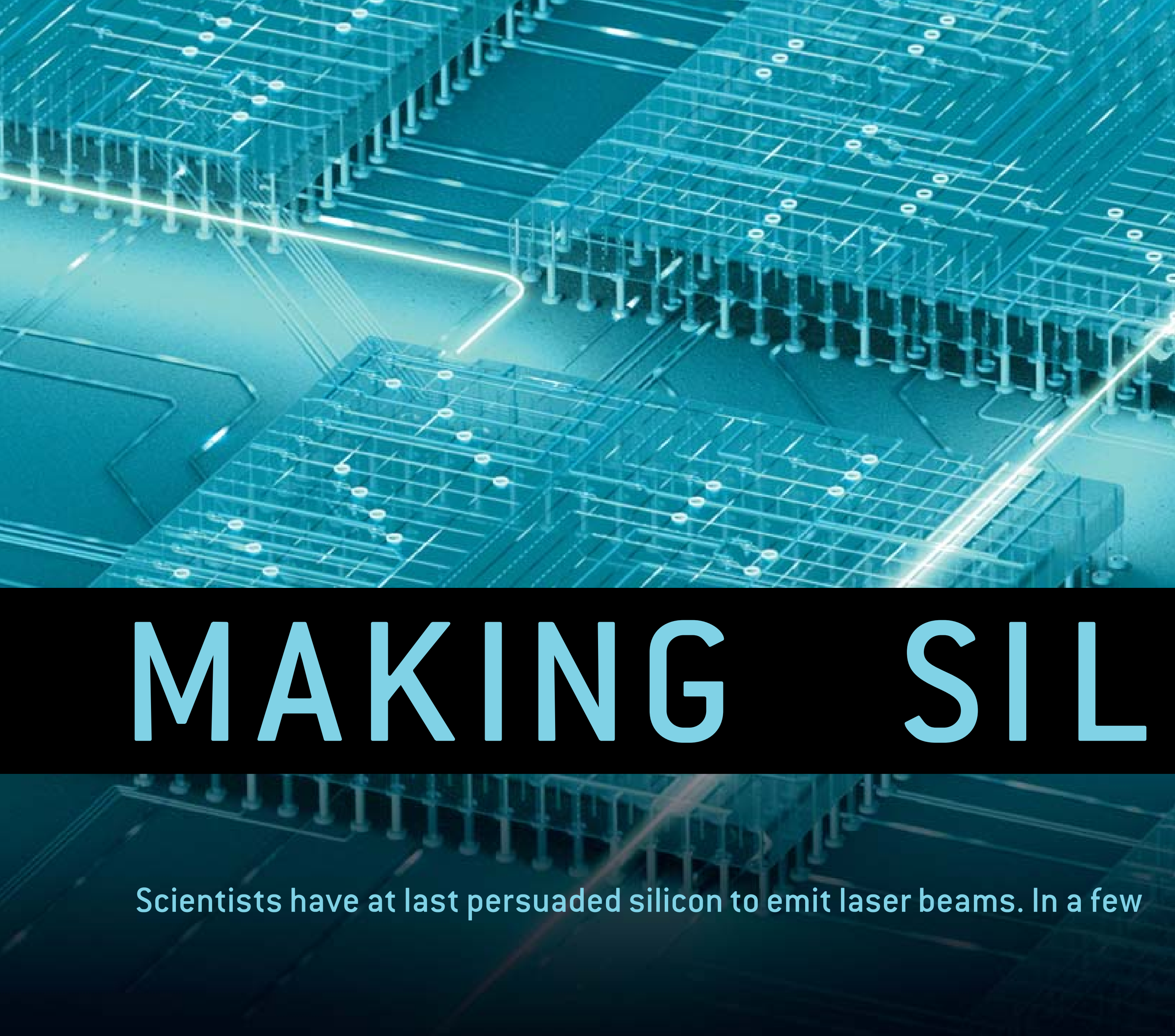
MORE TO EXPLORE

The Changing Atmosphere. Thomas E. Graedel and Paul J. Crutzen in *Scientific American,* Vol. 261, No. 3, pages 58–68; September 1989.

Climate Change 2001: The Scientific Basis. Edited by J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell and C. A. Johnson. Cambridge University Press, 2001. Available online at www.ipcc.ch/

Methane Emissions from Terrestrial Plants under Aerobic Conditions. Frank Keppler, John T. G. Hamilton, Marc Brass and Thomas Röckmann in *Nature,* Vol. 439, pages 187–191; January 12, 2006.

Methane Finding Baffles Scientists. Quirin Schiermeier. Ibid., page 128.



MAKING SIL

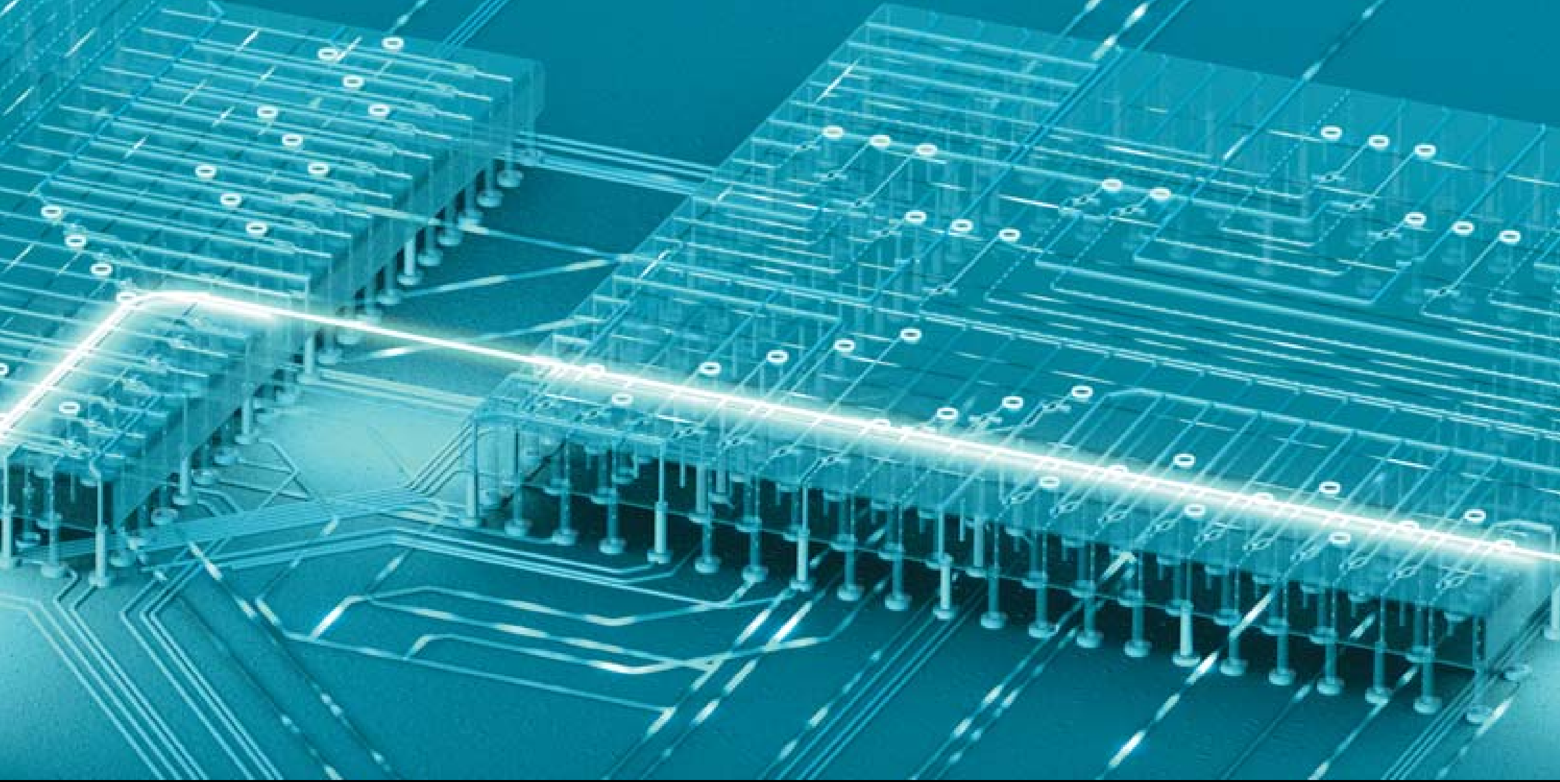
Scientists have at last persuaded silicon to emit laser beams. In a few

LOW-COST SILICON CHIPS ENABLE ENGINEERS to manipulate streams of electrons so that they can perform the myriad functions that make our computers, cellular phones and other consumer electronics so useful. If integrated silicon circuits could similarly create and control beams of light, they could make possible a range of inexpensive new technologies suited to many other applications. But for decades, silicon's very nature has thwarted scientists' dogged efforts to transform the material into a source of the necessary concentrated light.

Now several research groups, starting with mine, have coaxed silicon to produce laser light. The advance could have enormous implications for electronic devices that incorporate lasers and optical amplifiers, which currently rely on lasing materials that are far more costly and less common than silicon.

Replacing traditional copper interconnects and cables with optical conduits could raise data-trans-

U.C.L.A. (*laser glow*); BRYAN CHRISTIE DESIGN



IC N LASE

years, computers and other devices will handle light as well as electrons

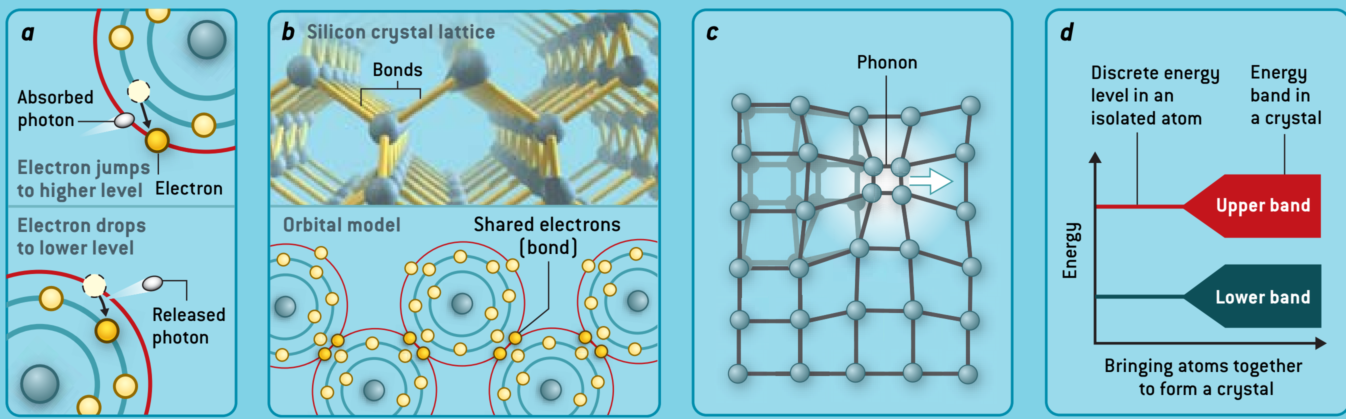
BY BAHRAM JALALI

fer speed limits by orders of magnitude beyond the capabilities of current technology. For example, cable modems, the workhorses of household Internet connections, are currently limited to a transfer rate of about one megabyte a second. Optical devices built on silicon chips could effortlessly transfer oversize digital files, such as high-definition videos, at rates as high as 10 gigabits a second—a 10,000-fold improvement. Compact sensors that include integrated circuits with silicon lasers could combine the capabilities of a diagnostic lab-on-a-chip with wireless communications to detect pollutants, chemical warfare agents or explosives as parts of extensive environmental monitoring and security networks. In a promising military applica-

LASER BEAM (*red and white glow in title*) was the first ever produced by a silicon device; its infrared light is invisible to the eye and is false-colored here. Silicon lasers that are integrated into microchips (*background*) could make low-cost computing with light practical.

THE CHALLENGE OF GETTING SILICON TO LASE

Silicon offers enormous promise for low-cost computing with light, but its very nature makes it an unlikely lasing medium.



The lasing process is based on the quantum behavior of electrons in the outer orbitals of the atoms in a suitable material. An outer-shell electron in a single atom is energized (or "pumped") after absorbing a photon—an elementary quantum unit of light—which raises it to a higher orbit and energy level (*a*). An energized electron releases a photon when it drops down to a lower level.

In a solid, atoms form bonds by sharing these outer electrons (*b*). To achieve light amplification, the prerequisite to lasing, external energy sources pump the shared electrons to higher energy levels. As the energized electrons release photons, those photons, in turn, stimulate further photon emissions—amplifying light. Photons can also be amplified when they collide with excited phonons—which are quantized atomic vibrations of the crystal lattice (*c*).

When single atoms bond together into crystals, the character of the shared electrons' energy levels changes to broader bands (*d*) because of the effect on the electromagnetic environment of many nearby atoms. Thus, a pumped electron in a crystal jumps from one band to another.

When the energies and momenta of electrons in a common lasing medium such as gallium arsenide are plotted on a graph, the energy bands line up vertically because they share the same momenta (*e*). (A band describes the possible quantum states that electrons can take; each state has a quantity that can be identified with a classical momentum and that must be conserved during collisions.) The bands in silicon, in contrast,

tion, silicon lasers might be able to mislead the infrared sensors of heat-seeking antiaircraft missiles to provide a cost-effective countermeasure against them.

Laser Primer

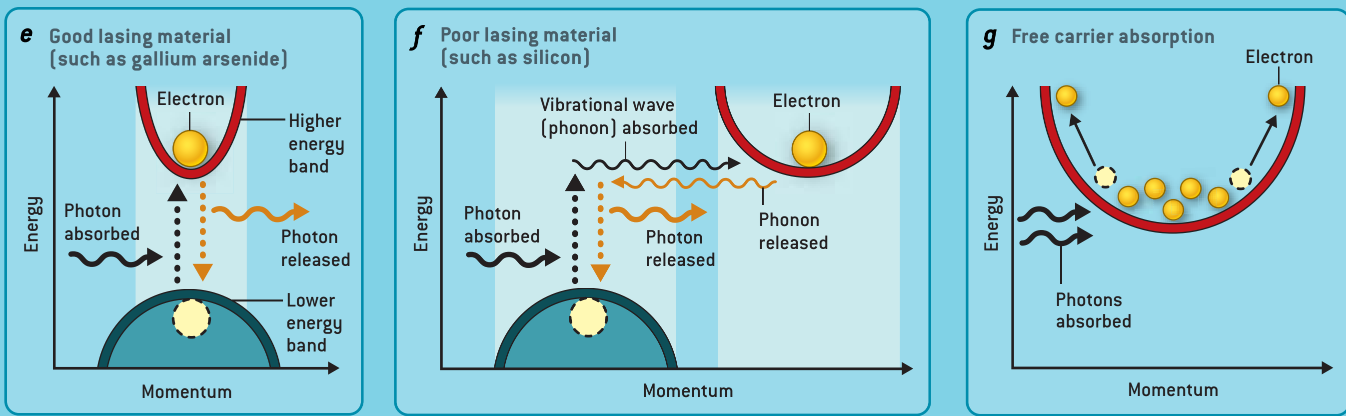
WHY HAS IT TAKEN so long to teach silicon this new trick? Unlike materials commonly used as host media for lasing, such as the gallium arsenide used in DVD players, silicon is not naturally organized to promote the two-step process that results in a coherent light beam. It cannot emit light efficiently when energized (the first requirement), and whatever light it does produce is incapable of amplifying light into a laser beam by "stimulating" it to create more photons. ("Laser" stands for *l*ight *a*mplification by *s*timulated *e*mission of *r*adiation.)

