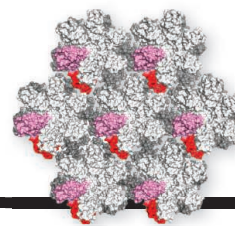# Science

$10
3 JULY 2015
sciencemag.org

AAAS

Floral fragrance
The biochemistry behind rose scent
pp. 28 & 81

CONTENTS

3 JULY 2015 • VOLUME 349 • ISSUE 6243

95 & 99
Native HIV-1 capsid protein hexamers

30 & 46
Squared off for stability

35 & 66
A nickel-pinching enzyme

ON THE COVER

In roses (*Rosa* x *hybrida*), scent is a highly desirable trait. Monoterpenes are the major rose scent compounds, representing up to 70% of the fragrance content in some varieties. Combined transcriptomic and genetic approaches show that monoterpene biosynthesis in roses differs from what is known in other plants and involves a specialized Nudix hydrolase protein localized in the cytoplasm of rose petal cells. See pages 28 and 81. *Photo: © Beneta Markaityte/National Geographic My Shot/ National Geographic*

PHOTOS: (ELEPHANT) © ART WOLFE/WWW.ARTWOLFE.COM; (SEAHORSE) COURTESY OF DOMINIQUE ADRIAENS, UGENT

Science

The beyond-two-degree inferno

I n the history of humankind, there is a dearth of examples of global threats so far-reaching in their impact, so dire in their consequences, and considered so likely to occur that they have engaged all nations in risk mitigation. But now with climate change, we face a slowly escalating but long-enduring global threat to food supplies, health, ecosystem services, and the general viability of the planet to support a population of more than 7 billion people. The projected costs of addressing the problem grow with every year that we delay confronting it. In recognition of the shared risks we face and the collective action that will be necessary, an international meeting of stakeholders will convene in Paris next week (www.commonfuture-paris2015.org), ahead of the United Nations Climate Change Conference (COP21) in December, to discuss solutions for both climate mitigation and adaptation.

The time for debate has ended. Action is urgently needed. The Paris-based International Energy Agency recently announced that current commitments to cut CO_2 emissions [known as Intended Nationally Determined Contributions (INDCs)] from the world's nations are insufficient to avoid warming the entire planet by an average of more than 2°C above the preindustrial level. This is a target viewed as the boundary between climate warming to which we can perhaps adapt and more extreme warming that will be very disruptive to society and the ecosystems on which we depend (see Gattuso *et al.* on p. 45). To set more aggressive targets, developed nations need to reduce their per-capita fossil fuel emissions even further, and by doing so, create roadmaps for developing nations to leapfrog technologies by installing low-CO_2–emitting energy infrastructure rather than coal-fired power plants as they expand their energy capacity.

The European Union (EU) is leading the way with the most aggressive INDC target for reduction: a cut of 40% below 1990 levels of CO_2 emissions by 2030. The United States has pledged reductions of 26 to 28% below 2005 levels by 2025, with California independently choosing to match the EU's more ambitious goal. All eyes are on China and India, two of the largest total emitters of CO_2, both of which have yet to submit their proposed INDCs in advance of COP21. Unfortunately, Piyush Goyal, India's Minister of State for Power, Coal, and New and Renewable Energy, intends to double his nation's coal production by the year 2019 to meet domestic energy requirements. China appears to be taking the opposite track, recognizing its vulnerability to climate change and investing heavily in renewable energy.* Like California, China is betting that good environmental policy will make for good fiscal policy by being in the vanguard of the clean energy economy.

I applaud the forthright climate statement of Pope Francis, currently our most visible champion for mitigating climate change, and lament the vacuum in political leadership in the United States. This is not the time to wait for political champions to emerge. Just as California has decided to go it alone, every sector (transportation, manufacturing, agriculture, construction, etc.) and every person need to do whatever is possible to reduce carbon pollution by conserving energy, adopting alternative energy technologies, investing in research, and capturing CO_2 at the source.

In Dante's Inferno, he describes the nine circles of Hell, each dedicated to different sorts of sinners, with the outermost being occupied by those who didn't know any better, and the innermost reserved for the most treacherous offenders. I wonder where in the nine circles Dante would place all of us who are borrowing against this Earth in the name of economic growth, accumulating an environmental debt by burning fossil fuels, the consequences of which will be left for our children and grandchildren to bear? Let's act now, to save the next generations from the consequences of the beyond-two-degree inferno.

– **Marcia McNutt**

"where [would]...Dante...place all of us who are borrowing against this Earth...?"
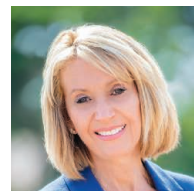
Marcia McNutt
Editor-in-Chief
Science *Journals*

*www.bbc.com/news/science-environment-33040965.

10.1126/science.aac8698

PHOTOS: (TOP RIGHT) STACEY PENTLAND PHOTOGRAPHY; (INSET) © GIAN BERTO VANNI/CORBIS

NEWS

> **" This is a tough day. This isn't where I wanted to be on a Sunday afternoon. "**
> **NASA Associate Administrator of Human Exploration Bill Gerstenmaier**
> after an unmanned SpaceX rocket carrying International Space Station (ISS) supplies exploded after launch, the third ISS resupply mission to fail in a year.

IN BRIEF

Protesters arrested trying to halt telescope's construction

A Thirty Meter Telescope protester chants during his 24 June arrest on Hawaii's Mauna Kea.

An attempt to restart construction on the Thirty Meter Telescope (TMT), intended to be one of the world's largest telescopes, was stalled on 24 June after state authorities escorting construction vehicles clashed with protesters blockading the road to the summit of Hawaii's Mauna Kea volcano. The protesters, who have said the $1.4 billion TMT would desecrate sacred land, blockaded the road, culminating in the arrests of 11 men and women, including several protest organizers. Officers from Hawaii's Department of Land and Natural Resources and construction workers turned back from the summit shortly after noon after finding the road obstructed by boulders. Hours after posting bail, the protesters headed back to Mauna Kea with plans to continue their blockade. In a statement calling the arrests "unfortunate," Mike McCartney, chief of staff to Hawaii Governor David Ige (D), said that the boulders in the roadway represent "a serious and significant safety hazard and could put people at risk." He said the construction would be on hold while teams cleared the roadway and the state reassesses how to proceed.
http://scim.ag/TMTarrests

AROUND THE WORLD

Ebola back in Liberia

MONROVIA | More than 7 weeks after being declared free of Ebola, Liberia has reported a new infection with the virus, raising worries about a resurgence. The body of a 17-year-old boy who died at home on 28 June tested positive for the deadly virus, Liberian deputy minister of health Tolbert Nyenswah announced on 29 June. The boy had been ill for a few days; his body was buried using appropriate safety precautions, Nyenswah said. His contacts are being monitored, and Nedowian, the village where he lived, has been put under quarantine. It's unclear how the patient got infected. Nedowian is close to Liberia's capital, Monrovia, and far from the borders with Sierra Leone and Guinea, where Ebola still occurs. The outbreak in West Africa has sickened more than 27,000 people and killed more than 11,000, according to the World Health Organization.

Conflict of interest overhaul

WASHINGTON, D.C. | The Smithsonian Institution is overhauling its ethics policies in the wake of a controversy surrounding one of its researchers. This past February, environmental advocates revealed that aerospace engineer Willie Wei-Hock Soon of the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, an outspoken critic of mainstream climate science, had not revealed to several journals that he had received funding from a major energy company. Smithsonian officials launched two reviews, and on 26 June announced four reforms. They include creating new processes that will allow officials to review all funding agreements for potential conflicts, as well as to collect financial disclosure statements from some 150 researchers not covered by existing policies. The institution also will require all of its researchers to disclose funding sources to journals, regardless of journal policies.

Supreme Court axes mercury regs

WASHINGTON, D.C. | U.S. government efforts to reduce the health risks posed by

Pink salmon in the Campbell River in British Columbia.

Acidification hurts freshwater fish

Rising CO_2 in the world's waters isn't just a threat to saltwater fish: Pink salmon raised in water infused with elevated CO_2 showed stunted growth, became more fearless in dangerous situations, and had a diminished sense of smell, scientists reported 29 July in *Nature Climate Change*. Scientists at the University of British Columbia in Vancouver, Canada, observed the differences after raising pink salmon in tanks with water containing the equivalent of 450 parts per million of CO_2 (a little more than current atmospheric levels), 1000 parts per million, or 2000 parts per million. The study is the first to identify these effects in a freshwater species, and the results "sound a warning," wrote ecologist Philip Munday of the ARC Centre of Excellence for Coral Reef Studies in Townsville, Australia, in a commentary accompanying the study. Munday has previously found similar effects on ocean fish. Approximately 40% of all fish species spend at least part of their lives in fresh water. **http://scim.ag/FWfishacid**

mercury from U.S. coal power plants will have to wait a bit longer. On 29 June, the U.S. Supreme Court struck down a 2011 Environmental Protection Agency (EPA) regulation—decades in the making—that aimed to reduce emissions of the toxin. On a 5 to 4 vote, the justices ruled that EPA had erred by failing to fully consider the costs of the rule early in its regulatory process. The agency had argued that it didn't need to consider all the costs and benefits until later, when it set specific standards. EPA said in a statement that it is disappointed with the ruling, but noted that most U.S. coal plants "are already well on their way to making emissions reductions." **http://scim.ag/suprmerc**

Bill would limit embryo editing

WASHINGTON, D.C. | A spending bill in the U.S. House of Representatives could restrict research on making heritable genetic changes to human embryos. The draft bill, released on 17 June, includes language that would bar the Food and Drug Administration (FDA) from reviewing such research, which can include mitochondrial replacement therapy to prevent rare inherited diseases or using genome-editing technologies such as CRISPR-Cas9 on embryos. The amendment is "the simplest way to slow down the FDA on this so everyone has a chance to review the ethics," says Representative Andy Harris (R–MD). But some scientists argue that the measure is excessive and that local Institutional Review Boards already provide ethical oversight. Preventing FDA from reviewing clinical trial proposals to evaluate potential cures for genetic

BY THE NUMBERS

70

Total number of impact craters larger than 6 kilometers in diameter anywhere on Earth (*Earth and Planetary Science Letters*).

35%

Amount of marine protected areas exposed to light pollution from ships and coastal development in 2012 (*Conservation Letters*).

42%

Average increase in fine particulate matter (PM2.5) across the United States on the evening of 4 July relative to days before and after (*Atmospheric Environment*).

diseases "would set a new standard for congressional stupidity and inhumanity," says Henry Miller, a physician and molecular biologist at Stanford University in Palo Alto, California. The committee is expected to vote on the bill in July.

FINDINGS

Ancient bobcat had human burial

About 2000 years ago, a group of Native American traders and hunter-gatherers buried something unusual in a sacred place. In the outer edge of a large mound of dirt in what is today southern Illinois, these Hopewell people interred the body of a young bobcat. Just a few months old and wearing a necklace of bear teeth and marine shells, the bobkitten is the only animal ever found in these funeral mounds, which were normally reserved

Ancient Native Americans buried these bone pendants and shell beads together with the bobcat.

for humans. Its burial is also the only known ceremonial burial of a wild cat in the archaeological record, researchers contend this month in the *Journal of Midcontinental Archaeology*. They say the

villagers may have brought the animal in from the wild and tamed it, transforming it into a member of their own society. If so, the find could shed light on how cats, dogs, and other creatures were domesticated.

Still looking for Amelia

On 2 July 1937, aviator Amelia Earhart and navigator Fred Noonan vanished over the Pacific Ocean during an attempt to fly around the world. Partial skeletal remains found in 1940 on Gardner (or Nikumaroro) Island in Kiribati were suspected to be Earhart's until an investigation by British officials that year concluded the bones belonged to a middle-aged male. In 1998, a private historic group had the British analysis re-evaluated and discredited. But last month in *The Journal of Archaeological Science*, Sydney University in Australia anthropologist Richard Wright and biological anthropologist Pamela Cross of Bradford University in the United Kingdom reaffirmed the original findings, noting that the historic group had "shamefully denigrated" the expertise of David Hoodless, who conducted the key 1940 analysis, Wright says. Furthermore, they found, the 1998 authors had merely used four cranial measurements—not sufficient to determine ancestry.

NEWSMAKERS

Misconduct report disputed

Surgeon **Paolo Macchiarini** has responded to a May report that concluded he was guilty of scientific misconduct as part of his clinical testing of artificial tracheas. The report, conducted by an independent investigator and requested by the Stockholm-based Karolinska Institute after whistleblowers lodged complaints, concluded that papers Macchiarini and his colleagues published did not accurately describe the condition of patients (*Science*, 29 May, p. 954). In his 23-page response, Macchiarini says that he has gathered additional clinical records that corroborate his description of the patients. The additional records are not public, however, because of patient confidentiality. Thirty people, including other researchers involved in the work as well as those who originally brought the misconduct allegations, have filed responses; a Karolinska spokesperson says the submissions total nearly 1000 pages. Karolinska's vice chancellor will assess the responses and decide how to respond to the report. http://scim.ag/Maccresp

Urgenda's lawyers celebrate after the ruling.

Dutch court orders government to cut more CO$_2$

A court has ordered the Dutch government to dramatically intensify its fight against climate change. Ruling in a civil case brought against the government by environmental group Urgenda, the district court in The Hague ruled that by 2020, the Netherlands must cut CO$_2$ emissions by 25% from 1990 levels. Current government policies would lead to a 17% reduction. The case framed global warming as a human rights violation that the Dutch government must do more to prevent. Environmental groups hailed the ruling as a legal landmark that could inspire similar action elsewhere. But the court didn't specify which measures the government must take to meet the target, and the verdict immediately triggered discussions about whether a 25% reduction in 5 years is feasible and whether it might hurt the Dutch economy. The government, which was also ordered to pay Urgenda's legal bills, estimated at €13,522, has not said whether it will appeal. **http://scim.ag/Dutchclimch**

Archaeologists have unearthed ancient tools, jewelry, and even bedding left by early humans in South Africa's Sibudu Cave.

ARCHAEOLOGY

Development threatens home of early humans

A rare record of ancient culture, left by early *Homo sapiens* in a South African cave, is at risk

By **Michael Balter**

Early humans were so drawn to South Africa's Sibudu Cave, a rock shelter chiseled into a cliff face high above a river, that they lived there nearly continuously for tens of thousands of years. They left the first recorded evidence of many crucial technologies—including bedding, bows and arrows, and snares to catch small animals—and their artifacts offer crucial clues as to how early human culture developed. But now, 21st-century humans seek to live next door to the formerly secluded cave, and archaeologists say a priceless record of our ancestors is in danger.

A huge industrial and housing development is planned right next to Sibudu, which is located in KwaZulu-Natal province about 40 miles north of Durban. Within the last few weeks, archaeologists and their supporters have filed two appeals against a decision by provincial officials to approve the 621-hectare development. The plans call for heavy and light industry, two schools, and a mixture of houses and apartments comprising about 2700 housing units. The project envisions only a 200- to 300-meter "buffer zone" between the structures and Sibudu Cave itself, a proximity that archaeologists say could be disastrous; the developer's land extends to about 20 meters from the cave.

"In fact, there would be no real buffer zone," says Lyn Wadley of the University of the Witwatersrand, Johannesburg, in South Africa, who led the first major excavations at Sibudu starting in 1998. She fears that residents gathering firewood along the nearby river or teenagers seeking a hangout would trample the cave's fragile stratigraphic layers.

Archaeologists say that Sibudu is one of the world's most important sites for understanding how early *Homo sapiens* developed culture and technology; the site is under consideration as a UNESCO World Heritage Site and has been nominated for South African national heritage status. Its thick deposits reveal one of the longest and most complete sequences of prehistoric human occupation anywhere in the world. The strata cover a period between 38,000 years to 77,000 years ago and may extend as far back as 100,000 years. Excavations so far have uncovered everything from sophisticated stone and bone tools to symbolic artifacts such as beads made of seashells, and have shown that early humans had cognitive abilities such as planning well in advance. (See *Science*, 10 June 2011, p. 1260, and http://scim.ag/sibudubeds.)

Research at Sibudu, which was discovered in the 1960s but not fully excavated until the 1990s, has spawned more than 100 peer-reviewed publications, including one just last week in *PLOS ONE* by current dig director Nicholas Conard, an archaeologist at the University of Tübingen in Germany. That paper reported three shifts in stone tool manufacturing strategy at Sibudu about 58,000 years ago, demonstrating the cultural flexibility of the humans there during this time. The site "gives us a unique picture of the cultural evolution of our species in a crucial moment of its development," says Francesco d'Errico, an archaeologist at the University of Bordeaux in France. "It

Too close for comfort?

A planned development sits right next to Sibudu Cave.

still has huge potential if it is preserved."

So when the owner of the vast tract of land adjacent to Sibudu, KDC Projects and Developments of Ballito, South Africa, announced plans in 2010 to develop their site, Wadley and others fought hard to stop it. Much of the 621 hectares for the development is currently planted with sugarcane, but the housing units and factories would bring a large influx of people—some 1800 residing in the section right next to Sibudu—and give a huge boost to the now sparse vehicle traffic in and around the site.

But the project passed an environmental impact assessment earlier this year, and was then quickly approved in April by KwaZulu-Natal's Department of Economic Development, Tourism and Environmental Affairs. "I wasn't able to convince any of these people that Sibudu was important and that the development will be a threat to it," Wadley laments. One KwaZulu-Natal official, who asked not to be identified, comments that "it is really difficult in this country to stop development."

Charlotte Mbali, an educator and filmmaker at the University of KwaZulu-Natal in Durban and an activist in a support organization called Friends of Sibudu (FOS), says provincial officials "clearly have not been paying attention to the archaeologists." Mbali, Wadley, and others are asking for a much larger conservation zone around the cave. FOS has proposed creating a 51-hectare tourist park around Sibudu that would include a visitor center and heritage trails (see www.sibudu.net).

This would require KDC to sell some of its land to a trust recently set up by Conard and others to protect the site. But Patrick Conway, a co-owner of KDC and the operations manager for the project, says KDC is not interested in selling. "We have enormous respect for the cave and its importance," Conway says, saying that the planned buffer zone is sufficient to protect the site. He stresses that Sibudu's supporters "need to take responsibility" for the security of the site. "It's not on our land. They should fence it. They haven't done anything to protect it," he says.

Wadley counters that in practical terms, the buffer zone is really only 200 meters at most. And Conard notes that until now, Sibudu "was off the beaten track and did not need protection" beyond what was provided by two caretakers who live just across the river.

"The local residents and the archaeologists have a long-term relationship based on trust and mutual respect," Conard says. "That will be hard, if not impossible, to maintain" if new people move in "who do not care that Sibudu is a place of the ancestors." ∎

BIOMEDICAL RESEARCH

Spending bills put NIH on track for biggest raise in 12 years

House and Senate also want to resurrect canceled National Children's Study

By **Jocelyn Kaiser**

At last, U.S. biomedical researchers may be getting some relief. Last week, a Senate panel approved a $2 billion increase for the National Institutes of Health (NIH) in 2016, or a 6.6% raise, to $32.1 billion. Although a House of Representatives panel has approved only about half as large an increase, Congress appears to be on track to give NIH its first major budget boost in more than a decade. The bills, which set spending levels for the 2016 fiscal year that begins 1 October, also give an unexpectedly hefty boost to Alzheimer's disease research and revive the recently canceled National Children's Study.

"The Senate mark is obviously the best … action we have seen for NIH in more than 12 years," says Pat White, president of ACT for NIH, a Washington, D.C., group that lobbies for biomedical research funding. But some groups are concerned about provisions in the two bills to either slash or eliminate funding for the Agency for Healthcare Research and Quality (AHRQ), which supports studies of evidence-based medicine.

The Senate's draft bill, approved by an appropriations subcommittee on 25 June, would double the $1 billion NIH increase requested by the Obama administration. The House version, approved by its full appropriations panel on 24 June, would provide $100 million more than the White House request. Eventually, the House and Senate will have to reconcile the numbers in the two bills; usually they meet somewhere in the middle. But this year's budget debate will be complicated by tight spending caps imposed by an earlier budget agreement; the White House and many lawmakers want the caps removed, but so far have reached no deal.

Still, advocates for biomedical research funding "are starting from a pretty good position," says Jennifer Zeitzer, deputy director of the office of public affairs of the Federation of American Societies for Experimental Biology in Bethesda, Maryland.

Lawmakers want much of the new NIH funding to go to specific projects. Both bills match the president's request for $100 million for NIH's piece of a federal antibiotic resistance initiative and $200 million for President Obama's new Precision Medicine Initiative. The multiagency BRAIN neuroscience project would get $135 million in new funding in the Senate bill, matching the president's request, while the House version would provide a total of $160 million.

> *Biomedical research advocates "are starting from a pretty good position."*
>
> **Jennifer Zeitzer**, Federation of American Societies for Experimental Biology

A boost for biomedical research

Both houses of Congress want to give NIH more money, with Alzheimer's research and the BRAIN neuroscience initiative among the bigger winners ($ billions).

	2015	PRESIDENT'S REQUEST	HOUSE PANEL	SENATE PANEL
National Institutes of Health	$30.1	$31.1	$31.2	$32.1
Alzheimer's research*	$0.587	$0.638	$0.887	$0.937
BRAIN	$0.065	$0.135	$0.160	$0.135

*estimated

Congress also wants to make a major push on Alzheimer's disease research. In the Senate bill, the National Institute on Aging would receive a 25%, $350 million increase to its $1.2 billion budget; the committee expects much of the new money to go to Alzheimer's research. The House bill directly earmarks a $300 million increase for Alzheimer's—51% above current levels, and well above the president's request for a $51 million increase. The new money is in part the result of a relentless lobbying push by Alzheimer's groups. And it would be well spent, argues Robert Egge, executive vice president of the Alzheimer's Association in Washington, D.C., because advocates and researchers have developed an ambitious national plan for research. The plan, he says, will ensure "a strategic and responsible ramp up to the funding levels deemed necessary by experts."

One surprise in the bills is a call to revive the National Children's Study (NCS), a huge study that aimed to follow the health of 100,000 U.S. children for 21 years. NIH terminated the NCS, which Congress had previously mandated, last December after it encountered numerous problems; the agency had spent more than $1.3 billion since 2000 on planning (*Science*, 19 December 2014, p. 1441).

Now, lawmakers want a new version of the NCS. They were "disappointed" by NIH's decision to end the study, a report accompanying the House bill states, and they want NIH to work with pediatric groups to come up with a new 10-year research plan that builds on pilot data the NCS had collected. NIH has already begun thinking about "creating a synthetic cohort from existing longitudinal studies," NIH Principal Deputy Director Lawrence Tabak said at a recent meeting. To fund the effort, both bills call for providing $165 million annually—the same amount previously provided for the NCS.

The American Academy of Pediatrics in Washington, D.C., is "pleased" with the moves, a spokesperson told *Science*. "Congress made the right call for children's research."

Meanwhile, one loser in the bills is AHRQ, which has become a target for Republicans opposed to Obama administration health care policies. It would lose 35% of its $364 million budget under the Senate bill, and is completely eliminated by the House bill. The possible termination is "very troubling," says Dave Moore, senior director, government relations for the Association of American Medical Colleges in Washington, D.C. House appropriators have previously tried, and failed, to abolish AHRQ. ∎
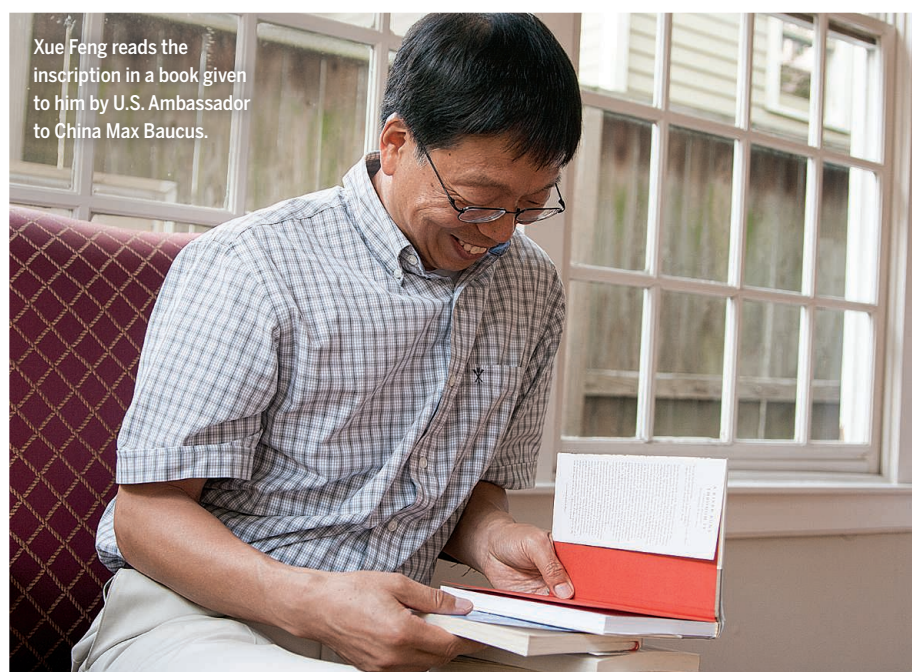
Xue Feng reads the inscription in a book given to him by U.S. Ambassador to China Max Baucus.

Q&A

Geologist reflects on life behind bars in China

Chinese-American specialist on oil industry had run afoul of China's secrecy laws

By **Catherine Matacic**

When Xue Feng landed his first job after academia in 2000, he was delighted. His new employer, Englewood, Colorado–based IHS, had high ambitions for the young geologist: overhaul how the company—a corporate intelligence firm—gathered oil and gas data on China. Feng dove into his assignment with gusto—so much so that in 2005, when he was 40 years old, he suffered a mild heart attack. By 2005, he had snared a rare, unclassified database of 30,000 oil wells in China from a private broker. The database promised substantial profits for IHS and—in a country that tightly guards such data on national security grounds—substantial risks.

Disaster struck on 20 November 2007. Feng, who had just left IHS for Houston-based C&C Reservoirs, was on a business trip in Beijing when he was abducted from his hotel room. Chinese security personnel interrogated him and charged Feng, a Chinese-born U.S. citizen, with selling state secrets. His chief crime: arranging for IHS's purchase of the oil well database, which the Chinese government declared a state secret in 2007. In 2010, Feng was convicted and sentenced to 8 years in prison, including the nearly 3 years he had already spent in detention.

Back in the United States, Feng's former Ph.D. adviser, University of Chicago geologist David Rowley, campaigned for his early release. Rowley met with U.S. embassy staff in Beijing and spurred prominent activists to petition both governments. But even U.S. President Barack Obama's personal request to Chinese leaders in 2009 wasn't enough. Feng was finally released in April—10 months before his sentence was set to expire—and immediately deported to the United States, where he rejoined his wife and two children in Houston. Feng spoke with *Science* about his time in prison and what other researchers working abroad might glean from his experiences. This transcript has been edited for clarity and brevity.

Q: Your work at IHS involved gathering a lot of sensitive data for clients. Were you ever worried about the risks involved?
A: Honestly speaking, I did not worry. I was doing honorable work in bridging the

international oil industry with a country that needed investment, and in helping my firm revamp its products. The profession of geologist seemed the safest in the world.

Was I misled about the potential danger working for IHS? Now I feel I was. It was the responsibility of the corporate executives, lawyers, and human resources people to worry about the sort of risks that they imposed on their employees. (IHS had no comment in response to questions from *Science*.)

Q: What happened on 20 November 2007?
A: At 12 o'clock that night, somebody knocked on my door. I told them I was sleeping. They said, "We have come to fix the electricity." I opened the door, and a bunch of guys without uniforms—about two dozen of them—came into the room. They told me they had some questions, and

Xue Feng's journal, which he smuggled out of prison by disguising it as a letter to his wife.

I needed to go with them. I said, "Okay, it's no problem. I have to go back [to Houston] tomorrow. Can we do this quickly?" And they said, "It's up to you. You've got to do it quickly." They took me to a secret location. I just wanted to finish the questions, because in my mind I never did anything wrong. I had such peace of mind. I even had a nap on the long drive.

The questioning started that night. It lasted probably 10 months. They interrogated me for 199 rounds.

Q: What was prison like?
A: I don't want to go into the details because a lot of things are very difficult to bring up. In the beginning it was only me. After 3 months, they moved me into a tiny detention center. I stayed in rooms [with] one or two or three other people. The detention center is a department of hell. In Chinese, people say there are 18 floors in hell. We were probably on the lowest level,

but slowly we took the elevators up, up, and up.

Q: How did you cope?
A: To keep my brain from being totally scrambled by [the interrogations], I mentally wrote a book for my children on all I'd learned. And I read the Chinese classics. In the first 3 years I wasn't allowed to wear my eyeglasses, so I had to squeeze my extremely near-sighted eyes to get clear vision, and that squeezing led to a piercing headache that annoyed me without stop.

After the interrogations ended, I started to do a lot of intentional thinking to keep my mind occupied and to avoid a mental cave-in. Sometimes, when it was too much to bear, I raised my head, saying to myself with contempt, "You can all go to hell. I will go to Texas, yeah, back to my home." In that mental picture, the orange-colored lights of my porch were like a lighthouse, guiding a lost sailor.

Q: Did you befriend any of your fellow prisoners?
A: In prison, people don't talk about friendship, because many people there have probably been betrayed by close friends. In my case, I just tried to help people. When I felt angry, I did volunteer work, unclogging the toilet drain. When I felt frustrated, I translated for inmates who needed sentence reduction papers written up. Helping others in need helped boost my morale. I didn't look for anything in return.

Q: Do you have any advice for foreign researchers working abroad on sensitive projects?
A: Who would want to hear from someone who is arguably one of the unluckiest of all? However, I did learn a few things from my ordeal.

First, only associate yourself with those people whose character and ability you admire. Do not compromise. Second, science may be without borders, but scientists do have nationalities. When the issue shifts from pure science to technology, the statement is even truer.

Before you travel abroad, you have to be cautioned of potential perils lurking there. Geopolitical rivalry sometimes makes scientific and technological enterprise extremely complicated in reality. Only when you get caught in geopolitical crossfire do you realize how hard it is for an obscure scientist to convince a mighty nation to do the right thing for you. ∎

Small scopes log an ever-changing sky

New surveys shoot celestial "movies" on a tight budget

By **Daniel Clery**

The universe is full of motion and change. Supernovae burst into life; variable stars wax and wane; exoplanets reveal their presence by passing in front of and dimming their parent star. Yet astronomers rely mostly on still images, identifying changes in the sky by persistence or chance. Now, a few teams of astronomers are asking: What if we could capture the whole sky regularly enough to create a movie of the heavens?

Taking advantage of powerful data processing and storage and, in some cases, cheap off-the-shelf optics, several telescope systems are now doing just that. The latest, known as the Evryscope, began observing in the past few weeks, snapping roughly a quarter of the entire sky every 2 minutes. The project's leader, Nicholas Law of the University of North Carolina, Chapel Hill, says he wants to look for exoplanets around rare widely scattered types of stars. "It became clear we could do an awful lot more [if we monitored] everything going on in the sky," he says.

Typical of the new breed of all-sky survey, Evryscope is run by a small team that targets a particular scientific question on a modest budget. Generally relying on small telescopes, the surveys don't look deep into space, but they do aim to capture everything that is happening in our celestial neighborhood.

Like many other exoplanet hunters, the Evryscope team finds its quarry by spotting "transits": the slight dimming of starlight that occurs when an orbiting planet passes in front of its star as viewed from Earth. Most transit searches have focused on "normal" stars similar to our sun and looked deep into one patch of sky to find many targets. But the Evryscope team targets rarer stars such as white dwarfs, and so must watch a large portion of the sky to find a few thousand candidate stars. Because these old, burned-out remnants of normal stars are very small—about the size of Earth—a planet

The Evryscope, here being installed in Chile, will snap the sky every 2 minutes.

passing in front of one will likely obscure it completely, for perhaps only 2 minutes. Capturing such a fleeting event anywhere in the sky is a huge challenge. "No one has seen a transit on a white dwarf," Law says.

Sited at the Cerro Tololo Inter-American Observatory in Chile, the roughly $250,000 Evryscope looks more like an architectural folly than a telescope. Its white hemispherical dome is dotted with portholes for its 27 off-the-shelf 7-centimeter telescopes. This setup generates 780 megabytes of data per minute—enough to fill more than 21,000 DVDs in a year. As a result, the raw data can't be streamed straight to the astronomers but must be stored on site and processed to reduce its volume. The team is currently in discussions with other observatories around the world to host more Evryscopes and provide all-sky coverage.

Gáspár Bakos of Princeton University and his Hungarian-made Automated Telescope Network (HATNet), along with HAT-South in the southern hemisphere, have also been searching for exoplanets, but their focus is Neptune-sized bodies. So far, their network of 31 small telescopes at five locations around the globe watch only part of the sky for transits, but Bakos and his team now want to add a wide-field instrument dubbed HATPi. They have begun construction of a "hedgehog" arrangement of multiple lenses and cameras at the Las Campanas Observatory in Chile, but they still require funding for 64 CCDs, the detectors that capture starlight—"our most expensive items," he says. HATPi will target bright nearby stars to tease out the slightest dips in star brightness. "We will get very high precision light curves from all over the sky," Bakos says.

For Kris Stanek of Ohio State University, Columbus, the quarry is supernovae: intensely bright stellar explosions that occur when a burned-out star collapses or a white dwarf gets an injection of new material and reignites. Supernovae are so rare that only a comprehensive sky census can find large numbers, Stanek says. Over the past 5 years, he and a team at Ohio State have built up the All-Sky Automated Survey for Supernovae (ASAS-SN). "We started with the science—the need to observe the whole sky—and found the cheapest way to do it," Stanek says. They managed the feat for less than $1 million using four 14-centimeter telescopes at Cerro Tololo and another four at Haleakalā in Hawaii. Together, the eight telescopes can image the whole sky every 2 or 3 nights, picking up fainter objects than Evryscope.

The small team has bagged more than 150 supernovae per year. Anything interesting they discover is disseminated within 2 hours through a Web-based service called The Astronomer's Telegram so that others can follow up with more detailed observations. "Every day we have something new and exciting," Stanek says. The team wants to add eight more telescopes at new sites to capture the whole sky every night, and to make all their data freely available, but at the moment the cost is prohibitive. "Taking data is the easy part," Stanek says. "Data reduction and recognition is much harder."

Some transient-event surveys employ instruments designed for other uses. The Palomar Transient Factory has repurposed the 67-year-old Samuel Oschin Telescope at the Palomar Observatory near San Diego, California, replacing its glass plates with CCDs. This $2 million project has bagged 700 supernovae in a single year, although it now views only 1/40 of the sky every night. But a planned $18 million upgrade with a giant new camera will enable it to image the whole available sky every night and automatically trigger follow-up observations by other robotic telescopes. "It's become a giant software algorithm program and stopped being an astronomy program," jokes Shri Kulkarni, director of optical observatories at the California Institute of Technology in Pasadena.

Not all the sky movies are low-budget productions. The $80 million Panoramic Survey Telescope and Rapid Response System (Pan-STARRS)—an instrument designed to search for threatening near-Earth asteroids—surveys the entire sky visible over its perch in Hawaii several times a month. With its 1.8-meter mirror, it is a prolific supernova hunter, having spotted more than 800 this year so far, and contributes to many other fields. A second mirror is currently being commissioned.

Even bigger will be the Large Synoptic Survey Telescope (LSST), whose 8.4-meter mirror and 3-billion-pixel camera will survey all the available sky every few days, spotting much fainter and more distant objects than other surveys. LSST's director, Steve Kahn of Stanford University in Palo Alto, California, says its 10-year survey, due to begin around 2022, will log 20 billion galaxies and 17 billion stars. "It will turn the universe into a catalog," he says. ■

Supernovae reported in 2015

All-sky surveys (bold) play a major role in scouring space for exploding stars.

TELESCOPE	SUPERNOVAE
Pan-STARRS	801
Catalina Real-Time Transient Survey	155
Gaia (ESA satellite)	113
ASAS-SN	89
Optical Gravitational Lensing Experiment (OGLE-IV)	81
Dark Energy Survey (DECam)	64
Palomar Transient Factory	56
Others	233

PUBLIC ATTITUDES

Politics doesn't always rule

Ideology is just one factor shaping views on science issues

By **Jeffrey Mervis**

Ideology is not the dominant factor in shaping what Americans think about most science-related issues, according to a new poll by the Pew Research Center. Although a person's political views are a strong predictor of their attitudes on climate change and a handful of energy issues, their gender, age, religion, race, or education play a larger role on many other controversial topics.

The Washington, D.C.–based think tank surveyed 2002 U.S. adults last summer on 22 issues ranging from global warming and offshore drilling to the safety of genetically modified (GM) foods, the use of animals in research, and the value of the International Space Station. A previous report based on the same survey found striking differences in what scientists and the public think about many topics, including GM foods and animal research.

The new analysis suggests that an oft-repeated claim that Republicans are "anti-science" is simplistic. "Sometimes politics is at the center of the story," says Cary Funk, the lead author and associate director for science research at Pew, "and sometimes politics has very little to do with the way people think about science issues."

The ideological divide is clearest when it comes to climate change. Almost 80% of respondents identifying themselves as liberal Democrats believe that global temperatures are rising "due to human activity," compared with only one in 10 conservative Republicans. At the same time, 57% of conservatives say "there is no solid evidence the Earth is getting warmer," a stance taken by only 10% of liberals. Six of seven liberals Democrats also favor stricter emissions limits on power plants, compared with only one in three conservative Republicans.

Public attitudes toward government support of basic research, a core issue for scientists, also seem to divide along party lines. Some 82% of liberal Democrats said that federal funding "is essential ... to ensure scientific progress," whereas only 43% of conservative Republicans held that view. On the other hand, some 55% of conservatives think "private investment will be enough" to support research, compared with only 16% of liberals.

But ideology takes a back seat on many other issues. Age is a much stronger factor than politics when people are asked whether childhood vaccinations should be compulsory. Some 41% of those under 30, compared with 20% of those over 65, think parents should make the call, whereas conservatives are only slightly more likely than liberals to hold that position.

Gender seems to be key on the use of animals in research: Whereas Americans are roughly split on the issue, 62% of women oppose the practice, compared with only 37% of men. Similarly, although 60% of respondents across the political spectrum support human space exploration, there's a big gap based on gender—only 42% of women back space exploration versus 66% of men.

Age and gender are both important factors when discussing nuclear power. Although the general population is roughly split, some 56% of those under 30 oppose building more plants, compared with only 39% of those over 65. Likewise, 59% of women oppose expanding nuclear energy, versus only 43% of men.

The poll also casts doubt on many scientists' belief that people would be more supportive of scientists' views on controversial issues if they just knew more about the topic. In only three of the 22 topics was a person's level of education or general scientific knowledge (as judged by answers to six questions) a significant factor in their views. One was animal research, where only 31% of persons with a postgraduate degree oppose the practice, compared with 56% of those with a high school education. (The other two issues where education appears to shape a person's stance were nuclear power and GM foods—in both cases, more knowledge leads to greater acceptance.)

But many scientists may be unhappy to learn that only 28% of U.S. adults think researchers have a "clear understanding of the health effects of GM crops." Respondents with a graduate degree were just as doubtful as those who never attended college.

On another issue, 69% of each group favors bioengineering to create a liquid fuel to replace gasoline. And college and high school graduates are equally worried about the planet's ability to feed a growing population. ∎

Six factors that shape views on hot science topics

Age and other factors can be a more important predictor of opinion than ideology.

	IDEOLOGY	AGE	EDUCATION	GENDER	RACE, ETHNICITY	RELIGION
Earth is warming due to human activity	Strong	Strong	Medium	Medium	Weak	Weak
Growing world population will be a major problem	Medium	Weak	Weak	Weak	Strong	Strong
Building more nuclear power plants	Medium	Strong	Strong	Medium	Weak	Weak
Private funding of science is enough to ensure progress	Strong	Weak	Medium	Weak	Weak	Weak
Humans have evolved due to natural processes	Medium	Weak	Medium	Weak	Weak	Strong
Childhood vaccines should be required	Medium	Medium	Weak	Weak	Weak	Medium
Safe to eat foods grown with pesticides	Medium	Weak	Weak	Medium	Weak	Weak
Use of animals in research	Weak	Strong	Strong	Strong	Weak	Medium
Astronauts essential for space program	Weak	Weak	Medium	Medium	Weak	Weak

■ Strong impact   ■ Medium impact   ■ Weak impact

Pandemrix triggers antibodies to a receptor that's involved in sleep regulation.

INFLUENZA

Narcolepsy link to pandemic flu vaccine becomes clearer

Pandemrix may have caused autoimmune reactions

By **Gretchen Vogel**

The 2009 H1N1 influenza pandemic left a troubling legacy in Europe: More than 1300 people who received a vaccine to prevent the flu developed narcolepsy, an incurable, debilitating condition that causes overpowering daytime sleepiness, sometimes accompanied by a sudden muscle weakness in response to strong emotions such as laughter or anger. The manufacturer, GlaxoSmith-Kline (GSK), has acknowledged the link, and some patients and their families have already been awarded compensation. But how the vaccine might have triggered the condition has been unclear.

In a paper in *Science Translational Medicine* (*STM*) this week, researchers offer a possible explanation. They show that the vaccine, called Pandemrix, triggers antibodies that can also bind to a receptor in brain cells that help regulate sleepiness. The work strongly suggests that Pandemrix, which was given to more than 30 million Europeans, triggered an autoimmune reaction that led to narcolepsy in some people who are genetically at risk.

"They put together quite a convincing picture and provide a plausible explanation for what has happened," says Pasi Penttinen, who heads the influenza program at the European Centre for Disease Prevention and Control in Stockholm. "It's really the kind of work we've been waiting for for 5 years." But the results still need to be confirmed in a larger study, the authors and other narcolepsy researchers say. A 2013 paper in *STM* by another group, documenting a different type of vaccine-triggered autoimmune reaction, was retracted after the results proved irreproducible (*Science*, 1 August 2014, p. 498).

Narcolepsy, a mysterious malady that affects roughly one in 3000 people in Europe, most often appears in childhood or adolescence. Patients lose certain brain cells in the hypothalamus, leading to a deficiency of hypocretin, a molecule that helps regulate the sleep-wake cycle. Researchers suspect an autoimmune reaction is to blame because many people who develop narcolepsy—and just about everyone with the vaccine-associated form—have a specific variant in a gene in the HLA family, which helps the body distinguish its own proteins from those made by microbial invaders.

When they heard about the rise in narcolepsy in 2010, neuroscientist Lawrence Steinman of Stanford University in Palo Alto, California, and rheumatologist Sohail Ahmed, who at the time was global head of clinical sciences at Novartis's vaccines and diagnostics division in Siena, Italy, began scouring databases for proteins expressed in the brain that might resemble those in the vaccine.

Their search turned up a suspect: a piece of a receptor for hypocretin resembles part of the H1N1 influenza nucleoprotein—which binds to the virus genome and plays a key role in its replication. "That was really an 'Aha' moment," says Ahmed, who is now at GSK, which bought part of Novartis's vaccine division. The flu vaccine is designed to trigger antibodies to influenza's surface proteins, but if it elicits antibodies to the nucleoprotein as well, those might latch on to the hypocretin receptor, and eventually lead to death of the cells, the researchers thought.

Some of the same researchers already had circumstantial evidence that the nucleoprotein might be an important player. In December, Outi Vaarala of the University of Helsinki and her colleagues had reported that Pandemrix contained a much higher level of nucleoprotein than a GSK vaccine called Arepanrix, which was associated with a much smaller risk for narcolepsy. The group also found that children with narcolepsy had an altered immune response to the nucleoprotein in Pandemrix.

In the new work, the researchers added serum from Finnish narcolepsy patients who had received Pandemrix to cells that were engineered to display human hypocretin receptor 2 on their surface. Antibodies from the patients bound to these cells in 17 of 20 samples. Serum from Italians who had been vaccinated with a different pandemic vaccine from Novartis, called Focetria, did not have such antibodies. The researchers also showed that Focetria, which has not been linked to narcolepsy, had a much lower concentration of nucleoprotein than Pandemrix did.

If a vaccine can trigger narcolepsy, can the flu virus itself—either H1N1 or other influenza strains? Cases of narcolepsy increased after the 2009 pandemic in China, where Pandemrix wasn't used, and researchers discovered that serum from some H1N1 influenza patients does bind to the hypocretin receptor. But the jury is still out, says virologist Ilkka Julkunen of the University of Turku in Finland, a co-author on the paper.

In the meantime, the new insights might eventually lead to ways to help narcolepsy patients. Researchers could explore whether somehow unblocking the receptor might allow the hypocretin system to recover, says Noni MacDonald, a pediatric vaccine researcher at Dalhousie University in Halifax, Canada. The study also suggests ways to make influenza vaccines even safer, for instance by keeping the nucleoprotein levels low or removing the specific region of the protein that mimics the receptor. "We should do everything to understand the disease mechanism," Julkunen says, "so that this kind of situation would never be repeated." ∎
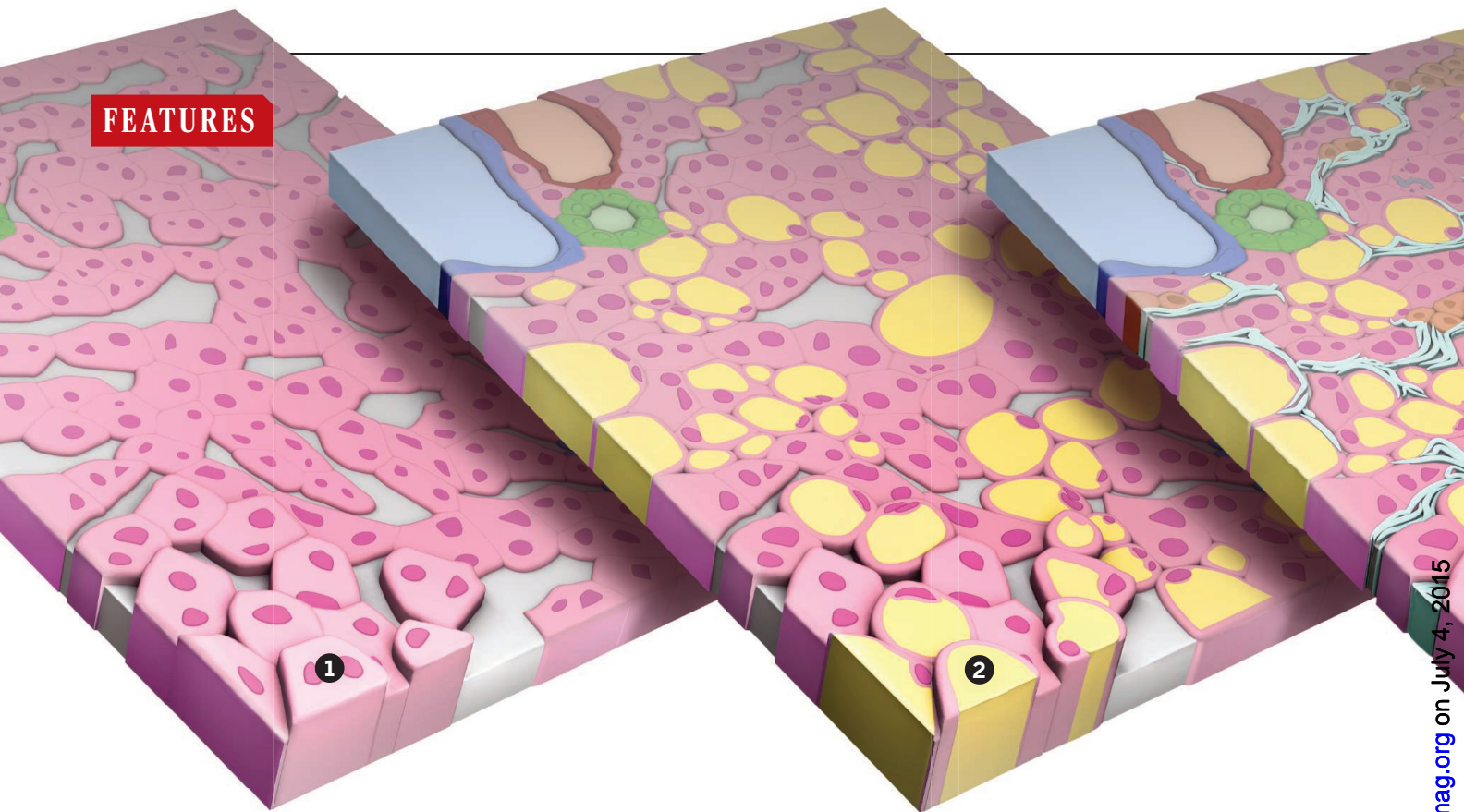
THE LIVER'S WEIGHTY PROBLEM

As obesity rates soar, a sometimes fatal liver disease is becoming epidemic *By* **Mitch Leslie**

The patient who walked into Joel Lavine's office at the University of California, San Diego (UCSD), medical center one day in the mid-1990s didn't know how sick he really was. He was morbidly obese. A brownish blemish known as acanthosis nigricans sprawled over the nape of his neck and into his armpit, signaling that he probably had developed insulin resistance, a condition in which cells don't respond normally to the hormone that controls blood sugar.

A biopsy revealed striking damage to the patient's liver: so much fat crammed into the cells that it squashed their nuclei and other contents. Cirrhosis, or severe liver injury, was beginning as scar tissue ousted healthy cells. The patient essentially had the liver of a middle-aged alcoholic. Yet he was only 8 years old.

To Lavine, a pediatric hepatologist, it was clear the boy was suffering from non-alcoholic steatohepatitis (NASH). A condition usually associated with obesity, NASH results from excess fat in the liver and, as the name indicates, doesn't stem from the alcohol abuse that causes many cases of severe fatty liver disease. Because NASH can destroy the liver, patients can require

a liver transplant or even die.

Although Lavine had come across a few kids with NASH earlier in the 1990s when he worked at Boston Children's Hospital, he found "it was rampant" in San Diego among the large Hispanic population. A dozen children suspected of having the illness were coming to his office every week for diagnosis. "It was clear we had an epidemic of sorts," says Lavine, who is now at Columbia University's College of Physicians and Surgeons.

That epidemic has now spread throughout the country as our livers pay the price for our calorie-rich diets and sedentary
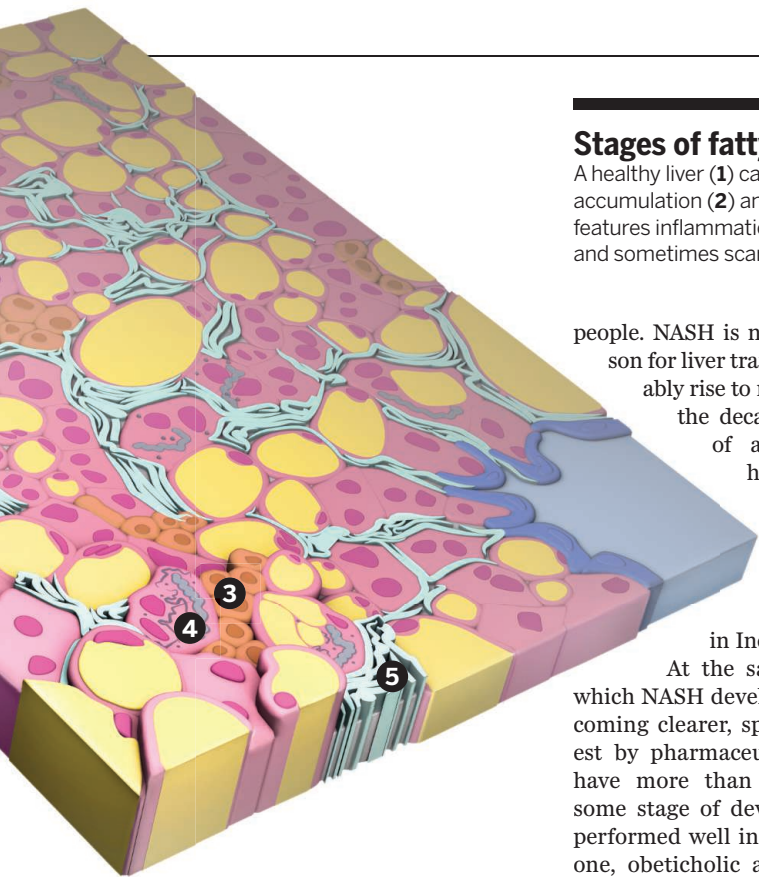
Stages of fatty liver disease

A healthy liver (**1**) can progress to fat accumulation (**2**) and then to NASH, which features inflammation (**3**), cell swelling (**4**), and sometimes scarring (**5**).

people. NASH is now the number two reason for liver transplants, and it will probably rise to number one by the end of the decade as a new generation of antiviral drugs controls hepatitis C, currently the largest cause of liver failure, says hepatologist Naga Chalasani of the Indiana University School of Medicine in Indianapolis.

At the same time, the paths by which NASH develops and worsens are becoming clearer, sparking a surge of interest by pharmaceutical companies, which have more than 20 potential drugs in some stage of development. Several have performed well in early clinical trials, and one, obeticholic acid, recently impressed researchers by reducing the amount of liver scarring, a serious consequence of NASH. "We are at a pivotal time" in the history of NASH, Lavine says. "In 5 years we will have at least one treatment approved," Ratziu predicts.

NASH WENT UNRECOGNIZED for decades, mainly because doctors confused it with alcoholic steatohepatitis, or ASH, which also involves fat buildup in the liver but results from heavy drinking. Although some people who had fat-flooded livers denied being alcoholics, doctors were convinced they were boozing in secret, says hepatologist Stephen Caldwell of the University of Virginia in Charlottesville. The assumption was "if you had liver disease, you

were a drinker," he says.

Then in 1980, pathologist Jurgen Ludwig of the Mayo Clinic in Rochester, Minnesota, and colleagues put the "N" in "NASH," describing a cluster of people with fatty, damaged livers who were not alcoholics. The large number of kids subsequently found to have NASH helped dispel any lingering doubts that the illness wasn't alcohol-related, Lavine says.

Extra fat in the liver is now fairly easy to spot with imaging techniques. Gauging how many people have the severe form of fatty liver disease, NASH, is trickier. "The only way to know if your patient has NASH is a liver biopsy," says gastroenterologist Stephen Harrison of the Brooke Army Medical Center in San Antonio, Texas. Although he describes the procedure as "a nick in the skin" that doesn't involve surgery, people often refuse, which makes it hard to determine NASH's prevalence; researchers' estimates generally run between 2% and 5% of the population.

Ultimately, says molecular biologist Jay Horton of the University of Texas Southwestern Medical Center in Dallas, fatty liver "is a disease of caloric excess." The imbalance between calories consumed and burned triggers a complex series of changes that transform the organ's character. Many of the diet-derived fatty acids in the bloodstream make their way to the liver, which directs them to other parts of the body. "The liver is your traffic cop" for these building blocks of fats, says Elizabeth Parks, a nutritional physiologist at the University of Missouri in Columbia. But the organ itself typically holds onto little fat. For instance, Parks says, a fairly fit 70-kg man will carry about 14 kg of body fat—and only 125 g will reside in the liver. "It's such as small pool, and there are lots of options that the liver uses to get rid of it," she says.

In some people, however, the liver becomes a hoarder and starts stockpiling

lifestyles. About 20% to 30% of people in the United States who don't abuse alcohol carry extra fat in their livers, a precursor to NASH. The condition is even turning up in unlikely places such as rural India. "It's becoming an important problem everywhere," says hepatologist Vlad Ratziu of the Pitié-Salpêtrière Hospital in Paris.

That's worrying, doctors and researchers say, because although such liver fat accumulation is usually benign, NASH develops in about 30% of people with fat in the liver. It can lead to liver failure, liver cancer, and death, and no drug treatments exist yet. A new liver is the only option for many

The culprits

The liver stores excess fat in fatty liver disease, and these five contributors help determine how much.
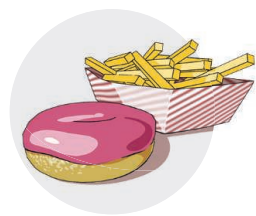
Diet
Fats that we eat can travel to the liver, which may stash some of them.
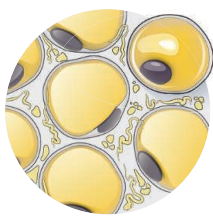
Adipocytes
These fat-storage cells release fatty acids, which can travel to the liver.
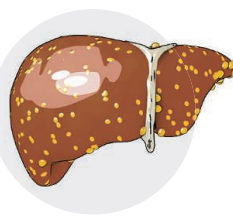
Fat producer
In fatty liver disease, the liver can turn up its production of fatty acids.
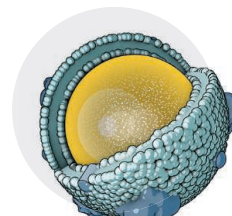
VLDL
The liver sheds some—but not enough—excess fat by producing these particles.

Insulin resistance
This condition causes adipocytes to spill fats and the liver to make them.

triglycerides, a fat variety, sometimes accumulating so much that the organ "can look like butter," Lavine says. In addition to dietary fatty acids, fat in the liver can come from two other sources, researchers have found. Outside the liver, fat-storing cells known as adipocytes continually release some of their contents, which end up in the organ. And the liver itself synthesizes some fat.

All three sources ramp up in fatty liver disease. Parks and her team recently showed how the liver contributes to its own woes. They fed volunteers meals tagged with radioactive compounds, enabling them to track the synthesis of free fatty acids, precursors of triglycerides. People who had fatty liver disease manufactured three times more free fatty acids in their livers than did healthy subjects, the group reported last year in *Gastroenterology*.

Of the many metabolic abnormalities that could drive accumulation of fat in the liver, insulin resistance seems key, Chalasani says. It develops when cells become less responsive to the hormone, usually because of a surfeit of nutrients from food, and it promotes fat buildup in the liver in several ways. For example, insulin normally prompts adipocytes to stop releasing fatty acids. But as Parks and her team showed in their study last year, insulin-resistant adipocytes continue to spill fatty acids into the blood. The liver can usually rid itself of some fat by exporting triglycerides, but insulin resistance prompts the organ to synthesize more triglycerides, overwhelming the disposal mechanism.

Other factors, from the types of microbes living in the intestine to genetics, may also induce fat accrual in the liver. Having a particular variant of the metabolic gene *PNPLA3* boosts the risk of developing a fatty liver, for instance.

A RAFT OF HEALTH PROBLEMS typically accompanies the fat buildup, including obesity and diabetes. But the liver's extra heft doesn't normally kill, Harrison says. Chances are, "you are going to grow old with your fatty liver," he says.

The odds are worse for people who progress to NASH. About 25% of them eventually develop cirrhosis that devastates the liver. Compared with individuals with more benign fatty livers, NASH patients typically show two distinctive characteristics. The inflammation in their livers is more intense as white blood cells infiltrate the organ. And hepatocytes, the most abundant liver cells,

swell to as much as twice their normal size, a condition known as ballooning that heralds their demise.

These two phenomena can segue to an insidious problem: fibrosis, or the buildup of scar tissue rich in the protein collagen. As scarring progresses, it takes over more and more of the liver, and fewer and fewer healthy cells remain, sometimes resulting in liver failure. Cirrhosis can also trigger liver cancer, although researchers haven't worked out exactly how.

The biggest question—why about one-third of people with fatty liver disease develop NASH—remains unanswered, notes UCSD hepatologist Rohit Loomba. Some scientists posit that the fat itself is poisonous to liver cells, whereas other evidence points to inflammation and oxidative stress.

Whether scar tissue accumulates in a fat-laden liver may depend on the organ's ability to heal itself. Hepatologist Anna Mae Diehl of the Duke University Medical Center in Durham, North Carolina, and her colleagues have found evidence that repair goes awry in the livers of NASH patients because of excessive activity by a molecular signaling pathway normally prominent in development. In an embryo, a protein called hedgehog and its partners help sculpt the brain, kidneys, intestines, and other parts of the body. The pathway all but shuts down in the livers of healthy adults, but it can ramp up again if the liver is injured. The pathway then spurs certain liver cells to transform into cells that produce scar tissue. The ballooned hepatocytes that accrue in the livers of NASH patients could be the culprit, Diehl and her team suggest. The bloated cells pour out proteins that activate the hedgehog pathway, the researchers revealed in a 2011 paper in *The Journal of Pathology*.

DOCTORS NOW HAVE LITTLE TO OFFER NASH patients apart from the recommendation to eat less and exercise more. If a patient worsens to the point where the liver starts to shut down, a transplant becomes a possibility. But at the moment, "the ability to rescue someone who is getting really sick is limited," Caldwell says.

That grim outlook may soon change. Many biotech and pharmaceutical companies have launched NASH drug development efforts, enticed by the potentially huge demand for a chronic medication—a $35 billion annual market by one estimate. The drugs researchers have tested against NASH or hope to test include vitamin E, the diabetes treatment pioglitazone, and a

raft of new compounds. "What's interesting is that almost all of them have a different mode of action," Ratziu says.

One drug, aramchol, combines a synthetic fatty acid and bile acid; it is now in a phase II trial to determine how well it reduces fat accumulation in the liver. In contrast, cenicriviroc, also in a phase II trial, originated as an antiviral candidate because it blocks certain receptor molecules on immune cells that viruses use as entryways. But it also may stem inflammation and liver fibrosis.

At least two phase III clinical trials of potential NASH treatments are likely to begin this year, although one them is already starting with a strike against it. In March, the French pharmaceutical firm Genfit announced that GFT505, which stimulates cellular receptors that promote insulin sensitivity and spur fatty acid breakdown, didn't reduce the amounts of liver fat or fibrosis overall in a phase II trial of 274 NASH patients. But the most serious cases did show improvement, and Ratziu says the company plans to launch a phase III trial in 2500 patients with advanced NASH.

The most promising NASH treatment so far, several researchers say, is obeticholic acid. The liver produces bile acids to help the intestine absorb fats, which help manage fat and sugar metabolism. Obeticholic acid is a modified bile acid, and it prods a cellular receptor that promotes sensitivity to insulin and reduces the amount of triglycerides in the blood.

Last November, hepatologist Brent Neuschwander-Tetri of the Saint Louis University School of Medicine and colleagues reported the results of the FLINT trial, a phase II study that found that obeticholic acid reduced the amount of fibrosis in NASH patients, suggesting the drug can reverse the disease. "That's something that's never been seen for a drug before," says Lavine, who was on the trial's steering committee.

Still, researchers—and drug company investors—remain cautious. As Loomba points out, none of the drugs tested so far has benefited more than 50% of NASH patients. And although obeticholic acid reduced fibrosis, it also increased the levels of LDL cholesterol, which promotes cardiovascular disease. That's worrying, Chalasani says, because "these people are already at greater risk for coronary artery events and cardiovascular disease." Even so, a phase III trial of obeticholic acid in 2500 patients with severe liver fibrosis is scheduled to start later this summer. And the scientific community studying NASH is hopeful for the first time. Preventing or even reversing NASH, Neuschwander-Tetri says, finally seems "a reachable goal." ∎

> *"In 5 years we will have at least one treatment approved."*
>
> **Vlad Ratziu**,
> Pitié-Salpêtrière Hospital

OF MICE AND MEN

Researchers are adding human DNA to mice to pinpoint sequences that helped define our species *By* **Elizabeth Pennisi**

n 2006, computational whizzes at the University of California, Santa Cruz (UCSC), announced that they had glimpsed the DNA that made us human. They had compared all of the vertebrate genomes sequenced to date and come up with a list of about 50 DNA regions that were identical in many animal species but had changed in humans. Those sequence changes must have been important in the evolution of humans, the team concluded, contributing to our big brains, bipedalism, broad diet, and other traits that have made our species so successful.

By 2008, almost two dozen bioinformatics studies had added hundreds of other uniquely human genetic sequences to the list, and the pursuit of such DNA continues to this day. Yet almost a decade later, not much progress has been made in demonstrating that these DNA sequences—genes, regulatory sequences, and other genome elements—actually shaped human origins. "It's pretty embarrassing how little we know

about the specific changes in our genome" related to our evolution, says Greg Wray, an evolutionary biologist at Duke University in Durham, North Carolina. "You could imagine [their roles], but they were just sort of 'just so' stories."

Now, with the help of genetic engineering and the humble laboratory mouse, researchers are starting to pinpoint just how some of these sequences might have helped give rise to our uniquely human features. Several groups have gotten clues by inserting pieces of human DNA into mouse embryos, declaring evidence for an evolutionary role if a piece of DNA simply functions in the brain, a limb, or some other structure where humans differ from chimps or other animals. But those studies fail to go the distance. Those studies, however, did not breed mice to have the new DNA incorporated into their genomes, a step that makes it possible to assess how the human sequence affects an organism's anatomy and function.

By making that effort, Wray and Duke developmental neurobiologist Debra Silver

The blue in the brain of this mouse embryo reflects the activity of a reporter gene driven by the HARE5 enhancer, implicated in human brain evolution.

have recently watched how the human version of a regulatory element called HARE5 altered mice, boosting their brain size by 12%. HARE5 "could easily be a huge component in how the human brain expanded," says Mary Ann Raghanti, a biological anthropologist at Kent State University in Ohio.

TO WRAY, "the most interesting question is the origin of our species." The lure of that question drove him and other researchers to their computers a decade ago when sequencing genomes came into its own. First, they compared the human and mouse genomes, and, soon afterward, looked for meaningful differences between humans and other primates, especially our closest relative, the chimp. In 2005, Evan Eichler, a geneticist at the University of Washington, Seattle, and colleagues compared human and chimp genomes to find DNA duplicated in us but not

in our closest relative. They determined that 33% of the duplications in humans were not present in chimps. Four years later, Eichler's team went further, pinning down about 1000 instances in which the human genome has extra copies of genes and regulatory regions called enhancers.

A gene called *SRGAP2*, found as a pair in chimps but in six copies in humans, particularly intrigued Eichler. The initial duplication occurred about 3.4 million years ago, judging from the number of differences between it and the original gene, and a second duplication a little more than a million years later created a shorter gene with a new function. When inserted into the developing brains of mouse embryos, this truncated gene causes cells in the rodent organ to migrate farther and sprout more spines,

they measured how active these genes are in fetal human brain stem cells, the champion was a truncated copy of a gene called *ARHGAP11A*, which was also on Eichler's list. Huttner and his colleagues inserted the copy into the brains of developing mouse embryos and the number of cortex cells nearly doubled. The rodent brains sometimes also developed folds normally seen only in human brains, the group reported online 19 February in *Science*. The truncated copy, they note, is found in the genomes of Neandertals and another ancient human group, the Denisovans, but not in chimps, bolstering suspicions that it played a key role in human evolution.

Humans may also have gained brain by losing some sequences. David Kingsley, an evolutionary geneticist at Stanford Uni-

most intriguing is a stretch of DNA identified 9 years ago by David Haussler at UCSC and Katherine Pollard, now a biostatistician at the Gladstone Institutes in San Francisco, California. They called it HAR1, for human accelerated region 1, because the human sequence has diverged dramatically, whereas that DNA has remained the same in other animals examined. They determined that HAR1 encodes a standalone RNA, rather than a protein, and that it is expressed in the developing human brain as nerve cells are forming connections and organizing into layers in the cortex. To many evolutionary biologists, this was exciting, if circumstantial, evidence that the sequence had a role in human brain evolution.

Whereas many of the DNA changes in the spotlight appear to have shaped our brain, some may have influenced other human structures. One of the regulatory DNA deletions Kingsley identified may have caused the human penis to lose the spines seen on the penises of mice, chimps, and many other mammals. And another of Pollard's sequences, a DNA segment that is second only to HAR1 in the number of changes between the human and animal versions, may play a role in the development of the human forelimb, according to a team led by Shyam Prabhakar, now a geneticist at the Genome Institute of Singapore, and James Noonan, now at Yale University. Like Pollard, they and colleagues had pinpointed this sequence as a possible driver of human evolution. When they inserted this regulatory sequence, called HACNS1, into mouse embryos in 2008, they discovered the human version, but not the chimp one, was active in developing rodent forelimbs.

Seven years later, however, Prabhakar and Noonan still don't know what gene or genes HACNS1 controls, having not yet fully incorporated this DNA into rodents by creating transgenic mice. "Introducing large segments of human sequence into the mouse genome … is slow, not easily scalable, and expensive," Noonan says. That's been "the limitation for the field overall."

THAT'S THE STEP WRAY finally took. Rather than looking for glimmers of function by inserting a sequence into a mouse embryo, he and Duke graduate student Lomax Boyd set out in 2010 to try to create

Building a bigger brain

Human brains differ from those of chimps in part because variations in a piece of regulatory DNA called an enhancer cause the human *FZD8* gene to be more active over a larger region of the developing brain, as shown when each species' enhancer is engineered to turn on the mouse version of the gene.