In a laser, an external energy source, usually light or an electric current, "pumps" the electrons in the atoms of the host medium to a higher energy level, which physicists call an upper (or excited) state. When these atoms return to their normal (ground) state, the extra energy is released as photons of light (elementary quantum units of electromagnetic radiation that exist simultaneously as a wave and a particle). Albert Einstein dubbed this process "spontaneous emission": a phenomenon that yields photons that travel outward in all directions at random, giving rise to low-intensity, diffused light—much like that cast by a fluorescent bulb. When one of these emitted photons passes through a group of previously pumped electrons in the host material, it triggers, or stimulates, the electrons to discharge their surplus energy all at once, a concept first proposed in a paper published by Einstein in 1917. The resulting photons travel together in the same direction in synchrony, forming a highly directional light beam. As the beam travels through other excited atoms in the medium, its photons, in turn, stimulate the emission of yet more photons in a cascade. The effect

Overview/*Silicon Lasers*

- Scientists have long sought a silicon chip that can handle light as deftly as it does electrons, but silicon does not produce light easily, especially concentrated laser light. Such an advance could lead to ultrafast digital data transfers, new sensor networks and many other innovations.
- After many years of work, researchers are finally making silicon lase by using several different materials-based techniques. The birth of a novel hybrid technology—silicon electrophotonics—is at hand.

JEN CHRISTIANSEN (*this page and opposite page*); KEN EWARD *BioGrafx* (*b*)



have different momenta, which means that energy from an absorbed photon alone is not enough to cause an electron to jump to a higher band (*f*). Instead the electron must wait until a phonon with the right extra momentum shows up to broker the transfer of energy. Unfortunately, these electrons often lose their surplus energy as heat before a suitable phonon arrives, which leads silicon to emit light inefficiently.

Silicon's low emission efficiency allows an intraband phenomenon called free carrier absorption to obstruct light amplification and lasing. When a passing photon interacts with an energized electron (a free carrier) in an upper band, one of two competing processes can occur. Either the photon can stimulate the emission of another photon, causing the electron to fall to a lower band, or the electron can simply absorb the photon, which merely moves it higher in the same band (*g*), an event that does not produce another photon and thus does not promote light amplification and lasing.

Gallium arsenide's upper bands hold comparatively few electrons. When depicted on a graph, the upper band is narrow, with steeply sloped sides. Because gallium arsenide has a high emission rate (it amplifies light efficiently because its bands line up), its photon emissions easily outpace its absorptions, and thus the material amplifies light. The wider, less-steep-sided upper bands in silicon require more electrons to fill them up. Silicon, with its low emission rate (caused by its indirect lineup) and high free carrier absorption rate, is unable to amplify light.

is analogous to the way the mass of an avalanche grows as it courses down a snow-covered mountainside.

Einstein's prediction of stimulated emission did not garner much interest until the 1950s, when physicists began to realize its potential applications in optical or photonic devices. In 1958 Charles Townes and Arthur Schawlow proposed partially surrounding a light-amplifying material with mirrors to reflect some of the photons it generated back inside. They showed that the stimulation process then would feed on itself (as in a chain reaction). This approach, once fully developed, could create a powerful light flow with a well-defined wavelength—a laser beam. Just two years later Theodore Maiman demonstrated the first operational laser, made by optically pumping a ruby crystal with a powerful lamp.

Silicon has proved considerably less pliant than ruby crystals or other subsequently developed lasing media. In semiconductors—materials whose electrical performance lies midway between an excellent conductor such as copper and an insulator such as rubber or certain ceramics—electrons exist in energy bands, which are ranges of energy levels, or states, that electrons can occupy.

The energy band describes the range of levels that electrons are "forbidden" or "allowed" to inhabit according to quantum theory; the gap between permitted bands (the band gap) is a range of energy levels that the electron is forbidden to occupy. An electron in an atom's outer orbital can take on energy (letting it jump to a higher band) by absorbing a photon or can release energy (dropping back down) by emitting one. Physicists classify these interactions as a kind of scattering event.

Imagine the energy bands as a series of buckets in which electrons sit [*see box above*]. Normally, almost all the electrons remain in the lower energy band, or bucket, leaving the upper band nearly empty. But if a photon with energy equal to or greater than the band gap collides with an electron, it can raise the electron to the upper band. The electron hops from the lower bucket to a higher one. Light absorption, the name for this effect, is the basis for the way solar cells convert light into electricity.

For the material to produce photons, it must receive enough energy to pump many electrons from the lower band to the upper band, causing a so-called population inversion (compared with the usual distributions present in the bands). It is unnecessary to pump the entire electron population; only the part near the top of the lower band need be affected. Engineers often excite electrons directly by forcing electric current through a semiconductor diode. Illuminating the substance with an external light source, as Maiman did, can also pump electrons.

AS PHOTONS TRAVEL THROUGH EXCITED ATOMS, THEY STIMULATE THE EMISSION OF YET MORE PHOTONS IN A CASCADE.

The electrons in the upper band eventually release energy, thereby emitting photons. When the resulting photons speed through a semiconductor that has many electrons in the upper level (an inverted electron population), they stimulate emissions of still other photons. In the best case, semiconductor emissions match the energy absorbed.

Although electrons and photons trade energy in these emission and absorption (scattering) processes, the system's total energy is conserved—that is, energy credits equal energy debits, as required by the law of conservation of energy. But absorption and emission will occur only if momentum is also conserved, according to the law of conservation of momentum. Momentum, which for a photon traveling (as a wave) in a crystal is determined directly from its wavelength, can be thought of as a tendency for a photon to continue to travel in the same direction. Being packets of pure energy, photons do not have much momentum to contribute in scattering collisions, so the transfers work best when the upper and lower bands (the starting and ending points of the interband transactions) have the same momentum. This equality of momentum occurs in commonly used lasing materials such as gallium arsenide and indium phosphide, whose energy bands lie directly on top of each other when plotted on a graph comparing energy and momentum. Such a direct lineup allows energy to be traded directly between an electron and a photon [*see box on preceding two pages*]. Whether a substance has this so-called direct lineup is intrinsic to the arrangement of atoms in the material's crystal lattice.

Silicon, however, naturally has an indirect lineup as a result of a less than optimal atomic crystal structure—that is, the material suffers from a large difference in the momentum between its upper and lower bands. (A band describes the possible quantum states that electrons can take; each state has a quantity that can be identified with a classical momentum and that must be conserved during collisions.) Thus, electrons cannot easily exchange energy with a photon and still conserve momentum. Instead they must wait until a vibrational wave of the silicon's crystal lattice, called a phonon, with just the right momentum shows up to provide the necessary extra momentum to broker the transfer of energy. Unfortunately, the electrons in silicon often lose their pump energy as heat while waiting for a suitable phonon to arrive. As a result, silicon exhibits a low emission efficiency; only about one excited electron in a million will successfully release a photon. Common lasing media such as gallium arsenide, in comparison, feature emission efficiencies some 10,000 times larger.

An indirect band gap limits the efficiency of a silicon laser, but it does not prohibit lasing by itself. Two other factors, also intrinsic to silicon, are involved. The first is free carrier absorption, a process that happens within a given energy band. Imagine a group of electrons (the free carriers) that have been pumped to an upper band. When a passing photon interacts with an energized electron, two events can occur—one favorable, the other unfavorable. The photon can cause the electron to drop down to a lower band by stimulating emission of another photon, which feeds the process of light amplification. Or the photon can be absorbed by the electron, which then merely moves higher in the upper band, a process that fails to generate another photon and thus does not result in amplification of light. The rates at which these two competing effects occur depend on the number of pumped electrons that reside in the upper energy band.

The upper bands (or buckets) in good lasing materials such as gallium arsenide are narrow and have steep sides, so they tend to hold relatively few electrons. In contrast, silicon features wider, less-steep-sided upper bands that require more electrons to fill up. When pumped, silicon has a large tendency to support free carrier absorption. Because gallium arsenide has a high emission rate (it amplifies light efficiently because its bands line up), its total photon emissions easily outpace its absorptions. Silicon, with its low emission rate (caused by its indirect lineup) and high free carrier absorption rate, is unable to amplify light.

An esoteric process known as Auger recombination also impedes silicon lasing. In this phenomenon, rather than emitting light, an electron in the upper band loses its energy to other electrons that subsequently give up the excess energy as heat. The amount of wasted light energy depends on the number of electrons present in the upper band. Silicon undergoes more Auger recombination than does gallium arsenide because it needs more electrons to be pumped into the upper band to overcome its low light-emission efficiency.

Teaching Silicon to Lase

IN THE PAST FIVE YEARS, researchers have begun to find ways around silicon's built-in hurdles. One method to enhance light emission takes advantage of an intriguing phenomenon called quantum confinement, which occurs when electron movement is constrained in one or more directions. In a three-dimensional restriction, called a quantum cage, an electron becomes agitated when the cage's size shrinks. This effect occurs as a result of the Heisenberg uncertainty principle, which states that localizing an electron makes its velocity, and hence its momentum (which equals mass times velocity), more random. This condition effectively relaxes the momentum con-

THE AUTHOR

BAHRAM JALALI is professor of electrical engineering at the Henry Samueli School of Engineering and Applied Science at the University of California, Los Angeles. He is also trustee of the California Science Center. Jalali spends much of his leisure time kayaking and sailing with his three children off the southern coast of California.

servation requirement governing electron-photon energy transfer, which boosts the rate of light emission by the semiconductor.

To make a quantum cage for silicon, researchers can create a thin film of silica (silicon dioxide) glass with tiny pieces of crystalline silicon embedded in it. These nanocrystals, which can be pumped by illuminating them with an external light source, are only a few atoms wide, so they can achieve quantum confinement. In 2000 Lorenzo Pavesi's group at the University of Trento in Italy first reported evidence of optically amplified silicon nanocrystals. The physics community initially greeted the result with skepticism, but Philip Fauchet of the University of Rochester and others subsequently confirmed the effect. Although this approach has yet to produce a laser, it inspired other innovations that have yielded encouraging results.

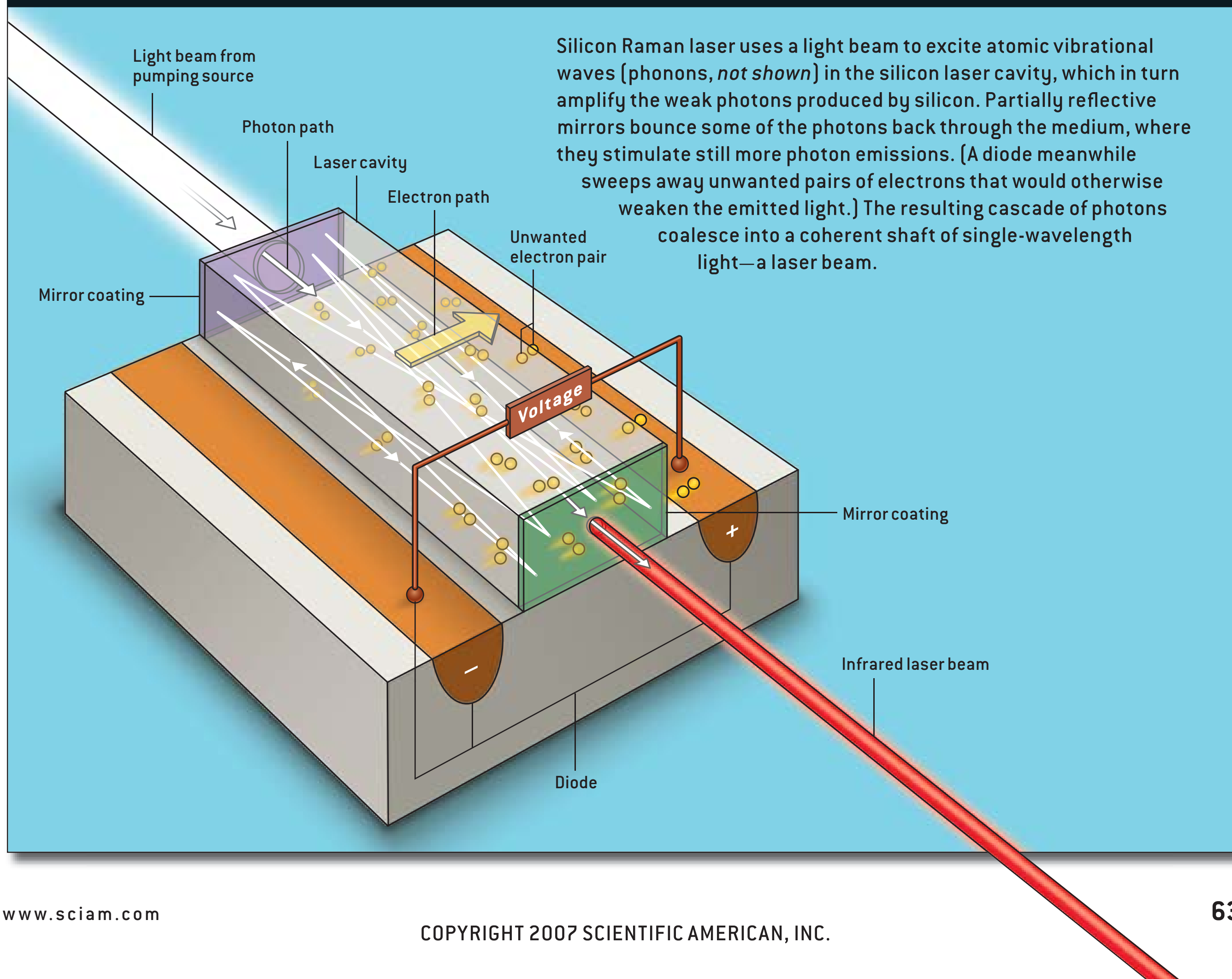
One advance that exploits quantum confinement makes use of rare earth elements, such as erbium, that scientists know to be good light emitters. Device makers routinely add erbium to the glass in optical fibers to create light-pumped amplifiers and lasers for telecommunications networks. Francesco Priolo of the University of Catania in Italy and Salvatore Coffa of Geneva-based STMicroelectronics have led the research into this method as a means for improving silicon's optical performance. Coffa's group has demonstrated light-emitting diodes (LEDs) that operate at room temperature with efficiencies as high as those of gallium arsenide devices.

The STMicroelectronics LED is a metal-glass-semiconductor sandwich in which a voltage that is maintained between the metal and the semiconductor accelerates electrons through the glass. As they move through, these electrons pump the electrons of erbium atoms in the glass, causing them to emit light. In this case, the quantum confinement in nanocrystals plays a relatively modest role, that of enhancing the conductivity of the glass so that the voltage required to establish the electron flow is reduced. Although LED technology is extremely useful, it produces diffused light (via spontaneous emission) rather than laser light generated by stimulated emission. The STMicroelectronics researchers do, however, anticipate demonstrating true lasing in erbium-doped silicon sometime soon.

Most recently, James Xu's group at Brown University observed lasing at low temperatures (–230 Celsius—too low for common use) in a piece of nanostructured silicon [*see illustration on page 65*]. They produced this effect by first forming an array of closely placed physical holes (110 nanometers apart) on the surface of a thin film of silicon and then pumping them optically. Xu and his team attribute the laser emissions they moni-

RUNNING A SILICON LASER

Silicon Raman laser uses a light beam to excite atomic vibrational waves (phonons, *not shown*) in the silicon laser cavity, which in turn amplify the weak photons produced by silicon. Partially reflective mirrors bounce some of the photons back through the medium, where they stimulate still more photon emissions. (A diode meanwhile sweeps away unwanted pairs of electrons that would otherwise weaken the emitted light.) The resulting cascade of photons coalesce into a coherent shaft of single-wavelength light—a laser beam.



JEN CHRISTIANSEN

SURPRISINGLY, OUR SILICON DEVICE CONVERTED PUMP ENERGY TO LIGHT NEARLY AS EFFICIENTLY AS CONVENTIONAL LASERS.

tored to electrons localized on lattice defects that occur naturally on crystal surfaces of the silicon nanostructures. They further ascribed the enhanced emissions to the quantum-based uncertainty in momentum produced by the tight localization of electrons. These structures create exciting possibilities for nanoscale silicon lasers, which exploit not just optical lasing in silicon but also the element's ability to function as sophisticated mirrors and filters that can manipulate the generated light. Such devices could be useful in future communications networks [see "Photonic Crystals: Semiconductors of Light," by Eli Yablonovitch; Scientific American, December 2001].

Silicon Learns to Lase

PUMPING ELECTRONS into the upper energy band of a semiconductor crystal is not the only way to amplify light. Researchers are also following other routes toward silicon lasers. For example, if one adds energy to the phonons in a crystalline semiconductor, a weak beam of light traveling through the lattice can pick up this energy and become amplified. Feeding some of the intensified light back into the crystal then produces a laser.