Human

Human enhancer

DNA

Mouse embryos

FZD8

Chimpanzee

Chimpanzee enhancer

Enhancer activity

FZD8
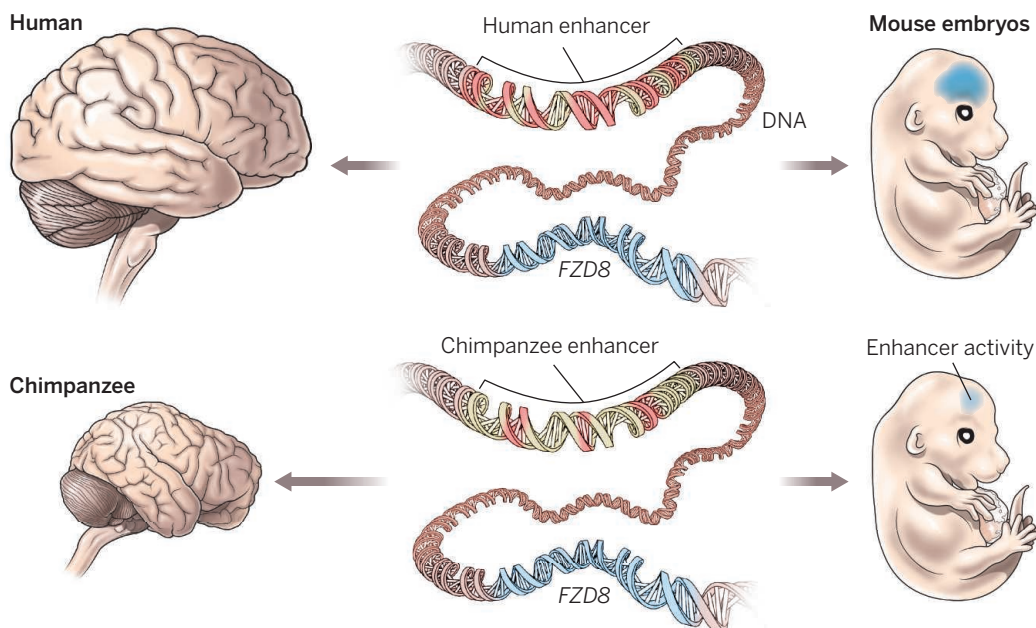
which can increase the number of connections they can make, Franck Polleux, a neuroscientist at Columbia University, and his colleagues reported in 2012 in *Cell*.

Recently, suggestive evidence has emerged that another human-specific duplicated gene also helped increase our brain power. Wieland Huttner, a developmental neurobiologist at the Max Planck Institute of Molecular Cell Biology and Genetics in Dresden, Germany, and his colleagues identified 56 genes that humans have in multiple copies but mice lack completely. When

versity in Palo Alto, California, and his colleagues found 500 stretches of regulatory DNA that chimps have but humans do not. One sits near a tumor suppressor gene that keeps cell growth in check. Kingsley is now testing the effect of the deletion in mice, but he suspects that it disables the tumor suppressor and removes the brakes on cell division, allowing extra brain growth in humans.

Along with duplications and deletions, investigators are studying sequences that humans share with other animals—but that show signs of dramatic change. One of the

a mouse strain with the human DNA permanently integrated into the rodent's genome. Although new gene-editing technologies promise to streamline such an experiment today, at the time they knew that they faced years of hard work. Because his expertise was in bioinformatics rather than in creating transgenic mice, Wray approached Silver, a new Duke faculty member and a transgenic mouse and brain development expert, about working on the project.

To increase the group's odds of success, Boyd only considered DNA sequences that multiple lists had pinpointed as likely to have been important for human evolution. He also narrowed his focus to sequences that appeared to act as enhancers, because these gene regulators seem to play an outsized role in evolution. For each candidate enhancer, he combed the scientific literature for nearby genes that might be under its control and that either had roles in brain growth or encoded proteins that regulate the activity of other brain-related genes. After about a year, the Duke team had hand-picked nine candidates.

To learn more about their function, Boyd and Silver hooked each candidate enhancer to a reporter gene that would produce a blue color whenever the enhancer was active, then injected the complex into mouse embryos—the type of experiment other labs had done with their sequences of interest. "We use the mouse embryo sort of like a test tube," Silver says. One enhancer, HARE5, produced a bright, consistent pattern in just the brain of the mouse embryos. In the few mammals studied, HARE5 sits near a gene called *FZD8*, which is part of a well-known pathway that controls mammalian brain growth.

Then came the "money shot," as Wray calls it—seeing whether the human and chimp versions of HARE5 affected the embryonic mouse brain differently by creating transgenic mouse strains that had one or the other. To remain unbiased, Boyd was blinded to the type of embryo he was analyzing, and he hung pictures of the embryos up in the lab so colleagues could opine on what they saw.

After comparing hundreds of modified embryos at various points during development, Boyd and Silver became convinced that the two HARE5s differed in their timing and level of activity. To confirm, Boyd linked the human and chimp enhancers to genes that produced differently colored fluorescent tags. The team then created a transgenic mouse with the chimp enhancer pairing and a strain with the human DNA, then bred the resulting adults so the offspring carried both types of HARE5s, allowing a direct comparison. The human version turned on
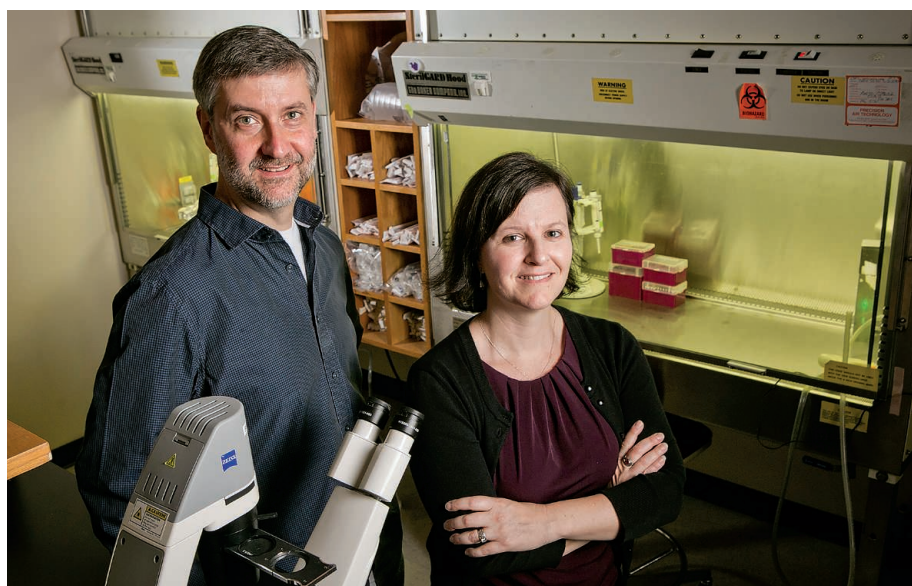
At Duke University, Debra Silver and Gregory Wray used mice to demonstrate how one piece of DNA likely contributed to human evolution.

much earlier and over a broader expanse of embryonic brain than the chimp enhancer.

Finally, the group engineered several strains of mice to carry the human enhancer attached to the mouse version of the *FZD8* gene, the enhancer's suspected target. It was these mice whose brains ended up about 12% larger than those of normal mice or of strains with the chimp enhancer. By studying thin slices of the mouse fetal brains, Silver's team identified the reason. They found that the stem cells that give rise

---

"They have found a smoking gun in the human genome that connects a regulatory element with a proposed pathway for increasing brain size."

Todd Preuss, Yerkes National Primate Research Center

---

to neurons divide 23% faster and generate more neuron progeny when they are equipped with the human HARE5.

The finding, announced 16 March in *Current Biology*, fits with a leading theory about how the human brain evolved. In 1995, Yale University's Pasko Rakic had proposed that the bigger human brain could have resulted from a simple alteration of the cell cycle in progenitor cells—if these precursors divided more often before transitioning to nerve cells, the brain would be much bigger. Rakic and others consider the Duke evidence

strong support for the idea. "They have found a smoking gun in the human genome that connects a regulatory element with a proposed pathway for increasing brain size," says Todd Preuss, a neuroanatomist at the Yerkes National Primate Research Center in Atlanta. Eichler adds: "The molecular dissection is the finest I've seen."

Not everyone is satisfied that the Duke team's work is done. In their experiments, the researchers left the mouse HARE5 intact and added the other versions of the enhancer, but Kingsley suggests a better experiment would have been to completely replace the mouse sequence with the human or chimp version and monitor the effects on *FZD8* and the rodent brain. Boyd tried to do that swap but failed—possibly because the mouse HARE5 is located on a part of the chromosome that was difficult to manipulate using gene-targeting approaches available at the time.

Kingsley would also like to have seen cognitive testing of the big-brained mice. "It would be nice to know if you could make not only a bigger brain, but an animal that's also smarter." Eichler adds that Silver and Wray could also look for people with brain-size abnormalities to see if they have mutations in HARE5. "That would really be proof of what they are trying to show," agrees Nenad Sestan, a Yale neurobiologist.

Even with such proof, however, Wray and Silver's mice will never reveal the whole story of how our brains diverged from those of chimps and other animals. "The evolution of the human brain did not occur by a single gene or even a few genes," Eichler notes. "Rather it was a concert of evolutionary changes." ■

INSIGHTS

LETTERS

Edited by **Jennifer Sills**

Postdocs reimagined

In April, we asked young scientists if the idea of the postdoc position is obsolete. If so, what should replace it, and if not, how can it be improved? We received a record response of more than 300 submissions. About a third of respondents felt that postdocs are indeed obsolete. The rest deemed them necessary, if imperfect. Many felt that creating a permanent staff scientist position, with full salary and benefits, would help the plight of postdocs, either by replacing the postdoc entirely or by serving as a long-term option after the completion of a postdoc. Others suggested that a system in which postdocs had dedicated funding and were not beholden to a PI would foster more creativity and minimize exploitation. Tailoring postdocs to a broader array of career paths was another common theme. Many responses reflected the feeling that postdocs deserve more respect and recognition. A sample of the responses describing these and other ideas can be found below. To allow for as many voices as possible, in some cases we have printed excerpts of longer submissions (indicated by ellipses) and lightly copyedited original text for clarity. To read the complete versions, as well as many more, go to http://scim.ag/NG15R. Follow *Science*'s NextGen VOICES survey on Twitter with the hashtag #NextGenSci.

NEXTGEN VOICES

Postdoctoral scientists are far from being obsolete; rather, they are the drivers of research. They made it through a Kryptonian Ph.D. thesis and emerged ever-more-powerful and fearless to continue in science. They act as mentors and as students. They constitute the

chain link between the "drowning under grant deadlines, conference talks, endless responsibilities," stressed-out professor, and the "desperate to finally finish his Ph.D. and get out of here" stressed-out Ph.D. student. They think of projects, perform experiments, analyze data, and write papers. They generate beautiful hypotheses, as well as the ugly data that annihilate these same hypotheses…. They stand in a limbo state: They have already committed themselves to science, but the science world has not committed to them….

Nikolaos Konstantinides

Department of Biology, New York University, New York, NY 10003, USA. E-mail: nk1845@nyu.edu
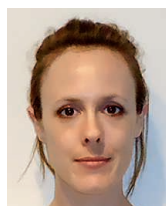
There should be no postdoc positions. Recent graduates should be able to compete directly for full-time research and faculty positions. It's absurd to expect an individual with a Ph.D. to work into their mid- to late 30s without full employment benefits at income levels that are often well below the median levels for their area. The current system does not offer fair compensation at precisely the time that one needs to start and support a family. The average number of years spent working as a postdoc is 5; the average number of years spent at any job in the United States is also 5. Why is the postdoc neither fully compensated nor fully employed?

Christina Dinkins

Bethesda, MD 20892, USA. E-mail: senadinkins@gmail.com

If anything, the concept of a grad student is obsolete. We do not need Ph.D.s anymore. Once upon a time, pre-Internet age, Ph.D.s were rare and conferred expertise, recognition, and prestige. This is no longer the case. The Ph.D. system…should be replaced with a system that allows researchers with Bachelor's or Master's degrees to be free agents in host… (academic or corporate) labs until they are ready to apply for faculty positions…. The free-agent system would be more dynamic and eliminate the postdoc problem. It would also force academic research to be more competitive in terms of pay, conditions, and personal development.

Martin P. Stewart

Department of Chemical Engineering and Koch Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: martstew@mit.edu

The goal of any academic career track after the Ph.D. should be to prepare a researcher for a leadership research position…. Currently, postdocs apply for positions in the labs of tenured PIs, where their primary work includes pursuing research topics assigned to them by their superior, managing students, preparing project proposals, and applying for grants within the PI's field…To facilitate more independence in postdocs,…these postdoc positions should be financed directly from the government or university, and not through PIs….This would allow for the postdocs to (i) at least partially pursue their own scientific questions, (ii) try working under more than a single PI and therefore experience different managerial practices, and (iii) see whether independent research suits them. The last is especially important, because setting your own scientific goals is much more difficult than following someone else's.

Jernej Zupanc

Seyens Ltd., Ljubljana, 1000, Slovenia. E-mail: jernej.zupanc@seyens.com

…A 1-year pre-assistant professor position could replace the postdoc position. The postdoc would be offered a small lab area and university funds to lead a small group of jointly supervised graduate students who are interested in the proposal topic. Applying for external funding and being fully responsible for a small lab model would not only help the postdoc to mature for his next career step as an assistant professor, but would stimulate new avenues of research for graduate students.

Esraa Elsanadidy

Department of Chemistry, University of Connecticut, Storrs, CT 06269, USA and Department of Chemistry, Faculty of Science, Tanta University, Tanta, Egypt. E-mail: esraa_elsanadidy@yahoo.com

…We need to start funding people, and not only projects. Today, funding agencies are giving money for projects. This creates a situation in which people continuously have to move to their next

project, putting strain on relationships and families…Is this the system we want to support, where there is only room for those willing to sacrifice everything to keep a low-paying job, without any job security whatsoever and the need to move every 2 to 4 years?

Magnus V. Persson

Leiden Observatory, Leiden University, Leiden, 2333CA, Netherlands. E-mail: magnusp@strw.leidenuniv.nl

…Under the current system, if a postdoc substantially contributes with experiments and grant writing for a grant that is awarded to the PI, he/she benefits from the experience (especially if the postdoc is pursuing an academic position) and in securing funds for the lab. Beyond that, the postdoc should receive credit, via CV or Biosketch, for contributions made to the grant. It would also be beneficial for a postdoc who is transitioning to be able to take a portion of funds from grants to which he/she contributed substantially.

Bernardo A. Mainou

Department of Pediatrics Infectious Diseases, Vanderbilt University, Nashville, TN 37211, USA. E-mail: bernardo.mainou@vanderbilt.edu

…Creativity is crucial in science; therefore, young scientists should be exposed to an open arena, where their ideas are not constrained, guided, or limited by the structure of well-established research groups…Research institutions should create positions for young scientists without expecting them to join a given research group. This would create new research areas for the institution, and push new scientists to be more effective while fighting to get their own funding for research and equipment….

Nicolas Bambach

Biomicrometeorology Group, Department of Land, Air, and Water Resources, University of California, Davis, Davis, CA 95616, USA. E-mail: nbambach@ucdavis.edu

…PIs have been able to build their reputations purely on the achievements of postdocs, because they represent the most skilled, cheap labor force. From the point of view of a postdoc, it's the most challenging time in trying to build a base for one's interests, publish, and most important, be able to walk that tightrope of what can be taken with them and what the PI would like

to retain in the lab.... To give the postdoc more autonomy, I would strongly recommend that there be more funding opportunities that allow for postdocs to be more in charge of their projects and what they would like to pursue. This also serves as training to be able to manage money and plan experiments according to a fixed budget.

Suchitra D. Gopinath

Translational Health Science and Technology Institute (THSTI), NCR Biotech Science Cluster, Faridabad, 121001, India. E-mail: sgopinath@thsti.res.in

...There should be different kinds of postdocs available for each career track. We need more postdoc positions to help recent graduates transition into industry research, teaching, or administrative positions.

Kierstyn Schwartz

Working Bugs, LLC, East Lansing, MI 48823, USA. E-mail: kierstyn.schwartz@gmail.com

The postdoc employment should have a very concrete purpose, agreed upon before the start. For example: "I want to be ready to join a certain branch of industry," "I want to be able to lead a team and raise funds," or "I want to become a great tutor/teacher." A grant-giving institution would then limit the number of postdocs with each goal according to market demand (for example, 50% for the industry goal, 30% for the team lead goal, and 20% for the teacher goal). At the end of the postdoc period, the success should be evaluated and if the goal was not reached, consequences should follow; the university should have to pay the past postdoc salary from its own money, or future grant money should be reduced.

Michael Böttger

SKF, Department of Technology and Solutions, Steyr, 4401, Austria. E-mail: michael.boettger@skf.com

With an increasing number of Ph.D. graduates aiming to pursue nonacademic career tracks, there is an urgent need to revamp the traditional postdoc position, the majority of which do not prepare graduates for jobs outside academia. A systemic integration of career support into Ph.D. programs in all universities is critical to ensure that maximum returns are channeled from the national investment in graduate education. An experiential 1-year career-rotation program (not unlike first-year lab rotations) should follow graduation. Ph.D. graduates would have the opportunity to pursue three to four internships in different settings that value a terminal degree in their field of study. Examples of organizations for which the graduates might intern include science societies, policy think-tanks, law firms, regulatory affairs teams, nonprofit organizations, government agencies, and biotech start-ups. The graduates could be placed in these internships through a matching process facilitated by the university and funded by stakeholders in the process. Such a transition program would be transformative for the educational ecosystem because it would help students identify and gain a foothold in the career track of their choice, allow employers access to a pool of highly qualified job candidates, and help participating universities bolster their student success metrics.

Meera Govindaraghavan

College of Agriculture and Life Sciences, Cornell University, Ithaca, NY 14850, USA. E-mail: govind.meera@gmail.com

No more than two postdocs for a given person, up to 3 years each, should be allowed. This would force universities to hire people more permanently after this maximum 6-year period, rather than exploiting them with numerous postdocs.

Maciej Bilicki

Leiden Observatory, Leiden University, NL-2333 CA Leiden, Netherlands. E-mail: bilicki@strw.leidenuniv.nl

If by "postdoc", we mean a 2- to 3-year research position meant exclusively as a bridge between the Ph.D. program and a professorship, then I think the postdoc is obsolete. I would make two changes to improve it: (i) Make postdocs into permanent, non-faculty positions. Both individuals and institutions would benefit. A scientist who has to reapply for his/her job every 2 to 3 years—especially if that means moving to a new institution—has much less time to develop expertise....(ii) Remove the stigma associated with the postdoc—i.e., the idea that if you postdoc for more than 3 years, you must not be faculty material.... Why treat postdocs like lepers if they take more than 3 years to discover something exciting enough to rocket them into a faculty position?

Lisa Neef

GEOMAR Helmholtz Centre for Ocean Research, Kiel, Kiel, 24116, Germany. E-mail: lneef@geomar.de

...The problem is that the postdoc is currently an open-ended and unregulated career stage. This situation leaves young scientists exposed, because investigators at all levels need lots of active postdocs for their labs to be competitive. Academic science has accidentally evolved into a pyramid structure, where the vast majority of postdocs cannot expect to find long-term sustainable careers.... A pragmatic solution would be to limit laboratory size to 10 or fewer individuals, and to create sustainable and independent staff scientist positions....

Tomás Ryan

Picower Institute for Learning and Memory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: tr1@mit.edu

The postdoc position should be replaced by an academic research scientist position that is treated as a career.... Although research scientists should be held accountable for the progress of their research, the labs in which they work should be run with more oversight. Lead investigators should receive managerial training, and an overarching university body should evaluate the management of labs to ensure that research scientists receive appropriate guidance and regular feedback. Within labs, a substructure of collaborative research groups could provide new hires with direct supervision and would also provide senior lab members with the opportunity to develop managerial skills needed for their future success as lead investigators. However, each research scientist need not aim to run an independent lab, as one could rise to a senior research scientist or director level position within a lab, so long as he/she continues to be productive in his/her own research and contributes to the development of others in the lab. A key strength of this structure

is that it values contribution to the overall success of the group rather than solely rewarding individual accomplishments.

Nidhi Ruy

Medical and Scientific Affairs, Synapse Medical Communications, New York, NY 10017, USA. E-mail: nidhig01@gmail.com

...After completing my Ph.D. in Canada, I started a postdoc in Australia. The contrast in postdocs between these two countries is appreciable. In Australia, I am treated as a full and equal staff member and colleague. I can co-supervise graduate students and apply for grants as a chief/principal investigator, and I am paid a salary that allows me to adequately support my family. Postdoc positions are very different in Canada, where postdocs are often lumped in with graduate students, are not recognized for graduate student supervision, cannot apply for many of the grants full faculty can, and usually earn much lower salaries. Failing to treat postdocs as colleagues and provide them with reasonable funding opportunities and salaries risks losing their energy and expertise....

Matthew Mitchell

School of Geography, Planning, and Environmental Management, University of Queensland, Brisbane, QLD, 4072, Australia. E-mail: m.mitchell@uq.edu.au

In China, obtaining a postdoc after getting a Ph.D. used to be regarded highly, but this is no longer the case because postdocs in Chinese universities do not receive enough recognition. Recruitment advertisements often emphasize "overseas research experience preferred." A potential solution would be to strengthen global collaboration, which would help domestic postdocs improve their competitiveness by proposing original ideas, designing innovative experiments, and achieving leading results together with international colleagues....

Bo Cao

College of Life Sciences, Shaanxi Normal University, Xi'an, Shaanxi, 710119, China. E-mail: bocao@vip.qq.com

People usually know what is valuable after they lose it. If postdocs vanish, scientific advancement would inevitably suffer... The postdoc position is an indispensable bridge between beginners and experienced mentors...I suggest that scientific journals

include in the contents a special section that exclusively publishes research articles on studies conducted by postdocs in the capacity of leading investigators, thereby highlighting the contribution of postdocs. Such a simple change not only encourages postdocs to immerse themselves more into research, but also instantly eliminates the unrealistic impression that they do the same work as what students do....

Chun-Wai Ma

Department of Physiology, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China. E-mail: cwma2010@hku.hk

Postdocs who are raising a family while training for a scientific career need more support.... Postdocs pay serious penalties in dollars and, for mothers, in professional reputation when they have a child, which are exacerbated for underrepresented minorities. Increasing support for postdoc parents would increase diversity at the postdoc and faculty stages, especially in fields where postdocs are longer and more common.... Child care financial aid, eligibility for pre-tax dependent care payment plans, and backup care programs are needed to support postdoc parents. Postdoc fellowship policies should also be adjusted to support new parents. Postdocs on fellowships are often ineligible for maternity leave. And parenting a newborn does not stop when postdocs

return to work; a 1-year extension on the "postdoc clock" of eligibility for NIH's Pathway to Independence award, akin to the 1-year extension faculty get on their "tenure clock," is necessary.

Katherine L. Thompson-Peer, on behalf of P-Value* at UCSF

Department of Physiology, University of California, San Francisco, CA 94158, USA. E-mail: katherine.thompson-peer@ucsf.edu; pvalueUCSF@gmail.com

Postdocs should be treated like actual employees. ... We don't get the benefits of employees with less education, less training, and less experience. When you have been in higher education for more than 10 years, it would seem reasonable to get employee health (not student health or no health) insurance and retirement benefits. When a research technician straight out of college gets 6 to 12% retirement benefits, and a postdoc gets zilch, there seems to be a problem...!

Jared M. Fischer

Department of Molecular and Medical Genetics, Oregon Health and Science University, Portland, OR 97239, USA. E-mail: fischerj@ohsu.edu

I think the problems facing research make the postdoc position more important, not obsolete.... I suspect that a good number of fresh ideas in research would not have come about if there were one less step between the lab bench and grant writing...Perhaps there should be grant mechanisms for grad students funding a short postdoc, with the stipulation that it must be a complete departure from what the student worked on in grad school.

Phil Spear

Department of Pediatrics, University of Colorado, Denver, Anschutz, Aurora, CO 80045, USA. E-mail: interkin3tic@gmail.com

...The increasing rarity of tenure-track positions means that postdocs need to be prepared for jobs that are significantly less attractive. One must honestly answer the question "What am I willing to give up for a career in academia?" before proceeding toward a postdoc.

Kriti Charan

School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853, USA. E-mail: kc636@cornell.edu

Source of a scent. *The Soul of the Rose*, painting by John William Waterhouse (1908).

BIOCHEMISTRY

The flowering of a new scent pathway in rose

The monoterpene-based scent of roses is generated by an unusual biosynthetic route

By **Dorothea Tholl**[1] *and* **Jonathan Gershenzon**[2]

Floral scent is an important trait of ornamental roses that has provided sensual pleasures for humans since antiquity. However, most modern rose cultivars used for cut flowers have little fragrance as a result of breeding preferences for traits such as color and longevity. Restoring scent attributes by breeding or biotechnological means (*1*) requires a detailed understanding of the biosynthesis of rose scent. Rose fragrance consists of hundreds of volatile compounds with diverse biosynthetic origins whose amounts vary among the different rose varieties. Genomic approaches over the past 15 years have identified several genes and enzymes involved in rose scent production (*2*), but these efforts have not succeeded in elucidating the biosynthetic steps in the formation of geraniol and other monoterpenes that constitute one of the major groups of rose fragrances. On page 81 of this issue, Magnard *et al.* (*3*) describe the discovery of an unexpected enzymatic pathway to these scent compounds.

The biosynthesis of monoterpenes, the 10-carbon representatives of the enormous class of terpenoid natural products, usually requires the enzymatic action of terpene synthases. As one of the most celebrated groups of catalysts in nature (*4*), terpene synthases convert the 10-carbon prenyl diphosphate intermediates—in most cases, geranyl diphosphate—to a highly diverse range of cyclic and acyclic products by a carbocation-mediated reaction mechanism that is initiated by the ionization of the diphosphate group. Such terpene synthase activities have been previously associated with the formation of sesquiterpenes, 15-carbon terpenes, in rose scent (*5*). However, the enzymatic machinery leading to the biosynthesis of the 10-carbon monoterpenes, such as geraniol, has remained a mystery. Geraniol, an alcohol, has been demonstrated to be made in basil from geranyl diphosphate via a terpene synthase that simply cleaves the diphosphate moiety and allows the resulting carbocation to be quenched by water (*6*). Although using a terpene synthase to carry out such a simple reaction might be considered a waste of its catalytic potential, the ubiquity of terpene synthases in monoterpene formation has led to the belief that this is the sole route of geraniol biosynthesis in all plants.

Surprisingly, Magnard *et al.* demonstrated that rose flowers use an alternative route to produce geraniol by employing a diphosphohydrolase of a completely unexpected enzyme family (see the figure). Roses lack the full contingent of modern genetic and genomic tools available for model organisms such as *Arabidopsis thaliana*; instead, the authors had to painstakingly accrue evidence for the role of this enzyme in scent production. To sniff out how geraniol is formed in roses, they first compared a rose variety with a typical rose scent to one with low scent that produced very small amounts of geraniol and other typical monoterpene scent constituents. The gene with highest differential expression in the scented versus unscented varieties did not encode a geraniol-forming terpene synthase or any other obvious enzyme of monoterpene biosynthesis, but instead was a gene of

the Nudix family, which usually encodes enzymes that cleave nucleoside diphosphates from a variety of other chemical moieties (7). Expression of this gene correlated with the presence of geraniol and other monoterpenes in a survey of cultivars with different scent profiles. Moreover, it colocalized with a quantitative trait locus (QTL) for geraniol formation in a full-sib family of hybrid lines derived from a cross between two genotypes with different scent compositions. In addition, this Nudix hydrolase gene was found to be expressed almost exclusively in petals, and when knocked down by stable and

to be hydrolyzed to geraniol by a rose petal–derived phosphatase activity. Expression of the Nudix protein in *Nicotiana benthamiana* further supported its role in geraniol biosynthesis in planta.

The authors' success in elucidating this novel biosynthetic route to geraniol in rose flowers raises the thorny question of why such an unusual monoterpene biosynthetic pathway evolved in the first place. One answer may come from enzyme kinetics. With a Michaelis-Menten constant (K_M) for geranyl diphosphate that is two orders of magnitude lower than that of the geraniol-

from the methylerythritol 4-phosphate pathway (8), but the lack of photosynthesis in rose petals may result in diminished flux through the methylerythritol 4-phosphate pathway, so that terpene formation must rely on the mevalonate pathway located in the cytosol. In fact, rose petals were one of the first plant tissues shown to readily use the mevalonate pathway in monoterpene formation (9). The cytosolic location of the Nudix protein requires the provisioning of geranyl diphosphate in this cellular compartment by synthesis in situ or possibly by transport from plastids, the usual site of geranyl diphosphate formation in plants. It will be interesting to find out how roses accomplish this feat, one very seldom documented in plant metabolism.

Other unanswered questions revolve around when the pathway to geraniol originated in roses and how widely distributed it is in other plants. The recruitment of Nudix hydrolase into this pathway may have been a recent event in the evolution of roses, or this type of enzyme may be much more widely involved in monoterpene formation in plants. There are now an increasing number of examples in the plant kingdom where the same types of natural products are produced by different pathways in different plant lineages, and thus arose independently (10, 11). The value of monoterpenes in flowers—for example, as pollinator attractants (12) or as defenses against pathogens—may have led to the repeated evolution of different biosynthetic pathways for them.

An unexpected source of scent. The formation of geraniol and other geraniol-derived monoterpenes typical of the rose proceeds from geranyl diphosphate (GPP) to geranyl monophosphate (GP), catalyzed by NUDX1. GP is hydrolyzed to geraniol, and this product and other downstream monoterpene metabolites are then volatilized (blue arrows). The pathway is presumably situated in the epidermal cells of rose petals (depicted are conical cells of the adaxial epidermis). The subcellular compartmentalization of the pathway is not fully understood. Although NUDX1 is localized to the cytosol, its substrate GPP might be produced from the 5-carbon precursors, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), formed in the cytosol or in plastids. OPP, pyrophosphate. MEP, methylerythritol 4-phosphate.

The value of rose monoterpenes for humans is driven by the enormous popularity of roses in the floral industry and as sources of essential oils for natural fragrances. The discovery of a Nudix hydrolase involved in the formation of geraniol now provides a reliable molecular marker for a major group of floral scent compounds in roses that can be exploited to enhance fragrance in these iconic flowers. ∎

REFERENCES

1. N. Dudareva, A. Klempien, J. K. Muhlemann, I. Kaplan, *New Phytol.* **198**, 16 (2013).
2. G. Scalliet *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5927 (2008).
3. J.-L. Magnard *et al.*, *Science* **349**, 81 (2015).
4. J. Degenhardt *et al.*, *Phytochemistry* **70**, 1621 (2009).
5. I. Guterman *et al.*, *Plant Cell* **14**, 2325 (2002).
6. Y. Iijima *et al.*, *Plant Physiol.* **134**, 370 (2004).
7. A. G. McLennan, *Cell. Mol. Life Sci.* **63**, 123 (2006).
8. A. Hemmerlin, J. L. Harwood, T. J. Bach, *Prog. Lipid Res.* **51**, 95 (2012).
9. M. J. O. Francis, M. O'Connell, *Phytochemistry* **67**, 1110 (1969).
10. J. Jirschitzka *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 10304 (2012).
11. R. Sohrabi *et al.*, *Plant Cell* **27**, 874 (2015).
12. K. J. R. P. Byers, J. P. Vela, F. Peng, J. A. Riffell, H. D. Bradshaw Jr., *Plant J.* **80**, 1031 (2014).

transient RNA interference, the result was a reduction in monoterpene levels. After heterologous expression in *Escherichia coli*, the encoded protein was found to have a clear role in geraniol formation by converting geranyl diphosphate in vitro to geranyl monophosphate, which was in turn found

forming terpene synthase known from basil, the Nudix hydrolase may be a much more efficient catalyst than a terpene synthase for forming this simple monoterpene. This may be particularly advantageous in roses, where monoterpene scent compounds are produced in relatively large amounts.

Another factor that may have led to the recruitment of the Nudix protein for monoterpene formation is its cytosolic location. Monoterpenes are thought to be manufactured principally in plastids, using substrates

¹Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA. ²Department of Biochemistry, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany. E-mail: tholl@vt.edu

BIOMECHANICS

When it's hip to be square

The square shape of the seahorse tail helps it resist mechanical deformation

By **Miriam A. Ashley-Ross**

Most animals and plants approximate a cylinder in shape, and where junctions occur (as with branches of trees or limbs on animals), those corners are "faired," meaning smoothly curved so that one surface grades into the next (*1*). When living organisms deviate from the norm, there's usually a good biomechanical reason: a clue to some specific problem that needs to be solved. Among their suite of unusual characteristics, seahorses possess a true oddity: a prehensile tail with a square, rather than round or elliptical, cross-sectional shape. On page 46 of this issue, Porter *et al.* (*2*) report that there are distinct mechanical advantages to being square. Using three-dimensional (3D) printing to construct physical models, the team demonstrates that the multiplated anatomy of the square seahorse tail shows greater resistance to

Department of Biology, Box 7325, Wake Forest University, Winston-Salem, NC 27109, USA. E-mail: rossma@wfu.edu

Getting a grip on the seahorse tail. Although the seahorse uses its tail for anchoring itself, the square cross-section shape apparently serves another function. It makes the tail harder to puncture when bitten by a predator.

mechanical deformation than a similar model that has a round cross section.

Rather than the small surface scales typical of most fish, seahorses have a partially rigid armor of expanded subdermal bony plates, often with spiky projections, that completely encase the body. Only the tail has the square cross-sectional shape, suggesting some specialized function (see the first figure). It is tempting to assume that the square shape might be related to its prehensile capabilities (*3*, *4*); the size of the plates decreases toward the tip of the tail, allowing for coiling into a tight logarithmic spiral when bent in a ventral direction (*5*), and the flat ventral surface might provide greater grip (*2*), although this latter notion remains to be experimentally tested. However, close relatives of seahorses (pipefishes and seadragons) lack tail prehension, but still have square tails (*6*, *7*). Thus, the shape seems not to have evolved for grasping.