In 2002 and 2003, with support from the Defense Advanced Research Projects Agency, my group at the University of California, Los Angeles, showed that a silicon chip could generate and amplify light using this technique. In 2004 we demonstrated the first silicon laser. As with Maiman's apparatus, we pumped our device optically, which is usually a rather inefficient process. Surprisingly, however, our silicon device converted pump energy to light nearly as effectively as today's conventional lasers. Shortly afterward, we embedded the laser device in a diode and succeeded in switching it on and off electrically.

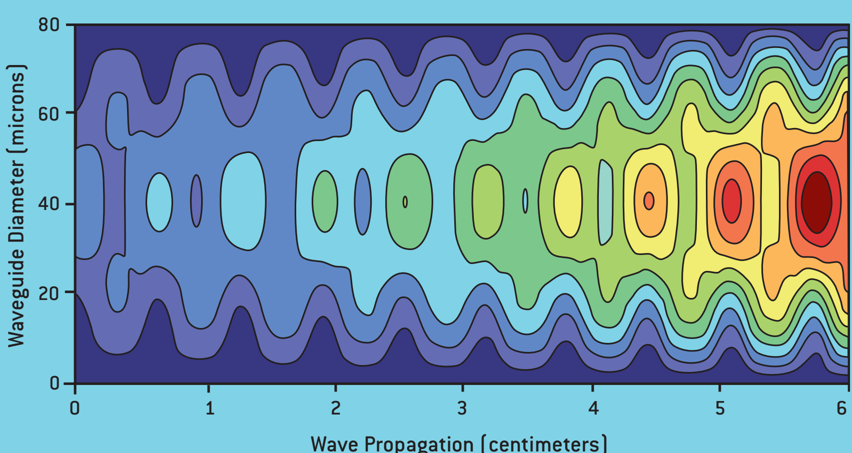
Scientists call the interaction of light with phonons the Raman effect. They employed it extensively in the late 1960s and the 1970s to probe the physical properties of many materials, including silicon. More recently, engineers have harnessed the effect to make optical fibers function as amplifiers and lasers. Because several kilometers of fiber are needed for this purpose, earlier researchers had failed to see it as a practical route to a silicon laser chip. Our team, however, realized that everyone overlooked that the Raman effect in silicon can be 10,000 times larger than in optical fibers, which are made of glass. This much bigger response follows from the well-ordered atomic structure in a silicon crystal (finally, an inherent property of silicon that assists its ability to lase). The random atomic arrangement in the amorphous glass of optical fibers keeps the Raman effect small.

A Raman laser requires optical pumping. To avoid generating electrons in the silicon's upper energy band that would preclude light emission (the free carrier absorption problem), our group excited the silicon using infrared light with a wavelength of 1,500 nanometers. This technique kept the photon energy less than the band-gap level—thus, it remained insufficient to elevate an electron into the upper band. Occasionally, however, two photons will pool their energies and manage to boost an electron into the upper band. Although these kinds of pumped electrons are relatively few, they sap the system of energy.

Raman-based lasers are not the only ones subject to this kind of energy loss. In 2006 Alexander Gaeta and Michal Lipson of Cornell University demonstrated a potentially useful device that amplifies light by mixing it with a more powerful light beam. This amplifier, and its yet to be demonstrated laser counter-

Silicon Laser Image Amplifier

In an optical fiber (or waveguide) with a cross section that is much larger than the wavelength of some incoming light, any pattern in the light goes in and out of focus as it travels down the light pipe as a result of constructive and destructive interference among the lightwaves reflecting off the waveguide's walls. This focusing effect combines with optical amplification to simultaneously focus and amplify an image as the light passes down the waveguide (*warmer colors toward the right*). Researchers at the University of California, Los Angeles, and Northrop Grumman are jointly developing a device in which the Raman effect (the interaction of photons and phonons) amplifies an optical image as it propagates through a thick silicon waveguide. This image amplifier should improve the sensitivity of laser-based remote sensing and imaging systems that scientists use for environmental monitoring.



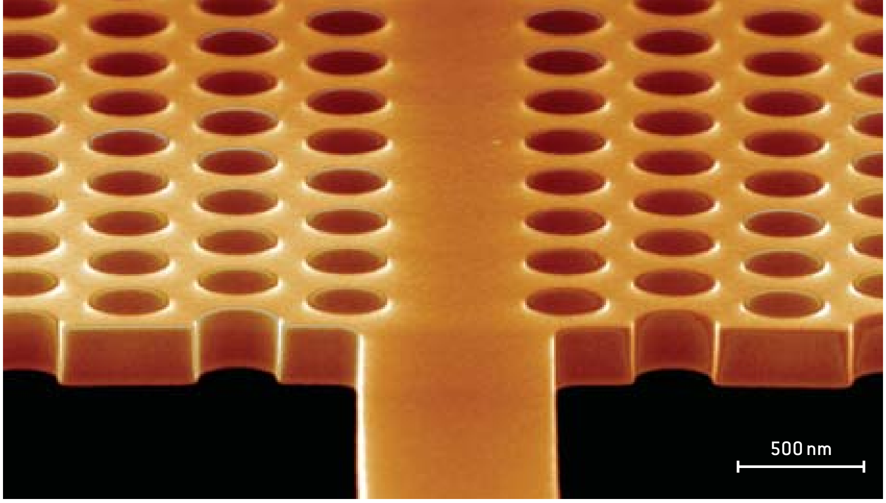
U.C.L.A./NORTHROP GRUMMAN

part, will experience the same losses as a Raman-based system.

To avoid such losses, our first laser operated in a pulsed mode that did not allow electrons to accumulate and drain the system of energy. For continuous laser operation, one can apply an electric field (created by an adjacent diode) to sweep lingering electrons away. Researchers at the Chinese University of Hong Kong suggested trying this approach, and Haisheng Rong and his coworkers at Intel demonstrated it in 2005. Recent investigations indicate that this method is only partially effective because the rate at which electrons can be removed is limited by the maximum velocity the particles can attain in silicon (one-thousandth the speed of light). It also requires significant electrical power to achieve. Fortunately, we know some tricks that can improve the efficiency of the silicon laser: bombarding the silicon with protons or adding small amounts of platinum tends to force the electrons to return rapidly to the lower energy band rather than soaking up scarce photons via free carrier absorption.

These procedures reduce the number of electrons in the upper band, which minimizes their reabsorption of light. Removing the electrons solves only part of the problem, though. The devices still lose pump energy when these electrons are generated unintentionally. By borrowing a trick that underpins the operation of solar cells, my team showed in 2006 that silicon Raman lasers can generate electrical power by harvesting the lost pump energy. Electrons that are generated by the unintentional two-photon absorption flow through the silicon to generate electricity. We learned that we could arrange the electron flow in such a way that the device's power consumption, given by the product of its electric current and its voltage, is negative, which means it actually generates power. The collected electrical power can then drive electronic circuits that reside on the same chip.

This difficulty vanishes altogether if one starts with an optical pumping wavelength that is longer than about 2,300 nanometers, as my research group later demonstrated. The resulting photon energy is so low that even a pair of photons does not possess enough energy to raise an electron into the upper band, which would be unhelpful in a Raman laser. We have found that silicon becomes an excellent lasing medium, arguably one of the best, when pumped with infrared wavelengths of 2,300 to about 7,000 nanometers (at which point other forms of deleterious effects start to appear). This spectrum lies beyond the reach of existing semiconductor lasers, so silicon laser technology permits the development of new applications. Among all laser materials, silicon offers one of the best combinations of thermal conductivity (to pass on unwanted heat) and resistance to damage from high levels of optical power, making it ideal for generating superintense laser beams.



LASING AT LOW TEMPERATURE was demonstrated by James Xu's research group at Brown University in a thin film of silicon similar to that shown above. The surface of the team's device features nanoscale holes placed only 110 nanometers apart. Lasing occurs because electrons are quantum-confined in electron cages at the silicon surfaces.

IBM RESEARCH

Scientists have also developed a promising hybrid approach to producing a silicon-based laser that relies on adding a piece of gallium arsenide or indium phosphide to the top of a silicon substrate. The silicon research community has traditionally resisted hybrid techniques because the addition of other materials changes the electrical properties of silicon, so those materials are viewed as contaminants. Recent encouraging results, however, obtained by groups working at the University of Michigan at Ann Arbor and separately by a team of investigators at Intel and the University of California, Santa Barbara, have led to renewed interest in this approach. If research can overcome the problems of material incompatibility, this method may provide another near-term commercial path to a silicon-based laser.

The unrelenting pursuit of the silicon laser has finally begun to pay off. At last, the field seems to have reached the critical mass that will allow silicon to challenge traditional laser materials. This progress should make the convergence of electronics and photonics all but inevitable. Although it is too early to know the precise trajectory that this new electrophotonic technology will take, the new applications that will be made possible by silicon lasers are likely to have a dramatic impact on our everyday lives. SA

MORE TO EXPLORE

Subtle Is the Lord: The Science and the Life of Albert Einstein. Abraham Pais. Oxford University Press, 1982.

The Incredible Shrinking Transistor. Yuan Taur in *IEEE Spectrum,* Vol. 36, No. 7, pages 25–29; July 1999.

Demonstration of a Silicon Raman Laser. Ozdal Boyraz and Bahram Jalali in *Optics Express,* Vol. 12, No. 21, pages 5269–5273; October 2004.

Computing at the Speed of Light. W. Wayt Gibbs in *Scientific American,* Vol. 291, No. 5, pages 80–87; November 2004.

Optical Gain and Stimulated Emission in Periodic Nanopatterned Crystalline Silicon. S. G. Cloutier, P. A. Kossyrev and J. Xu in *Nature Materials,* Vol. 4, No. 12, pages 887–891; December 2005.

Optical Interconnects: The Silicon Approach. Edited by L. Pavesi and G. Guillot. Springer Series in Optical Sciences. Springer, 2006.

Silicon Photonics: The State of the Art. Graham Reed. John Wiley & Sons, 2007.

Spice Healer

By Gary Stix

TURMERIC (*Curcuma longa*) grows rhizomes, tuberous underground stems, from which the spice of the same name is produced.

Searching for new drugs by milling through ancient folk pharmacopoeia or by just picking a plant while walking in the woods has a decidedly checkered history. Many well-established therapeutic compounds originated in trees, shrubs, mollusks, even dirt. Aspirin came from willow bark, cholesterol-lowering statins from a mold, and the antimalarial artemisinin from a shrub used in traditional Chinese medicine. Yet after raising $90 million during the 1990s in a much publicized bid to tap indigenous knowledge for new drug leads, Shaman Pharmaceuticals had to lower its sights until it was doing nothing more than selling its products as nutritional supplements before finally shutting its doors for good a few years ago.

Now the trend may be reversing itself again. Recently a number of natural compounds—such as resveratrol from red wine and omega-3 fatty acids from fish oil—have begun to receive close scrutiny because preliminary research suggests they might treat and prevent disease inexpensively with few side effects.

DANIELA NAOMI MOLNAR

Turmeric, an orange-yellow powder from an Asian plant, *Curcuma longa,* has joined this list. No longer is it just an ingredient in vindaloos and tandooris that, since ancient times, has flavored food and prevented spoilage.

A chapter in a forthcoming book, for instance, describes the biologically active components of turmeric—curcumin and related compounds called curcuminoids—as having antioxidant, anti-inflammatory, antiviral, antibacterial and antifungal properties, with potential activity against cancer, diabetes, arthritis, Alzheimer's disease and other chronic maladies. And in 2005 nearly 300 scientific and technical papers referenced curcumin in the National Library of Medicine's PubMed database, compared with about 100 just five years earlier.

Scientists who sometimes jokingly label themselves curcuminologists are drawn to the compound both because of its many possible valuable effects in the body and its apparent low toxicity. They ponder how the spice or its derivatives might be used, not just as a treatment but as a low-cost preventive medication for some of the most feared ailments. As a treatment, it also has some enticing attributes. Because curcumin targets so many biological pathways, it could have benefits for cancer therapy: malignant cells may be slow to acquire resistance to it and so might have to go through multiple mutations to avoid the substance's multipronged attack.

Can an ingredient in curry treat diseases from Alzheimer's to cancer?

But is the compound ready for widespread use? Some work offers grounds for caution. Among the more than 1,700 references to curcumin in PubMed are studies showing how a compound that can affect so many biological pathways can sometimes hit the wrong switch and actually help to foster disease.

Long Medical History

KNOWN AS *HALDI* IN HINDI, *jiang huang* in Chinese, *manjal* in Tamil (and just plain "yuk" as the yellow stain on a white T-shirt from the splatting of ballpark mustard), turmeric has a medicinal history that dates back 5,000 years. At that time it was a key medicament for wound healing, blood cleansing and stomach ailments in India's Ayurvedic system of medicine.

The first record in PubMed of research on the biological activity of curcumin dates back to 1970, when a group of Indian researchers reported the effects of the compound on cholesterol levels in rats. The pace of studies picked up in the 1990s; one of the leaders was Bharat Aggarwal, a former scientist at Genentech who, before turning to curcumin, had taken another approach to seeking cancer treatments. That work led him circuitously to the compound.

In the 1980s Aggarwal and his team at Genentech were the first to purify two important immune molecules—tumor necrosis factor (TNF) alpha and beta—that have been identified as potential anticancer compounds. These molecules can, in fact, kill cancer cells when deployed in localized areas, but when circulated widely in the bloodstream, they take on different properties, acting as potent tumor promoters. The TNFs activate an important protein, nuclear factor kappa B (NF kappa B), which can then turn on a host of genes involved in inflammation and cell proliferation.

This link between inflammation and the unchecked proliferation of cancer cells prompted Aggarwal to return to his roots. In 1989 he moved to the University of Texas M. D. Anderson Cancer Center and began looking for compounds that might quell inflammation and have an anticancer effect. Remembering from his youth in India that turmeric was an anti-inflammatory in the Ayurvedic literature, he decided to give the spice a try. "We took some from the kitchen and threw it on some cells," he remembers. "We couldn't believe it. It completely blocked TNF and NF kappa B."

Aggarwal has gone on to publish studies showing that blocking the NF kappa B pathway with curcumin inhibits the replication and spread of various types of cancer cells. This work has served as a jumping-off point for early, small clinical trials at M. D. Anderson using curcumin as an adjunct therapy to treat pancreatic cancer and multiple myeloma. Trials are beginning or under way elsewhere for prevention of colon cancer and Alzheimer's disease, among others. And early cell-based or animal studies have shown that curcumin may act against a range of inflammatory diseases, including pancreatitis, arthritis, inflammatory bowel disease, colitis, gastritis, allergy and fever. It has also shown some promise for diabetes and autoimmune and cardiovascular diseases.

So far the large clinical trials needed to prove efficacy against cancer and other diseases have yet to be conducted. But Aggarwal has nonetheless become an aggressive champion for a spice that Vasco da Gama brought back to Europe from his voyages eastward. Aggarwal's chapter in a new textbook that he co-edited is entitled "Curcumin: The Indian Solid Gold."

M. D. Anderson, a world-leading cancer institution, has also begun to promote the use of curcumin more than would be expected for a treatment that has not gone through the rigors of full clinical trials. The "frequently asked questions" section on its Web site recommends buying curcumin from a specific wholesaler, for which Aggarwal has served as a paid speaker. That company even issued a press release declaring that its product is the "ingredient of choice" of M. D. Anderson.

The FAQ section suggests that cancer patients gradually work up to a daily dose of eight grams a day, some 40 times the amount consumed in the average Indian diet. Most pharmaceuticals, in contrast, are meted out in milligrams. At one point, the Web site had even asserted: "By the end of eight weeks, a significant improvement is expected." Asked whether he was worried that any side effects might emerge at a dosage of eight grams, Aggarwal said that small clinical trials at other institutions have dosed up to 12 grams and that patients would have notified him if any untoward effects had occurred with the dosage recommended by M. D. Anderson. The researcher, who takes a curcumin pill every day, shuns the caution typical of investigators before well-controlled, large-scale clinical trials have been conducted. "People take a lot of other supplements, and I don't think you need anything else if you're taking this," Aggarwal says.

Does Curcumin Abet Cancer?

THE M. D. ANDERSON FAQS and the stream of press releases from various institutions on the wonders of curcumin ignore a small portion of the literature that points to a dark side: the possibility that this spice may sometimes actually encourage the survival of cancer cells. In 2004 Yosef Shaul in the department of molecular genetics at the Weizmann Institute of Science in Rehovot, Israel, was studying an enzyme, NQO1, that regulates the amount of a well-known protein called p53. When p53 levels increase in cells, the protein institutes a defensive maneuver for the organism by inducing cancerous or damaged cells to stop dividing or even to commit suicide.

Shaul and his colleagues had found that an anticoagulant, dicoumarol, and related compounds blocked NQO1, which prevented p53 from doing its job. The researchers wondered what would happen if they exposed p53 in normal and myeloid leukemia cells to antioxidants such as curcumin and resveratrol. To their surprise, curcumin, by inhibiting the same enzyme, stopped p53 from sending aberrant cells to the gallows, a finding that was reported in 2005 in the *Proceedings of the National Academy of Sciences USA*. A few other researchers have published similar results. Aggarwal responds to this body of work by pointing to studies that show the opposite, that curcumin actually activates p53.

Clinical researchers will now have to address whether Shaul's work in cell cultures relates to what happens when a person ingests the compound. The curcumin concentrations used by the Weizmann team in cell cultures—measuring 10 to 60 μM (micromolar)—are roughly comparable to levels reached in some of the test-tube experiments conducted at M. D. Anderson. But because curcumin is absorbed poorly from the gut into the bloodstream and is also broken down in the body rapidly, a patient consuming eight grams would probably end up with a concentration in blood plasma no higher than about 2.0 μM, Shaul notes, although that level could range higher in the gastrointestinal tract and in the liver. It could also remain elevated if researchers develop various means of increasing the concentration of curcumin in the bloodstream.

M. D. Anderson's FAQs might convey the impression of certitude by prescribing an eight-gram dose. But the low presence of curcumin in the blood—and the corresponding need to elevate the amount consumed if the substance does indeed fight disease—is a challenge that will continue to nag curcumin researchers. The animal studies that investigators cite as suggestive of curcumin's diverse benefits have generally used less than the equivalent of eight grams in humans, and blood concentrations have usually been in the nanomolar range. "We don't know how to explain how such low concentrations of curcumin can be beneficial in animals tested," Shaul states.

Dose is everything for a new drug—any therapeutic agent, including aspirin, turns toxic at high levels. For most new pharmaceuticals, the best dose for achieving the desired blood plasma levels is usually found through round after round of preclinical trials in cell cultures and mice. Yet drug companies are not battling one another to be the first to conduct these tests on curcumin. They have a preference for highly targeted therapeutics: hitting a specific receptor, for instance, may treat disease while lowering side effects, whereas a drug with multiple ac-

Recent Studies Show Possible Benefits from Curcumin . . .

CONDITION	FINDINGS	INSTITUTION	PUBLICATION
Rheumatoid arthritis	An extract of turmeric root inhibited joint inflammation and destruction in rats	University of Arizona College of Medicine	*Arthritis and Rheumatism*, November 2006
Alzheimer's disease	In test-tube studies, curcumin helped immune cells degrade components of Alzheimer's plaques	U.C.L.A. and the Veterans Administration	*Journal of Alzheimer's Disease*, October 9, 2006
Colon cancer	In cell cultures, curcumin blocked the activity of a hormone tied to development of colon cancer	University of Texas Medical Branch at Galveston	*Clinical Cancer Research*, September 15, 2006
Colorectal polyps	A combination of curcumin and the plant compound quercetin reduced the size and number of precancerous lesions in five patients	Johns Hopkins University and Cleveland Clinic	*Clinical Gastroenterology and Hepatology*, August 2006
Cognitive impairment	More than 1,000 elderly subjects from Singapore who reported eating curry at least occasionally had better scores on a cognitive test than did those who rarely or never ate the dish, an effect that might be attributed to curcumin	National University of Singapore and other institutions	*American Journal of Epidemiology*, November 1, 2006

. . . But Some Research Suggests Possible Cancer-Promoting Effects

CONDITION	FINDINGS	INSTITUTION	PUBLICATION
Myeloid leukemia	Curcumin at high doses in cell culture spurs degradation of a protein, p53, that prevents replication of cancer cells or induces their death	Weizmann Institute of Science, Rehovot, Israel	*Proceedings of the National Academy of Sciences USA,* April 12, 2005
Colon cancer	Curcumin inactivates p53's tumor suppressor role in colon cancer cells	University of Utah	*Carcinogenesis,* September 2004
Breast cancer	Curcumin inhibits several chemotherapeutic drugs from inducing cell death both in cell culture and in animal models	University of North Carolina at Chapel Hill	*Cancer Research,* July 1, 2002

FOODPIX

tions could, in theory, increase the chance that an unwanted effect will occur. Another reason is the nettlesome issue of property rights for folk medicines.

Turmeric is a poster child for one of the most noted intellectual-property cases on biopiracy, which pitted an Indian government–supported research organization against a 1995 patent issued to the University of Mississippi for the use of the spice for wound healing. The U.S. Patent and Trademark Office invalidated the patent after the Indian Council for Scientific and Industrial Research questioned whether one criterion for patentability—that an invention be new—had been met. The council objected by pointing to a 1953 Indian journal article about the spice and by offering a citation about turmeric's healing properties from an ancient Sanskrit text.

The patent office has subsequently issued patents for specific uses for curcumin as an isolate. But the rejection means that drug companies will never obtain a "product" patent with a much broader scope that would help them to fend off competitors for drugs based on the spice. A few small companies are still trying to exploit the substance's promise by changing its chemical composition to enhance activity and, by creating a novel compound, to bolster intellectual-property protection.

AndroScience in San Diego plans to enter the first phase of clinical trials this year with a drug candidate for acne based on compounds derived from curcumin that were discovered in collaboration with the University of North Carolina at Chapel Hill. Similarly, Curry Pharmaceuticals in Research Triangle Park, N.C., is trying to raise financing to move curcumin derivatives from Emory University into clinical trials. But in an age of targeted pharmaceuticals, venture capitalists, leery of side effects, have been hesitant to back new drugs that act on multiple pathways. For his part, Aggarwal, even though he is a co-founder of Curry Pharmaceuticals and holds patents on curcumin, asserts that chemists may have trouble improving on nature: modifying curcumin may only introduce unwanted side effects in patients, he says.

If the multitude of developmental hurdles can be overcome and safety can be assured, curcumin might provide an inexpensive alternative to mainstream pharmaceuticals. Based on positive results in rodents, Greg Cole of the University of California, Los Angeles, and the Veterans Administration, is organizing a clinical trial in humans to test whether curcumin can prevent the buildup of amyloid plaques that burden the brains of Alzheimer's patients. If successful, he and his collaborator (and wife), Sally Frautschy, plan to come up with formulations that could be mixed in cooking oil (to enhance bioavailability) and eaten as part of a meal to impede plaque accumulation—a recipe that might be affordable for both rich and poor in an aging world. SA

MORE TO EXPLORE

Inhibition of NAD(P)H:Quinone Oxidoreductase 1 Activity and Induction of p53 Degradation by the Natural Phenolic Compound Curcumin. Peter Tsvetkov et al. in *Proceedings of the National Academy of Sciences USA,* Vol. 102, pages 5535–5540; April 12, 2005.

Multiple Biological Activities of Curcumin: A Short Review. Radha K. Maheshwari et al. in *Life Sciences,* Vol. 78, pages 2081–2087; 2006.

The Molecular Targets and Therapeutic Uses of Curcumin in Health and Disease. Edited by Bharat B. Aggarwal, Young-Joon Surh and Shishir Shishodia. Springer (in press).

DIGITAL TV AT LAST?

Analog TV broadcasting is set to end in two years, but its legacy could make the digital transition anything but smooth

BY MICHAEL ANTONOFF

February 17, 2009, is D-day—

BRIAN MARANAN PINEDA

when the term "digital divide" will take on a whole new meaning unrelated to computer access. That is when the nation's 1,700 analog television stations will shut down in the long-promised changeover to all-digital broadcasting. Cable and satellite viewers or those whose TV has a digital tuner will be able to watch *CSI* and *American Idol* unaware that anything has changed. But the 21 million households using a conventional set with rabbit ears or a rusty roof antenna—typically people who are poor, elderly or living in rural America—will turn on their TVs and see … nothing.

Not since color made black-and-white sets passé has a new technology so compelled consumers to replace their primary entertainment appliance. In contrast to the variable radio waves of analog broadcasting, digital TV (or DTV) uses electrical pulses to transmit information precisely and efficiently. Thus, DTV offers startlingly sharp pictures, capable of revealing individual blades of grass on a field or the writing on a ransom note held in a fictional detective's hand. DTV also will enable new interactive features, thanks to its ability to pack enormous amounts of information in scarce bandwidth. For instance, it could let viewers call up stats on a ballplayer during the game or view a recipe from a cooking show. Broadcasters will be able to offer even more channels in the available bandwidth, too. The freed-up space will make room for better communications for public safety agencies during emergencies such as hurricanes and for next-generation cellular services.

Several thorny technology issues remain, however. Even after the end of TV broadcasting as we have known it, analog TV's legacy will be with us for years to come. While analog-retentive households contemplate the grim reality of no TV at all, others who seem to be digitally outfitted will still experience stumbling blocks. Cable systems, which are already set up for DTV, may not actually be able to deliver purely digital signals, because many of their customers' older "cable ready" sets still have analog-only tuners. And even in areas where digital is already available, it has experienced service hiccups: picture problems and even clogged bandwidth. Perhaps it is more accurate to say that long after 2009, viewers will encounter a series of digital divides, between analog and digital and among different kinds of digital.

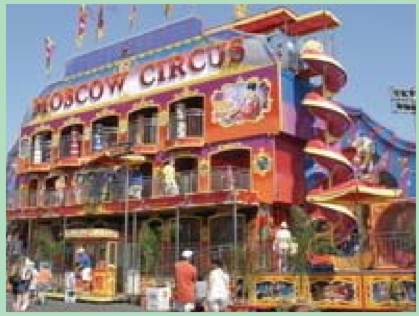
The idea of switching the country to digital TV first gained momentum in Congress during the 1980s, when the legislature became aware that Japan already had high-definition TV, albeit produced by an analog system. (There has been much confusion over the difference between DTV and high-definition television, or HDTV. HDTV is simply the sexier subset of DTV, and it has a few special requirements: more bandwidth to deliver it, extra capacity to store it, a higher-resolution screen to show it and a larger display to appreciate it.) The U.S. government proceeded to run a series of tests with competing manufacturers designed to determine a next-generation broadcast system. The initial results showed that high-definition broadcasting was not practical without compression—and the most important thing to know about digital is that, unlike analog, it compresses video extremely efficiently. So the idea of retrofitting analog broadcasting for greater resolution was abandoned.

The Advanced Television System Committee, working with an alliance of companies that initially had incompatible systems, ultimately succeeded in unifying a standard. The consortium authorized 18 DTV formats, two of which—1080i and 720p—are high-definition and meant for viewing on wide screens. The 1080i format puts 1,080 lines on the screen by transmitting the odd lines, then the even lines, and interlacing them. Hence the "i." The second puts 720 lines on the screen one after the other—that is, progressively. Hence the "p." (In contrast, analog broadcasts offer up to 480 horizontal lines.) Debates over the merits of each are confined to broadcast engineers and videophiles, who cite the spatial advantage of displaying more lines (1080i) versus the temporal advantage of showing fast motion (720p). Most viewers do not care.

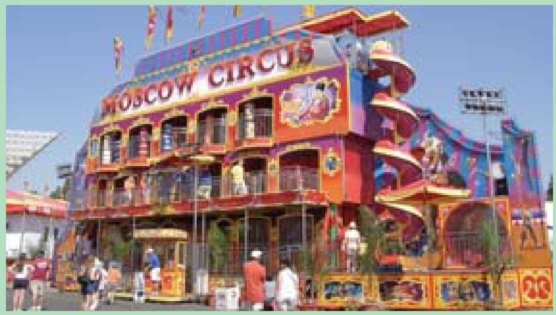
The first publicly available high-definition broadcasts in the U.S. began in 1998, as HDTV sets started going on sale. In the late 1990s stations were expected to give up their analog frequencies by 2006. But that plan fell through because too few people bought the expensive digital TVs or the set-top boxes for receiving the signals and because high-definition programming rolled out slowly as well. NBC, for example, offered only one series in high definition, *The Tonight Show*, beginning in 1999. It took another five years before the network showed all its prime-time dramas and sitcoms in 1080i. Broadcasters had little incentive to give up their analog frequencies. Compared with DVD, a home video format introduced just a year earlier that became the most successful new product launch in consumer electronics history, HDTV barely registered in sales or public awareness.

HDTV OFFERS UP to 10 times as many pixels for sharper resolution. It also has a wider aspect ratio of 16:9 compared with 4:3 for conventional sets.

EXISTING TV



HIGH-DEFINITION TV

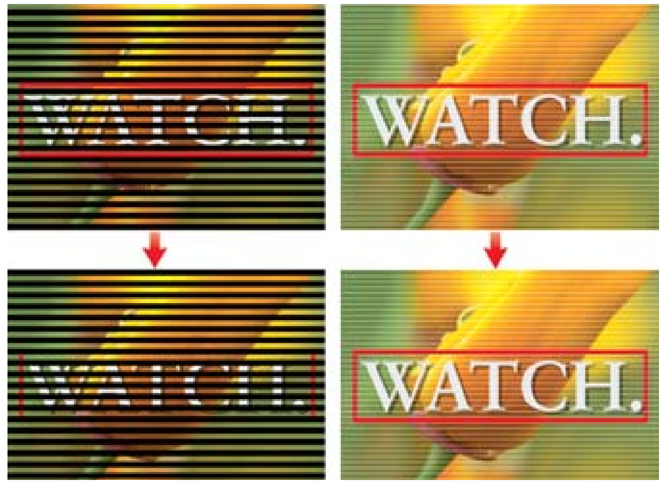


WAYNE E. BRETL *Zenith Electronics Corporation*

Overview/*Digital TV*

- After many years of anticipation, the end of conventional analog broadcasting now appears to be set for February 17, 2009; after that date, all broadcasts will be digital.
- Digital broadcasts offer consumers advantages such as crystal-clear pictures and new information services. Frequencies no longer used for analog broadcasting will be available for communication by emergency response teams or auctioned for other uses, such as advanced cellular systems.
- Analog's legacy, however, will continue to dog the new digital era for several years past the cutoff.

In March 2005, when manufacturers met in Washington, D.C., at an HDTV summit, there was still no definitive cutoff for analog broadcasting. But in early 2006 President George W. Bush signed a law compelling a transition in just three years. What changed? The name of the legislation was a clue. It was called the Deficit Reduction Act of 2005. The government expects to raise billions of dollars when it auctions off much of the 108 megahertz (MHz) of "returned" analog frequencies. Other factors that make analog's demise more likely this time are that DTV tuners are now mandated in all new TV sets with a screen size of 25 inches or larger (and will be required in all sets, regardless of screen size, by March 1 of this year), most prime-time network series and all major sports events are broadcast in high-def, and most cable systems carry from four to a dozen high-def channels (although that number still represents a mere fraction of their overall channel lineups).



TWO HIGH-DEFINITION STANDARDS work on different principles. The 1080i format (*left*) puts 1,080 lines on-screen by transmitting the odd lines, then the even ones and interlacing them (hence the "i"). In contrast, 720p beams 720 lines one after the other, or progressively (for "p").

SLIM FILMS

The Long Good-bye

SOME ARGUE THAT the cessation of over-the-air analog TV by federal law will be far more disruptive than when color elbowed black-and-white aside in the 1960s. After all, those silvery tubes continue to work even today. That will not be possible with conventional TV sets unless they have a digital converter box. Adding boxes, which are expected to cost about $60 for a no-frills model, to every TV in any given household could be an expensive proposition.