Porter *et al.* suggest that a square tail provides superior resistance against compressive injury (i.e., a bite from a predator), and that the improved armor function may have been the driving force in its shape evolution. To test the hypothesis that the square shape is mechanically superior, they made computer-aided design (CAD) models of simplified seahorse vertebral columns and subdermal plates. They then 3D printed them at greater-than-life size, joining the pieces with springs and elastic bands to mimic the natural connections between the bony elements. They also printed models in which the L-shaped corner plates were replaced with quarter-circles, and tested both versions by subjecting them to impacts (striking them with a rubber mallet) and compression in one and two dimensions.

As they predicted, the square version demonstrated greater resistance to compression (higher forces required to make the "tail" deform by a given amount), and tolerated more actual compression, in both one and two dimensions, without damage. This difference appears to result from the ability of its overlapping plates to linearly slide over one another (see the
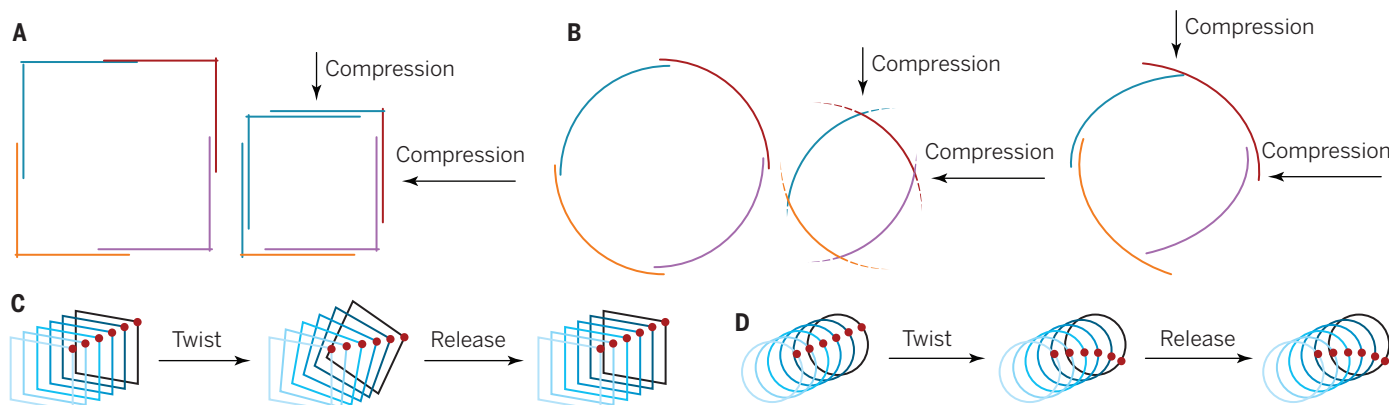
Consequences of shape when loading arrays of rings. (**A**) A square ring changes size but not shape when biaxially compressed because of the ability of the corner plates to slide over one another. (**B**) A round ring (left image) cannot retain its shape when biaxially compressed, because the plates interfere with one another. In order for the array to compress to the same extent as the square ring, the rounded corner plates would have to interpenetrate one another (center image). Actual rings deform when forced together (right image). (**C**) When released after being twisted, the array of square rings returns passively to its original shape. (**D**) The array of round rings does not return to its original orientation upon release from torsion.

second figure, panel A), which does not alter the shape of the structure. In contrast, when the round tail was compressed, no such linear sliding is possible because of the shape of the plates; rather, the plates are forced out of register with one another (see the second figure, panel B), disrupting their load-bearing orientation and thus their resistance to deformation.

The square array was also more resistant to torsion (long-axis twisting) than the round array (see the second figure, panels C and D). Finally, the square array naturally returned to its initial alignment (all of the square segments in register with one another) after an impact, whereas the round tail segments would remain skewed with respect to one another. Torsion resistance and the passive return property may provide protection to the delicate spinal cord, which is easily injured by stretching.

Current research in bioinspired robotics is moving away from strictly rigid elements (*8*) and toward soft actuators, but these soft robots are more susceptible to damage (*9*). Thus, the combination of mechanical strength with flexibility exemplified by the seahorse tail is attractive, though in some cases it appears that a strict interpretation of the anatomy is less than optimal. In this regard, in Porter's presentation of this work at the annual meeting of the Society for Integrative and Comparative Biology in January (*10*) (which I had the pleasure to hear), he pointed out in his introduction that he had hoped to use the seahorse tail as the model for a steerable catheter (but with a round cross section). By the end of the talk, he expressed his chagrin that the round tail failed to match the performance of the square one. Nonetheless, in situations where the sharp corners are not a drawback, a flexible armor based on the seahorse tail may prove advantageous, bending smoothly while still affording protection to structures within—proving that it's not always a bad thing to be "square." ∎

REFERENCES

1. S. Vogel, *Comparative Biomechanics: Life's Physical World* (Princeton Univ. Press, Princeton, NJ, ed. 2, 2013).
2. M. M. Porter, D. Adriaens, R. L. Hatton, M. A. Meyers, J. McKittrick, *Science* **349**, aaa6683 (2015).
3. M. E. Hale, *J. Morphol.* **227**, 51 (1996).
4. T. Praet *et al.*, *Int. J. Num. Met. Biomed. Eng.* **28**, 1028 (2012).
5. M. M. Porter, E. Novitskaya, A. B. Castro-Ceseña, M. A. Meyers, J. McKittrick, *Acta Biomater.* **9**, 6763 (2013).
6. R. H. Kuiter, *Seahorses and Their Relatives* (Aquatic Photographics, Seaford, Australia, 2009).
7. C. Neutens *et al.*, *J. Anat.* **224**, 710 (2014).
8. R. Playter *et al.*, *Proc. SPIE* **6230**, 1 (2006).
9. R. F. Shepherd *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20400 (2011).
10. www.sicb.org/meetings/2015
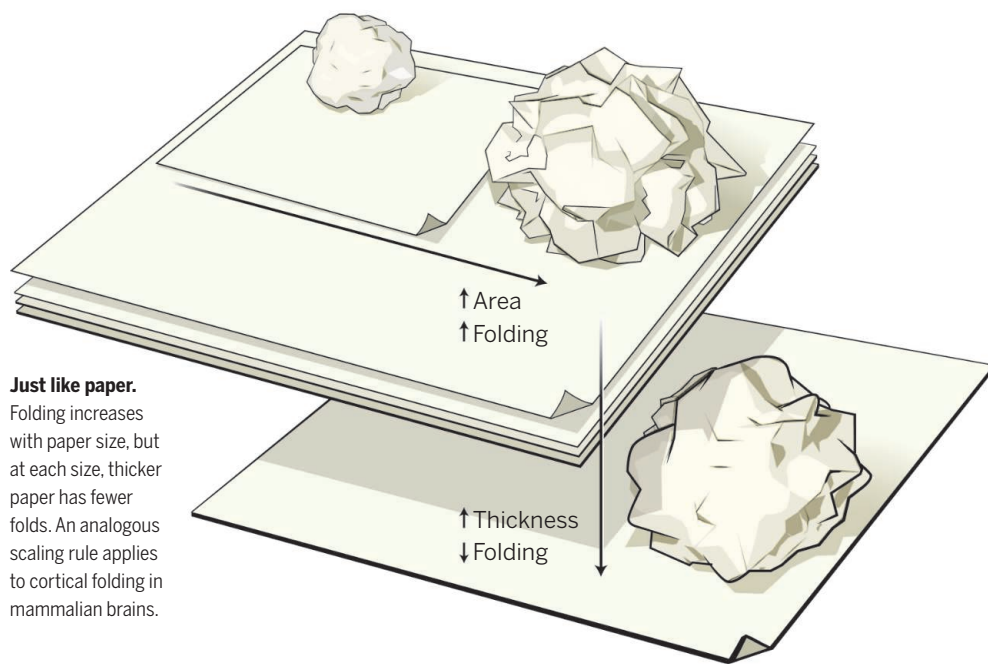
Just like paper. Folding increases with paper size, but at each size, thicker paper has fewer folds. An analogous scaling rule applies to cortical folding in mammalian brains.

BRAIN EVOLUTION

Knowing when to fold them

Crumpled paper is a possible model for cortical folding in mammalian brains

By **Georg F. Striedter**[1] *and* **Shyam Srinivasan**[2]

Almost 100 years ago, in an influential book titled *On Growth and Form*, D'Arcy Thompson called for the integration of biology with mathematics and physics (*1*). Fulfillment of his vision has been slow, but more and more scientists are heeding Thompson's call, attempting to explain problems of "growth and form" through physics and mathematics. An excellent example of this trend is the report by Mota and Herculano-Houzel (*2*) on page 74 of this issue. They provide a novel mathematical description of how the degree of neocortical folding in mammalian brains varies with other biological parameters, notably cortical thickness and surface area, and they offer the crumpling of a paper ball as a physical model to explain the observed scaling rules.

Many scientists have looked for scaling rules that govern folding of the mammalian neocortex, but a universal scaling rule has remained elusive. For example, if you graph the total surface area of the neocortex (following the surface deep into all folds) against its exposed surface area (which you can measure by covering the cortex with a piece of cellophane), you get two different scaling laws. In species with smooth cortices (such as mice and rats), the best-fit line has a slope of 1.0 because the total and exposed areas are by definition the same. By contrast, for species with folded cortices (such as the human brain), total surface area increases faster than the exposed surface area (i.e., the slope is substantially greater than 1). Thus, you need at least two lines to fit the data points (*3*). Moreover, several species are outliers given those scaling rules. Dolphins and whales, for example, exhibit more cortical folds than other mammals for the same cortical surface area (*4*).

Mota and Herculano-Houzel manage to bring these data together under a single, universal scaling law by including cortical thickness in their equation. Specifically, they discovered that the exposed cortical surface area is proportional to the total surface area multiplied by the square root of cortical

[1]*Department of Neurobiology & Behavior, University of California, Irvine, CA 92697, USA.* [2]*Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA, and Kavli Institute for Brain and Mind, University of California, San Diego, La Jolla, CA 92037, USA. E-mail: georg.striedter@gmail.com*

thickness (squared correlation coefficient = 0.997). Another way of saying this is that the degree of cortical folding increases with total cortical area (i.e., larger brains tend to be more folded) but increases more rapidly for thinner cortices. Other scientists had noted previously that, as brains become larger and more folded, surface area increases faster than thickness (5), but none had captured the relationship as precisely.

To explain their observed scaling rule, Mota and Herculano-Houzel crumpled sheets of paper varying in both thickness and area. Remarkably, these paper balls obey the same general scaling law as the neocortex (see the figure): Their degree of folding increases with total surface area, but folding increases faster for thinner sheets. Moreover, the scaling exponent is nearly identical for crumpled paper and folded cortices. On the basis of these observations, Mota and Herculano-Houzel modeled both kinds of structures as self-avoiding (i.e., nonintersecting) surfaces that minimize their effective free energy, which means that the structures move to

---

> **"...paper balls obey the same general scaling law as the neocortex..."**

---

minimize the balance of internal and external stresses until, eventually, equilibrium is reached. In the case of the paper balls, the crumpled balls relax after being compressed. In the case of the cortex, the stresses are mainly internal; according to Mota and Herculano-Houzel, they result mainly from the tension generated by axons that connect the growing gray matter to (and through) the underlying white matter (6).

Although the physics of paper balls has been explored in several previous studies (7), the analogy to cortical folding is a new scientific proposal. Most existing models of cortical folding involve the tangential expansion of an outer layer that is bonded to a more slowly growing central core (or deeper layer); swelling composite hydrogels are of this class (8), as are some computational models (9). Folding in these other models results not from external compression but from differential tangential expansion, which causes buckling of the more rapidly expanding outer layer; axonal tension is thought to play at best a minor role. Although Mota and Herculano-Houzel's proposal is sure to reinvigorate a long debate about the importance of axons in shaping cortical folds (10), axonal ten-

sion is not essential to their model (after all, the crumpled paper sheets do not have little strings attached to them). The critical element of almost all current models is that folding minimizes overall stress, however it is created.

The principal problem with the paper ball analogy is that it describes the folding of a structure that no longer grows, whereas the cortex folds during development. The scaling laws likewise only describe relationships between adult parameters. For example, Mota and Herculano-Houzel note that cortical folding occurs in any taxonomic group for which total cortical surface area increases faster than the square of cortical thickness. This scaling rule sounds like a developmental rule, but it describes only adult relationships and may result from a variety of different developmental mechanisms. For instance, cortical area may increase faster than the square root of cortical thickness if species with larger cortices expand the surface area of their cortical progenitor zone but retain the ancestral program for how to transform a patch of progenitor zone into adult gray matter (11). Alternatively (or in addition), species with larger cortices may change neuronal migration in such a way that a patch of ventricular zone gives rise to gray matter with an expanded surface area, relative to the ancestral condition (12, 13). Either way, cortical surface in the adult would expand without increasing cortical thickness, thus fulfilling this adult scaling rule.

The larger message is that adult scaling laws do not reveal what evolutionary changes in development generate those laws. Mota and Herculano-Houzel have beautifully described some adult morphological relationships—what D'Arcy Thompson called "form"—but much more work is needed to explain what types of changes in development—or "growth"—gave rise to them. ■

REFERENCES

1. D. W. Thompson, *On Growth and Form* (Cambridge Univ. Press, Cambridge, ed. 2, 1942).
2. B. Mota, S. Herculano-Houzel, *Science* **349**, 74 (2015).
3. M. A. Hofman, *J. Hirnforsch.* **32**, 103 (1991).
4. P. R. Manger, M. Prowse, M. Haagensen, J. Hemingway, *J. Comp. Neurol.* **520**, 2430 (2012).
5. J. W. Prothero, J. W. Sundsten, *Brain Behav. Evol.* **24**, 152 (1984).
6. D. C. Van Essen, *Nature* **385**, 313 (1997).
7. A. S. Balankin, D. S. Ochoa, I. A. Miguel, J. P. Ortiz, M. A. Cruz, *Phys. Rev.* **81**, 061126 (2010).
8. T. Tallinen, J. Y. Chung, J. S. Biggins, L. Mahadevan, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 12667 (2014).
9. R. Toro, Y. Burnod, *Cereb. Cortex* **15**, 1900 (2005).
10. G. Xu et al., *J. Biomech. Eng.* **132**, 071013 (2010).
11. P. Rakic, *Science* **241**, 170 (1988).
12. L. Ronan, P. C. Fletcher, *Brain Struct. Funct.* 10.1007/s00429-014-0961-z (2014).
13. G. F. Striedter, S. Srinivasan, E. S. Monuki, *Annu. Rev. Neurosci.* 10.1146/annurev-neuro-071714-034128 (2015).

10.1126/science.aac6531

Landscapes in the lab

A table-top experiment can probe the processes of landscape evolution

By **Scott W. McCoy**

Rivers dissect much of Earth's surface into conspicuous networks of valleys and hillslopes. Viewed from an airplane, these erosional networks might appear fractal; a close look at a part presents a view indistinguishable from the whole (1). But meter-scale topographic measurements reveal that the extent of landscape dissection by downcarving streams is finite and limited by the size of smooth, undissected hillslopes that separate adjacent valleys (1, 2). Within a drainage basin, hillslopes commonly have a characteristic size (see the figure), yet across different landscapes this scale can vary from meters to a kilometer (1, 3). What imposes this scale on a landscape, and why does it change from one landscape to another? On page 51 of this issue, Sweeney *et al.* (4) demonstrate using controlled laboratory experiments that landscape scale is set by a competition between river incision that cuts valley networks and diffusive hillslope processes that fill them in. Their experiments highlight that the extent of valley incision is an emergent, dynamic landscape characteristic that depends on a delicate balance of forces shaping the landscape.

More than a century ago, Davis (5) and Gilbert (6) proposed that the limit of landscape dissection, and the point at which valleys transition to hillslopes, emerges from a competition between hillslope processes that smooth the landscape and channel processes that cut valleys. Hillslope processes that cause soil creep are biotic (for example, animal burrowing or tree throw) or abiotic (for example, frost heave or rain splash). Channel processes derive their incisional power from the concentrated action of flowing water and debris.

The evolution of soil-mantled landscapes can be modeled using an advection-diffusion equation derived by combining conservation of mass with physically based equations for sediment transport by creep and surface water flow (7). It has been shown (2, 3) that the process competition of Davis can be encoded in one of the equation's nondimensional

groups, its Péclet number, $Pe = KL^2/D$, which captures the relative vigor of advective channel processes, represented by the coefficient K, and diffusive hillslope processes, represented by D, at a given length scale, L. The framework predicts that in landscapes where Pe is larger at a chosen L, stream incision dominates erosion over an expanded fraction of the terrain; valley networks should be more extensive, and hillslopes should be smaller. Conversely, if Pe is smaller, stream incision dominates erosion over a reduced fraction of the terrain; hillslopes should be larger and the valley network less extensive. Perron *et al.* (*3*) showed, using a variety of soil-mantled field sites, that the theory pre-

rates of diffusive versus advective transport by changing the relative amount of water applied in the form of large drops versus mist. Precise control over imposed initial conditions, boundary conditions, and material properties, coupled with high temporal and spatial resolution topographic measurements, allowed them to document the emergence of hillslopes and valleys, as well as calculate time-averaged values for the process rate parameters, D and K. Sweeney *et al.* clearly demonstrate that the extent of valley incision increases with increasing values of Pe and, hence, is set by a balance between diffusive hillslope processes and advective channel processes, thereby con-

the benefits of controlled experimentation and the ability to manipulate a single variable while all others remain constant outweigh these scaling uncertainties (*8*).

Sweeney *et al.* add important experimental evidence demonstrating that the advective-diffusion equation is a useful description of how soil-mantled landscapes evolve and that the relative intensities of channel versus hillslope processes encoded in the Péclet number determine fundamental landscape properties. What is less clear, however, is the functional form of relationships that relate these important rate parameters, D and K, to measurable properties of process mechanics, climate,
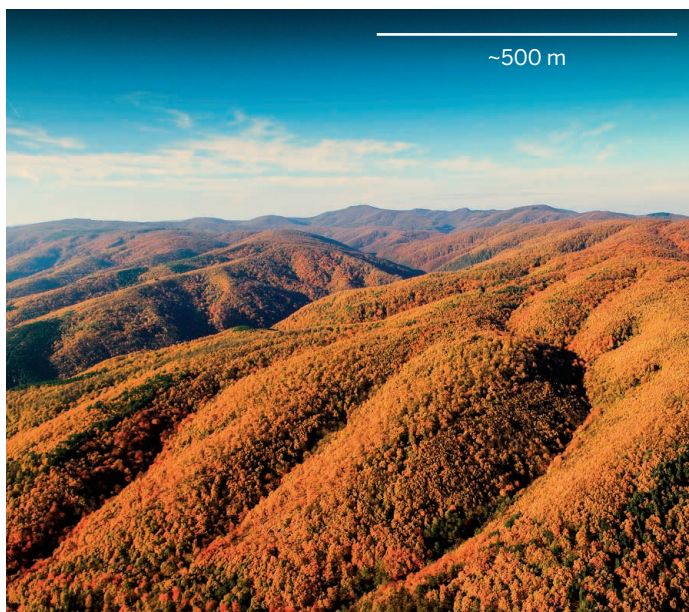
A matter of scale. Sweeney *et al.* (*4*) show that the difference in scale that can arise between landscapes is because each landscape has a unique balance between the vigor of river incision that cuts valley networks and diffusive hillslope processes that fill them in.

PHOTOS: (LEFT) ELENA CHUKHLEBOVA/THINKSTOCK; (RIGHT) ROMULIC-STOJCIC/HORIZONTAL/LUMI IMAGES/CORBIS

dicted the characteristic size of hillslopes and valley spacing.

Before the experiments of Sweeney and colleagues, the advection-diffusion theory for the evolution of soil-mantled landscapes had not been subjected to tests in a controlled laboratory environment. Using a novel setup, Sweeney *et al.* added both diffusive and advective sediment transport processes into a controlled landscape evolution experiment. To produce hillslope disturbance mimicking soil creep, they bombarded the experimental landscape with large, energetic water drops that dislodged sediment upon impact. To produce channel incision without inducing additional hillslope transport, the landscape was misted. They then manipulated the relative

firming the theoretical predictions.

One implication of their results is that perturbations to this process-rate balance, driven by changes in climate or land use, for example, will change the reach of erosional valley networks. Thick, fertile soils and infrastructure are commonly located on soil-mantled hillslopes. Network expansion can transform fertile farmland into barren badlands by stripping away nourishing soil.

A topic of discussion regarding the experiments from Sweeney *et al.* will likely be the degree to which these miniature experiments are accurate models of real landscapes. A dynamic scale model for landscape evolution has never been realized (*8*), and the experiments conducted by Sweeney *et al.* are no exception. Scaling issues arise when squeezing an entire erosional network into a 0.5 m by 0.5 m box such that some care and caution may be needed before the results are declared definitive. But

substrate, land use, life, and tectonics. Quantifying these relationships is necessary for geomorphology to be a truly predictive science. The experiments presented by Sweeney *et al.* point the way to opportunities in manipulated, controlled experimentation that should help to reveal controls on rate parameters as well as further probe the mechanics of landscape evolution. ■

REFERENCES

1. D. R. Montgomery, W. E. Dietrich, *Science* **255**, 826 (1992).
2. J. T. Perron, J. W. Kirchner, W. E. Dietrich, *J. Geophys. Res.* **113**, 1 (2008).
3. J. T. Perron, J. W. Kirchner, W. E. Dietrich, *Nature* **460**, 502 (2009).
4. K. E. Sweeney, J. J. Roering, C. Ellis, *Science* **349**, 51 (2015).
5. W. M. Davis, *Science* **20**, 245 (1892).
6. G. K. Gilbert, *J. Geol.* **17**, 344 (1909).
7. A. D. Howard, *Water Resour. Res.* **30**, 2261 (1994).
8. C. Paola, K. Straub, D. Mohrig, L. Reinhardt, *Earth Sci. Rev.* **97**, 1 (2009).

University of Nevada Reno, Department of Geological Sciences and Engineering, Reno, NV 89557, USA. E-mail: scottmccoy@unr.edu

Ivory pyre. A ranger from the Kenya Wildlife Service walks past 15 tons of elephant tusks burned in Nairobi, Kenya, on 3 March 2015.

CONSERVATION

Can DNA foil the poachers?

Forensic data help to identify elephant poaching hotspots

By **A. Rus Hoelzel**

Protecting wildlife helps to sustain ecosystems and the services they provide, but beyond that, wildlife species have an intrinsic value. It is hard to imagine a world without iconic species such as the African elephant. However, elephants carry with them an appendage of great commercial value: their tusks. Conservation efforts struggle to keep pace with illegal takes by poachers fueled by the promise of great financial gain. On page 84 in this issue, Wasser *et al.* (*1*) provide an example of how modern technologies are making life harder for the poachers.

The effective conservation and manage-

*School of Biological and Biomedical Sciences, Durham University, South Road, Durham DH1 3LE, UK.
E-mail: a.r.hoelzel@durham.ac.uk*

ment of wildlife is a complex and challenging problem, requiring detailed knowledge about the species, its environment, and the threats it faces. Conservation success sometimes comes from reduced anthropogenic pressures, such as the diminishing market for oil from northern elephant seals after hunting left this species too rare to sustain the industry (*2*). At the same time, legislation and enforcement are essential and were part of the process of recovery for the northern elephant seal, which is now protected in both Mexico and the United States. A dramatic and controversial example of success that relied heavily on enforcement was the essentially paramilitary approach of the Kenya Wildlife Service, sanctioned by the Kenyan government of the late 1980s, which reversed an epidemic of African elephant poaching in Kenya.

However implemented, enforcement and

effective management rely critically on good information about the nature and extent of the human impact. For example, since the 1940s, the International Whaling Commission has set catch quotas and received catch level reports from participating countries. In 1993, a Russian official announced that catch levels by Soviet vessels had been grossly underreported for decades and that protected species (where the quota was zero) had also been taken (*3*). These activities negatively affected various species, and at least one struggles to recover (*3*). Permitted whale takes are now cataloged on genetic registers and commercial products are routinely checked, helping to reveal illegal catches when they occur (*4*).

However, a register on its own cannot address the problem of poaching directly. For that, illegal takes need to be identified and prevented—a very difficult task. In an exten-

sive program of crime scene identification, Wasser *et al.* now apply DNA typing to exactly this objective.

The crime is the renewed epidemic of elephant poaching across Africa. The problem is great. Tens of thousands of elephants are illegally killed every year and many tons of ivory seized by enforcement agencies (see the photo)—51 tons in 2013. As Wasser *et al.* note, that scale of known illegal ivory suggests a take of over 50,000 elephants that year, compared to an estimated 434,000 to 683,000 African elephants left in the world in 2012 (*5*). Consequently, populations are in decline (*6*), implying an extinction risk if poaching continues at this rate. Wasser *et al.* use DNA forensics on a continental scale to match seized hauls of ivory to the geographic locations of poaching activity. If accurate and sufficiently fast, these data could greatly enhance the potential for effective enforcement.

> "*Wasser et al.'s results allow the park systems where poaching occurs to be identified, thus providing local authorities with invaluable information toward more effective enforcement.*"

There are two essential aspects to their method. First, the authors assessed the population genetics of the source populations by genotyping 1350 African elephants (including both savannah and forest elephants) from 71 locations across the range of the species. They then genotyped the seized materials and assigned individual genotypes back to putative source populations. These two aspects are interdependent and require accurate data and diverse genetic markers. Furthermore, there must be detectable genetic differentiation among geographic populations; otherwise, there will be no signal from which to identify sample origin. Forest and savannah African elephants are distinct at least at the subspecies level (*7*). For each, population-level differences provide the potential for good resolution, although the scale of population structure is fairly large in this highly mobile species (*8, 9*). By using the genetic data to assign reference individuals back to their known source populations, Wasser *et al.* demonstrate an accuracy of ~300 to 500 km for their method.

There is a devil in the details. Sampling large numbers of animals to build a full ref-erence set of genotypes is difficult for wildlife species and often relies on noninvasive sampling—for example, from excrement—but this DNA may be quite degraded. The animal parts used to assign particular animals back to their source population also typically provide degraded or low-concentration DNA. However, there is good precedent for success with processed materials. For example, testing for illegal whale takes often involves processed meat (*4*), and shark species have been identified from fin soup and supplement pills made from shark cartilage (*10*).

All the same, careful controls need to be undertaken to take account of DNA degradation to ensure genotype accuracy. Some data may be lost when quality is insufficient. Wasser *et al.* provide a very useful analysis of this in their supplementary tables, where they show the proportion of loci not lost to quality control (often less than half) for the different sample sets. This has further implications because the resolution of population assignments depends on the level of information provided in the genetic screen; if loci are lost to quality control, assignments become less precise.

In many cases, Wasser *et al.*'s results allow the park systems where poaching occurs to be identified, thus providing local authorities with invaluable information toward more effective enforcement. Of course, the finer the resolution and the faster the result, the more likely it is that enforcement can result in prevention before the poachers move on to new pastures. Fortunately for wildlife, technology also moves on. It is becoming increasingly realistic and affordable to screen hundreds or even thousands of loci at once in an automated array system that works well with degraded DNA. For example, such a system has been developed to monitor wolf populations in Europe (*11*), achieving both high precision and fast turnaround times. The DNA forensics approach exemplified by Wasser *et al.*'s study can be extended with these further innovations and integrated with other effective approaches (*6*). Hopefully this will start to turn the tide for the African elephant and other threatened wildlife. ∎

REFERENCES

1. S. K. Wasser *et al.*, *Science* **349**, 84 (2015).
2. B. C. Busch, *The War Against the Seals: A History of the North American Seal Fishery* (McGill-Queen's Press, 1987).
3. Y. V. Ivashchenko *et al.*, *Mar. Fish. Rev.* **73**, 1 (2011).
4. C. S. Baker *et al.*, *Biol. Lett.* **6**, 647 (2010).
5. www.elephantdatabase.org
6. G. Wittemyer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13117 (2014).
7. A. L. Roca *et al.*, *Science* **293**, 1473 (2001).
8. J. B. Okello *et al.*, *J. Hered.* **99**, 443 (2008).
9. L. S. Eggert *et al.*, *Conserv. Biol.* **28**, 107 (2014).
10. A. R. Hoelzel, *Conserv. Genet.* **2**, 69 (2001).
11. R. H. S. Kraus *et al.*, *Mol. Ecol. Resour.* **15**, 295 (2015).

ENZYMOLOGY

It costs more than a nickel

Lactate racemase contains a complex nickel cofactor

By **Deborah Zamble**

Biology class teaches us that enzymes are proteins, but many enzymes rely on supplementary nonprotein cofactors for catalytic activity. Some enzyme active sites contain elaborate auxiliary components that provide chemical versatility well beyond the capabilities of naked proteins. Such enzymes often require a lot of extra effort to prepare, entailing synthesis and/or mobilization of multiple scarce components. On page 66, Desguin *et al.* (*1*) show that the catalytic center at the active site of lactate racemase contains both a prosthetic group derived from nicotinic acid (vitamin B_3) and a nickel ion. This organometallic cofactor bears several features that have not been seen before in nature and raises questions about both the biosynthesis and the catalytic reactivity.

Although nickel is used by only a few enzymes, they cover a diverse array of chemical reactions (*2*–*5*). The first enzyme found to require at least one nickel ion cofactor was urease, which catalyzes the hydrolysis of urea into ammonia and carbon dioxide. Today, the list includes enzymes that produce or use carbon monoxide, acetyl–coenzyme A, hydrogen gas, or methane, as well as several that provide protection from toxic by-products. Each of these nickel enzymes has a distinctive active site, and many include additional components or protein modifications.

Last year, Desguin *et al.* showed that lactate racemase (LarA) also uses nickel for activity (*6*). LarA allows microorganisms to metabolize either the L- or D-isomer of lactate. It can also produce D-lactate, which provides resistance to the antibiotic vancomycin when it is incorporated into the peptidoglycan layer of bacterial cell walls. In *Lactobacillus plantarum*, *larA* is in the multigene *larABCDE* operon (*7*), which is downstream of a second operon encoding a putative nickel membrane transporter (*6, 8*). Desguin *et al.* showed that nickel supplementation and several other Lar proteins are required to generate active LarA. Metal analysis indicated that two of these accessory proteins can bind nickel, and ac-
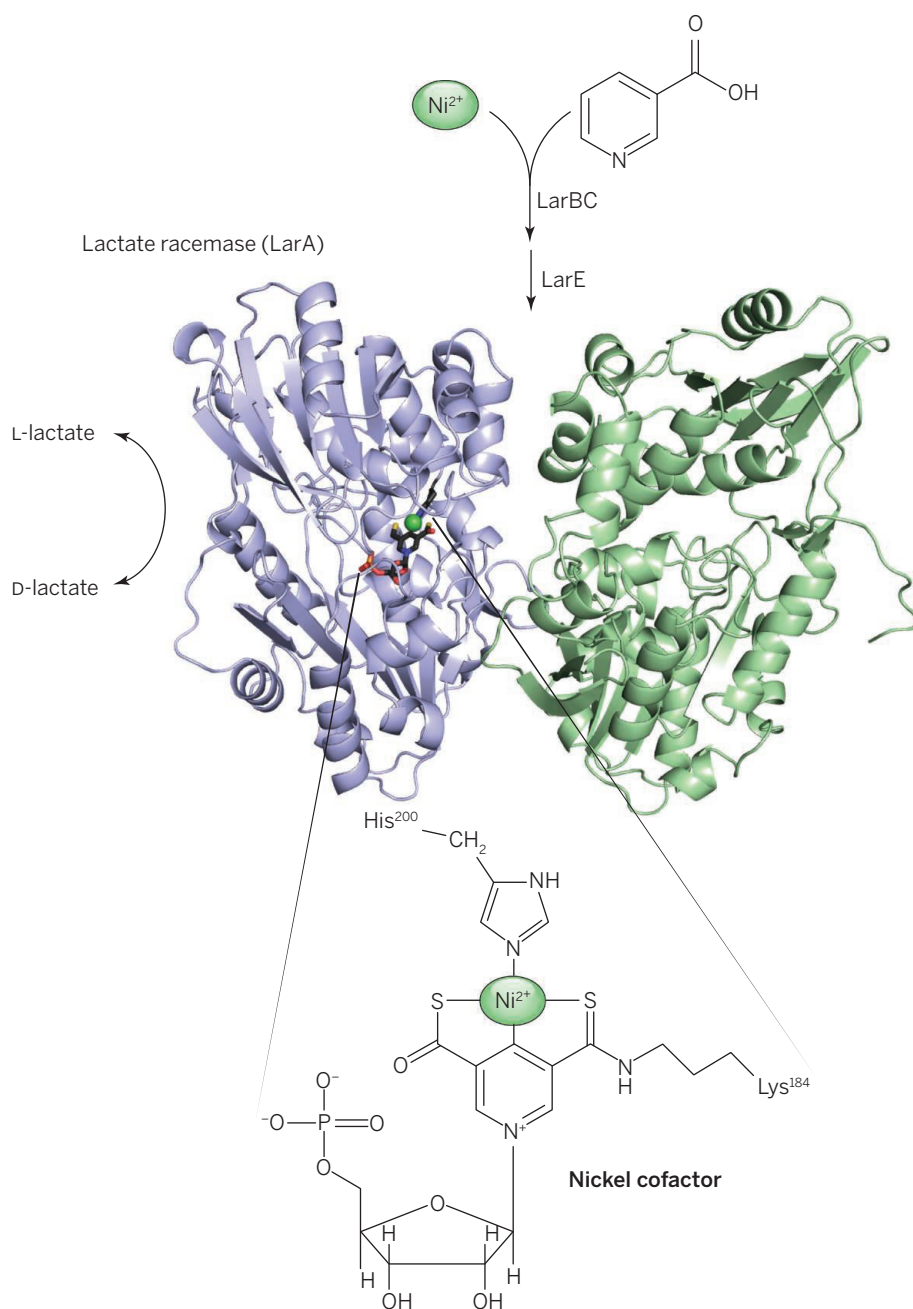
The active site of lactate racemase. Biosynthesis of the nickel enzyme lactate racemase (LarA) requires the activity of several accessory proteins (LarB, LarC, and LarE). Desguin *et al.* report the structure of this unusual metal center, which includes a tailored nicotinic acid.

tivation experiments suggest that LarE is first loaded up by LarBC and then delivers nickel to LarA (*6*). The requirement for auxiliary factors to produce metalloenzymes is not uncommon (*9*). For instance, many nickel enzymes are assembled by a suite of accessory proteins, including one or more metallochaperones that deliver the metal cofactor (*10–12*).