According to the National Association of Broadcasters, the 21 million homes that receive only over-the-air signals own a total of 45 million TV sets. If you add that figure to the 28 million second and third TVs in homes with cable or satellite where only the primary set is hooked up to the pay service, some 73 million TVs nationwide could go dark on D-day. (Because many of those secondary sets are attached to DVD or video game players, they will not necessarily become useless just because they cannot receive broadcasts.)

To take the financial sting out of the changeover, Congress has authorized the U.S. Commerce Department's National Telecommunications and Information Administration to spend up to $1.5 billion on subsidies. Starting in 2008, those who cannot afford a box will be eligible for two coupons worth $40 each toward their purchase. To record an over-the-air signal, older VCRs and DVRs will also have to be attached to a conversion box, although by July 1, 2007, any new device sold that includes a tuner (DVD player/recorder, DVR, computer TV card) must include a DTV tuner.

Even if your TV can receive over-the-air digital signals, that does not guarantee you can see the pictures. Analog offers what is called "graceful degradation": people in fringe reception areas can at least see something, even if the picture ghosts or fades in and out. DTV is not as forgiving. You either get it, or you do not.

Analog TV transmits a complete picture 30 times a second, a lavish waste of bandwidth from a digital perspective. Digital compression, on the other hand, is based on the premise that it is only necessary to transmit reference frames and intermittent changes. That is how the same 6 MHz allocated to a single analog TV channel can accommodate up to six standard-definition digital channels or perhaps one HDTV channel. But it also explains why momentary signal losses can be so disruptive to DTV reception. Without the luxury of analog redundancy, a DTV set may freeze up on the last reference frame it received, or the picture may break into a less than gorgeous mosaic or just go black. There is nothing graceful about any of those glitches, especially if they happen as the killer is revealed during a crime show.

And you do not have to be living in a valley 80 miles from the broadcast tower to feel abandoned by a DTV signal. Multipath reflection continues to be a problem within a city; the receiver gets multiple versions of the same broadcast because it bounces off buildings. The U.S. uses a transmission standard called 8-VSB (for 8-level vestigial sideband), which is more susceptible to signal woes than the European standard COFDM (coded orthogonal frequency division multiplexing).

Today's receivers can better compensate for this problem than ones manufactured at the dawn of DTV, thanks to more powerful processors. Newer receivers are using higher-performance chips capable of calculating a greater quantity of reflected signals and discarding them more quickly than the processors in earlier-generation DTV receivers [see "The Multipath to Clarity," by Philip Yam; SCIENTIFIC AMERICAN, June 2005]. But viewers may still have to reposition an indoor

THE AUTHOR

MICHAEL ANTONOFF has been covering consumer electronics technologies for magazines, including *Video, Sound & Vision* and *Popular Science,* for more than 20 years. In high school he assembled a Thinking Boy radio transmitter good for two feet and promoted it in speech class as the Electric Kazoo. Antonoff notes that the spelling of his name starts out as analog but ends up in the digital realm.

antenna when they switch to a channel coming from a different direction—a frustrating exercise that recalls the days before most homes had rooftop antennas or cable. While multipath is DTV's Achilles' heel, the problem is restricted to a small portion of the 15 percent of American households not subscribing to cable or satellite service. Although there is something to be said for commercially supported "free" TV, more HDTV programming (including premium movie, sports and nature channels) is already available by subscription only than from over-the-air DTV, and the gap will likely increase.

Dual Broadcasts

UNTIL D-DAY, over-the-air stations will continue to simulcast analog and digital signals on their two allotted frequencies of 6 MHz apiece—the federal government has given stations the extra space for digital during the transition period. The DTV frequency lets broadcasters do something they could not do with analog: multicast several programs at the same time. Whereas a station might choose to allot the entire 6 MHz to high definition, it could instead transmit several channels in standard-definition digital. One standard-definition channel could be accommodated around the clock—even when the station is carrying the network's HDTV programs—without a significant degradation in quality, the National Association of Broadcasters claims.

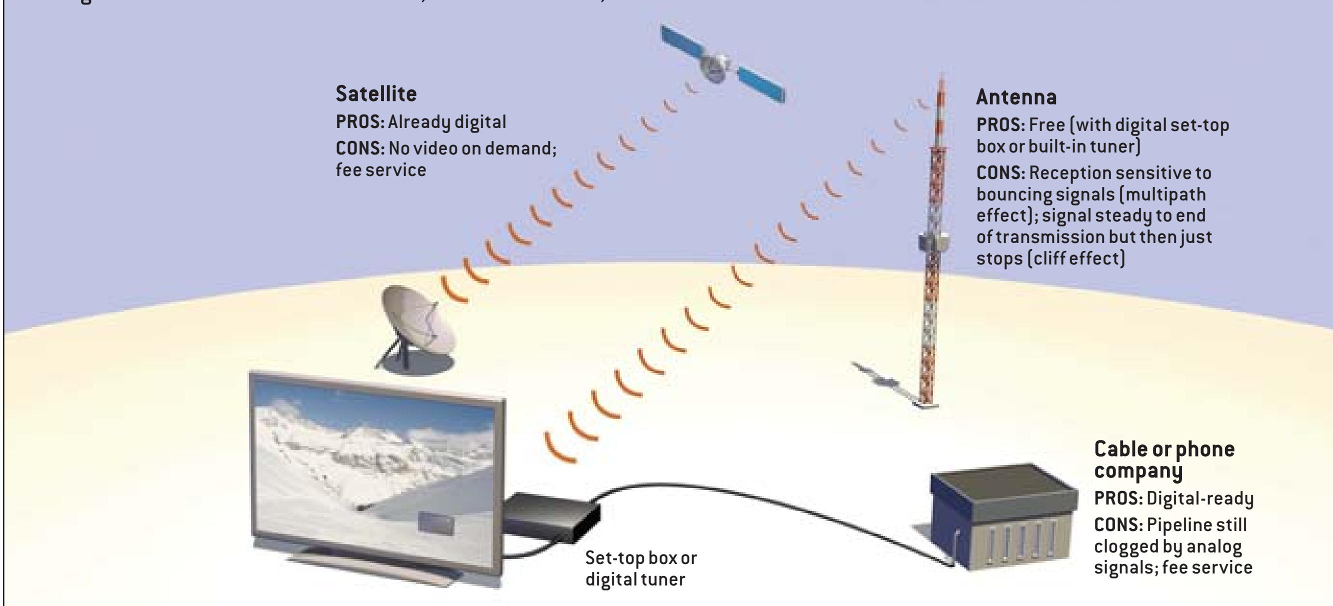
Those consumers whose signals travel by cable still will not escape analog after 2009. Operators would like to go all-digital because viewers prefer watching the high-definition version of a channel if available. Not only does an HDTV program's wide-screen format fit their new TVs' 16:9 aspect ratio, not to mention their old pair of eyes, better than standard definition's 4:3 shape, but they also can pick out much more detail. If *CSI*'s mantra is to follow the evidence, then HDTV's two million pixels (on a 1920 × 1080 screen) point the way. The rising demand for high definition has, however, already become a problem for cable and satellite operators, who have a finite amount of bandwidth to divide among various channels and services.

Cable systems can handle 750 MHz of bandwidth, with hundreds of channels, video on demand, telephone services, broadband Internet access and interactive programming guides vying for carriage. But even the most advanced cable systems offer only about a dozen high-definition channels to their customers today. Why is that the case? The biggest hog is analog TV channels, which cable operators must continue to carry well past 2009, according to CableLabs, the research and development arm of the cable TV industry. Those linear (continually transmitting, one-way-only) analog channels occupy as much as 550 MHz of the coaxial cable entering homes. Cable companies cannot deep-six all their analog channels and move everything into the bandwidth-efficient digital realm because of the legacy of cable-ready TVs. Millions of viewers plug the cable directly into the analog tuners on their old tubes. Digital cable boxes leased by subscribers do bypass the tuners in their TVs.

So cable operators are faced with the costs of offering a box to everyone or multiple boxes to households that have several TVs. Or another option is that the companies may simply wait well into the next decade, when sufficient numbers of viewers will have finally replaced their long-lasting analog sets with ones containing DTV tuners as well as other so-called conditional-access systems, such as credit-

DIGITAL DELIVERY

Digital signals commonly travel to televisions in homes using three different routes—satellite, antenna or cable; telephone companies are beginning to offer TV service in select communities as well.



GEORGE RETSECK

TYPES OF TVS

TECHNOLOGY	RESOLUTION	DISPLAY	AUDIO
High-definition television (HDTV)	720p or 1080i	16:9	Dolby digital 5.1
Enhanced-definition television (EDT)	480p minimum	16:9 or 4:3	Dolby digital 5.1
Standard-definition television (SDTV)	480i or less	16:9 or 4:3	Dolby digital 5.1
Analog TV	Less than 480i	4:3	FM stereo

card-size CableCARDs or their software-only counterparts.

One way to manage the pipeline problem is to adjust the transmission techniques. Cable operators have begun employing a modified system architecture called switched broadcasting (also known as switched digital video). A hybrid fiber coaxial cable system brings an optical fiber trunk with seemingly limitless bandwidth to a neighborhood, but then branches to coaxial cable, with its 750-MHz limitation, for the last leg, serving 300 to 500 homes on a node that loops from the fiber. Conventional signal distribution carries every channel in parallel into each home, so that a parent might be tuning into ESPN-HD in one room while a child enjoys MHD (the high-def sibling of MTV) in her room. At any particular time, though, nobody is watching most of the channels on the node.

The new way is to stream only the channels that tuners on that node have requested at that moment. The beauty of the system is that node bandwidth, which was once allocated equally to popular and hardly watched channels alike, can now be freed up to make room for more high-definition channels as needed—because most people are likely to be choosing from among the same popular channels rather than from obscure options. Part of the reason the cable industry is now talking about a hybrid fiber-coaxial system is increasing competition from the telephone companies Verizon and AT&T, which have begun rolling out fiber to homes (or groups of homes) in select communities. In a couple of years many more households are expected to be able to get TV service from their "phone" company.

Unlike cable's fiefdoms, Direct Broadcast Satellite (DBS) companies (DirecTV and DISH Network) operate nationally and have been digitally efficient all along. Every subscriber uses an external tuner/decoder to receive channels, bypassing the TV tuner. But DBS faces its own bandwidth constraints as channels overall have proliferated, the number of network HDTV affiliates has swelled, and subscribers have increasingly had their local channels beamed to them by satellite.

DBS has been coping in three ways. First, it has added satellites. Subscribers using older equipment who want to get more recently added channels may need to buy another dish or replace a round dish with an elliptical one so they can receive signals from more than one position in the sky. Second, it has increased the use of spot beams as a way to make the most of the available frequencies. Instead of raining down identical bits on the entire continent, a spot beam narrowly focuses channels meant for a particular metropolitan area. Other spot beams are pointed at other cities. Although there is redundancy transmitting *Grey's Anatomy* on each spot beam, the technology helps to preserve a network like ABC's affiliate system, because local commercials and programs are contained in the signal just as they are if the viewer received the program over the air or by cable.

Last, satellite systems are changing over from the MPEG-2 compression scheme to MPEG-4, which accommodates about twice as many channels; viewers may need to replace their set-top box with one that can decompress MPEG-4. MPEG-4 coding is more efficient because various objects in the same scene can be scaled for different levels of spatial detail. For example, the full resolution for an important foreground object, such as a football, can be maintained, whereas a less important object, such as a group of fans in the background, can be updated at a lower rate. The concept is similar to the perceptual coding technique used for MP3 audio compression, in which, at a particular moment, one instrument masks another; the encoding process discards what is imperceptible anyway to save bits and consume less bandwidth.

The Last Leap

NO MATTER HOW the signal arrives, consumers who want to record their shows may have one more digital hurdle after the transition. A VCR's analog-only tuner will no longer be able to record over-the-air TV channels, although it may be compatible with analog channels still being carried on cable. Cable and satellite subscribers are more likely to lease or buy a high-definition-capable set-top box with a built-in hard drive to record programs at the highest resolution. Alternatives include CableCARD hard-drive recorders such as a new high-def model from TiVo, HDTV sets with built-in hard drives, and—if viewers want to play their old VHS cassettes in the same device—DVHS machines that can also record over-the-air HDTV programs.

Of course, occurrences of digital-picture losses and the inconvenience of getting new set-top boxes are worth putting up with, HDTV advocates say, because conventional TV cannot hold a candle to the amazing clarity of high definition or the services it will provide. When February 2009 rolls around, one thing is certain: television will never be the same again. SA

MORE TO EXPLORE

Defining Vision: How Broadcasters Lured the Government into Inciting a Revolution in Television. Updated and expanded. Joel Brinkley. Harvest Books, 1998.

CableLabs Web site from Cable Television Laboratories, Inc., a nonprofit research and development consortium of the cable TV industry: **www.cablelabs.com**

CED Broadband Direct, daily newsletter about broadband technology: **www.cedmagazine.com**

Countdown to DTV Transition, a Web site from the Federal Communications Commission: **www.dtv.gov**

HDTV Magazine Web site, including technical articles and program highlights: **www.hdtvmagazine.com**

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

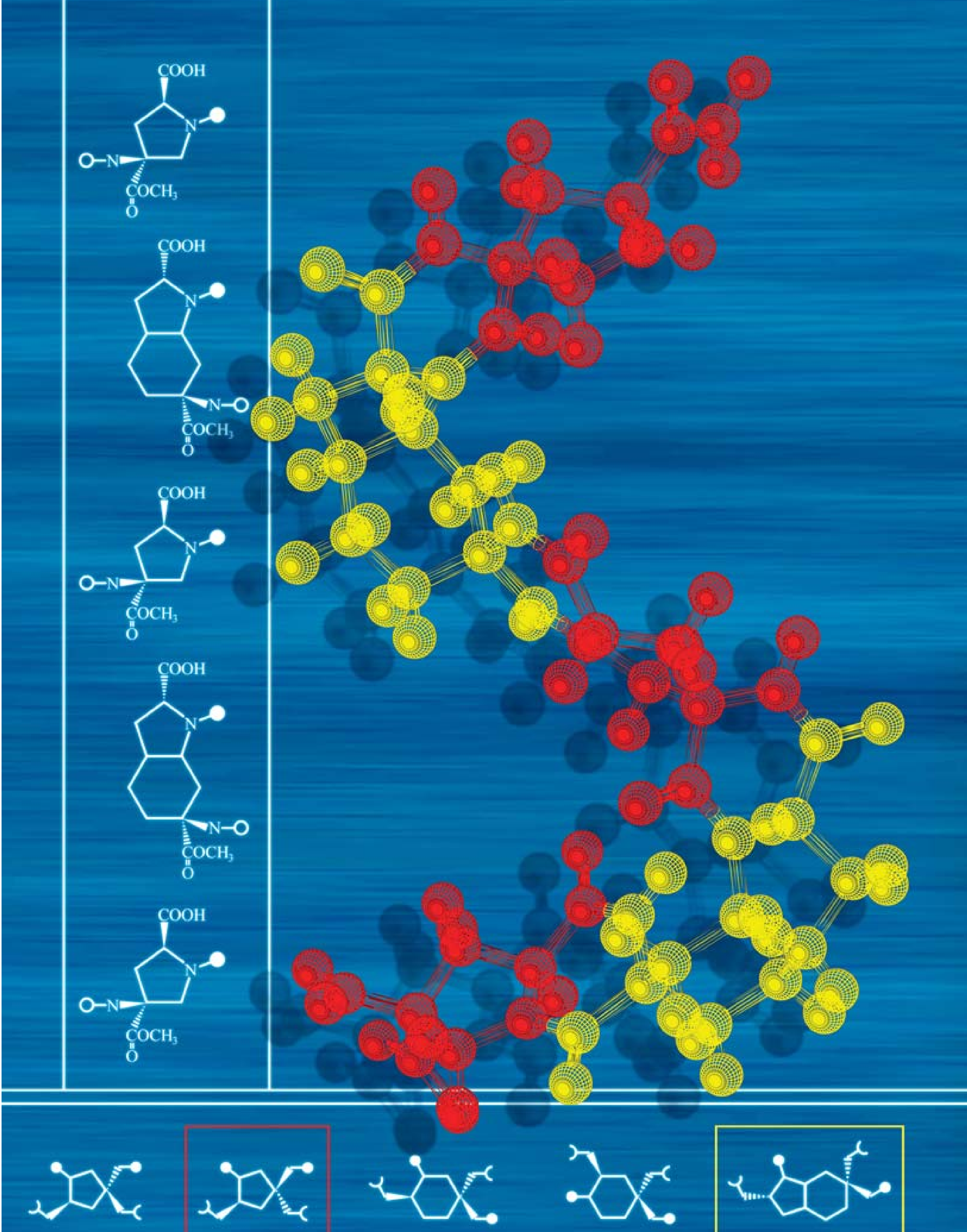
Proteins, the fundamental nanomachines of life, have provided scientists like me with many lessons in our own efforts to create nanomachinery. Proteins are large molecules containing hundreds to thousands of atoms and are typically a few nanometers (billionths of a meter) to tens of nanometers across. 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

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*


COOH
N
N
COCH$_3$
O
COOH
N
N
COCH$_3$
O
COOH
N
N
COCH$_3$
O
COOH
N
N
COCH$_3$
O
COOH
N
N
COCH$_3$
O

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 up is one of the most significant unsolved challenges of science and engineering (the "protein folding problem").

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.

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 novel molecular building blocks, called bis-amino acids, that join 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.

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 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.

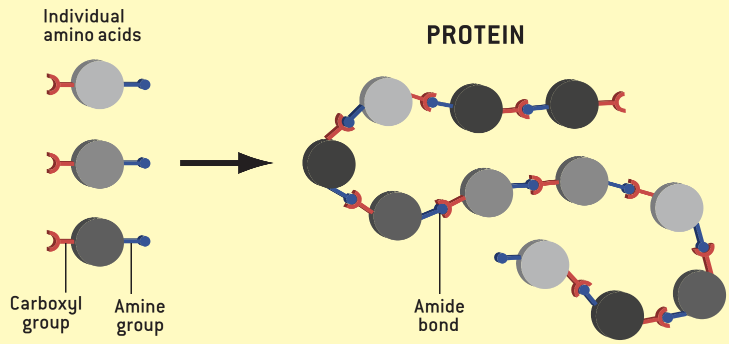
These new approaches to constructing macromolecules that held a specific shape were inspiring, but they seemed to trade one folding problem for another. The difficulty is that natural proteins and these new molecules involve chains of molecules connected by single bonds that leave the structure with 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*

HOW BIS-PEPTIDES DIFFER FROM PROTEINS

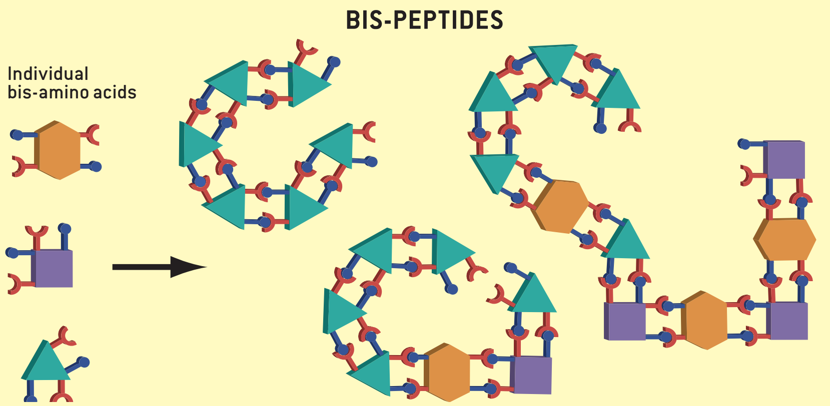
NATURAL PROTEINS

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.



MELISSA THOMAS

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.

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.

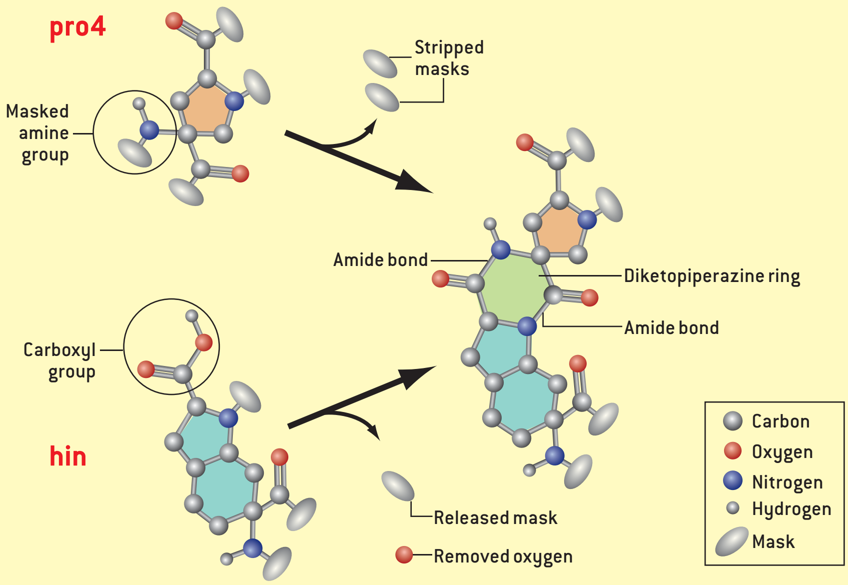
In real chemical terms, each monomer would consist of a rigid molecule of mostly carbon atoms with two amino acid groups integrated into it, and the amines and carboxyls of both amino acids would be available for bonding to other monomers. Two monomers would join by having an amino acid group on each 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

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.)



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), 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

THE AUTHOR

CHRISTIAN E. SCHAFMEISTER is assistant professor of chemistry at the University of Pittsburgh, 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

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 bottom 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.

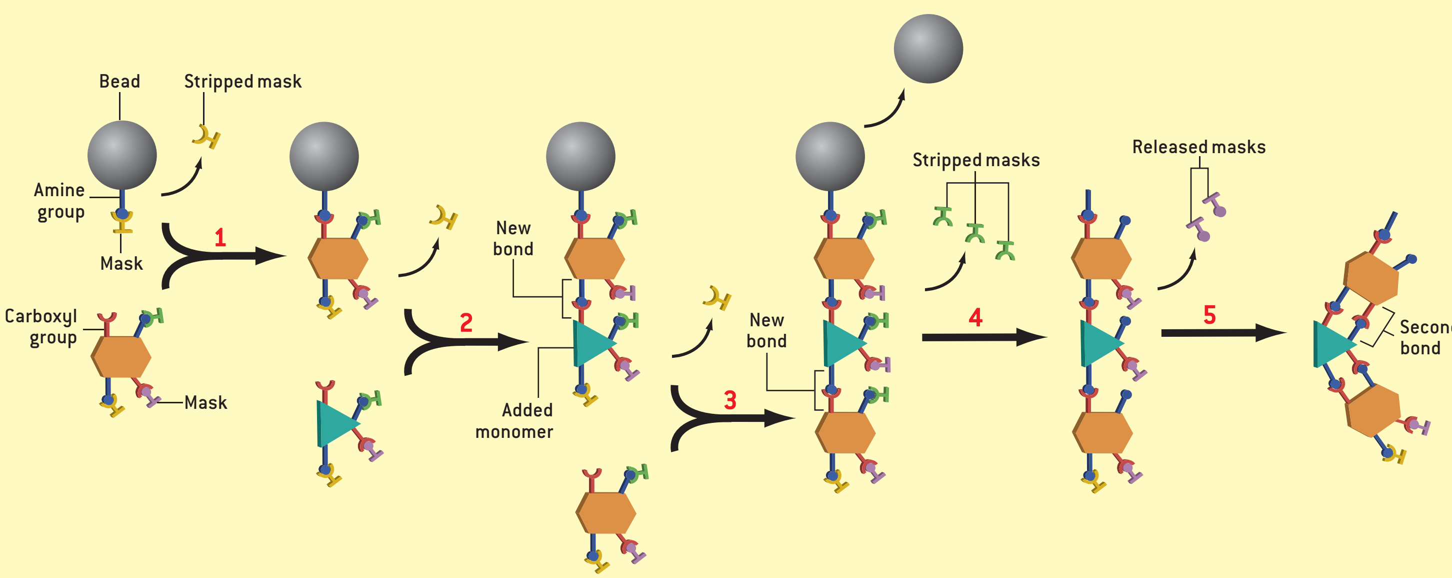
The pro4 group of bis-amino acids are like Lego bricks that have relatively small tilts, so we can use them to make rodlike and gently curving shapes, which could function like struts to hold chemical groups apart at specific distances. Many useful functions of proteins, however, come about because of 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

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 to 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*).



reaction that forms the rigidifying second amide bond was very rapid between pro4 monomers but was 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.

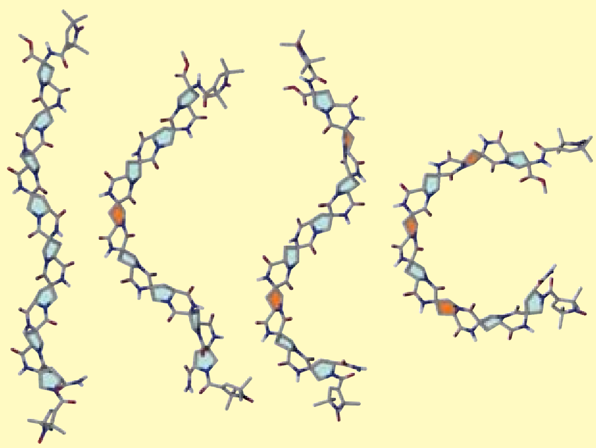
We have not, however, been able to determine whether each bis-peptide was binding two pockets of one Ctx or binding with pockets on two different Ctx molecules and thus creating a cross-linked network of Ctx molecules. Cross-linking Ctx would not be a useful way to fight cholera, because it would be 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 page 82B*]. 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 IBM's "Millipede" drive [see "The Nanodrive Project," by Peter Vettiger and Gerd

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.

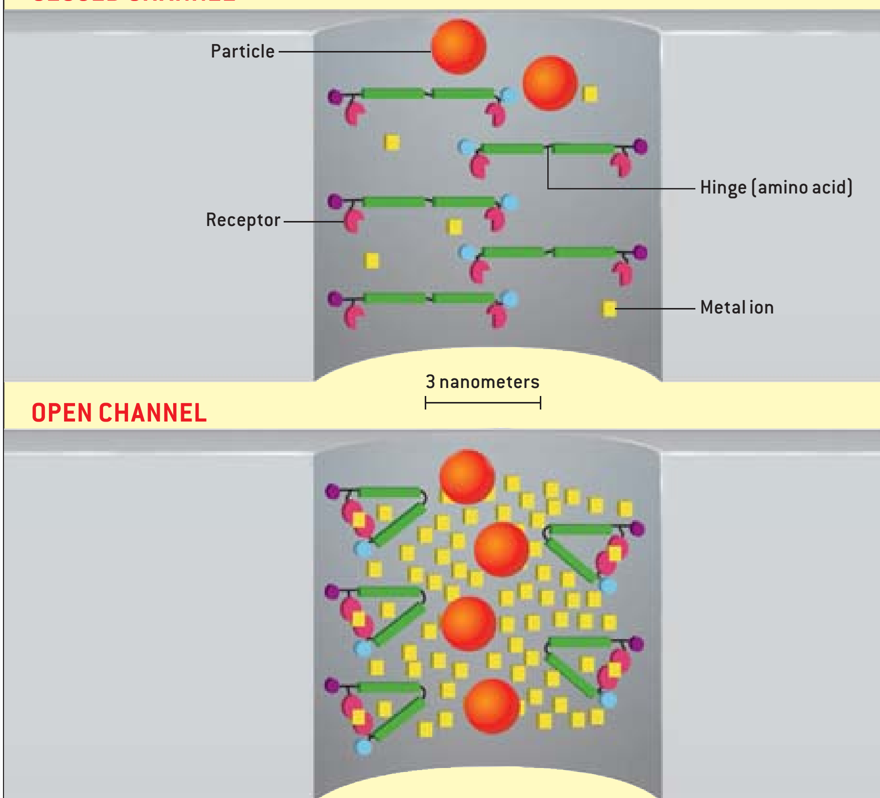


CHRISTIAN E. SCHAFMEISTER

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 (*red*), causing the actuators to fold over, opening the channel (*bottom*).

CLOSED CHANNEL



Binnig; SCIENTIFIC AMERICAN, January 2003]. 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 can develop artificial enzymes that break down plant cellulose into ethanol or that use light energy to combine water and carbon dioxide to create ethanol, it 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

WORKINGKNOWLEDGE

SATELLITE RADIO

Song Beams

Satellite radio can send the same 100 channels of music, talk and sports to you at any street corner in the nation. Yet someone next to you may receive a different set of channels, and a third person who does not subscribe cannot pick up a thing. How can the service blanket the country yet be so discriminating?

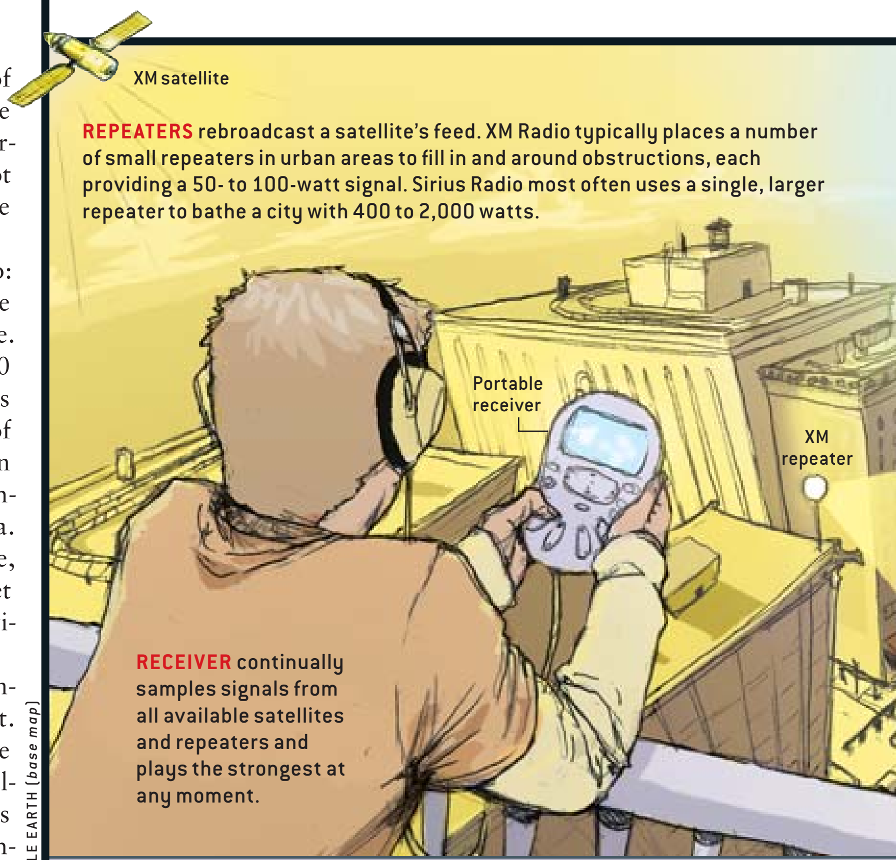
Three companies provide the world's satellite radio: XM Satellite Radio and Sirius Satellite Radio for the U.S.; and WorldSpace for Africa, Asia and Europe. XM uses two geostationary satellites and about 800 low-power ground repeaters scattered around cities where tall structures can block the satellites' line of sight [*see upper illustration*]. Sirius has three birds in highly inclined, elliptical orbits and about 100 high-power repeaters, each bathing a metropolitan area. Both architectures provide equally reliable service, says Dan Goebel, a senior research scientist at the Jet Propulsion Laboratory in Pasadena, Calif., who previously designed amplifiers for ground repeaters.

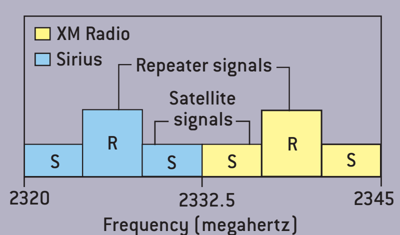
The consumer's receiver, however, "is the most innovative part of the system," Goebel points out. Whether the unit is portable or installed in a car, the antenna receives signals from all of a company's satellites and repeaters. Processors inside sample the feeds and play whichever is strongest, continuously switching as warranted.