Departments of Chemistry and Biochemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada.
E-mail: deborah.zamble@utoronto.ca

But as the structure of LarA now reported by Desguin *et al.* makes clear, production of this enzyme involves far more than the insertion of a nickel ion. The authors use a combination of mass spectrometry, isotopic labeling, and x-ray crystallography to define a new metal center that is unusual in several ways (see the figure). First, the nicotinic moiety is highly tailored to have a thioamide at C3 of the pyridine ring, which covalently links the cofactor to a lysine side chain, as well as a thiocarboxylic acid at C5. Second, the nickel forms a metal-carbon

bond with C4 of the ring. The two sulfurs of the cofactor also coordinate the metal, resulting in a tridentate coplanar ligand that is dubbed a pincer complex. The fourth ligand in the distorted square planar coordination site is supplied by the only protein component, a histidine residue.

How is this cofactor made? The fact that the Lar accessory proteins are sufficient for recombinant expression of this enzyme suggests that they accomplish the complete biosynthesis. Working out the mechanisms of these proteins is sure to reveal some interesting chemistry. Perhaps a more basic question, however, is why an organism would go to the trouble of making such a fancy cofactor. The proposed catalytic mechanism involves hydride transfer and is vaguely reminiscent of a classic alcohol dehydrogenase. Why does the nicotinamide need to be modified? Is it to convert it into a good nickel ligand? Furthermore, metal carbon bonds are rare in biology, and this is the first one found to include a nicotinic-based cofactor. Does this nickel center extend the function of the enzyme from the lactate dehydrogenase half reaction (which would work with only one stereoisomer) to the full racemase? Perhaps small-molecule synthetic chemistry, where such organometallic complexes are more commonly used as catalysts, can shed light on the reactivity of this enzyme. In particular, the pincer ligand has become a platform for small-molecule metal complexes with wide-ranging reactivity and applications (*13*).

The answers to these questions about the synthesis and function of this new cofactor may have implications that reach far beyond lactate racemase. Bioinformatic analysis has shown that although the *larA* gene for lactate racemase is typically accompanied by the genes encoding the accessory proteins, the reverse is not the case (*6*). The genomes of many species encode these maturation factors in the absence of a clear lactate racemase homolog. Future work will reveal whether this fascinating cofactor is in fact a multipurpose tool, and what nature can do with it. ∎

REFERENCES

1. B. Desguin *et al.*, *Science* **349**, 66 (2015).
2. M. Can *et al.*, *Chem. Rev.* **114**, 4149 (2014).
3. W. Lubitz *et al.*, *Chem. Rev.* **114**, 4081 (2014).
4. M. J. Maroney, S. Ciurli, *Chem. Rev.* **114**, 4206 (2014).
5. Y. Sheng *et al.*, *Chem. Rev.* **114**, 3854 (2014).
6. B. Desguin *et al.*, *Nat. Commun.* **5**, 3615 (2014).
7. P. Goffin *et al.*, *J. Bacteriol.* **187**, 6750 (2005).
8. P. T. Chivers, *Metallomics* **7**, 590 (2015).
9. A. W. Foster *et al.*, *J. Biol. Chem.* **289**, 28095 (2014).
10. A. M. Sydor, D. B. Zamble, *Met. Ions Life Sci.* **12**, 375 (2013).
11. H. de Reuse *et al.*, *Front. Cell. Infect. Microbiol.* **3**, 94 (2013).
12. B. Zambelli *et al.*, *Acc. Chem. Res.* **44**, 520 (2011).
13. G. van Koten, D. Milstein, Eds., *Organometallic Pincer Chemistry* (Springer, New York, 2013).

Ten things we have to do to achieve precision medicine

Major but surmountable hurdles should be addressed now to hasten the advent of precision medicine.

By **Isaac S. Kohane**

On 30 January 2015, President Obama announced funding for an Initiative in Precision Medicine (IPM) (*1*) less than 3 years after a National Academy of Sciences (NAS) committee report (*2*) made clear just how such an initiative could accelerate progress in medical care and research. The core concept of this initiative is that by harnessing measurements of multiple modalities—not just clinical and genomic evaluations, but environmental exposures, daily activities, and many others, we can develop a much more comprehensive view of the patient's state and its trajectory over time. By understanding precisely, across all these modalities, what the distinguishing features of specific subgroups of patients are, we can better individualize therapies.

POLICY

The combination of clinical findings and genomic sequencing is already resulting in life-saving cancer treatments, even though we have accumulated only a tiny fraction of the data sets envisaged by the NAS report and IPM. Today, we engage in preventative medicine—including the use of pharmaceuticals with pleiotropic, often unwanted, effects (such as statins)—on the basis of a handful of biomarkers. With more precise molecular characterizations of outcomes over time within populations, selection of primary prevention regimens can become more data driven, and unwanted effects could be minimized. What do we have to do to make the early promising results more generalizable, as well as available to clinical and public health practice? It is not merely a matter of scaling-up the molecular measures and clinical characterizations of large populations. Below are enumerated 10 large, but surmountable, challenges that will have to be addressed if precision medicine is to realize its full potential (see the figure).

1. *Linkage*. Data gathered from health care institutions, research projects, and

Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA.
E-mail: isaac_kohane@harvard.edu

daily events must be linked to the right individuals across their lifetimes (*3*). Even in the unlikely event that the United States would adopt a universal health identifier, many health-pertinent data types—such as diet, environmental exposure, or social Web annotations—are not linked to health care databases. Consequently, enabling the mechanism for such linkages that are both accurate and socially accepted is currently an unmet challenge.
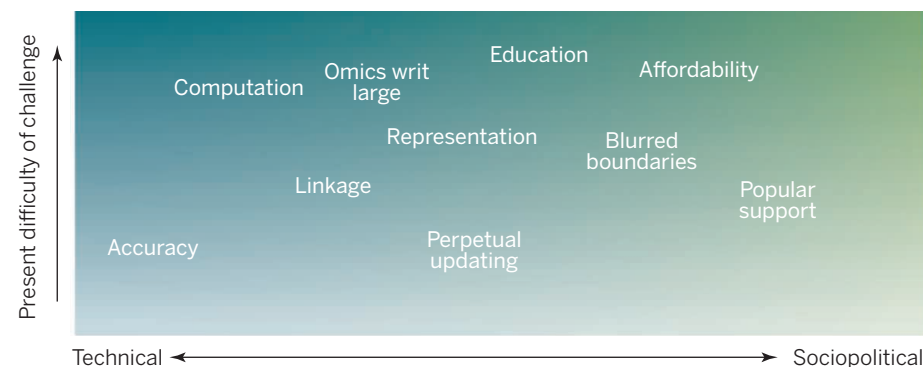
Moving toward precision medicine. Ten challenges for achieving precision medicine are qualitatively ordered on the *x* axis by how much they are intrinsically technical versus sociopolitical challenges. The *y* axis qualitatively orders the difficulty each challenge currently presents if we are to attain the widely articulated goals for precision medicine.

2. *Accuracy*. Regardless of data type, precision medicine requires some form of ongoing assessment of the accuracy and reproducibility of the data gathered in the "information commons" envisaged in the NAS report. For example, continued methodological progress is required to overcome the current gross discordance in identification of insertions and deletions in whole-genome sequences when different technologies are used (*4*). Similarly, the bias in clinical annotations (i.e., billing codes) designed to maximize reimbursement for services must be overcome. The initiatives that many publishers have promoted in this vein (*5*) are necessary steps in this direction, but a culture and institutional mechanisms for ongoing assessment of data accuracy are still in early development.

3. *Blurred boundaries*. If there is no clear demarcation between research and clinical care, methodological and ethical problems

can result (*6*). However, precision medicine often creates opportunities to benefit individual patients from the recent rapid advances in biology and high-throughput measurement. Increasingly, society as a whole and patient advocacy groups are questioning the current separation between clinical care and discovery-driven research. Yet, unlike routine clinical laboratory measurements, most omic assays are not performed in laboratories approved under the Clinical Laboratory Improvement Amendments and may be more prone to considerable variance in practice and error (*6*). Crossing the clinical care–research boundary therefore requires new models of data sharing, data verification, and vetting.

4. *Popular support*. If concerns around the privacy risk of broad data sharing and data integration are not addressed, then public concern might outweigh public support for precision medicine. The public conversation for President Obama's IPM therefore has to extend far beyond the usual community of scientists, health care providers, and technologists. Most likely, those patient advocacy groups that have seen the benefit of precision medicine or keenly experience its absence will be the most articulate and credible drivers of public support [e.g., both the Multiple Myeloma Research Foundation (www.themmrf.org/) and the Simons Foundation for autism research (http://sfari.org/) have built open-access databases of genomic and other data to promote the research effort]. Integral to this support is the development of a public consensus for appropriate procedures for consenting and clarity regarding personal control of disclosure of personal data for broader sharing (*7*).

5. *Omics writ large*. The success of high-throughput methods in capturing various personal data at the molecular level (such as genomes, transcriptomes, and proteomes) only emphasizes how much

remains to be done in obtaining comprehensive measurements of other equally relevant health measures (8). Given that for many common diseases, environmental exposures are responsible for the majority of risk, how are we going to obtain systematic, comprehensive measures of these exposures over a lifetime? It may be that we require a higher level of coordination between the various agencies responsible for monitoring the relevant data (e.g., Environmental Protection Agency, Occupational Safety and Health Administration, and Health and Human Services). Here again, public support by patient advocacy groups might be the most effective in driving these agencies toward more comprehensive data acquisition and sharing.

6. *Perpetual updating.* A knowledge-diffusion process that is far more nimble and timely than the current publication process is going to be needed. For example, every year, new data (sometimes conflicting) appear regarding the pathogenicity of mutations (9). If timeliness matters, recomputation and re-evaluation of the interpretation of these data cannot wait for investigators to add errata to their publications. Who could provide such recurrent and frequent reevaluations? It remains uncertain whether semiautomated and crowd-sourced efforts will be appropriate (9). A dedicated and trusted third party, akin to the National Institute of Standards and Technologies, has yet to emerge in this domain.

7. *Computation.* Computer-driven, real-time decision support at the point of care is an absolute requirement for the success of precision medicine (10). Even narrow specialists will not be able master all the rare and common molecular findings associated with different therapeutic subgroups and, even less, all their interdependencies. Perpetual and timely updates of this knowledge therefore requires the implementation of automated decision-support procedures that fit into the episodic and often brief encounters between patients and their care providers. Unfortunately, the capability of commercially available electronic health record systems to represent and compute with omic data is limited at best (10). Conversely, the genome-scale interpretations provided by direct-to-consumer or sequencing companies are not integrated with the highly relevant context provided by the clinical record. The opportunity remains sizable for innovative entrants into the market to automate decision support for precision medicine.

8. *Affordability.* Affordability of precision medicine ultimately will determine whether it is the preserve of the wealthy or the few or whether financial incentives can be made sufficient for continued discovery work. Already, there is an important debate about the cost of impressively effective but also terribly expensive targeted treatments (11). Similar debates about the cost of precision diagnostics also have emerged (12). Without large decreases in cost, precision medicine will result in a massive increase in the already large share of the overall gross national product claimed by health care in most developed countries.

9. *Representation.* Representative precision medicine is not only the expression of hope for social justice, but it is mathematically necessary if we are to avoid making gross diagnostic and therapeutic mistakes. Individuals coming from different ethnicities have different frequencies of multiple characteristics from various factors—from

---

> *"Capturing...differences will be at the core of precision medicine's success in our multiethnic society."*

---

genomic variants to diet. Capturing these differences will be at the core of precision medicine's success in our multiethnic society. Failure to do so will result in over- and underdiagnosis in multiple populations (13, 14)—the antithesis of precision medicine.

10. *Education.* Education of the providers of health care merits a close reexamination so as to find the means to create a workforce that is best able to translate the knowledge gained from precision medicine to clinical care. This includes competence in quantitative reasoning (15) and in the ability to access just-in-time knowledge (16). It also requires a greater recognition that, in many instances, patients will be precision medicine experts (17) and, therefore, can be instrumental in collaboratively applying unique knowledge resources in precisely their own medical challenges.

Who will participate in addressing these challenges to drive precision medicine from conception to practice? At a minimum, the stakeholders must include the public, individual patients, and organizations representing patients. Then, there are (i) the health care system that will deliver precision diagnostics and therapeutics and (ii) the scientists who will develop the principles, infrastructure, and specific insights for precision medicine. Whether at the level of local institutional review boards or national regulatory authorities, as intimated above, a full implementation of precision medicine will require, at the very least,

adjusting the existing oversight mechanisms. Most useful will be public examples of "information altruism" that are consistent with current regulations (7). Furthermore, with the entry of giants of the information age, such as Apple and Google, into healthcare (18), concurrent with the emergence of a new generation of profit and nonprofit biotechnology companies, industry will have a perspective on precision medicine that may not always resemble that of the other more seasoned, and perhaps more jaded, stakeholders. Even with the minimal and conservative view presented here, it is apparent that implementation of precision medicine will require an extensive national conversation. Fortunately, with IPM and the efforts of advocacy organizations, there now is a concrete program with which to productively drive the conversation. ■

REFERENCES AND NOTES

1. Office of the Press Secretary, "Fact Sheet: President Obama's Precision Medicine Initiative" (White House, Washington, DC, 2015); https://www.whitehouse.gov/the-press-office/2015/01/30/fact-sheet-president-obama-s-precision-medicine-initiative.
2. National Research Council, *Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease* (National Academies Press, Washington, DC, 2011); www.nap.edu/catalog/13284/toward-precision-medicine-building-a-knowledge-network-for-biomedical-research.
3. G. M. Weber, K. D. Mandl, I. S. Kohane, *JAMA* **311**, 2479 (2014).
4. J. O'Rawe, T. Jiang, G. Sun, Y. Wu, W. Wang, J. Hu, P. Bodily, L. Tian, H. Hakonarson, W. E. Johnson, Z. Wei, K. Wang, G. J. Lyon, *Genome Med.* **5**, 28 (2013) *Genome Med.* **5**, 28 (2013).
5. M. McNutt, *Science* **346**, 679 (2014).
6. L. G. Dressler, S. Smolek, R. Ponsaran, J. M. Markey, H. Starks, N. Gerson, S. Lewis, N. Press, E. Juengst, G. L. Wiesner, *Genet. Med.* **14**, 215 (2012).
7. I. S. Kohane, R. B. Altman, *N. Engl. J. Med.* **353**, 2074 (2005).
8. E. E. Schadt, J. L. Bjorkegren, *Sci. Transl. Med.* **4**, 115rv1 (2013).
9. H. L. Rehm, J. S. Berg, L. D. Brooks, C. D. Bustamante, J. P. Evans, M. J. Landrum, D. H. Ledbetter, D. R. Maglott, C. L. Martin, R. L. Nussbaum, S. E. Plon, E. M. Ramos, S. T. Sherry, M. S. Watson, *N. Engl. J. Med.* **372**, 2235 (2015).
10. D. R. Masys, G. P. Jarvik, N. F. Abernethy, N. R. Anderson, G. J. Papanicolaou, D. N. Paltoo, M. A. Hoffman, I. S. Kohane, H. P. Levy, *J. Biomed. Inform.* **45**, 419 (2012).
11. P. Curl, I. Vujic, L. J. van 't Veer, S. Ortiz-Urda, J. G. Kahn, *PLOS ONE* **9**, e107255 (2014).
12. A. Barzi, S. Sadeghi, M. W. Kattan, N. J. Meropol, *J. Natl. Cancer Inst.* **107**, djv005 (2015).
13. S. A. Tishkoff, K. K. Kidd, *Nat. Genet.* **36** (suppl.), S21 (2004).
14. S. Dagogo-Jack, *Nat. Rev. Endocrinol.* **6**, 589 (2010).
15. A. K. Manrai, G. Bhatia, J. Strymish, I. S. Kohane, S. H. Jain, *JAMA Intern. Med.* **174**, 991 (2014).
16. T. H. Davenport, J. Glaser, *Harv. Bus. Rev.* **80**, 107, 126 (2002).
17. S. Mnookin, in *The New Yorker*, 21 July 2014, p. 32; http://demystifyingmedicine.od.nih.gov/DM15/m4d21/reading01.pdf.
18. E. Fry, Apple and Google lead the tech industry into the health care minefield, (INTHEBLACK, 2015); http://intheblack.com/articles/2015/03/02/apple-and-google-lead-the-tech-industry-into-the-health-care-minefield.

In *Anxious*, Joseph LeDoux argues that the emotional experience of fear may come from an unexpected place in the brain.

NEUROSCIENCE

Facing our fears

New insights into the brain's fear circuitry could yield more effective treatments for anxiety disorders

By **Stephen Maren**

"Stayin' Alive" is much more than a catchy pop song from the 1970s. It is an evolutionary anthem sung by organisms throughout the animal kingdom when faced with danger. When threatened by a bear in the woods or a drunken patron at the local watering hole, the mammalian brain stirs into action, orchestrating complex behaviors to defend against harm (*1*). Survival is the goal. In addition to defensive behaviors, threatening experiences also yield feelings (at least in humans) including fear, terror, panic, and trepidation. For many individuals, the melody of a fearful experience lingers long after the song is over: Roughly 40 million people in the United States suffer from persistent and debilitating anxiety (*2*).

In *Anxious*, renowned neuroscientist Joseph LeDoux tackles the neural origins of fear and anxiety with an eye to developing empirically informed treatments for anxiety disorders. His expertise in this area is broad and deep. LeDoux is not only a pioneer in

the neurobiological analysis of fear in animals but also a scholarly and accessible writer. In *Anxious*, he systematically builds on his earlier works (*3*, *4*), covering with aplomb a vast literature on emotion, memory, attention, and consciousness. With that said, *Anxious* is a significant and important departure from the author's earlier views on the neural underpinnings of fear.

According to LeDoux, emotions, such as fear, are conscious feelings that occur when we become aware that our brain has detected danger. He carves the emotional brain into two parts. The first is concerned with implicit (nonconscious) threat detection, which directs defensive behaviors engineered to ward off danger. The second deals with an explicit (conscious) appraisal of the experience, which yields feelings. Similar to how the flavor of a gumbo emerges from the marriage of many different ingredients, LeDoux argues that the conscious feeling of fear is assembled from information outside of awareness. And although many organisms exhibit fearful behavior, LeDoux contends that only humans experience the conscious feeling of fear. This assertion is not universally held by all fear researchers (*5*).

LeDoux argues that by failing to distinguish between evolutionarily old systems (centered in the amygdala) and newer sys-

tems (centered in the prefrontal cortex), neuroscientists and psychologists have conflated nonconscious "threat" circuits in the brain with circuits that underlie the conscious experience of fear. This leads LeDoux to the jarring conclusion that the amygdala, a decades-old centerpiece of the brain's "fear circuit" (a term LeDoux himself coined), does not underlie the conscious feeling of fear.
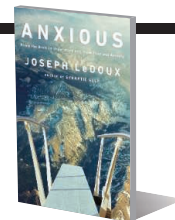
This is a controversial notion. Evidence in both animals and humans supports the idea that core emotions, including fear, can exist outside of conscious awareness (*5*–*7*). Indeed, epileptic seizures originating in the amygdala often yield autonomic arousal and summon spontaneous feelings of dread (*8*). Still, LeDoux presents a compelling case that threat detection and the emotional experience of fear are driven by different systems in the brain.

What does this reconceptualization of the emotional brain mean for the treatment of clinical disorders of fear and anxiety? LeDoux proposes that behavioral therapies should be two-pronged, independently targeting the implicit threat system and the explicit fear system. Aberrant implicit memories, he argues, would be best treated in the absence of conscious awareness, such as with subliminal presentations of threat stimuli. In contrast, feelings of fear and anxiety should be treated by focusing on the mindful reassessment of the origins of fear.

In *Anxious*, LeDoux challenges the reader to think differently about the neural origins of fear and its disorders. In doing so, he offers a masterful synthesis of animal and human work and a novel roadmap for future work in both the laboratory and the clinic.

The reviewer is at the Department of Psychology and Institute for Neuroscience, Texas A&M University, College Station, TX 77843, USA. E-mail: maren@tamu.edu

REFERENCES

1. S. Maren, *Annu. Rev. Neurosci.* **24**, 897 (2001).
2. R. C. Kessler et al., *Arch. Gen. Psychiatr.* **62**, 617 (2005).
3. J. LeDoux, *The Emotional Brain: Mysterious Underpinnings of Emotional Life* (Simon & Schuster, New York, 1996).
4. J. LeDoux, *Synaptic Self: How Our Brains Become Who We Are* (Penguin Books, New York, 2002).
5. J. Panksepp, *Affective Neuroscience: The Foundations of Human and Animal Emotions* (Oxford Univ. Press, New York, 1998).
6. R. C. Bolles, M. S. Fanselow, *Behav. Brain Sci.* **3**, 291 (1980).
7. P. Winkielmen, K. C. Berridge, *Curr. Dir. Psychol. Sci.* **13**, 120 (2004).
8. G. F. Henriksen, *Epilepsia* **14**, 39 (1973).

10.1126/science.aab3289

SCIENTIFIC METHODS

Eureka!

Recognizing and rewarding scientific creativity

By **Benjamin F. Jones**

Georges Köhler was lying in bed one evening when he suddenly realized that fusing a cancerous myeloma cell with a lymphocyte antibody cell might combine the immortality of the former with the immunological sensitivity of the latter. And so was born one of the great biomedical tools of the modern age: the monoclonal antibody.

If you have ever wondered what transpires in the moment of great scientific insight, you will find Albert Rothenberg's *Flight from Wonder* remarkable. Rothenberg examines the psychology of creativity, mainly through interviews with 45 Nobel laureates. The book is narrow and deep, focusing on the individual mental process. Many describe creativity as an exercise in combining existing material in new ways. Rothenberg shows us what that means.

Building on his previous research, Rothenberg articulates three processes that represent common forms of creative thinking. The first, the Janusian process, brings oppositional elements together and is named for the two-faced Roman god. Einstein's thought experiment for general relativity, in which a person in free fall is at once accelerating and yet meaningfully at rest, provides an example. Here, creativity emerges in the conflict between distinct ideas. Opposition, or even paradox, can mean inspiration.

Köhler's monoclonal antibodies demonstrate a different kind of combination, where functionally separate (sep) material can be joined concomitantly (con). In such a sep-con process, like in good food, the individual flavors remain intact but together create a novel and potent whole. The third process, termed "homospatial," relies on visual imagery, where distinct visualizations are overlaid in the mind. For example, Joseph Murray, who developed kidney transplantation, conceived of the operation by mentally overlaying pianos, which are shaped like kidneys, in a room, representing the abdomen.

Broader characteristics of Nobelists also

The reviewer is in the Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. E-mail: bjones@kellogg.northwestern.edu

emerge, including passion, self-confidence, openness, courage, and domain expertise. Iconoclastic tendencies and aggressive dispositions are observed in some instances as well. Indeed, the competition between James Watson and Linus Pauling over the structure of DNA and the ensuing efforts to diminish the contributions of Rosalind Franklin suggest that science is not simply a sequence of inspirational moments.

Moving from individual creativity to the broader scientific process, Roberta Ness's *The Creativity Crisis* assesses the economic, sociological, and institutional factors that underlie the current scientific funding model in the United States. She comes to some uncomfortable conclusions. Here, great minds are obstructed at nearly every turn by bad rules, bad norms, and the bad incentives that result. Her central thesis is that caution has undermined creativity. In her view, there is too little funding for basic research and that funding is too conservatively allocated, leav-

Georges Köhler, 1984

ing high-risk, high-reward avenues underexplored. The National Institutes of Health (NIH) review process, universities' growing interest in commercial income, unpredictable federal support, regulatory burdens, sociological norms, patent law, and the tenure review process all come under fire.

It's not a happy picture, but is it the right one? Ness's book is so broad that it often cannot do sufficient justice to its own arguments. Yet there is empirical evidence underneath many points, even where that depth is left more to the bibliography than the prose.

One of Ness's more compelling arguments is that federal agencies should increasingly fund "people not projects." She presents the Howard Hughes Medical Investigator (HHMI) program—where promising people are funded over longer periods, giving them the flexibility and time to attack transformational ideas—as a promising alternative to shorter-term, project-oriented grants. In support of this view, Pierre Azoulay and colleagues have shown that, controlling for individual potential, the HHMI system appears to produce higher-impact research (*1*).

Further challenges emerge in the changing nature of science itself. Late in his career, Einstein noted that "knowledge has become vastly more profound in every department of science. But the assimilative power of the human intellect is and remains strictly limited. Hence it was inevitable that the activity of the individual investigator should be confined to a smaller and smaller section." To compensate for individual narrowness, one may seek collaboration, which may explain the rise of team science, but collaboration may be an imperfect solution. Very little in Rothenberg's study suggests that the spark of great ideas naturally occurs across minds. Anita Williams Woolley and colleagues have shown that, beyond the intelligence of team members, team success correlates with members' social sensitivity (*2*). One then wonders whether the psychological processes and characteristics that drove great scientific insights in the 20th century remain as much in force today.

The shift to team science also raises questions about the reward system used to support the scientific process. Ness rightly asks whether it makes sense to reward individuals through vehicles like the Nobel Prize in a world of increasing collaboration.

Society has put more and more effort into research and development in recent decades, but the payoffs have not obviously accelerated. Innovation is getting harder and—crisis or no—our scientific systems may need to be retooled to unleash creative potential. As Ness notes, the time has come to "launch experiments." Let's follow her advice and use the scientific method to study science itself. If we experiment systematically with funding, training, intellectual property, and promotion models, we can learn what works.

REFERENCES
1. P. Azoulay *et al.*, *RAND J. Econ.* **42**, 3 (2011).
2. A. W. Woolley *et al.*, *Science* **330**, 6004 (2010).

10.1126/science.aaa7008

PHOTO: KEYSTONE/AP IMAGES

RESEARCH

Landscape evolution in a sandbox
Sweeney et al., *p. 51*
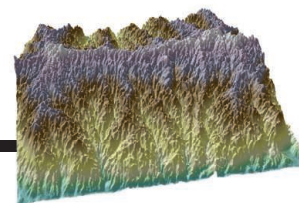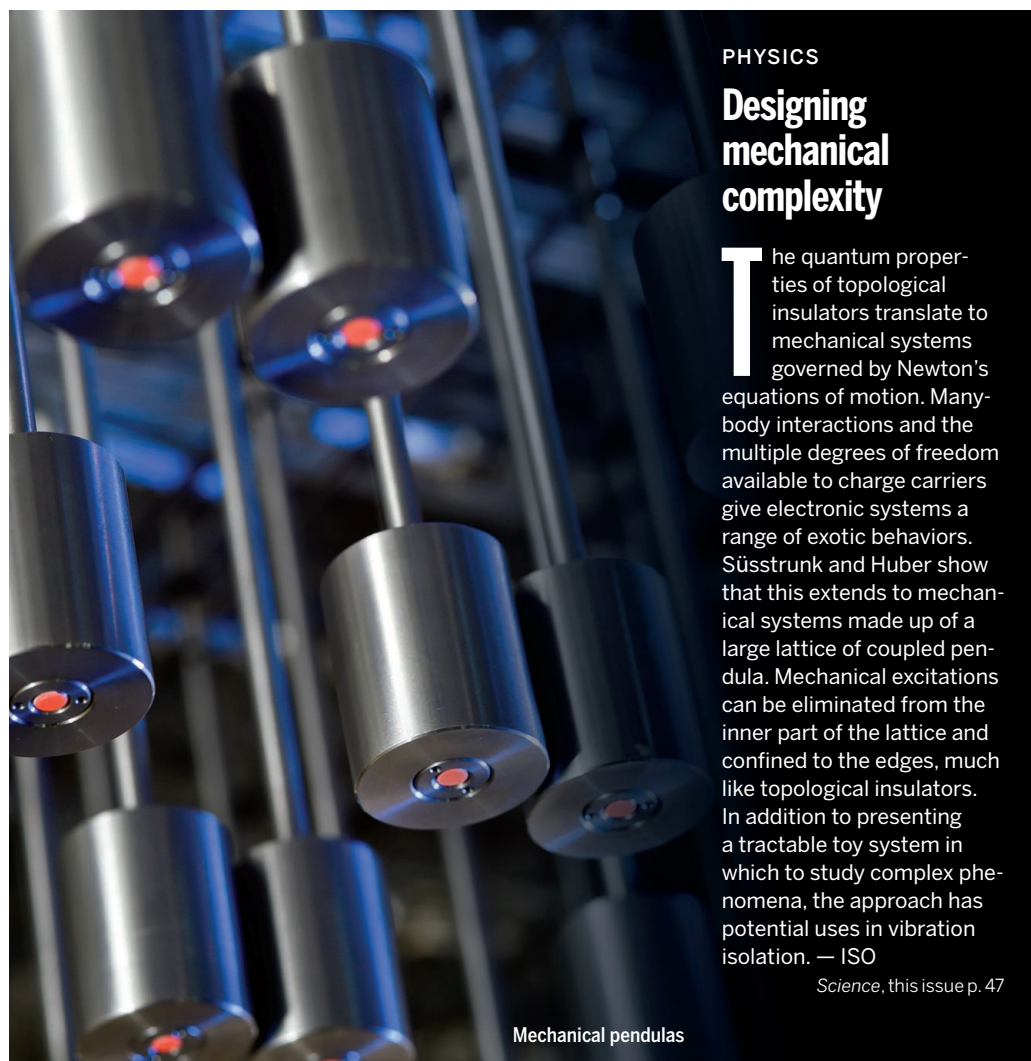
PHYSICS

Designing mechanical complexity

The quantum properties of topological insulators translate to mechanical systems governed by Newton's equations of motion. Many-body interactions and the multiple degrees of freedom available to charge carriers give electronic systems a range of exotic behaviors. Süsstrunk and Huber show that this extends to mechanical systems made up of a large lattice of coupled pendula. Mechanical excitations can be eliminated from the inner part of the lattice and confined to the edges, much like topological insulators. In addition to presenting a tractable toy system in which to study complex phenomena, the approach has potential uses in vibration isolation. — ISO

Science, this issue p. 47

Mechanical pendulas

BRAIN STRUCTURE

The best way to fold a mammalian brain

As mammalian brains grew larger through evolution, the organization and folding of brains changed too. In a series of statistical analyses comparing a large number of mammalian species, Mota and Herculano-Houzel found that brain folding is not simply a phylogenetic consequence of brain mass increase (see the Perspective by Striedter and Srinivasan). The exposed surface of the cortex scales across all mammals and across individuals as a single power law of the product of total cortical surface and the square root of cortical thickness. — PRS

Science, this issue p. 74; see also p. 31

INFLUENZA

Immunological mistaken identity

The incidence of narcolepsy increased after the global vaccination campaign against the 2009 A/H1N1 influenza pandemic, but only in some countries. Ahmed *et al.* examined the protein content of the vaccines used by country. One notable peptide within an influenza protein (nucleopeptide A) shared residues with human hypocretin receptor 2, which has been linked to narcolepsy. Patients who developed narcolepsy after being vaccinated produced antibodies that cross-reacted to both the influenza and the hypocretin receptor 2 epitopes. The vaccine used in countries without new narcolepsy cases contained lower levels of nucleoprotein A. — ACC

Sci. Transl. Med. **7**, 294ra105 (2015).

METALLOPROTEINS

Nickel pincers as enzyme cofactors

Organometallic nickel complexes long synthesized in the laboratory exist naturally in enzymes as well. Desguin *et al.* determined the structure and metal-binding residues of the Ni-containing active site in bacterial lactate racemase (see the Perspective by Zamble). A dithiodinicotinic acid mononucleotide derivative cofactor binds Ni through sulfur and carbon bonds, resembling synthetic nickel pincer complexes. Genes encoding accessory proteins involved in the synthesis of this cofactor are widely distributed in other bacteria, suggesting its involvement in other enzymes. — NW

Science, this issue p. 66; see also p. 35

PLANT VOLATILES

Stop to smell the roses

Some roses smell beautiful, yet others only look beautiful. Magnard *et al.* leveraged this distinction to study the

biosynthesis of geraniol, a monoterpene alcohol in rose scent (see the Perspective by Tholl and Gershenzon). Enzymes known for geraniol synthesis in other plants, such as basil, did not seem to provide that function for roses. Instead, a diphosphohydrolase, which functions in the cytoplasm of cells in rose petals, generates the geraniol emitted by fragrant roses. Identification of the enzyme and its gene enables marker-assisted breeding to put the perfume back into beauty. — PJH

Science, this issue p. 81; see also p. 28

PLANT GENOMICS
Signatures of adaptation in the field

Adaptation to the environment is critical for the survival of all species. For crops, this can be confounded or enhanced by farmers and breeders selecting for particular traits. To determine the associations between genes and local environments, Lasky *et al.* performed a global genetic survey of nearly 2000 regional varieties of domesticated sorghum. Regional environmental stresses such as climate and soil type were major determinants of adaptation. Enhancing sorghum (a staple crop for 500 million people in sub-Saharan Africa and Asia) and other crops may

Sorghum genomes show signs of adaptation

be possible based on marker-assisted selection of adaptive traits. — BJP

Sci. Adv. 10.1126/sciadv.1400218 (2015).

CONSERVATION
Focused on protecting a few

The illegal ivory trade threatens the persistence of stable wild elephant populations. The underground and covert nature of poaching makes it difficult to police. Wasser *et al.* used genetic tools to identify the origins of elephant tusks seized during transit (see the Perspective by Hoelzel). The majority of source animals were part of just a few wild elephant populations in Africa—and just two areas since 2006. Increased focus on enforcement in a few such areas could help interrupt poaching activities and restore wild elephant populations. — SNV

Science, this issue p. 84; see also p. 34

POROUS MATERIALS
Laser patterning polymer membranes

Porous materials are useful for membranes, filters, energy conversion, and catalysis. Their utility often depends on the ability to finely control both the pore sizes and their connectivity. Tan *et al*. prepared porous thin films of block copolymers mixed with phenol-formaldehyde resins (resols) on silicon substrates using a simple laser process. On exposure to ultraviolet light, rapid heating of the substrate causes polymerization of the resols and decomposition of the block copolymer. This method allows direct patterning of the films on a local scale, with tunable pore sizes and size distributions. — MSL

Science, this issue p. 54

Many *Dictyostelium discoideum* band together to form a stalk and fruiting body

EVOLUTION
Trench warfare keeps cheaters rare

When starving, thousands of normally solitary *Dictyostelium discoideum* amoebae band together to form a slug, which then differentiates into a stalk and fruiting body. The cells of the stalk die, allowing the cells at the top of the fruiting body to form spores and disperse. "Cheater" amoebae avoid forming the stalk and dying. To determine whether cheating provides amoebae with a selective advantage in nature, Ostrowski *et al*. studied ~150 positions in amoebae genomes that influence cheating. The genomes maintained a balanced mixture of sequence variation at many of these positions. This indicates that cheaters remain rare because they are engaged in a form of evolutionary "trench warfare" with cooperator amoebae, which results in a stalemate between the two. — GR

Curr. Biol. 10.1016/j.cub.2015.04.059 (2015).