The key to coverage anywhere, anytime, therefore lies in three levels of diversity, according to Terry Smith, senior vice president of technology at Sirius. There is spatial diversity, he explains, "because one satellite or repeater can reach a listener if another one can't." Satellites and repeaters broadcast at slightly different wavelengths, creating frequency diversity that a receiver can choose from. And the signals are sent at slight delays, creating temporal diversity—overlap that covers any fleeting dropouts.

Receivers sense all channels that a company broadcasts, but two people in adjacent driveways hear only the ones they paid for. When they subscribe, the satellite sends an activation code keyed to a receiver's unique subscription number. The code tells the receiver to block channels the subscriber has not signed up for. A neighbor cannot listen with a generic receiver because it does not hold a proper subscription number. *—Mark Fischetti*

MATT VINCENT (*top panel*); JEN CHRISTIANSEN (*bottom panel*); NASA/THE VISIBLE EARTH (*base map*)





DOWNLINK frequencies broadcast by XM and Sirius satellite transponders lie in the S-band. Each satellite spreads its 100 channels across 1,000 frequencies within four megahertz of spectrum at either end of the company's allotted range; repeaters use the middle four megahertz. Receivers recombine the frequencies into channels using spread-spectrum techniques. WorldSpace broadcasts in the L-band from 1,467 to 1,492 megahertz.

GEOSTATIONARY SATELLITES, fixed above the equator, broadcast XM Radio's full set of channels, their beams striking the continental U.S. at 30 to 40 degrees from the horizon. Sirius Radio uses three satellites in highly inclined, elliptical orbits, with a "look angle" greater than 60 degrees; each satellite, in effect, traces a figure-eight ground track (*dotted line*) and spends about 16 hours over North America, such that two are always broadcasting. WorldSpace has two geostationary satellites that shine three tightly focused, high-gain spot beams, each carrying a different set of channels for the countries it covers.

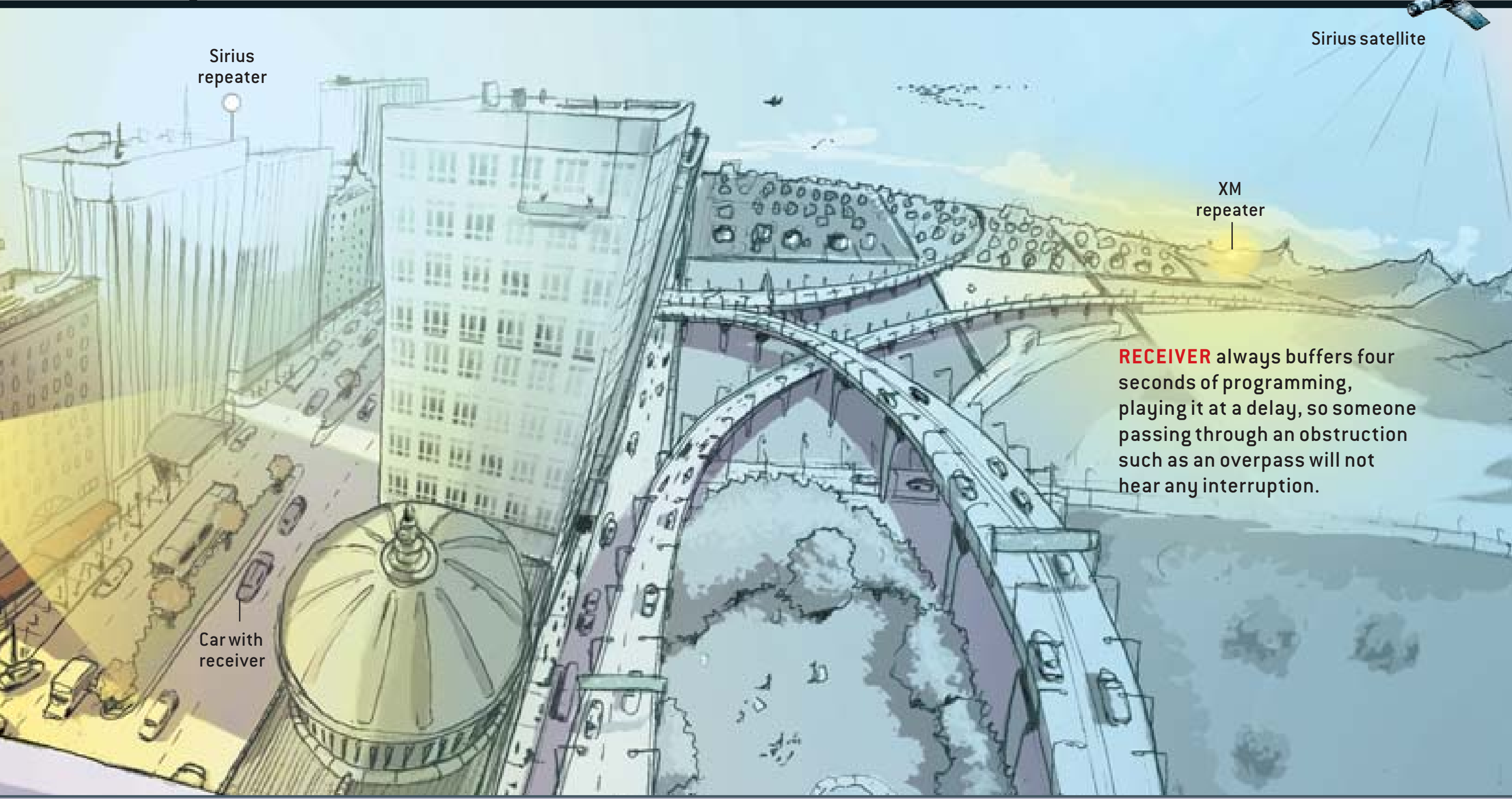
Have a topic idea? Send it to workingknowledge@sciam.com

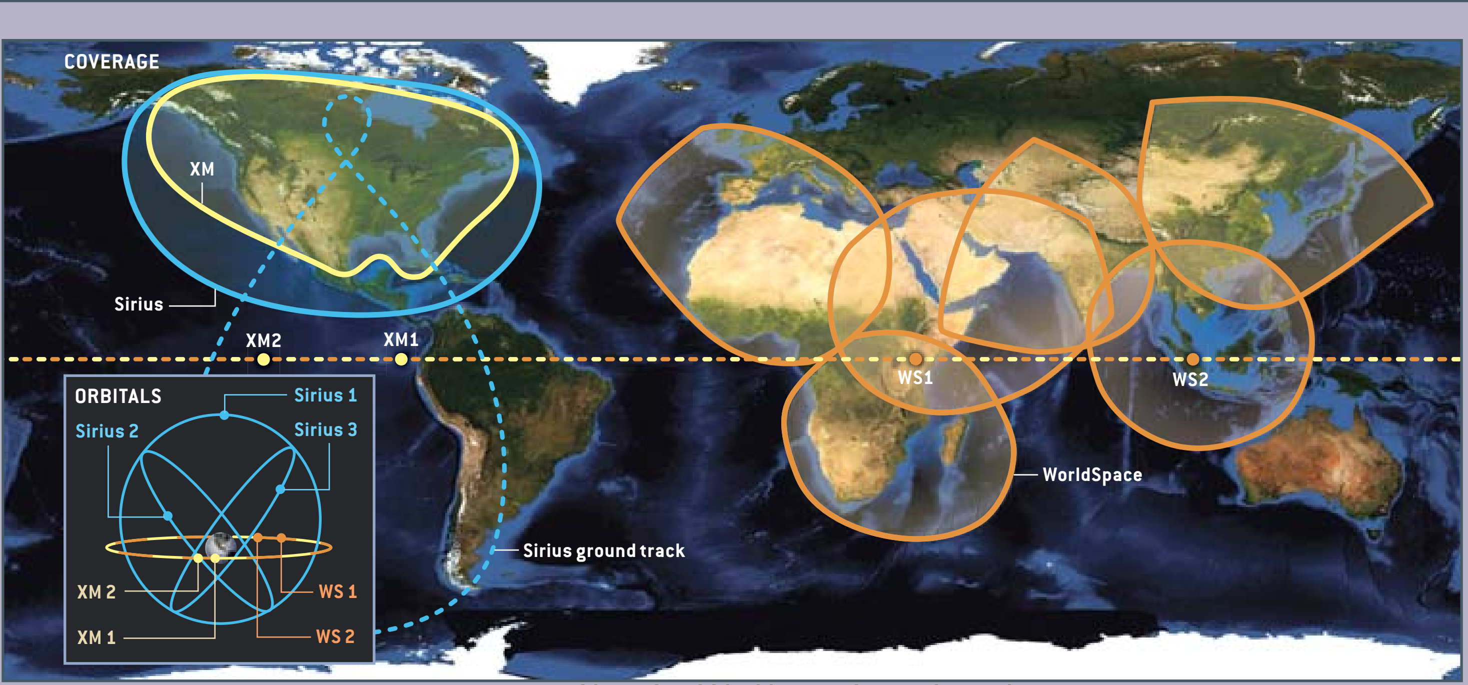
DID YOU KNOW . . .

COMPRESS: By describing the sound carried by a given channel with fewer bits of data, broadcasters can squeeze more channels (or higher-quality audio) into their slice of spectrum without altering customers' receivers. When Sirius and XM began service about five years ago, they offered around 60 channels. Better ways to compress the feed from studios up to a satellite have raised the number beyond 100, including local traffic reports. The industry continues to research how the human auditory system and brain sense sound patterns, to devise more space-saving compression algorithms.

FLOAT: Sirius's broadcast studios are on the 36th floor of a Manhattan skyscraper. The concrete floor was poured on a rubber bladder that was then inflated to isolate the slab from the building's beams. The cushion deadens vibrations from truck traffic, jackhammers and other city noise that could creep into the uplink satellite signal. Also, noise consumes signal bits but is difficult to compress because of its random nature.

IONIZE: Satellite radio and many television satellites are held steady in their orbits (against pull from the sun and the moon) by ion beams that fire twice a day for up to an hour. More than 30 communications satellites now utilize ion beams, which Dan Goebel of JPL says reduces the fuel they must carry aloft by a factor of 10 over common chemical propulsion schemes.





Power Walker

NOTHING COULD MATCH THE SEGWAY'S INITIAL HYPE, BUT HOW FAR HAS IT COME SINCE? **BY STEVEN ASHLEY**

The debut of the Segway Human Transporter in 2001 surely set the bar for anticipatory hype. Remember "Ginger"? Or was the code name "IT"? But sales of the "revolutionary," self-balancing two-wheeler never approached those of a truly paradigm-changing innovation, such as Apple's iPod. Even five years after the Segway's much ballyhooed introduction, fewer than 24,000 of them cruise the world's sidewalks, pathways, pedestrian malls, and (local laws permitting) bicycle lanes and streets. Still, engineer Dean Kamen's novel electric scooter has managed to attract a lively cult following.

Recently the Bedford, N.H., company rolled out its second-generation Segway Personal Transporter (PT), which comes in an urban/suburban model called the i2 and a beefier, wide-track cross-country version, the x2. The time seemed right to give this intriguing technology another look and perhaps to clarify its place in a wheeled-transportation spectrum dominated by cars, motorcycles, scooters, motorized carts and bicycles. I talked to Segway users, dealers and designers, then briefly drove the i2 and x2 myself.

MOBILE SECURITY FORCES on large corporate campuses or public facilities such as airports find the Segway a natural fit. It enables operatives to cover more ground and boosts their visibility as well.

MATT COLLINS

"Just step up like it's a stepladder," Ed Tsang urges, having just finished sweeping a clear path among the windblown autumn leaves strewn across the parking lot of his Segway dealership in Basking Ridge, N.J. I am mounting what vaguely resembles a modernized version of an old push-type lawn mower, except with a low-slung footpad in place of the rotating scythe blades. Further inspection confirms that the PT is more slickly designed than the original model, but the layout looks similar. I grab the T-shaped handlebar, step onto the platform very gingerly (almost as if it could somehow cut my foot) and finally board the vehicle. As if from a distance, Tsang's well-worn teaching litany just manages to penetrate my consciousness: "Relax … Don't look down; look forward … Stand up straight with your knees bent … Try not to rock back and forth …"

You are sure to start your first Segway ride a bit jumpy and unsure, seemingly primed to tip over. But after shifting about a little, you realize that all you really need do is "trust the machine," as Tsang advises. Make that leap of faith, and the gyro-stabilized device's intuitive control system does the rest, allowing you to direct it nearly effortlessly. So I stand tall and lean my body forward, and the i2 powers ahead smoothly and silently at walking speed. No doubt about it, a few moments gliding about on the uncannily responsive Segway shows that it is great fun, confirming the dealer's comment that "the Segway always exceeds initial expectations."

The trick to the Segway's uniquely satisfying ride experience is its user interface, says Doug Field, the company's chief technology officer. As in the original version, the automated control's nearly instantaneous sensory feedback mechanism relies on two safety-redundant microprocessor-based controller boards, five micromachined gyros, and a pair of tilt-sensing accelerometers that check the two-wheeler's condition 100 times a second. Then it deftly maintains balance by directing the electric motors to rotate the wheels just enough to keep the machine directly below the user's center of gravity. When you pitch forward slightly, you upset the balance, and the Segway rolls ahead a tad to keep you from falling over. Angle back some, and the machine compensates at once. Soon, standing in place is no chore at all. Tilting back a bit (even by extending your rear end a little) while traveling forward immediately brings the device to a smooth stop via a regenerative braking system that stores wheel-rotation energy in the battery.

To develop the machine's distinctive operation, Field says, Segway designers drew lessons from the familiar act of walking. Human environments are built around the assumption that people are standing and walking, he explains, adding that "walking is actually a form of controlled falling." During each step, your brain senses you are out of balance, because fluid in your inner ear shifts, so it triggers you to put your leg forward and stop the fall. Engineers conceived the Segway "as a way to graft wheels, motors and batteries onto a person so as to amplify walking in a way that mirrors walking, which is suited to our environment."

The new steering controls on the PT take this concept one step further. To turn the earlier model, the user had to rotate the left handgrip on the T-bar forward or rearward, like a motorcycle throttle. Although the skill is easy to learn, Segway engineers improved on this scheme by installing a "lean-steer" mechanism that allows users to corner by leaning into turns (shifting their weight in a slalomlike fashion) while tilting the handlebar—a much more natural motion. Push the bar all the way over, and the Segway rotates sharply in place, as one wheel advances while the other reverses.

The PT's new, larger lithium-ion battery pack (located under the footpad) can power both Segway models to a maximum speed of 12.5 miles an hour, the same as the initial versions. The i2 goes 16 to 24 miles on a charge, whereas the x2's range is 10 to 12 miles (comparable to that of previous models). Another useful upgrade is a wireless infokey fob that provides a readout for speed and distance traveled and enables a user to start a Segway remotely, control its top speed and rapidly engage the lock/security system.

As with most novel technologies, the Segway has faced a bumpy road. Service as a golf cart might seem like a surefire application, for example. But the two-wheelers have failed to prosper on the links because few golfers know how to use them, whereas anybody can drive a golf cart.

Current users come in three categories, says Barry Fulton, owner of Segway of Long Beach in California. The urban commuter wants to ride the short distance from home to work and back in business dress without having to drive and park a car. The recreational user just likes gliding around for fun or to do chores. This person is typically "a 55- to 75-year-old male with money who wants to let the kids play with it or ride it off his docked yacht." This category also encompasses some travel tourists, who, for instance, can rent a Segway from Fulton to explore the Long Beach shoreline and the HMS *Queen Mary* for $99 a day.

The third category comprises "commercial" customers, including police departments, security firms, warehouses and companies with large campuses. One of Fulton's clients uses Segways to replenish 85 delivery trucks, parked across a seven-acre plot of land, with vending-machine supplies. In the Midwest, another corporate customer is the security patrol for a large telecommunications firm that operates a 235-acre complex. Rather than walking beats, the crew Segways around the grounds. With Segways, the security chief has been able

LIGHT INDUSTRIAL applications are the focus for the Segway i2 Commercial Cargo model.

SEGWAY

to shrink his staff, because each member can cover a greater area, and the vehicles have also lengthened the careers of older workers, who were finding the constant walking increasingly difficult. The chief says that the machine's elevated driving position is helpful for handling crowds—each agent sees the surroundings better and is more visible to bystanders.

Despite the improved technology and growing public familiarity, the Segway remains a niche product. Why the resistance? The consensus among sellers and users alike comes down to sticker shock. At $4,000 to $5,500 apiece (a significant fraction of which pays for the batteries), the Segway is simply too pricey for those without a compelling need or use for it. As one observer notes, "Most potential buyers consider it as a fractional car, and at $5K, it just doesn't supply enough utility." Cut that price in half, and sales will spike, dealers claim, but there is little prospect that battery costs will drop precipitously anytime soon.

Gliding around on a Segway is a delight, but you have to wonder whether the technology is a solution chasing a problem. Still, it took a while for the Frisbee to catch on, but the unusual flying disk eventually found its place in the market. Maybe it's only a matter of time until the same thing happens to the Segway. SA

Why Aren't More Women Physicists?

TWO BOOKS LOOK FOR ANSWERS IN THE LIVES OF A FEW WHO SUCCEEDED **BY KAREN A. FRENKEL**

LA DAME D'ESPRIT: A BIOGRAPHY OF THE MARQUISE DU CHÂTELET
by Judith P. Zinsser
Viking, 2006 ($24.95)

OUT OF THE SHADOWS: CONTRIBUTIONS OF TWENTIETH-CENTURY WOMEN TO PHYSICS
edited by Nina Byers and Gary Williams
Cambridge University Press, 2006 ($35)

During the past 40 years, study after study has addressed why more women do not become scientists. The question is most apt for physics. Advanced physics degrees awarded to women have always lagged, hitting a nadir at under 5 percent from the 1950s to the 1970s. Progress has been made since; in 2003, 193 women (17.9 percent) received a Ph.D. in physics, according to the National Science Foundation. But physics still trails the other sciences.

The flip side of the question is: Why and how did those few prominent female physicists succeed? Historian Judith P. Zinsser's *La Dame d'Esprit* and the profiles of women physicists in *Out of the Shadows* unveil the scintillating lives of women who overcame discrimination and made major contributions that went largely unacknowledged.

Zinsser debunks legends about the life and loves of the vivacious and unorthodox 18th-century French noblewoman Marquise du Châtelet and analyzes her contribution to physics. Du Châtelet (1706–49) gained notoriety as Voltaire's lover. When government censors deemed his literature subversive, she offered him shelter at her château in Cirey. There they collaborated and commented on each other's manuscripts. Du Châtelet transformed a room into a laboratory for experimenting in the "natural sciences," as physics was then called, and built a small theater for Voltaire's plays.

According to Zinsser, du Châtelet's mother recognized her daughter's prodigious curiosity and fostered it by permitting her to question elders even though that was considered impolite. Other biographers report that du Châtelet's father considered her too plain for marriage, so he educated her with the best teachers.

Du Châtelet was in her 20s when she and Voltaire became lovers in 1733, and several biographers have highlighted their affair at the expense of the marquise's accomplishments. To trace her development as a mathematician and physicist, Zinsser plumbed archives all over Europe and the U.S.

Voltaire did introduce du Châtelet to the ideas of Descartes, Leibniz and Newton, which were being debated at the French Academy of Sciences, from which women were barred. When the academy held a competition for best essay on the nature of fire, the interpretations of the lovers differed, and both entered the contest—she under an assumed name. Leonhard Euler won, but the academy considered Euler's, Voltaire's and du Châtelet's essays so outstanding that it published all three. The couple then collaborated on *The Elements of the Philosophy of Newton,* introducing Newton's theories to readers with no background in higher mathematics.

As Zinsser interprets it, du Châtelet and Voltaire's relationship faltered as her scientific expertise exceeded his, as it did in her next book, *Institutions of Physics.* Not merely explaining the metaphysical theories of Descartes, Leibniz and Newton, du Châtelet synthesized them, Zinsser says, whereas academy members, including one of her mentors, only spoke of doing so. She explored a broad sweep of concerns, from "how one can know anything, to the origins of

MARQUISE DU CHÂTELET holds a mathematical compass, rather than the conventional fan. Portrait by Marianne Loir, circa 1740.

RÉUNION DES MUSÉES NATIONAUX/ART RESOURCE, NY the universe and the role of the divine."

Du Châtelet also devoted a chapter to the scientific method, explaining hypotheses as "probable propositions" that could be rejected because of one contradiction. One confirmation, however, was insufficient to prove a hypothesis. Rather "each non-contradictory result would add to the probability of the hypothesis and ultimately … we would arrive at a point where its 'certitude' and even its 'truth,' was so probable that we could not refuse our assent." If antievolutionists today understood the scientific method as well as this woman of the Enlightenment did, perhaps they wouldn't glibly say, "It's only a theory."

Du Châtelet's next major project was translating Newton's *Principia Mathematica* from Latin into French; she was the first person to undertake this task, and hers remains the definitive text. On its publication, an anonymous editor praised her commentary for being clearer than the original. In it, she walks us through Newton's three proofs of attraction, which incorporate his theory of fluids, the shape of the earth, its relation to the moon and sun, and the parabolic orbits of comets. Along the way, she updates Newton with new data from the first expedition to Lapland and Daniel Bernoulli's study of tides.

Several society women never forgave du Châtelet for winning Voltaire's devotion and described her as crude and ungainly. They ridiculed her when she fell madly in love with a younger poet and became pregnant at the age of 43. She died in 1749 shortly after giving birth. Despite her turbulent personal trials, du Châtelet had managed to find great joy in her lifetime and even wrote an essay on her search for happiness. In it, she explained that she sought fame as a man would but accepted having to find it in other ways, recommending cultivation of "mind and understanding."

Out of the Shadows celebrates the achievements of 40 *femmes d'esprit,* women physicists who worked in the 20th century. Editors Nina Byers and Gary Williams selected them from 200 nominations received after they posted a call on the Web. The book is filled with unexpected gems. Herta Ayrton stopped streetlamps from sputtering by perfecting the electric arc. Her patents transferred that technology to film projectors, and movies were called "flicks" because of the flickering arc. Dorothy Crowfoot Hodgkin persisted in solving the structure of insulin for 30 years, finally succeeding in 1970, when computers could perform the needed calculations.

One surprise is Freeman Dyson's essay on Mary L. Cartwright, a pure mathematician whom he credits with the foundations of chaos theory. He corrects the record as James Gleick tells it in *Chaos: Making a New Science,* which begins with the discoveries of Edward Lorenz in the 1960s. It "barely mentions the earlier discoveries of Cartwright and Littlewood (her collaborator)," Dyson writes, although he is careful to praise the book as an excellent popular account.

One hopes that professors will bring these female physicists' discoveries to light in their lectures, thereby demonstrating role models to female students. Perhaps *all* those attending—men and women—will perfect their "minds and understanding" so that none remain in the shadows. SA

Karen A. Frenkel is a science writer and documentary producer living in New York City.

Editors' note: A recent report from the American Psychological Association, Why Aren't More Women in Science? *edited by Stephen J. Ceci and Wendy M. Williams, looks in depth at barriers to women in science.* Passionate Minds: The Great Love Affair of the Enlightenment, *by David Bodanis (Crown Publishers, 2006), recounts the personal rather than the scientific aspects of the Marquise du Châtelet's story.*

The Marquise du Châtelet on Women's Education

"I FEEL THE FULL WEIGHT OF THE PREJUDICE which so universally excludes us from the sciences; it is one of the contradictions in life that has always amazed me, seeing that the law allows us to determine the fate of great nations, but that there is no place where we are trained to think. . . . Let the reader ponder why, at no time in the course of so many centuries, a good tragedy, a good poem, a respected tale, a fine painting, a good book on physics has ever been produced by women. Why these creatures whose understanding appears in every way similar to that of men, seem to be stopped by some irresistible force, this side of a barrier. Let the people give a reason, but until they do, women will have reason to protest against their education. . . .

If I were king . . . I would redress an abuse which cuts back, as it were, one half of human kind. I would have women participate in all human rights, especially those of the mind . . . . The new education would greatly benefit the human race. Women would be worth more and men would gain something new to emulate. . . . I am convinced that either many women are unaware of their talents by reason of the fault in their education or that they bury them on account of prejudice for want of intellectual courage. My own experience confirms this. Chance made me acquainted with men of letters who extended the hand of friendship to me. . . . I then began to believe that I was a being with a mind."

RICHARD HUTCHINGS *Corbis*

SOURCE: The Marquise du Châtelet, from the preface (written about 1735) to her translation of Bernard Mandeville's *The Fable of the Bees.*

The Kindest Cut

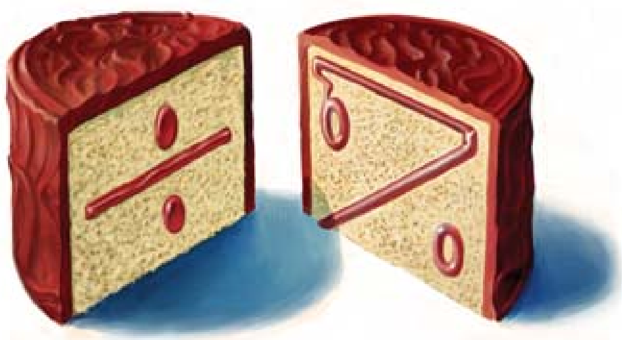
ALWAYS DO THE MATH BEFORE YOU DIVVY **BY STEVE MIRSKY**

In Mel Brooks's *The Producers*—also known as the Enron business plan, as noted in this space in October 2002—Max Bialystock and Leo Bloom conspire to sell 25,000 percent of the Broadway show *Springtime for Hitler.* In a moment of introspection, Max asks, "How much percentage of a play can there be altogether?" To which Leo gently responds, "You can only sell 100 percent of anything." Until recently, I accepted this statement as fact.

A mathematician, a political scientist and an economist walk into a bar. Wait, that's not right. A mathematician, a political scientist and an economist recently wrote a paper—amazingly, that *is* right—in which they point out that, under special circumstances, two people can split something up and both feel like they got more than half. No, the entire mirror factory isn't filled with smoke.

The paper, which appeared in the December issue of *Notices of the American Mathematical Society,* is entitled "Better Ways to Cut a Cake." The report does not deal with knife-sharpening technology. What it does deal with is the theory and method behind slicing up an object to maximize the satisfaction of those parties, possibly at a party, who will then receive the slices. "We use cake as a metaphor for dividing a heterogeneous divisible good, an item that people may have different preferences for," explains the mathematician, Michael Jones of Montclair State University. So the cake could be a stand-in for a tract of land that's part forest and part seashore. Or the cake could be an apartment that has small rooms with views and large windowless rooms. Or the cake could be a chicken with white meat and dark meat.

One thing the cake can't be is a pie. In fact, mathematics has a rich (but still light and moist) history of cake-cutting theory, which looks at ways to divide a space with a cut that goes completely across the space. And there is a literature



of pie-cutting theory, in which cuts begin at the center and travel radially outward. So it takes two cuts to completely sever a pie. Although, mathematically, as Jones points out, "In a sense, you can think of a cake as a pie in which one cut has already been made." Which sounds like something a wise man on a mountaintop would tell Betty Crocker.

Anyway, back to the Faber College homecoming parade. In other words, cut the cake. Traditionally, if two people are splitting a cake, the method is simple and was already ancient when Marie Antoinette was providing cake counseling to the masses: one slices, the other chooses. The slicer therefore wants to make the division equal, knowing he'll get stuck with the littler piece if he botches the job.

But this system can break down with certain cakes. "For example," Jones says, "if a cake is half chocolate and half vanilla, and one person likes chocolate a lot and the other person is indifferent, then there's a way to have both people, in their opinions, receive more than half the cake." The procedure, which involves equations that you are welcome to review at home over coffee and Danish, gets fairly complex, especially if more than two people are involved in the cutting. But you can see, in the chocolate-vanilla two-person example, that the chocolate lover will feel more than half-satisfied if he gets, say, 80 percent of the chocolate half, despite it being only 40 percent of the entire cake. Meanwhile his flavor-impartial buddy will be more than half-satisfied by getting the remaining 60 percent of the entire cake. And the cake maker will have an economic motivation to complicate his cakes and hike his prices.

The cake-cutting theory could find application in, for example, land division negotiations. It also indirectly pinpoints exactly where the Enron boys' scheme was half-baked. Because it turns out that you really can have more than 100 percent of the cake. But you have to at least *have* a cake. SA

ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

Why do cats have an inner eyelid, and what does it do?

—S. Huang, Brooklyn, N.Y.

Veterinary ophthalmologist Paul Miller of the University of Wisconsin–Madison offers this explanation:

The inner eyelid of cats is no biological curiosity—it plays a critical role in maintaining the health of the eye surface. In fact, this third eyelid—so called because it complements the upper and lower outer lids—is so important that among mammals and birds it is the norm. Species lacking one, such as humans and most other primates, are the true oddities in nature.

The third eyelid—more formally termed the palpebra tertia—is anatomically complex. It is a fold of tissue covered by a specialized mucous membrane (the conjunctiva). A dense population of lymphoid follicles populates the side in contact with the surface of the eye and the tear film, a thin layer of liquid. These follicles function as the lymph nodes of the eye, protecting the surface against invasion by microorganisms.

Between the two layers of conjunctiva is a dense T-shaped cartilage plate—the crossbar of which stiffens the free edge of the third eyelid—that is curved to conform to the surface of the cornea (the clear covering on the front of the eye). An accessory lacrimal gland, which produces a substantial portion of the tear film, surrounds the stem of the T.

When the cat is alert, the bulk of the third eyelid is hidden within the eye socket; only a small part is visible in the inner corner of the eye. When the feline is relaxed, asleep or blinking, however, a set of skeletal muscles retracts the eyeball, causing the third eyelid to move across the ocular surface and completely cover the cornea.

In doing so, the third eyelid acts much like a windshield wiper blade, removing debris from the surface and redistributing tears over the cornea. It is also believed to help protect the cornea from injury as cats move through tall grass or capture prey. The presence of an accessory tear gland allows for even greater rinsing of the ocular surface than occurs in primates, and the third eyelid also seems to hold the tear film against the cornea better than the outer eyelids do by themselves.

Although no one knows why humans lack a third eyelid, it is possibly because we do not typically capture prey by biting (as would a cat) or eat by rooting through vegetation (as would a horse). Thus, there may be no survival advantage for us in having this extra measure of protection for the eye.

Why do veins pop out when we are exercising?

Mark A. W. Andrews, professor of physiology at the Lake Erie College of Osteopathic Medicine, replies:

Bulging veins during exercise result not from some worrisome increase in venous blood pressure or volume but from processes pushing veins under the skin toward the surface.

At rest, the heart pumps blood from its left ventricle into the high-pressure arteries, where systolic blood pressure, the highest pressure exerted in these vessels, is normally recorded around 120 mmHg (millimeters of mercury), and diastolic pressure, the lowest such pressure, is recorded around 80 mmHg.

When exercise begins, the heart's rate and strength of contraction increase, quickly pumping blood into the arteries. Systolic blood pressure increases linearly with exercise intensity, rising to nearly 200 mmHg during high-intensity aerobic activity (and to more than 400 mmHg during weight lifting). Diastolic pressure and pressure in the veins, which return blood to the heart, actually tend to decrease during aerobic exercise (rising somewhat during weight lifting).

The rise in arterial blood pressure during exercise forces plasma fluid otherwise resting in the capillaries—the smallest blood vessels, which nourish and remove waste material from active cells—out through the thin vessel walls and into muscles and the compartments surrounding them. This process causes the swelling and hardening of muscles noticed during exercise. As a result of the muscle swelling, veins near the skin are pushed toward the surface, giving the appearance of bulging. SA

For a complete text of these and other answers from scientists in diverse fields, visit www.sciam.com/askexpert

MATT COLLINS