PHYSICS
Probing beyond the Standard Model

The Standard Model of physics presents an inventory of the known fundamental particles and how they interact in order to describe the world around us. But observations, such as that of the preponderance of matter over antimatter, suggest that the Standard Model is incomplete. Experiments with large particle colliders and measurements using smaller-scale precision atom-based techniques are under way in attempts to fill in the holes of the picture. Parker *et al*. describe one such cold

atom experiment, the search for an elusive permanent electric dipole moment in a cloud of laser-trapped radium atoms. Their initial results show promise for detecting such an effect, which should then provide constraints on establishing new physics beyond the Standard Model. — ISO

Phys. Rev. Lett. **114**, 233002 (2015).

MINERALOGY
The shocking history of the Moon

The mineral stishovite helps us determine the size of past meteorite impacts, because it forms only at the high pressure and temperature possible only in the most energetic meteorite strikes. Silica also needs to be present for it to form, which helps constrain what type of rock the meteorite hit. This rare mineral was first discovered in Arizona's Meteor Crater in 1962. Now, Kaneko *et al.*
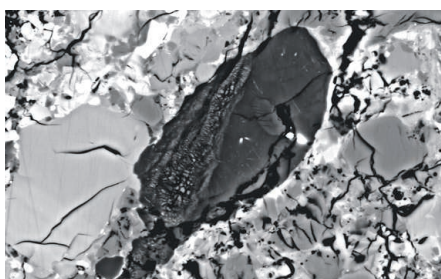
A lunar silica grain including stishovite

have found a small grain of stishovite in a Moon rock collected by the Apollo 15 mission. A more complete search for stishovite in lunar samples will provide a new path for reconstructing the impact history of the Moon. — BG

Am. Mineral. 10.2138/am-2015-5290 (2015)

CELL SIGNALING
Protein kinase C shapes calcium signals

The activation of hormone receptors on cells causes the concentration of free calcium in the cytoplasm to oscillate. Cells can encode information in the frequency, amplitude, and duration of such calcium oscillations, which may modulate the activation of downstream signaling pathways. Bartlett *et al.* compared calcium signals generated by hormone receptors or by a simpler mechanism that bypasses these receptors in rat hepatocytes, a type of liver cell. This allowed the authors to isolate the effects of the enzyme protein kinase C (PKC) on calcium oscillations, because PKC is only activated downstream of the receptor. PKC regulated both the shape and propagation rates of calcium waves — LBR

J. Biol. Chem. 10.1074/ jbc.M115.657767 (2015).

T CELL METABOLISM
Complementing metabolic demands

For T cells, fighting infections is taxing work. T cells must proliferate and produce inflammatory mediators that help rid the body of infection. Doing so requires that they change their metabolism; for instance, by increasing their glycolytic capacity. How T cells do this at a molecular level, however, remains incompletely understood. Kolev *et al.* studied how this occurs in human T helper 1 (T_H1) cells, which are important for clearing intracellular infections. When activated, T_H1 cells produced complement C3b, an ancient innate immune mediator, which signaled in an autocrine manner to drive increases in nutrient uptake, glycolysis, and oxidative phosphorylation. These data highlight how the immune system can repurpose its components for evolutionary advantage. — KLM

Immunity **42**, 1033 (2015).

MATERIALS SCIENCE
From solution to a solid bright idea

Semiconducting quantum dots are ideal for down-converting light, so that a single blue source can be partly changed into other colors to give an overall white emission. Hanson *et al.* note that the conversion efficiency seen for quantum dots in solution is rarely matched when they are put into solid-state devices. This is particularly true for the red part of the spectrum, for which the best converters depend on expensive rare earth materials. Using low-cost materials, they grew CdSe/CdS core/shell particles so that the particles were mostly shell material. Conversion efficiencies in solid-state devices approached unity. At high driving currents, thermal quenching was a problem but was partly solved by adding a spacer between the light source and quantum dot layer. — MSL

ACS Appl. Mat. Inter. 10.1021/ acsami.5b02818 (2015).

The rise of agriculture fundamentally changed human mobility

HUMAN ANATOMY
How agriculture shaped our bones

Human mobility declined after the Pleistocene, affecting human health and social organization—but what caused this decline? To find out, Ruff *et al.* examined skeletal remains from nearly 2000 individuals spanning a 33,000-year period from the Upper Paleolithic to the 20th century. They found that a decrease in the bending strength of leg bones accompanied the shift to an increasingly sedentary lifestyle that occurred in Europe during the Neolithic to Roman eras (approx. 7000 to 2000 years before the present). This implies a decline of mobility as agriculture came to dominate how people produced food. The remains did not reveal any further declines in the past 2000 years, even after agriculture became more mechanized. — AMS

Proc. Natl. Acad. Sci. U.S.A. **112**, 7147 (2015).

ALSO IN *SCIENCE* JOURNALS

Edited by **Nick Wigginton**

OCEANOGRAPHY
Carbon emissions and their ocean impacts

Anthropogenic CO_2 emissions directly affect atmospheric chemistry but also have a strong influence on the oceans. Gattuso *et al.* review how the physics, chemistry, and ecology of the oceans might be affected based on two CO_2 emission trajectories: one business as usual and one with aggressive reductions. Ocean warming, acidification, sea-level rise, and the expansion of oxygen minimum zones will continue to have distinct impacts on marine communities and ecosystems. The path that humanity takes regarding CO_2 emissions will largely determine the severity of these phenomena. — HJS

Science, this issue p. 45

ORGANIC CHEMISTRY
Elaborate amines from commodity olefins

Fine spatial control over the formation of carbon-nitrogen bonds is essential to the synthesis of many pharmaceutical and agrochemical compounds. Yang *et al.* synthesized a copper catalyst that shows exceptional discrimination between similar substituents in the addition of N-H bonds across the C-C double bonds of simple olefins. In particular, the catalyst favors one of two mirror-image products that differ in the relative orientations of methyl and ethyl groups. This selectivity fosters the use of readily available commodity olefins in the preparation of complex chiral amines. — JSY

Science, this issue p. 62

BIOMECHANICS
The curious tale of the square tail

Appendages in animals are typically round, but the seahorse tail has a square cross section.

Porter *et al.* hypothesize that this shape provides better functionality and strength than a round cross section (see the Perspective by Ashley-Ross). Three-dimensional printed models show that square cross section shapes behave more advantageously when subjected to compressive forces. By allowing greater deformation without damage and accommodating twisting deformations, square appendages passively return to their original configurations. The added flexibility of the square cross section enhances the tail's ability to grasp objects. — MSL

Science, this issue p. 46;
see also p. 30

SULFUR CHEMISTRY
An unexpected gaseous sulfur species

Sulfuric acid plays a central role in both industrial and atmospheric contexts. As such, the behavior of SO_3 mixtures in gas phases has been studied for over a century. In gas-phase experiments on wet SO_3 and formic acid, Mackenzie *et al.* discovered a previously unrecognized covalent adduct: formic sulfuric anhydride, or $HC(O)OSO_3H$. The combination of microwave spectroscopy and theoretical calculations reveals its structural properties. The compound may play a role in the nucleation of atmospheric aerosols by serving as an intermediate to H_2SO_4 formation. — JSY

Science, this issue p. 58

PRESYNAPTIC NETWORKS
Tracing cells that project to one neuron

Feature extraction is a prominent characteristic of cortical neurons involved in the early stages of sensory processing. Wertz *et al.* retrogradely marked an injected neuron and its direct inputs to reveal the network mechanisms that mediate

their response. Neurons within each presynaptic network layer of single direction-selective cells showed similar motion direction preferences. In some networks, layer-specific functional modules were identical to the orientation preference of the postsynaptic neuron. Presynaptic neurons, however, displayed a general bias toward the stimulus feature that elicited a response in the postsynaptic neuron. — PRS

Science, this issue p. 70

GREEN CHEMISTRY
Synthesizing more sustainable plastics

Zeolites can help synthesize cheaper plastic precursors from biologically sourced feedstocks. Producing sustainable plastics must compete with more cost-effective petrochemical-based synthesis routes. Dusseler *et al.* developed a zeolite-based strategy to catalyze the transformation of microbially produced lactic acid into lactide, a difficult-to-synthesize precursor of biodegradable polylactic acid plastics. The selectivity of nearly 80% is based on active site spatial confinement in the zeolite micropores. This step substantially simplifies current high-cost synthesis routes and generates nearly zero waste using current reactor technologies. — NW

Science, this issue p. 78

SELENOPROTEINS
Clearing out selenoprotein garbage

Our DNA consists of codons that code for 20 different amino acids. Another amino acid, selenocysteine, is also found in several human selenoproteins. Selenocysteine is incorporated through the recoding of a stop codon, but failures in this process result in premature termination of protein synthesis. Lin *et al.* showed that the

potentially dangerous truncated proteins formed in such cases are specifically degraded by a protein quality surveillance system. The surveillance system can specifically recognize the truncated ends of the various prematurely terminated selenoproteins and target their destruction. — GR

Science, this issue p. 91

DENGUE VIRUS
An antibody to lock dengue virus out

Mosquito-borne dengue virus (DENV) is a growing public health threat. Nearly 400 million people are infected annually, and no vaccine is currently available. Fibriansah *et al.* report that a human antibody (2D22) specific for DENV serotype 2, when given therapeutically, can protect mice from a lethal form of this virus. Structural analysis revealed that 2D22 binds across multiple DENV envelope proteins, which probably blocks the ability of these proteins to assemble into the orientation necessary for viral entry. The epitope where 2D22 binds to the virus may therefore represent a potential vaccine target. — KLM

Science, this issue p. 88

NEUROSCIENCE
Keeping immature neurons excited

After birth, signaling by the neurotransmitter GABA in the brain switches from excitatory to inhibitory. GABA mediates both responses by binding to ligand-gated ion channels that conduct Cl^-. The opening of these channels depends on the concentration of Cl^- in neurons. Friedel *et al.* identified phosphorylation events in the K^+-Cl^- cotransporter KCC2 that depended on the activity of the kinase WNK1. These phosphorylation events inhibited KCC2 activity and contributed to the

depolarizing effect of GABA-mediated signaling in immature rat neurons by maintaining a high intracellular Cl⁻ concentration. — NRG

Sci. Signal. **8**, ra65 (2015).

GEOMORPHOLOGY
Landscape evolution in a sandbox

The long-term response of hills and valleys to changes in climate depends on a variety of physical factors. Sweeney *et al.* performed tabletop erosion experiments as a function of rainfall and uplift: variables that are impossible to precisely control in nature (see the Perspective by McCoy). Ridge and valley spacing are set by the balance of sediment moving down hillslopes or being washed out of valleys by rivers. Landscapes therefore evolve as a response to climate change altering erosion rates. — BG

Science, this issue p. 51; see also p. 32

STRUCTURAL VIROLOGY
Retroviral capsids in their native form

Capsid proteins of retroviruses form protective lattices around viral RNA molecules. But what controls how individual full-length capsid proteins assemble to shield the viral genome? Obal *et al.* and Gres *et al.* report high-resolution crystal structures of the full-length capsid proteins from bovine leukemia virus and HIV-1, respectively. Both studies help reveal the dynamic nature of capsid protein assembly and how individual capsid proteins interact in the lattice. The findings may have relevance for antiretroviral drug design. — KLM

Science, this issue p. 95 and 99

REVIEW SUMMARY

OCEANOGRAPHY

Contrasting futures for ocean and society from different anthropogenic CO$_2$ emissions scenarios

J.-P. Gattuso* *et al.*

BACKGROUND: Although the ocean moderates anthropogenic climate change, this has great impacts on its fundamental physics and chemistry, with important consequences for ecosystems and people. Yet, despite the ocean's critical role in regulating climate—and providing food security and livelihoods for millions of people—international climate negotiations have only minimally considered impacts on the ocean. Here, we evaluate changes to the ocean and its ecosystems, as well as to the goods and services they provide, under two contrasting CO$_2$ scenarios: the current high-emissions trajectory (Representative Concentration Pathway 8.5, RCP8.5)

and a stringent emissions scenario (RCP2.6) consistent with the Copenhagen Accord of keeping mean global temperature increase below 2°C in the 21st century. To do this, we draw on the consensus science in the latest assessment report of the Intergovernmental Panel on Climate Change and papers published since the assessment.

ADVANCES: Warming and acidification of surface ocean waters will increase proportionately with cumulative CO$_2$ emissions (see figure). Warm-water corals have already been affected, as have mid-latitude seagrass, high-latitude pteropods and krill, mid-latitude bivalves, and

fin fishes. Even under the stringent emissions scenario (RCP2.6), warm-water corals and mid-latitude bivalves will be at high risk by 2100. Under our current rate of emissions, most marine organisms evaluated will have very high risk of impacts by 2100 and many by 2050. These results—derived from experiments, field observations, and modeling—are consistent with evidence from high-CO$_2$ periods in the paleorecord.

Impacts to the ocean's ecosystem services follow a parallel trajectory. Services such as coastal protection and capture fisheries are already affected by ocean warming and acidification. The risks of impacts to these services increase with continued emissions: They are

predicted to remain moderate for the next 85 years for most services under stringent emission reductions, but the business-as-usual scenario (RCP8.5) would put all ecosystem

services we considered at high or very high risk over the same time frame. These impacts will be cumulative or synergistic with other human impacts, such as overexploitation of living resources, habitat destruction, and pollution. Fin fisheries at low latitudes, which are a key source of protein and income for millions of people, will be at high risk.

OUTLOOK: Four key messages emerge. First, the ocean strongly influences the climate system and provides important services to humans. Second, impacts on key marine and coastal organisms, ecosystems, and services are already detectable, and several will face high risk of impacts well before 2100, even under the low-emissions scenario (RCP2.6). These impacts will occur across all latitudes, making this a global concern beyond the north/south divide. Third, immediate and substantial reduction of CO$_2$ emissions is required to prevent the massive and mostly irreversible impacts on ocean ecosystems and their services that are projected with emissions greater than those in RCP2.6. Limiting emissions to this level is necessary to meet stated objectives of the United Nations Framework Convention on Climate Change; a substantially different ocean would result from any less-stringent emissions scenario. Fourth, as atmospheric CO$_2$ increases, protection, adaptation, and repair options for the ocean become fewer and less effective.

The ocean provides compelling arguments for rapid reductions in CO$_2$ emissions and eventually atmospheric CO$_2$ drawdown. Hence, any new global climate agreement that does not minimize the impacts on the ocean will be inadequate. ∎
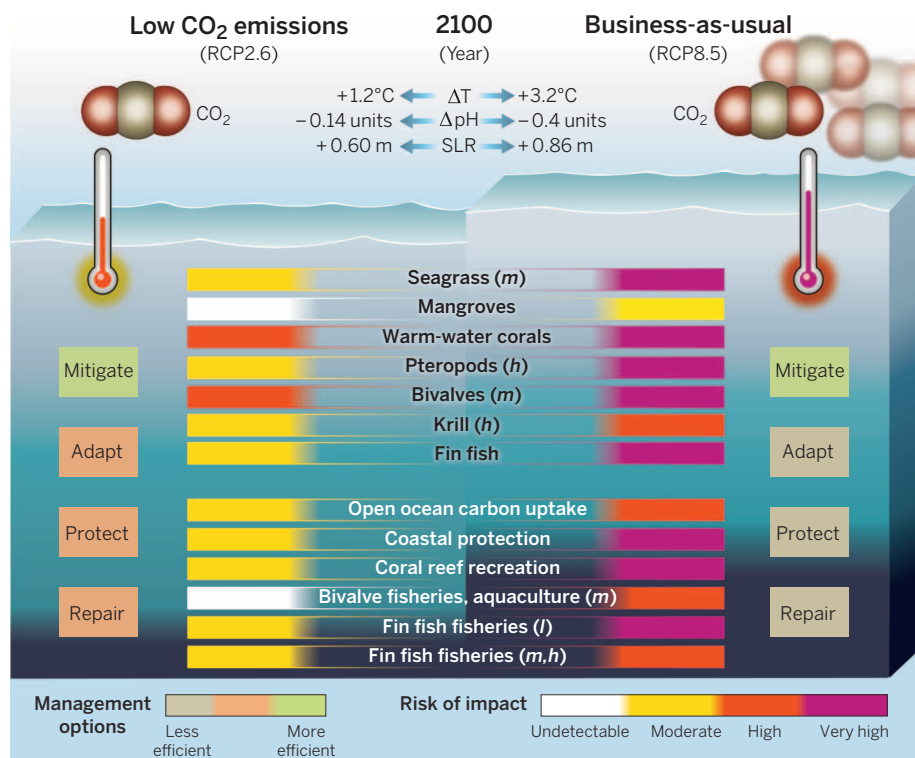
Changes in ocean physics and chemistry and impacts on organisms and ecosystem services according to stringent (RCP2.6) and high business-as-usual (RCP8.5) CO$_2$ emissions scenarios. Changes in temperature (ΔT) and pH (ΔpH) in 2090 to 2099 are relative to preindustrial (1870 to 1899). Sea level rise (SLR) in 2100 is relative to 1901. RCP2.6 is much more favorable to the ocean, although important ecosystems, goods, and services remain vulnerable, and allows more-efficient management options. *l*, *m*, *h*: low, mid-, and high latitudes, respectively.

OCEANOGRAPHY

Contrasting futures for ocean and society from different anthropogenic CO$_2$ emissions scenarios

J.-P. Gattuso,[1,2,3]* A. Magnan,[3] R. Billé,[4] W. W. L. Cheung,[5] E. L. Howes,[6] F. Joos,[7]
D. Allemand,[8,9] L. Bopp,[10] S. R. Cooley,[11] C. M. Eakin,[12] O. Hoegh-Guldberg,[13]
R. P. Kelly,[14] H.-O. Pörtner,[6] A. D. Rogers,[15] J. M. Baxter,[16] D. Laffoley,[17] D. Osborn,[18]
A. Rankovic,[3,19] J. Rochette,[3] U. R. Sumaila,[20] S. Treyer,[3] C. Turley[21]

The ocean moderates anthropogenic climate change at the cost of profound alterations of its physics, chemistry, ecology, and services. Here, we evaluate and compare the risks of impacts on marine and coastal ecosystems—and the goods and services they provide—for growing cumulative carbon emissions under two contrasting emissions scenarios. The current emissions trajectory would rapidly and significantly alter many ecosystems and the associated services on which humans heavily depend. A reduced emissions scenario—consistent with the Copenhagen Accord's goal of a global temperature increase of less than 2°C—is much more favorable to the ocean but still substantially alters important marine ecosystems and associated goods and services. The management options to address ocean impacts narrow as the ocean warms and acidifies. Consequently, any new climate regime that fails to minimize ocean impacts would be incomplete and inadequate.

Atmospheric carbon dioxide (CO$_2$) has increased from 278 to 400 parts per million (ppm) over the industrial period and, together with the increase of other greenhouse gases, has driven a series of major environmental changes. The global ocean (including enclosed seas) acts as a climate integrator that (i) absorbed 93% of Earth's additional heat since the 1970s, offsetting much atmospheric warming but increasing ocean temperature and sea level; (ii) captured 28% of anthropogenic CO$_2$ emissions since 1750, leading to ocean acidification; and (iii) accumulated nearly all water resulting from melting glaciers and ice sheets, hence furthering the rise in sea level. Thus, the ocean moderates anthropogenic climate change at the cost of major changes in its fundamental chemistry and physics. These changes in ocean properties profoundly affect species' biogeography and phenology, as well as ecosystem dynamics and biogeochemical cycling (1–3). Such changes inevitably affect the ecosystem services on which humans depend. The ocean represents more than 90% of Earth's habitable space, hosts 25% of eukaryotic species (4), provides 11% of global animal protein consumed by humans (5), protects coastlines, and more. Simply put, the ocean plays a particularly important role in the livelihood and food security of hundreds of millions of people.

The United Nations Framework Convention on Climate Change (UNFCCC) aims to stabilize atmospheric greenhouse gas concentrations "at a level that would prevent dangerous anthropogenic interference with the climate system ... within a time-frame sufficient to allow ecosystems to adapt naturally to climate change, to ensure that food production is not threatened, and to enable economic development to proceed in a sustainable manner" (6). According to the Copenhagen Accord (7), meeting these goals requires that the increase in average global surface temperature be less than 2°C over the preindustrial average. However, despite the ocean's critical role in global ecosystem goods and services, international climate negotiations have only minimally considered ocean impacts, especially those related to ocean acidification (8). Accordingly, highlighting ocean-related issues is now crucial, given that even achieving the +2°C target (set on global temperature) would not prevent many climate-related impacts upon the ocean (9).

This paper first summarizes the key findings of the Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) and, given the ongoing acceleration of climate change research, adds newer literature to assess the impacts of global change—including ocean warming, acidification, deoxygenation, and sea level rise—linking ocean physics and chemistry to biological processes, ecosystem functions, and human activities. Second, it builds on scenarios based on the range of cumulative fossil carbon emissions and the IPCC Representative Concentration Pathways (RCP) RCP2.6 and RCP8.5, contrasting two potential futures. RCP2.6 reflects the UNFCCC target of global temperature staying below +2°C, whereas RCP8.5 reflects the current trajectory of business-as-usual CO$_2$ emissions. Third, this paper provides a broad discussion of the options society has for addressing ocean impacts and ends with key messages that provide further compelling arguments for ambitious CO$_2$ emissions reduction pathways.

Changes in ocean physics and chemistry

Ocean changes resulting from anthropogenic emissions include long-term increase in temperature down to at least 700 m, increased sea level, and a decrease in Arctic summer sea ice (Fig. 1 and Table 1) (10). Other radiatively active agents—such as ozone, methane, nitrous oxide, and aerosols—do not affect the ocean as much as CO$_2$. Setting it apart, CO$_2$ accounts for two or more times the warming attributed to the non-CO$_2$ greenhouse gases by 2100 (11) and causes ocean acidification. The uptake of excess anthropogenic CO$_2$ by the ocean increases the partial pressure of carbon dioxide (Pco$_2$) and dissolved inorganic carbon while decreasing pH and the saturation state of seawater with respect to the calcium carbonate minerals aragonite and calcite, both being critical drivers of solubility of shells and skeletons (12). Rising global CO$_2$ also further exacerbates the nearshore biogeochemical changes associated with land use change, nutrient inputs, aquaculture, and fishing (13).

Both the magnitude and rate of the anthropogenic carbon perturbation exceed the extent of natural variation over the last millennium and

[1]Laboratoire d'Océanographie de Villefranche, CNRS–Institut National des Sciences de l'Univers, F-06230 Villefranche-sur-mer, France. [2]Sorbonne Universités, Université Pierre et Marie Curie, Univ Paris 06, Observatoire Océanologique, F-06230 Villefranche-sur-mer, France. [3]Institute for Sustainable Development and International Relations, Sciences Po, 27 rue Saint Guillaume, F-75007 Paris, France. [4]Secretariat of the Pacific Community, B.P. D5, 98848 Noumea Cedex, New Caledonia. [5]Nippon Foundation-UBC Nereus Program, University of British Columbia (UBC), Vancouver, BC V6T 1Z4, Canada. [6]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12, D-27570, Bremenrhaven, Germany. [7]Climate and Environmental Physics, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland. [8]Centre Scientifique de Monaco, 8 Quai Antoine Ier, MC-98000 Monaco, Principality of Monaco. [9]Scientific and Technical Committee, Prince Albert II of Monaco Foundation, 16 Boulevard de Suisse, MC-98000 Monaco, Principality of Monaco. [10]Institut Pierre Simon Laplace/Laboratoire des Science du Climat et de l'Environnement, UMR8212, CNRS–Commissariat à l'Énergie Atomique et aux Énergies Alternatives–Université de Versailles Saint-Quentin-en-Yvelines, Gif sur Yvette, France. [11]Ocean Conservancy, 1300 19th Street NW, 8th Floor, Washington, DC 20036, USA. [12]Coral Reef Watch, National Oceanic and Atmospheric Administration, College Park, MD 20740, USA. [13]Global Change Institute and Australian Research Council Centre for Excellence in Coral Reef Studies, University of Queensland, Building 20, St Lucia, 4072 Queensland, Australia. [14]School of Marine and Environmental Affairs, University of Washington, 3707 Brooklyn Avenue NE, Seattle, WA 98105, USA. [15]Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK. [16]Scottish Natural Heritage, 231 Corstorphine Road, Edinburgh EH12 7AT, Scotland. [17]IUCN, Rue Mauverney 28, CH-1196 Gland, Switzerland. [18]Environment Laboratories, International Atomic Energy Agency, 4a Quai Antoine 1er, MC-98000 Monaco, Principality of Monaco. [19]Program on Science, Technology, and Society, John F. Kennedy School of Government, Harvard University, 79 John F. Kennedy Street, Cambridge, MA 02138, USA. [20]Fisheries Economics Research Unit, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. [21]Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, UK.
*Corresponding author. E-mail: gattuso@obs-vlfr.fr
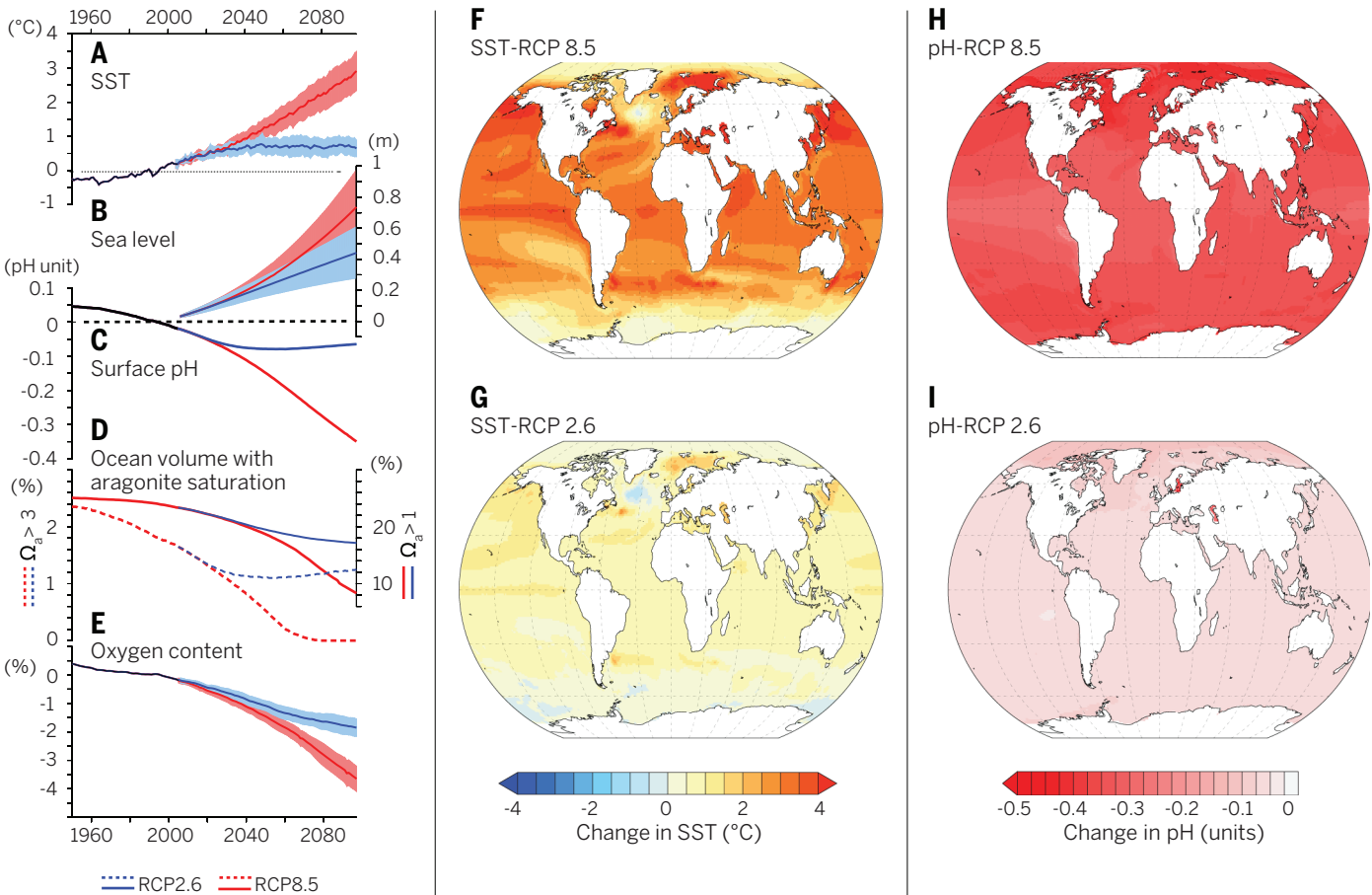
Fig. 1. Environmental changes over the industrial period and the 21st century for a business-as-usual scenario and a stringent emissions scenario consistent with the UNFCCC target of increase in global surface temperature by 2°C. (**A** to **E**) Changes in globally averaged (A) SST, (B) sea level, (C) sea surface pH (total pH scale), (D) ocean volume (in % of total ocean volume) with saturation state of calcium carbonate in aragonitic form (Ω_a) above 1 and above 3, and (E) dissolved oxygen. RCP8.5, red lines; RCP2.6, blue lines. Maps show the 21st century changes in SST (**F** and **G**) and in sea surface pH (**H** and **I**) for RCP8.5 (top) and RCP2.6 (bottom), respectively. All projected values represent ensemble mean values from the Coupled Model Intercomparison Project 5 [CMIP5 (*23*)].

Table 1. Changes in SST, pH, oxygen content, sea level, and ocean volume with respect to aragonite for CMIP5 models and several RCP emissions scenarios. After Bopp *et al.* (*23*) except sea level rise (*28*).

	ΔSST (°C)	ΔpH (units)	ΔO₂ content (%)	Sea level (m)	Vol. Ωₐ >1 (%)	Vol. Ωₐ >3 (%)
Changes relative to 1990–1999						
2090–2099 (RCP8.5)	2.73	−0.33	−3.48	0.67	9.4	0
2090–2099 (RCP4.5)	1.28	−0.15	−2.37	0.49	15	0.57
2090–2099 (RCP2.6)	0.71	−0.07	−1.81	0.41	17.3	1.22
1990s (1990–1999)	0	0	0	0	24	1.82
Preindustrial (1870–1899)	−0.44	0.07	–	–	25.6	2.61
Preindustrial (1870–1879)	−0.38	0.07	–	–	25.6	2.67
Changes relative to 1870–1899 (except sea level, relative to 1901)						
2090–2099 (RCP8.5)	3.17	−0.40	–	0.86	–	–
2090–2099 (RCP4.5)	1.72	−0.22	–	0.68	–	–
2090–2099 (RCP2.6)	1.15	−0.14	–	0.60	–	–
2010s (2010–2019)	0.83	−0.11	–	–	–	–
Past 10 years (2005–2014)	0.72	−0.10	–	0.19*	–	–
1990s (1990–1999)	0.44	−0.07	–	–	–	–
Preindustrial (1870–1899)	0	0	–	0	–	–

*Value for 2010 obtained from instrumental records.

over glacial-interglacial time scales (*14–16*). Variability of pH in coastal waters is considerably larger than that in the open ocean, partly driven by upwelling (*17*), freshwater input (*18*), eutrophication (*19*) and biogeochemical processes (*20*). Anthropogenic trends in biogeochemical variables—notably in pH, $P\text{co}_2$, and the saturation of calcite and aragonite—emerge from the noise of natural variability much faster than sea surface temperature (SST) (*21*). The combined changes in these parameters will be distinguishable from natural fluctuations in 41% of the global ocean within a decade (*22*), and the change in aragonite saturation over the industrial period has been more than five times greater than natural variability over the past millennium in many regions (*15*).

The condition of the future ocean depends on the amount of carbon emitted in the coming decades (Figs. 1 and 2A). The current suite of earth system models illustrate the contrast between future oceans under the high-carbon-emission, business-as-usual RCP8.5 versus the stringent emission-mitigation RCP2.6 (*23, 24*). The more stringent scenario allows less than one-sixth of 21st century emissions expected under business-as-usual. Between 2012 and 2100, compatible
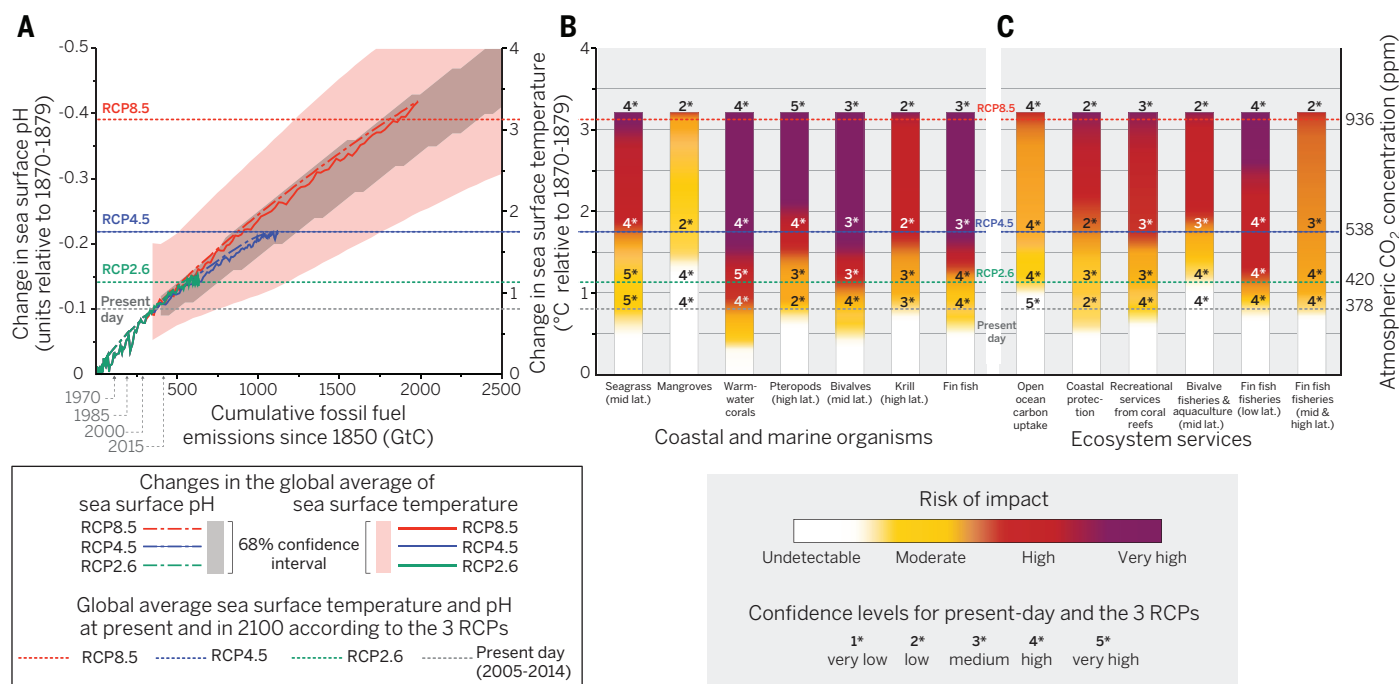
Fig. 2. Observed impact and risk scenarios of ocean warming and acidification for important organisms and critical ecosystem services. "Present-day" (gray dotted line) corresponds to the period from 2005 to 2014. Impact levels are for the year 2100 under the different projections shown and do not consider genetic adaptation, acclimatization, or human risk reduction strategies (mitigation and societal adaptation). RCP4.5 is shown for illustrative purposes as an intermediate scenario between the business-as-usual high-emissions scenario (RCP8.5) and the stringent reduction scenario (RCP2.6). (**A**) Changes in global average SST and pH versus cumulative fossil fuel emissions. Realized fossil emissions (26) are indicated for different years below the horizontal axis, whereas the lines are based on allowable emissions estimated from ensemble means of the CMIP5 simulations for the industrial period and the 21st century following RCP2.6, RCP4.5, and RCP8.5 (23). Cumulative emission of 1000 GtC causes a global SST change of about 1.7°C and a surface pH change of about −0.22 units. The colored shadings indicate the 68% confidence interval for pH (gray) and SST (pink) from observation-constrained, probabilistic projections using 55 multi–gas emissions scenarios (24). (**B**) Risk of impacts resulting from elevated CO_2 on key organisms that are well documented in the literature. (**C**) Risk of impacts resulting from elevated CO_2 on critical ecosystem services. The levels of confidence in the risk levels synthesize the author team's judgments (see materials and methods) about the validity of findings as determined through evaluation of evidence and agreement (157).

cumulative carbon emissions from fossil fuel use are 1685 gigatons of carbon (GtC) and 270 GtC for the two RCPs, respectively (10, 25). This is in addition to the 375 and 180 GtC already emitted by 2011 by fossil fuel and land use, respectively (25). Because carbon emissions were 10 GtC in 2013 (26), fast and massive emission reductions are required to keep global surface temperature below the 2°C target of the Copenhagen Accord. Carbon emissions would need to be even lower if the ocean absorbs less excess CO_2 than is currently predicted. Indeed, the ocean's effectiveness in absorbing CO_2 decreases with increasing emissions: the fraction of anthropogenic emissions absorbed by the ocean in the 21st century is projected to decline from 56% for RCP2.6 to 22% for RCP8.5 (27).

Ocean physics and chemistry will be quite different under these two emissions scenarios, although differences between the two trajectories will not be apparent until 2035. In 2100, the ocean will be much warmer and have a lower pH under RCP8.5 than under RCP2.6 (Fig. 1): The 21st century global mean change in SST differs by nearly a factor of 4 (mean ± 1 SD: 2.73 ± 0.72°C versus 0.71 ± 0.45°C), whereas global surface pH changes range from −0.33 ± 0.003 units to −0.07 ± 0.001 units). By 2100, the average global

increase in mean sea level relative to preindustrial is projected to be 0.86 m for RCP8.5 and 0.60 m for RCP2.6 (28). By 2300, it will be less than 1 m for RCP2.6 and from 1 to over 3 m for RCP8.5 (10). Generally, an increase in stratification, linked to sea-surface warming and freshening, is projected; this tends to slow ocean carbon uptake and nutrient supply to the surface (29).

CO_2 emissions also affect the deep ocean, although the responses are delayed by the surface-to-deep transport time and continue for centuries even after carbon emissions cease (30). The volume of ocean water that is supersaturated by more than a factor of 3 with respect to aragonite ($\Omega_a > 3$) is projected to completely vanish over the course of the century for RCP8.5 and to decrease from 2% to 1.25% of the ocean volume for RCP2.6 (Fig. 1 and Table 1). Conversely, the volume occupied by undersaturated water ($\Omega_a < 1$) that is corrosive to unprotected calcium carbonate shells and skeletons expands from 76% of the whole ocean volume in the 1990s to 91% in 2100 with RCP8.5 and to only 83% with RCP2.6. The whole ocean oxygen inventory is consistently projected to decrease (RCP8.5: −3.45 ± 0.44%; RCP2.6: −1.81 ± 0.31%) with largest changes in the subsurface mid-latitude regions. However, it remains unclear whether, and to what extent,

low-oxygen regions will expand and whether the observed expansion of oxygen minimum zones over recent decades resulted from direct anthropogenic perturbation or was caused by natural variability (31, 32).

Projections of ocean warming and acidification in coastal systems follow the general trends of global and regional IPCC models but have lower confidence values because of larger contributions of processes other than CO_2 uptake (3). Projected regional changes vary, with the largest sea-surface warming in the North Pacific, the tropical East Pacific, and in parts of the Arctic and the largest surface pH decrease in the Arctic (Figs. 1 and 3). By 2100, 69% of the surface ocean will warm by more than 1.5°C and acidify by more than −0.2 pH units relative to preindustrial under RCP8.5 as opposed to less than 1% under RCP2.6 (Fig. 3). The largest absolute decrease in aragonite saturation is projected for the tropical ocean, partly modulated by variability within coral reef sites (33, 34). Seasonally undersaturated conditions are already present in the northeastern Pacific and the California upwelling system (17) and in the Arctic Ocean (35) and expected for the Southern Ocean (36). pH reductions at the sea floor below 500-m depth, which includes biodiversity hot spots such as deep-sea canyons and seamounts, are
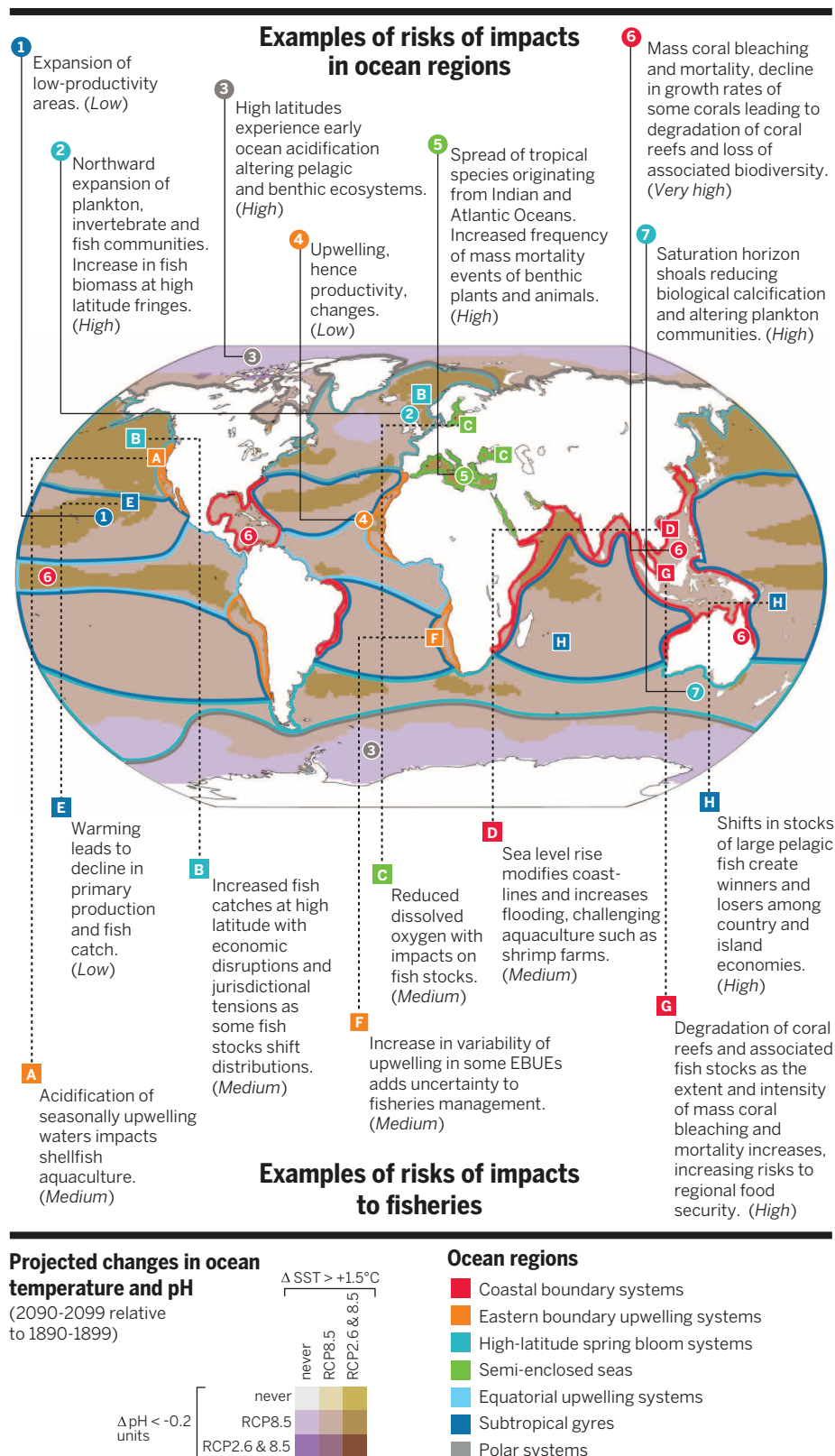
Examples of risks of impacts in ocean regions

1 Expansion of low-productivity areas. (*Low*)

2 Northward expansion of plankton, invertebrate and fish communities. Increase in fish biomass at high latitude fringes. (*High*)

3 High latitudes experience early ocean acidification altering pelagic and benthic ecosystems. (*High*)

4 Upwelling, hence productivity, changes. (*Low*)

5 Spread of tropical species originating from Indian and Atlantic Oceans. Increased frequency of mass mortality events of benthic plants and animals. (*High*)

6 Mass coral bleaching and mortality, decline in growth rates of some corals leading to degradation of coral reefs and loss of associated biodiversity. (*Very high*)

7 Saturation horizon shoals reducing biological calcification and altering plankton communities. (*High*)

Examples of risks of impacts to fisheries

E Warming leads to decline in primary production and fish catch. (*Low*)

A Acidification of seasonally upwelling waters impacts shellfish aquaculture. (*Medium*)

B Increased fish catches at high latitude with economic disruptions and jurisdictional tensions as some fish stocks shift distributions. (*Medium*)

F Increase in variability of upwelling in some EBUEs adds uncertainty to fisheries management. (*Medium*)

C Reduced dissolved oxygen with impacts on fish stocks. (*Medium*)

D Sea level rise modifies coastlines and increases flooding, challenging aquaculture such as shrimp farms. (*Medium*)

G Degradation of coral reefs and associated fish stocks as the extent and intensity of mass coral bleaching and mortality increases, increasing risks to regional food security. (*High*)

H Shifts in stocks of large pelagic fish create winners and losers among country and island economies. (*High*)

Projected changes in ocean temperature and pH
(2090-2099 relative to 1890-1899)

Δ SST > +1.5°C

ΔpH < -0.2 units

Ocean regions
- Coastal boundary systems
- Eastern boundary upwelling systems
- High-latitude spring bloom systems
- Semi-enclosed seas
- Equatorial upwelling systems
- Subtropical gyres
- Polar systems

Fig. 3. Regional changes in the physical system and associated risks for natural and human-managed systems. Projected changes in SST (ΔSST) and pH (ΔpH) in 2090–2099 relative to pre-industrial under the RCP2.6 and RCP8.5 scenarios are displayed in different colors on the map. The major ocean regions are indicated as well as examples of risks for natural systems and fisheries [modified from (*1*)]. Text in parentheses specifies the level of confidence (*157*).

projected to exceed 0.2 units (the likely bound of natural variability over the past hundreds of thousands of years) by 2100 in close to 23% of North Atlantic deep-sea canyons and 8% of seamounts under RCP8.5—including sites proposed as marine protected areas (*37*).

In summary, the carbon that we emit today will change the earth system irreversibly for many generations to come (*10*). The ocean's content of carbon, acidity, and heat as well as sea level will continue to increase long after atmospheric CO_2 is stabilized. These irreversible changes increase with increasing emissions (Fig. 2), underscoring the urgency of near-term carbon emission reduction if ocean warming and acidification are to be kept at moderate levels.

Effects on biological processes and ecosystems

Organisms and ecosystems are changing in response to ocean warming, acidification, and deoxygenation. The inherent difficulty of distinguishing climate signals from natural variability (*38*), and of accounting for genetic adaptation (*39*), makes documenting these shifts challenging, but nevertheless broad anthropogenic impacts are evident (Figs. 2B and 3).

Warming

Species' range shifts, usually following a shift in isotherms or temperature extremes, are a key consequence of ocean warming (*40*). Recent studies strongly reiterate that many species—including various invertebrates, commercially important fish species, and marine mammals—are undergoing phenological and geographical shifts as a result of warming (*41, 42*). Organisms move at different rates, up to 400 km per decade, as they track temperature changes and local climate velocities according to their ecological niches (*43, 44*). These shifts will continue with projected ocean warming (*42, 45*), causing potentially permanent changes to ecosystems, including local extinctions (*42*), while simultaneously producing novel assemblages (*46*). Responses to changing temperature depend on species' specific windows of thermal tolerance and are positively related to the degree of warming. Exceeding these limits can affect growth, body size, behavior, immune defense, feeding, and reproductive success (*2*), although species' individual tolerances vary. Globally, poleward range shifts of more than 800 species of exploited marine fish and invertebrates projected under RCP8.5 are 65% faster than those under RCP2.6 by mid–21st century relative to the years 2000s (*42*). There is medium confidence that animals adapted to a wide range of temperatures will cope better with future conditions, whereas tropical and polar specialists are at greatest risk (*2*). Changes are not synchronous across trophic levels; alterations in body sizes within food webs (*47*) and in food web composition (*48*) have been reported. Recent experimental studies suggest that some species may adapt to warming projected under RCP8.5 [e.g., (*49, 50*)], but biogeographical shifts restrict adaptive potential and the small number

of species- and population-scale studies limit the ability to generalize the importance of genetic adaptation in moderating impacts.

Reef-building corals are extremely vulnerable to warming (1, 2, 51). Warming causes mass mortality of warm-water corals through bleaching as well as through biotic diseases, resulting in declines in coral abundance and biodiversity. Coral reefs can recover from bleaching events when thermal stress is minimal and of short duration (52). However, ocean warming and acidification are expected to act synergistically to push corals and coral reefs into conditions that are unfavorable for coral reef ecosystems (53). There is limited agreement and low confidence on the potential for corals to adapt to rapid warming. Most coral species have clearly adapted to warm environments (54, 55) although the time scale of adaptation is likely to be long given the relatively lengthy generation times of corals [3 to 100 years (56)]. Recent studies have shown short-term acclimation and adaptation in some fast-growing species (57) and suggested that some genetic mechanisms may allow faster rates of change (58). It is, however, doubtful that corals will be able to adapt quickly enough to maintain populations under most emissions scenarios (56, 59, 60), especially where temperature keeps increasing over time (RCP4.5 and higher). Temperature is also an important determinant of deep-sea coral distribution, although less is known about how deep coral communities respond to thermal stress (61). The consensus is that adaptive responses of organisms will have little chance to keep current ecosystems unchanged if ocean temperature and chemistry are not stabilized, giving marine ecosystems the time needed to adapt to the new, stable environmental conditions.

Ocean acidification

Organisms producing calcium carbonate shells and skeletons experience the strongest negative impacts from ocean acidification (62). Responses to future levels of ocean acidification expected by 2100 under RCP8.5 include reduced calcification, reduced rates of repair, and weakened calcified structures, but responses are species-specific [e.g., (63)]. Reproductive success, early life-stage survival, feeding rate, and stress-response mechanisms may also be affected (2). Most studies have investigated the effects of ocean acidification on isolated organisms; far less is known about the effects on communities and ecosystems.

Few studies measure present-day acidification effects in natural settings. However, recent field observations show a decrease in coccolith thickness over the past 12 years in the Mediterranean (64) and dissolution of live pteropod shells in the California Current system and Southern Ocean, both areas that experience significant anthropogenic acidification (65, 66). Recent investigations have also begun to report community-level responses, for example, in phytoplanktonic (67, 68), bacterial (69), seagrass (70), and algal (71) communities. Decreases in net calcification, at least partly because of ocean acidification, have also been observed in a coral reef over 1975 to 2008

(72), and conditions are already shifting some coral reefs to net erosion (73).

Most studies have investigated phenotypically plastic responses in relatively short-term, single-generation experiments and therefore did not consider the potential for transgenerational response and genetic adaptation (74). Studies published since the AR5 have expanded on the longer-term responses to ocean acidification and have found that transgenerational and evolutionary responses can partly mitigate adverse effects, for example, in phytoplankton (75), planktonic crustaceans (76), sea urchins (77), and fish (78).

Deoxygenation

Expanding oxygen minimum zones benefits microbes and life forms adapted to hypoxia while restricting the ranges of most other species (2), with eutrophication from coastal pollution exacerbating the problem, resulting in organic matter increasing metabolic rates in deeper coastal areas (79). Moreover, higher temperatures increase species' sensitivity to hypoxia (80), limiting the depth distribution of fish and invertebrates not adapted to hypoxic conditions (81) and leading to community-level shifts to smaller Eukarya, Bacteria, and Archaea under conditions of diminished O_2 (82). Conversely, hypoxia-adapted species are likely to benefit, as illustrated by the range-expansion of a squid adapted to hypoxia (83).

Multiple drivers

Investigations of single drivers can produce misleading inferences about organismal responses in a multivariate natural environment because interactive (additive, synergistic, or antagonistic) effects often are not predictable from single-driver studies. This is a major source of uncertainty for projections (39), but several recent studies have better characterized interactions among some drivers. Changes in temperature and pH, such as those projected under RCP8.5 for the year 2100, can have synergistic negative effects on species growth, survival, fitness, calcification, and development (84–88). In some cases, hypoxic conditions can mediate negative effects of ocean acidification (89, 90); however, ocean acidification and hypoxia increase heat sensitivity and vice versa (2), and oxygen loss combined with warming is projected to contract metabolically viable habitats of marine animals on a global scale (91). Growing evidence also suggests that interactions with other environmental factors—such as irradiance, nutrient availability, geographic location, and species community composition—can strongly modulate the biological effects of warming, ocean acidification, and hypoxia (68, 92–95). Few studies addressed the potential for genetic adaptation to multiple drivers, but the phytoplankton *Emiliania huxleyi* can adapt to simultaneous warming and acidification (49). Other direct human impacts (such as fishing) can reduce the adaptive capacity of marine species and ecosystems to CO_2-related impacts. For example, fishing reduces species diversity, simplifies the trophic food web, and increases ecosystem sensitivity to climate change (96). Because relatively little

is known on the interacting effects of environmental factors and the complexity of the marine food web, it is premature to make ecosystem-wide projections. However, impacts on keystone species and ecosystem engineers of three-dimensional habitats are likely to shift whole communities (97).

Present-day impact and future risks

The observed impacts and future additional risks resulting from ocean warming and acidification vary by organism and ecosystem (Fig. 2B). Warm-water corals are already affected, as are mid-latitude seagrass, high-latitude pteropods and krill, mid-latitude bivalves, and finfish. If CO_2 levels are kept to the RCP2.6 scenario, by 2100 the risk of impact increases to "high" for warm-water corals and mid-latitude bivalves. Projections with RCP8.5 indicate very high risk of impact on most marine organisms considered, except mangrove. Avoiding very high levels of risk requires limiting the increase in global surface temperature between 1990 and 2100 to below 2°C and the increase in SST below ~1.2°C. These risks of impact, based on perturbation experiments, field observations, and modeling, are consistent with the paleorecord, which indicates mass extinctions triggered by carbon perturbation events such as at the Permo-Triassic boundary [at a rate slower than the present one (98)] or severe losses of deep-sea fauna during the last glaciation, attributed to oxygen depletion (99). Evolution in response to environmental changes that occurred much slower than those projected in the coming decades did not, therefore, prevent major large-scale alterations of marine ecosystems. Levels of confidence are generally medium to very high for RCP2.6 but significantly lower for RCP8.5, except for seagrass, warm-water corals, and pteropods, for which they remain high or very high (see supplementary materials).

Effects on ecosystem services and ocean-related human activities

Ocean warming, acidification, and deoxygenation alter earth-system-regulating processes (e.g., climate, heat distribution, weather, water flow, and waste treatment), habitat provision, and cultural services [e.g., recreation and leisure, inspiration, and cultural heritage (100)]. As a consequence, CO_2-driven global change is expected to result in economic impacts for humans through the alteration of ocean-derived resources and increasing risks to public health, human development, well-being, and security (101).

Ocean carbon uptake

Ocean uptake of anthropogenic CO_2 is a key service to society that moderates climate change, although it comes at the cost of ocean acidification. CO_2 uptake depends on multiple processes, many of which are sensitive to climate change [see above (102)], and the open ocean is projected to absorb a decreasing fraction of anthropogenic CO_2 emissions as those emissions increase. The fraction of 21st century emissions remaining in the atmosphere consequently increases from 30% for RCP2.6 to 69% for RCP8.5 (27). The

contribution of vegetated coastal ecosystems—including seagrasses, mangrove forests, and salt marshes—to contemporary carbon sequestration (*103*) is an order of magnitude less than that of the land biosphere and open ocean, and the coastal carbon sequestered is likely part of the natural carbon cycle rather than related to anthropogenic emissions. The projected loss of these habitats would not only reduce this relatively small uptake of CO_2, but would also release carbon previously stored and thus exacerbate CO_2-driven changes.

Coastal protection

Coastal habitats—including coral reefs, oyster beds, mangrove forests, salt marshes, kelp forests, and seagrass beds—protect human infrastructure, notably by reducing coastal wave energy, with additional benefits, such as limitation of coastal erosion and marine inundation (*104*, *105*). Nevertheless, the projected increases in coastal human settlements and sea level will combine to expose 0.2 to 4.6% of the global population to inundation annually at a cost to global gross domestic product of 0.3 to 9.3% (*106*). The value of coastal protection in terms of prevented damage can be very large. Coastal wetlands in the United States were estimated to provide U.S. \$23.2 billion year^{-1} in storm protection services (*107*). In contrast to human infrastructure, natural habitats can grow to keep up with sea-level rise, depending on the rate and local conditions, while offering other ecosystem services such as fish and timber (*104*, *108*). These habitats are, however, themselves affected by ocean warming and acidification, in combination with other human disturbances such as urbanization, deforestation, and dredging, making global projections difficult.

Capture fisheries

Ocean warming significantly affects provisioning services through its effects on marine capture fisheries (*109*). Warm-water species have increasingly dominated global fishery catches in recent decades, which can be attributed to a warming ocean (*110–114*). In addition, the maximum size of exploited fishes decreases with rising SST and decreasing oxygen level, ultimately reducing potential fish yield (*115*) in agreement with model predictions (*111*).

Human communities, especially in developing nations, that depend heavily on coastal fisheries resources for food, economic security, and traditional culture are at particular risk from shifts in ocean primary productivity and species ranges (*116–120*). For example, tropical fisheries yield is expected to decrease (*42*, *117*, *121*) in ways that vary among subregions and species (*120*). The loss of critical habitats, such as coral reefs and mangroves, will exacerbate the impacts on tropical fisheries and hence on vulnerable human communities. Substantial declines for tropical fisheries are projected, with robust evidence and strong agreement, even under RCP2.6 by mid–21st century.

Arctic fisheries may benefit from increased primary production, with projected revenue increasing by 14 to 59% by mid–21st century relative to the present day under a high-emissions scenario (*118*). Nevertheless, the Arctic faces increasing overall risk because it is a hot spot of ocean acidification and social vulnerability [including high economic and nutritional dependence on marine resources and limited employment and nutritional alternatives (*118*, *122*)]. Risk of impact on mid-latitude fisheries is more variable depending on the locations and exploited species, but it is expected to increase substantially under RCP8.5 because of the combination of ocean warming, acidification, and deoxygenation (*2*, *123*, *124*). Eventually, changes in the accessibility of marine resources will likely lead to increasing geopolitical and governance challenges for managing trans-boundary stocks and mitigating overexploitation (*125*, *126*), leading to additional economic and societal costs that will be felt unequally and will place heavier burdens on less-advantaged human communities.

Aquaculture

Climate and acidification-related impacts to aquaculture are expected to be generally negative, with impacts varying by location, species, and aquaculture method. Farmed species at higher trophic levels are expected to exhibit higher mortality rates and lower productivity under warming, with open and semi-open aquaculture and those in the tropics particularly at risk (*127*, *128*). A reduction of mussel production by 50 or 70% is projected in the United Kingdom under the RCP2.6 or RCP8.5 scenarios, respectively (*127*). Projected declines in oyster production resulting from warming are much lower, but ocean acidification increases the risk in upwelling areas, such as the Northeast Pacific (*129*). The global economic cost of losses in the capture and aquaculture of molluscs resulting from ocean acidification based on the high-emissions scenario RCP8.5 could be higher than U.S. \$100 billion by the year 2100 (*130*). Sea level rise will bring saline water into deltas and estuaries, where aquaculture commonly occurs (*131*), driving aquaculture upstream and destroying wetlands. Infectious diseases also pose a greater threat to aquaculture in a warmer ocean, with impacts observed, for example, in oysters and abalone aquaculture (*132*) and coastal fish farming (*133*). Risks are also generated by the increased mobility of invasive species (*46*).

Tourism

Decreases in the quality and abundance of coral reef cover are expected to negatively affect tourism (*1*, *3*). Loss of coral reefs to tourism under the RCP2.6 and RCP8.5 scenarios could cost between U.S. \$1.9 billion and U.S. \$12 billion per year, respectively (*134*). Coral reef losses due to ocean warming and acidification on the Great Barrier Reef place up to A\$5.7 billion and 69,000 jobs in Australia at risk (*135*). In addition, ocean acidification may cause an annual loss of reef ecosystem services that are valued up to U.S. \$1 trillion by 2100 (*136*). For about a quarter of countries with reef-related tourism, mainly less-developed countries, this kind of tourism accounts for more than 15% of gross domestic product (*137*) and is more sustainable than extractive livelihoods.

Human health

Ocean warming and acidification affect public health and security, although the impact pathways and associated costs are poorly understood. Hosts and parasites are likely to undergo poleward range shifts under climate change, and disease outbreaks of cholera (*138*) and other *Vibrio* infections (*139*) have already been linked to warmer conditions. The increased risk of pathogens and parasites in marine species and increased opportunities for pathogen transfer between hosts (*140*) can reduce food security (*141*). Increasing intensity and frequency of storm surges and sea-level rise may expand the geographical and seasonal ranges of bacteria, increasing human exposure to diseases (*132*). Inundation can also flood agricultural land in coastal regions, jeopardizing food security and harming human health (*142*).

Present-day impact and future risks

The impacts of ocean acidification and warming have already been detected in some key ecosystem services, such as coastal protection and capture fisheries (Figs. 2C and 3). The risks of impacts increase as a function of increased temperature and decreased pH but are still moderate by 2100 for most services with the RCP2.6 scenario. However, under RCP8.5, we find that the risks of impact will become high or very high by 2100 for all seven ecosystem services considered. Fin fisheries at low latitude will be affected sooner than other services; they will face very high risk at a CO_2 level corresponding to RCP2.6 in 2100. In addition, cumulative or synergistic impacts with other human-induced drivers, such as overexploitation of living resources, habitat destruction, and pollution, will likely exacerbate the risk of CO_2-related impacts.

Management options

Limiting the effects of ocean warming and acidification is critical considering the widespread risks of impacts facing natural and human systems, even under a stringent emissions scenario (RCP2.6; Fig. 2). A growing body of literature presents options for action in response to climate change and ocean acidification (*143–145*). Drawing on Billé *et al.* (*146*), these actions can be clustered in four groups (Fig. 4): reducing the drivers of climate change and ocean acidification (mitigate), building or maintaining resilience in ecosystems (protect), adapting human societies (adapt), and repairing damage that has already occurred (repair). At present, only one of these (reducing CO_2 emissions) addresses the fundamental problem; the others merely delay or decrease impacts (e.g., protecting reefs from major disturbances such as coral mining). Some actions rely on readily available technologies (e.g., sewage treatment plants to reduce exacerbating effects of coastal nutrient pollution) and socioeconomic mechanisms (e.g., coastal setback zones), whereas more engineering-intensive techniques are being developed and will require testing (e.g., removal of
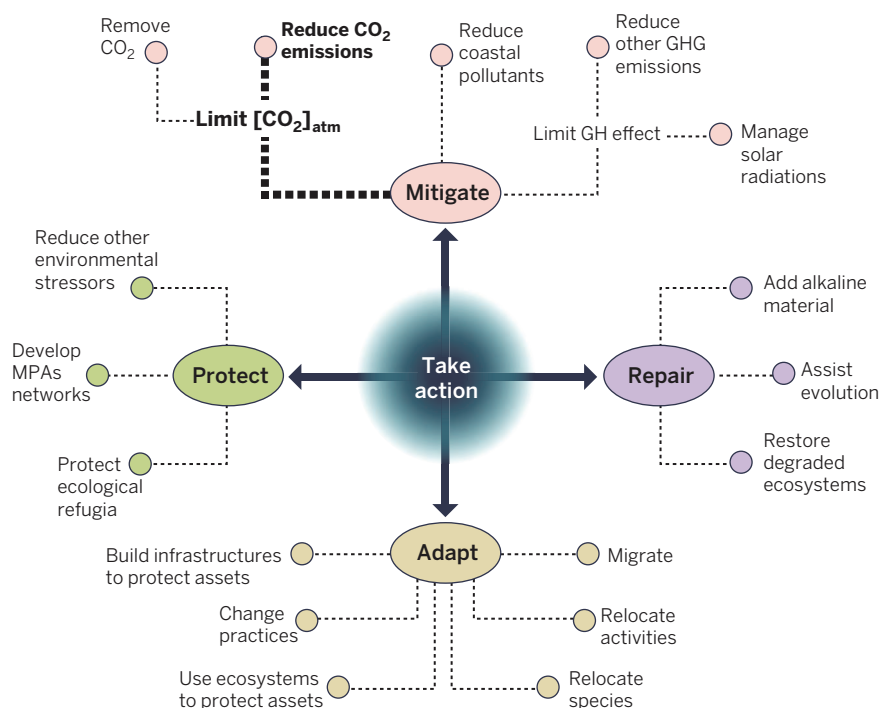
Fig. 4. Four clusters of actions against climate change, including ocean acidification. For each cluster, a nonexhaustive list of actions is shown. $[CO_2]_{atm}$ is concentration of atmospheric CO_2; GH, greenhouse; GHG, greenhouse gases; MPAs, marine protected areas. The mitigation pathway leading to CO_2 reductions is represented in bold, consistent with the consensus view that significant reductions in CO_2 emissions is presently the only actual "solution" to the ocean impacts of climate change and ocean acidification (see main text).

CO_2 from the atmosphere). These options interact. For example, reducing secondary environmental stressors so as to retain ecosystem resilience works over some range of PCO$_2$ values but is ultimately relevant only if ocean warming and acidification are drastically limited. One cannot manage coral reef resilience, for example, if there are no healthy reefs remaining (*46*). Importantly, some policy options are antagonistic: For example, solar radiation management could limit the increase of surface temperature but would reduce the incentive to cut greenhouse gases emissions, including CO_2, thereby providing no relief from ocean acidification (*147*).

A positive development is that a widening range of stakeholders are testing new practices or reviving old ones, including CO_2 extraction from seawater (*148*), assisted evolution of corals (*149*), coral farming (*150*), and customary local management (*151*). Such field tests provide useable information and tools for decision-makers and climate negotiators as to the costs, benefits, and timing of management options. Aquaculture, for example, has shown some potential to reduce the risk of impacts from climate change and ocean acidification through societal adaptation, such as improved monitoring and changing cultured species or farm locations (*127*, *152*). However, the cost of adaptation measures—such as real-time monitoring of water chemistry—can be prohibitive and not within the reach of most aquaculture operations, especially those in the

developing world. Ecosystem-based adaptation—or using ecosystems to reduce the vulnerability of people—appears to offer cost-efficient solutions bringing multiple co-benefits, especially for developing countries and marginalized communities (*153*). Stimulating ecosystem resilience by reducing the number and magnitude of local stressors and setting up marine protected areas (*154*) with strictly enforced no-take areas and limited pollutant inputs also stand out as tractable priorities. Moreover, some regions and local areas that are relatively less exposed to warming, hypoxia, and acidification could be climate change refugia, where more favorable environmental conditions would enable survival under CO_2-driven impacts (*155*). Thus, identifying these climate change refugia and conserving biodiversity there contribute to building resilience to climate change (*156*). Nevertheless, all of these options require appropriate policy frameworks and financial commitments to cover transaction and opportunity costs, surveillance, and enforcement and monitoring and likely offer only limited protection in the face of persistent climate change and ocean acidification.

As the ocean warms and acidifies, the range of protection, adaptation, and repair options—and our confidence in those options—dwindles, while the cost of remaining options skyrockets. Lower-emissions scenarios such as RCP2.6 leave society with a greater number of effective options for safeguarding marine ecosystems and the services

they provide. Therefore, actions that do not reduce carbon emissions are meaningful ocean management options only if the future climate regime entails ambitious national contributions toward the phaseout of global CO_2 emissions as well as a strong funding mechanism and a relevant framework to support on-the-ground implementation of these options.

Key messages

Maintaining ocean ecosystems and services depends in large part on the negotiation process toward a global climate agreement under the UNFCCC. In this regard, four key messages emerge from our analysis. First, the ocean strongly influences the climate system and provides important services to humans. Second, impacts on key marine and coastal organisms, ecosystems, and services from anthropogenic CO_2 emissions are already detectable, and several will face high risk of impacts well before 2100, even with the stringent CO_2 emissions scenario (RCP2.6). These impacts are occurring across all latitudes and have become a global concern that spans the traditional north/south divide. Third, the analysis shows that immediate and substantial reduction of CO_2 emissions is required in order to prevent the massive and effectively irreversible impacts on ocean ecosystems and their services that are projected with emissions scenarios more severe than RCP2.6. Limiting emissions to below this level is necessary to meet UNFCCC's stated objectives. Management options that overlook CO_2, such as solar radiation management and control of methane emission, will only minimize impacts of ocean warming and not those of ocean acidification. Fourth, as CO_2 increases, the protection, adaptation, and repair options for the ocean become fewer and less effective.

Given the contrasting futures we have outlined here, the ocean provides further compelling arguments for rapid and rigorous CO_2 emission reduction and eventual reduction of atmospheric CO_2 content. As a result, any new global climate agreement that does not minimize the impacts on the ocean will be incomplete and inadequate.

REFERENCES AND NOTES

1. O. Hoegh-Guldberg et al., "The ocean," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field et al., Eds. (Cambridge Univ. Press, Cambridge, 2014), pp. 1655–1731.
2. H.-O. Pörtner et al., "Ocean systems," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field et al., Eds. (Cambridge Univ. Press, Cambridge, 2014), pp. 411–484.
3. P. P. Wong et al., "Coastal systems and low-lying areas," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field et al., Eds. (Cambridge Univ. Press, Cambridge, 2014), pp. 361–409.
4. C. Mora, D. P. Tittensor, S. Adl, A. G. Simpson, B. Worm, How many species are there on Earth and in the ocean? *PLOS Biol.* **9**, e1001127 (2011). doi: 10.1371/journal.pbio.1001127; pmid: 21886479

5. Food and Agricultural Organization (FAO), *The State of World Fisheries and Aquaculture 2014* (FAO, Rome, 2014).

6. United Nations, *United Nations Framework Convention on Climate Change* (United Nations, New York, 1992).

7. Copenhagen Accord, *Decision 2/CP.15: Copenhagen Accord* (UNFCCC, Geneva, 2009).

8. E. R. Harrould-Kolieb, D. Herr, Ocean acidification and climate change: Synergies and challenges of addressing both under the UNFCCC. *Clim. Policy* **12**, 378–389 (2012). doi: 10.1080/14693062.2012.620788

9. IPCC, "Summary for policymakers," in *Climate Change 2014: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2014), pp. 1–32.

10. T. F. Stocker *et al.*, "Technical summary," in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2013), pp. 33–115.

11. K. M. Strassmann, G. K. Plattner, F. Joos, CO_2 and non-CO_2 radiative forcings in climate projections for twenty-first century mitigation scenarios. *Clim. Dyn.* **33**, 737–749 (2009). doi: 10.1007/s00382-008-0505-4

12. J.-P. Gattuso, L. Hansson, "Ocean acidification: Background and history," in *Ocean Acidification*, J.-P. Gattuso, L. Hansson, Eds. (Oxford Univ. Press, Oxford, 2011), pp. 1–20.

13. L. A. Levin *et al.*, Comparative biogeochemistry-ecosystem-human interactions on dynamic continental margins. *J. Mar. Syst.* **141**, 3–17 (2015). doi: 10.1016/j.jmarsys.2014.04.016

14. D. Lüthi *et al.*, High-resolution carbon dioxide concentration record 650,000-800,000 years before present. *Nature* **453**, 379–382 (2008). doi: 10.1038/nature06949; pmid: 18480821

15. T. Friedrich *et al.*, Detecting regional anthropogenic trends in ocean acidification against natural variability. *Nat. Clim. Change* **2**, 167–171 (2012). doi: 10.1038/nclimate1372

16. F. Joos, R. Spahni, Rates of change in natural and anthropogenic radiative forcing over the past 20,000 years. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1425–1430 (2008). doi: 10.1073/pnas.0707386105; pmid: 18252830

17. R. A. Feely, C. L. Sabine, J. M. Hernandez-Ayon, D. Ianson, B. Hales, Evidence for upwelling of corrosive "acidified" water onto the continental shelf. *Science* **320**, 1490–1492 (2008). doi: 10.1126/science.1155676; pmid: 18497259

18. J. Salisbury, M. Green, C. Hunt, J. Campbell, Coastal acidification by rivers: A new threat to shellfish? *Eos* **89**, 513 (2008). doi: 10.1029/2008EO500001

19. W.-J. Cai *et al.*, Acidification of subsurface coastal waters enhanced by eutrophication. *Nat. Geosci.* **4**, 766–770 (2011). doi: 10.1038/ngeo1297

20. A. V. Borges, N. Gypens, Carbonate chemistry in the coastal zone responds more strongly to eutrophication than to ocean acidification. *Limnol. Oceanogr.* **55**, 346–353 (2010). doi: 10.4319/lo.2010.55.1.0346

21. K. M. Keller, F. Joos, C. C. Raible, Time of emergence of trends in ocean biogeochemistry. *Biogeosciences* **11**, 3647–3659 (2014). doi: 10.5194/bg-11-3647-2014

22. K. B. Rodgers, J. Lin, T. L. Frölicher, Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosciences* **12**, 3301–3320 (2015). doi: 10.5194/bg-12-3301-2015

23. L. Bopp *et al.*, Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences* **10**, 6225–6245 (2013). doi: 10.5194/bg-10-6225-2013

24. M. Steinacher, F. Joos, T. F. Stocker, Allowable carbon emissions lowered by multiple climate targets. *Nature* **499**, 197–201 (2013). doi: 10.1038/nature12269; pmid: 23823728

25. P. Ciais *et al.*, "Carbon and other biogeochemical cycles," in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2013), pp. 465–570.

26. T. A. Boden, G. Marland, R. J. Andres, *Global, Regional, and National Fossil-Fuel CO_2 Emissions* (Carbon Dioxide Information Analysis Center, Oak Ridge, TN, 2013).

27. C. Jones *et al.*, Twenty-first-century compatible CO_2 emissions and airborne fraction simulated by CMIP5 earth system models under four representative concentration pathways. *J. Clim.* **26**, 4398–4413 (2013).

28. J. A. Church *et al.*, "Sea level change," in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2013), pp. 1137–1216.

29. T. Roy *et al.*, Regional impacts of climate change and atmospheric CO_2 on future ocean carbon uptake: A multimodel linear feedback analysis. *J. Clim.* **24**, 2300–2318 (2011). doi: 10.1175/2010JCLI3787.1

30. T. L. Frölicher, F. Joos, Reversible and irreversible impacts of greenhouse gas emissions in multi-century projections with the NCAR global coupled carbon cycle-climate model. *Clim. Dyn.* **35**, 1439–1459 (2010). doi: 10.1007/s00382-009-0727-0

31. S. Emerson, S. Bushinsky, Oxygen oxygen concentrations and biological fluxes in the open ocean. *Oceanography* **27**, 168–171 (2014). doi: 10.5670/oceanog.2014.20

32. V. Cocco *et al.*, Oxygen and indicators of stress for marine life in multi-model global warming projections. *Biogeosciences* **10**, 1849–1868 (2013). doi: 10.5194/bg-10-1849-2013

33. E. C. Shaw, B. I. McNeil, B. Tilbrook, R. Matear, M. L. Bates, Anthropogenic changes to seawater buffer capacity combined with natural reef metabolism induce extreme future coral reef CO_2 conditions. *Glob. Change Biol.* **19**, 1632–1641 (2013). doi: 10.1111/gcb.12154; pmid: 23505026

34. T. Cyronak, I. R. Santos, D. V. Erler, D. T. Maher, B. D. Eyre, Drivers of pCO_2 variability in two contrasting coral reef lagoons: The influence of submarine groundwater discharge. *Global Biogeochem. Cycles* **28**, 398–414 (2014). doi: 10.1002/2013GB004598

35. L. L. Robbins *et al.*, Baseline monitoring of the western Arctic Ocean estimates 20% of Canadian basin surface waters are undersaturated with respect to aragonite. *PLOS ONE* **8**, e73796 (2013). doi: 10.1371/journal.pone.0073796; pmid: 24040074

36. M. Mattsdotter Björk, A. Fransson, A. Torstensson, M. Chierici, Ocean acidification state in western Antarctic surface waters: Controls and interannual variability. *Biogeosciences* **11**, 57–73 (2014). doi: 10.5194/bg-11-57-2014

37. M. Gehlen *et al.*, Projected pH reductions by 2100 might put deep North Atlantic biodiversity at risk. *Biogeosciences* **11**, 6955–6967 (2014). doi: 10.5194/bg-11-6955-2014

38. P. W. Boyd, S. T. Lennartz, D. M. Glover, S. C. Doney, Biological ramifications of climate-change-mediated oceanic multi-stressors. *Nat. Clim. Change* **5**, 71–79 (2015). doi: 10.1038/nclimate2441

39. U. Riebesell, J.-P. Gattuso, Lessons learned from ocean acidification research. *Nat. Clim. Change* **5**, 12–14 (2015). doi: 10.1038/nclimate2456

40. E. Poloczanska, O. Hoegh-Guldberg, W. Cheung, H.-O. Pörtner, M. T. Burrows, "Observed global responses of marine biogeography, abundance, and phenology to climate change," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2014), pp. 123–127.

41. L. E. Chambers *et al.*, Phenological changes in the southern hemisphere. *PLOS ONE* **8**, e75514 (2013). doi: 10.1371/journal.pone.0075514; pmid: 24098389

42. M. C. Jones, W. W. L. Cheung, Multi-model ensemble projections of climate change effects on global marine biodiversity. *ICES J. Mar. Sci.* **72**, 741–752 (2015). doi: 10.1093/icesjms/fsu172

43. M. L. Pinsky, B. Worm, M. J. Fogarty, J. L. Sarmiento, S. A. Levin, Marine taxa track local climate velocities. *Science* **341**, 1239–1242 (2013). doi: 10.1126/science.1239352; pmid: 24031017

44. J. G. Hiddink, M. T. Burrows, J. García Molinos, Temperature tracking by North Sea benthic invertebrates in response to climate change. *Glob. Change Biol.* **21**, 117–129 (2015). doi: 10.1111/gcb.12726; pmid: 25179407

45. M. S. Wisz *et al.*, Arctic warming will promote Atlantic–Pacific fish interchange. *Nat. Clim. Change* **5**, 261–265 (2015). doi: 10.1038/nclimate2500

46. O. Hoegh-Guldberg, J. F. Bruno, The impact of climate change on the world's marine ecosystems. *Science* **328**, 1523–1528 (2010). doi: 10.1126/science.1189930; pmid: 20558709

47. J. P. Gibert, J. P. DeLong, Temperature alters food web body-size structure. *Biol. Lett.* **10**, 20140473 (2014). doi: 10.1098/rsbl.2014.0473; pmid: 25165457

48. A. Vergés *et al.*, The tropicalization of temperate marine ecosystems: Climate-mediated changes in herbivory and community phase shifts. *Proc. Biol. Sci.* **281**, 20140846 (2014). doi: 10.1098/rspb.2014.0846; pmid: 25009065

49. L. Schlüter *et al.*, Adaptation of a globally important coccolithophore to ocean warming and acidification. *Nat. Clim. Change* **4**, 1024–1030 (2014). doi: 10.1038/nclimate2379

50. N. J. Muñoz, A. P. Farrell, J. W. Heath, B. D. Neff, Adaptive potential of a Pacific salmon challenged by climate change. *Nat. Clim. Change* **5**, 163–166 (2015). doi: 10.1038/nclimate2473

51. J.-P. Gattuso, O. Hoegh-Guldberg, H.-O. Pörtner, "Coral reefs," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2014), pp. 97–100.

52. N. A. Graham, S. Jennings, M. A. MacNeil, D. Mouillot, S. K. Wilson, Predicting climate-driven regime shifts versus rebound potential in coral reefs. *Nature* **518**, 94–97 (2015). doi: 10.1038/nature14140; pmid: 25607371

53. O. Hoegh-Guldberg *et al.*, Coral reefs under rapid climate change and ocean acidification. *Science* **318**, 1737–1742 (2007). doi: 10.1126/science.1152509; pmid: 18079392

54. B. C. C. Hume *et al.*, *Symbiodinium thermophilum* sp. nov., a thermotolerant symbiotic alga prevalent in corals of the world's hottest sea, the Persian/Arabian Gulf. *Sci. Rep.* **5**, 8562 (2015). doi: 10.1038/srep08562; pmid: 25720577

55. R. N. Silverstein, R. Cunning, A. C. Baker, Change in algal symbiont communities after bleaching, not prior heat exposure, increases heat tolerance of reef corals. *Glob. Change Biol.* **21**, 236–249 (2015). doi: 10.1111/gcb.12706; pmid: 25099991

56. O. Hoegh-Guldberg, The adaptation of coral reefs to climate change: Is the Red Queen being outpaced? *Sci. Mar.* **76**, 403–408 (2012). doi: 10.3989/scimar.03660.29A

57. S. R. Palumbi, D. J. Barshis, N. Traylor-Knowles, R. A. Bay, Mechanisms of reef coral resistance to future climate change. *Science* **344**, 895–898 (2014). pmid: 24762535

58. M. Schweinsberg, L. C. Weiss, S. Striewski, R. Tollrian, K. P. Lampert, More than one genotype: How common is intracolonial genetic variability in scleractinian corals? *Mol. Ecol.* **24**, 2673–2685 (2015). doi: 10.1111/mec.13200; pmid: 25872099

59. C. A. Logan, J. P. Dunne, C. M. Eakin, S. D. Donner, Incorporating adaptive responses into future projections of coral bleaching. *Glob. Change Biol.* **20**, 125–139 (2014). doi: 10.1111/gcb.12390; pmid: 24038982

60. C. M. Eakin, Lamarck was partially right—and that is good for corals. *Science* **344**, 798–799 (2014). doi: 10.1126/science.1254136; pmid: 24855237

61. J. M. Roberts, A. J. Wheeler, A. Freiwald, Reefs of the deep: The biology and geology of cold-water coral ecosystems. *Science* **312**, 543–547 (2006). doi: 10.1126/science.1119861; pmid: 16645087

62. J.-P. Gattuso *et al.*, "Ocean acidification: Background and history," in *Ocean Acidification*, C. B. Field *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2014), pp. 129–131.

63. J. Meyer, U. Riebesell, Reviews and Syntheses: Responses of coccolithophores to ocean acidification: A meta-analysis. *Biogeosciences* **12**, 1671–1682 (2015). doi: 10.5194/bg-12-1671-2015

64. K. J. S. Meier, L. Beaufort, S. Heussner, P. Ziveri, The role of ocean acidification in *Emiliania huxleyi* coccolith thinning in the Mediterranean Sea. *Biogeosciences* **11**, 2857–2869 (2014). doi: 10.5194/bg-11-2857-2014

65. N. Bednaršek, G. A. Tarling, D. C. E. Bakker, S. Fielding, R. A. Feely, Dissolution dominating calcification process in polar pteropods close to the point of aragonite undersaturation. *PLOS ONE* **9**, e109183 (2014). doi: 10.1371/journal.pone.0109183; pmid: 25285916

66. N. Bednaršek *et al.*, Extensive dissolution of live pteropods in the Southern Ocean. *Nat. Geosci.* **5**, 881–885 (2012). doi: 10.1038/ngeo1635

67. U. Riebesell, J.-P. Gattuso, T. F. Thingstad, J. J. Middelburg, Arctic ocean acidification: Pelagic ecosystem and biogeochemical responses during a mesocosm study. *Biogeosciences* **10**, 5619–5626 (2013). doi: 10.5194/bg-10-5619-2013

68. S. Richier *et al.*, Phytoplankton responses and associated carbon cycling during shipboard carbonate chemistry

manipulation experiments conducted around Northwest European shelf seas. *Biogeosciences* **11**, 4733–4752 (2014). doi: 10.5194/bg-11-4733-2014

69. S. Endres, L. Galgani, U. Riebesell, K. G. Schulz, A. Engel, Stimulated bacterial growth under elevated pCO₂: Results from an off-shore mesocosm study. *PLOS ONE* **9**, e99228 (2014). doi: 10.1371/journal.pone.0099228; pmid: 24941307

70. S. L. Garrard *et al*., Indirect effects may buffer negative responses of seagrass invertebrate communities to ocean acidification. *J. Exp. Mar. Biol. Ecol.* **461**, 31–38 (2014). doi: 10.1016/j.jembe.2014.07.011

71. A. Ordoñez, C. Doropoulos, G. Diaz-Pulido, Effects of ocean acidification on population dynamics and community structure of crustose coralline algae. *Biol. Bull.* **226**, 255–268 (2014). pmid: 25070869

72. J. Silverman *et al*., Community calcification in Lizard Island, Great Barrier Reef: A 33 year perspective. *Geochim. Cosmochim. Acta* **144**, 72–81 (2014). doi: 10.1016/j.gca.2014.09.011

73. N. J. Silbiger, Ò. Guadayol, F. I. M. Thomas, M. J. Donahue, Reefs shift from net accretion to net erosion along a natural environmental gradient. *Mar. Ecol. Prog. Ser.* **515**, 33–44 (2014). doi: 10.3354/meps10999

74. J. M. Sunday *et al*., Evolution in an acidifying ocean. *Trends Ecol. Evol.* **29**, 117–125 (2014). doi: 10.1016/j.tree.2013.11.001; pmid: 24355315

75. K. T. Lohbeck, U. Riebesell, T. B. H. Reusch, Gene expression changes in the coccolithophore *Emiliania huxleyi* after 500 generations of selection to ocean acidification. *Proc. R. Soc. London Ser. B* **281**, 20140003–20140003 (2014). doi: 10.1098/rspb.2014.0003; pmid: 24827439

76. P. Thor, S. Dupont, Transgenerational effects alleviate severe fecundity loss during ocean acidification in a ubiquitous planktonic copepod. *Glob. Change Biol.* **21**, 2261–2271 (2015). doi: 10.1111/gcb.12815; pmid: 25430823

77. C. C. Suckling *et al*., Experimental influence of pH on the early life-stages of sea urchins II: Increasing parental exposure times gives rise to different responses. *Invertebr. Reprod. Dev.* **58**, 161–175 (2014). doi: 10.1080/07924259.2013.875951

78. P. L. Munday, Transgenerational acclimation of fishes to climate change and ocean acidification. *F1000Prime Rep.* **6**, 99 (2014). doi: 10.12703/P6-99; pmid: 25580253

79. R. J. Diaz, R. Rosenberg, Spreading dead zones and consequences for marine ecosystems. *Science* **321**, 926–929 (2008). doi: 10.1126/science.1156401; pmid: 18703733

80. H.-O. Pörtner, Oxygen- and capacity-limitation of thermal tolerance: A matrix for integrating climate-related stressor effects in marine ecosystems. *J. Exp. Biol.* **213**, 881–893 (2010). doi: 10.1242/jeb.037523; pmid: 20190113

81. A. Brown, S. Thatje, The effects of changing climate on faunal depth distributions determine winners and losers. *Glob. Change Biol.* **21**, 173–180 (2015). doi: 10.1111/gcb.12680; pmid: 25044552

82. D. Storch, L. Menzel, S. Frickenhaus, H.-O. Pörtner, Climate sensitivity across marine domains of life: Limits to evolutionary adaptation shape species interactions. *Glob. Change Biol.* **20**, 3059–3067 (2014). doi: 10.1111/gcb.12645; pmid: 24890266

83. J. S. Stewart *et al*., Combined climate- and prey-mediated range expansion of Humboldt squid (*Dosidicus gigas*), a large marine predator in the California Current System. *Glob. Change Biol.* **20**, 1832–1843 (2014). doi: 10.1111/gcb.12502; pmid: 24443361

84. J. D. Gaitán-Espitia *et al*., Interactive effects of elevated temperature and pCO₂ on early-life-history stages of the giant kelp *Macrocystis pyrifera*. *J. Exp. Mar. Biol. Ecol.* **457**, 51–58 (2014). doi: 10.1016/j.jembe.2014.03.018

85. C. J. Gobler, E. L. DePasquale, A. W. Griffith, H. Baumann, Hypoxia and acidification have additive and synergistic negative effects on the growth, survival, and metamorphosis of early life stage bivalves. *PLOS ONE* **9**, e83648 (2014). doi: 10.1371/journal.pone.0083648; pmid: 24416169

86. C. L. Mackenzie *et al*., Ocean warming, more than acidification, reduces shell strength in a commercial shellfish species during food limitation. *PLOS ONE* **9**, e86764 (2014). doi: 10.1371/journal.pone.0086764; pmid: 24489785

87. D. Madeira, L. Narciso, M. S. Diniz, C. Vinagre, Synergy of environmental variables alters the thermal window and heat shock response in the crab *Pachygrapsus marmoratus*. *Mar. Environ. Res.* **98**, 21–28 (2014). doi: 10.1016/j.marenvres.2014.03.011; pmid: 24836643

88. R. Rosa *et al*., Differential impacts of ocean acidification and warming on winter and summer progeny of a coastal squid (*Loligo vulgaris*). *J. Exp. Biol.* **217**, 518–525 (2014). doi: 10.1242/jeb.096081; pmid: 24523499

89. C. A. Frieder, J. P. Gonzalez, E. E. Bockmon, M. O. Navarro, L. A. Levin, Can variable pH and low oxygen moderate ocean acidification outcomes for mussel larvae? *Glob. Change Biol.* **20**, 754–764 (2014). doi: 10.1111/gcb.12485; pmid: 24343909

90. J. Mukherjee *et al*., Proteomic response of marine invertebrate larvae to ocean acidification and hypoxia during metamorphosis and calcification. *J. Exp. Biol.* **216**, 4580–4589 (2013). doi: 10.1242/jeb.094516; pmid: 24307710

91. C. Deutsch, A. Ferrel, B. Seibel, H.-O. Pörtner, R. B. Huey, Climate change tightens a metabolic constraint on marine habitats. *Science* **348**, 1132–1135 (2015). doi: 10.1126/science.aaa1605

92. S. Comeau, R. C. Carpenter, P. J. Edmunds, Effects of irradiance on the response of the coral *Acropora pulchra* and the calcifying alga *Hydrolithon reinboldii* to temperature elevation and ocean acidification. *J. Exp. Mar. Biol. Ecol.* **453**, 28–35 (2014). doi: 10.1016/j.jembe.2013.12.013

93. C. J. Hoppe *et al*., Iron limitation modulates ocean acidification effects on southern ocean phytoplankton communities. *PLOS ONE* **8**, e79890 (2013). doi: 10.1371/journal.pone.0079890; pmid: 24278207

94. G. W. K. Ko *et al*., Interactive effects of ocean acidification, elevated temperature, and reduced salinity on early-life stages of the pacific oyster. *Environ. Sci. Technol.* **48**, 10079–10088 (2014). doi: 10.1021/es501611u; pmid: 25014366

95. A. J. Poulton *et al*., Coccolithophores on the north-west European shelf: Calcification rates and environmental controls. *Biogeosciences* **11**, 3919–3940 (2014). doi: 10.5194/bg-11-3919-2014

96. R. I. Perry, R. E. Ommer, M. Barange, F. Werner, The challenge of adapting marine social–ecological systems to the additional stress of climate change. *Curr. Opin. Environ. Sustain.* **2**, 356–363 (2010). doi: 10.1016/j.cosust.2010.10.004

97. J. Brodie *et al*., The future of the northeast Atlantic benthic flora in a high CO₂ world. *Ecol. Evol.* **4**, 2787–2798 (2014). doi: 10.1002/ece3.1105; pmid: 25077027

98. M. O. Clarkson *et al*., Ocean acidification and the Permo-Triassic mass extinction. *Science* **348**, 229–232 (2015). doi: 10.1126/science.aaa0193; pmid: 25859043

99. S. E. Moffitt, T. M. Hill, P. D. Roopnarine, J. P. Kennett, Response of seafloor ecosystems to abrupt global climate change. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4684–4689 (2015). doi: 10.1073/pnas.1417130112; pmid: 25825727

100. A. Böhnke-Henrichs, C. Baulcomb, R. Koss, S. S. Hussain, R. S. de Groot, Typology and indicators of ecosystem services for marine spatial planning and management. *J. Environ. Manage.* **130**, 135–145 (2013). doi: 10.1016/j.jenvman.2013.08.027; pmid: 24076513

101. J. Mendler de Suarez, B. Cicin-Sain, K. Wowk, R. Payet, O. Hoegh-Guldberg, Ensuring survival: Oceans, climate and security. *Ocean Coast. Manage.* **90**, 27–37 (2014). doi: 10.1016/j.ocecoaman.2013.08.007

102. U. Riebesell, A. Körtzinger, A. Oschlies, Sensitivities of marine carbon fluxes to ocean change. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20602–20609 (2009). doi: 10.1073/pnas.0813291106; pmid: 19995981

103. C. M. Duarte, I. J. Losada, I. E. Hendriks, I. Mazarrasa, N. Marbà, The role of coastal plant communities for climate change mitigation and adaptation. *Nat. Clim. Change* **3**, 961–968 (2013). doi: 10.1038/nclimate1970

104. M. D. Spalding *et al*., The role of ecosystems in coastal protection: Adapting to climate change and coastal hazards. *Ocean Coast. Manage.* **90**, 50–57 (2014). doi: 10.1016/j.ocecoaman.2013.09.007

105. B. Ondiviela *et al*., The role of seagrasses in coastal protection in a changing climate. *Coast. Eng.* **87**, 158–168 (2014). doi: 10.1016/j.coastaleng.2013.11.005

106. J. Hinkel *et al*., Coastal flood damage and adaptation costs under 21st century sea-level rise. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3292–3297 (2014). doi: 10.1073/pnas.1222469111; pmid: 24596428

107. R. Costanza *et al*., The value of coastal wetlands for hurricane protection. *Ambio* **37**, 241–248 (2008). doi: 10.1579/0044-7447(2008)37[241:TVOCWF]2.0.CO;2; pmid: 18686502

108. E. B. Barbier, Valuing the storm protection service of estuarine and coastal ecosystems. *Ecosyst. Serv.* **11**, 32–38 (2015). doi: 10.1016/j.ecoser.2014.06.010

109. U. R. Sumaila, W. W. L. Cheung, V. W. Y. Lam, D. Pauly, S. Herrick, Climate change impacts on the biophysics and economics of world fisheries. *Nat. Clim. Change* **1**, 449–456 (2011). doi: 10.1038/nclimate1301

110. W. W. Cheung, R. Watson, D. Pauly, Signature of ocean warming in global fisheries catch. *Nature* **497**, 365–368 (2013). doi: 10.1038/nature12156; pmid: 23676754

111. W. W. L. Cheung *et al*., Shrinking of fishes exacerbates impacts of global ocean changes on marine ecosystems. *Nat. Clim. Change.* **3**, 254–258 (2012). doi: 10.1038/nclimate1691

112. I. Montero-Serra, M. Edwards, M. J. Genner, Warming shelf seas drive the subtropicalization of European pelagic fish communities. *Glob. Change Biol.* **21**, 144–153 (2015). doi: 10.1111/gcb.12747; pmid: 25230844

113. S. D. Simpson *et al*., Continental shelf-wide response of a fish assemblage to rapid warming of the sea. *Curr. Biol.* **21**, 1565–1570 (2011). doi: 10.1016/j.cub.2011.08.016; pmid: 21924906

114. D. Yemane *et al*., Assessing changes in the distribution and range size of demersal fish populations in the Benguela Current Large Marine Ecosystem. *Rev. Fish Biol. Fish.* **24**, 463–483 (2014). doi: 10.1007/s11160-014-9357-7

115. A. R. Baudron, C. L. Needle, A. D. Rijnsdorp, C. T. Marshall, Warming temperatures and smaller body sizes: Synchronous changes in growth of North Sea fishes. *Glob. Change Biol.* **20**, 1023–1031 (2014). doi: 10.1111/gcb.12514; pmid: 24375891

116. S. R. Cooley, J. T. Mathis, Addressing ocean acidification as part of sustainable ocean development. *Ocean Yearbook Online* **27**, 29–47 (2013). doi: 10.1163/22116001-90000153

117. M. Barange *et al*., Impacts of climate change on marine ecosystem production in societies dependent on fisheries. *Nat. Clim. Change* **4**, 211–216 (2014). doi: 10.1038/nclimate2119

118. V. W. Y. Lam, W. W. L. Cheung, U. R. Sumaila, Marine capture fisheries in the Arctic: Winners or losers under climate change and ocean acidification? *Fish Fish.*, (2014). doi: 10.1111/faf.12106

119. V. W. Y. Lam, W. Cheung, W. Swartz, U. R. Sumaila, Climate change impacts on fisheries in West Africa: Implications for economic, food and nutritional security. *Afr. J. Mar. Sci.* **34**, 103–117 (2012). doi: 10.2989/1814232X.2012.673294

120. J. D. Bell *et al*., Mixed responses of tropical Pacific fisheries and aquaculture to climate change. *Nat. Clim. Change* **3**, 591–599 (2013). doi: 10.1038/nclimate1838

121. W. W. L. Cheung *et al*., Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Glob. Change Biol.* **16**, 24–35 (2010). doi: 10.1111/j.1365-2486.2009.01995.x

122. J. T. Mathis *et al*., Ocean acidification risk assessment for Alaska's fishery sector. *Prog. Oceanogr.*, (2014). doi: 10.1016/j.pocean.2014.07.001

123. C. H. Ainsworth *et al*., Potential impacts of climate change on Northeast Pacific marine foodwebs and fisheries. *ICES J. Mar. Sci.* **68**, 1217–1229 (2011). doi: 10.1093/icesjms/fsr043

124. W. W. L. Cheung, J. Pinnegar, G. Merino, M. C. Jones, M. Barange, Review of climate change impacts on marine fisheries in the UK and Ireland. *Aquat. Conserv. Mar. Freshwater Ecosyst.* **22**, 368–388 (2012). doi: 10.1002/aqc.2248

125. J. S. Christiansen, C. W. Mecklenburg, O. V. Karamushko, Arctic marine fishes and their fisheries in light of global change. *Glob. Change Biol.* **20**, 352–359 (2014). doi: 10.1111/gcb.12395; pmid: 24105993

126. K. A. Miller, G. R. Munro, U. R. Sumaila, W. W. L. Cheung, Governing marine fisheries in a changing climate: A game-theoretic perspective. *Can. J. Agric. Econ.* **61**, 309–334 (2013). doi: 10.1111/cjag.12011

127. R. Callaway *et al*., Review of climate change impacts on marine aquaculture in the UK and Ireland. *Aquat. Conserv. Mar. Freshwater Ecosyst.* **22**, 389–421 (2012). doi: 10.1002/aqc.2247

128. M. Ruckelshaus *et al*., Securing ocean benefits for society in the face of climate change. *Mar. Policy* **40**, 154–159 (2013). doi: 10.1016/j.marpol.2013.01.009

129. A. Barton, B. Hales, G. G. Waldbusser, C. Langdon, R. A. Feely, The Pacific oyster, *Crassostrea gigas*, shows negative correlation to naturally elevated carbon dioxide levels: Implications for near-term ocean acidification effects. *Limnol. Oceanogr.* **57**, 698–710 (2012). doi: 10.4319/lo.2012.57.3.0698

130. D. Narita, K. Rehdanz, R. S. J. Tol, Economic costs of ocean acidification: A look into the impacts on global shellfish production. *Clim. Change* **113**, 1049–1063 (2012). doi: 10.1007/s10584-011-0383-3

131. S. S. De Silva, "Climate change impacts: Challenges for aquaculture," in *Farming the Waters for People and Food*, R. P. Subasinghe *et al.*, Eds. (FAO and Network of Aquaculture Centres in Asia-Pacific, Rome and Bangkok, 2012), pp. 75–110.

132. C. A. Burge *et al.*, Climate change influences on marine infectious diseases: Implications for management and society. *Annu. Rev. Mar. Sci.* **6**, 249–277 (2014). doi: 10.1146/annurev-marine-010213-135029; pmid: 23808894

133. J. Garai, "The impacts of climate change on the livelihoods of coastal people in Bangladesh: A sociological study," in *International Perspectives on Climate Change*, W. Leal Filho, F. Alves, S. Caeiro, U. M. Azeiteiro, Eds. (Springer, Switzerland, 2014), pp. 151–163.

134. P.-Y. Chen, C.-C. Chen, L. F. Chu, B. McCarl, Evaluating the economic damage of climate change on global coral reefs. *Glob. Environ. Change* **30**, 12–20 (2015). doi: 10.1016/j.gloenvcha.2014.10.011

135. Deloitte Access Economics, *Economic Contribution of the Great Barrier Reef* (Great Barrier Reef Marine Park Authority, Townsville, Australia, 2013).

136. L. M. Brander, K. Rehdanz, R. S. J. Tol, P. J. H. Van Beukering, The economic impact of ocean acidification on coral reefs. *Clim. Change Econ.* **03**, 1250002 (2012). doi: 10.1142/S2010007812500029

137. L. M. Burke, K. Reytar, M. Spalding, A. Perry, *Reefs at Risk Revisited* (World Resources Institute, Washington, DC, 2011), p. 114.

138. M. Pascual, X. Rodó, S. P. Ellner, R. Colwell, M. J. Bouma, Cholera dynamics and El Niño-Southern Oscillation. *Science* **289**, 1766–1769 (2000). doi: 10.1126/science.289.5485.1766; pmid: 10976073

139. C. Baker-Austin *et al.*, Emerging Vibrio risk at high latitudes in response to ocean warming. *Nat. Clim. Change* **3**, 73–77 (2013). doi: 10.1038/nclimate1628

140. S. Altizer, R. S. Ostfeld, P. T. Johnson, S. Kutz, C. D. Harvell, Climate change and infectious diseases: From evidence to a predictive framework. *Science* **341**, 514–519 (2013). doi: 10.1126/science.1239401; pmid: 23908230

141. T. L. F. Leung, A. E. Bates, More rapid and severe disease outbreaks for aquaculture at the tropics: Implications for food security. *J. Appl. Ecol.* **50**, 215–222 (2013). doi: 10.1111/1365-2644.12017

142. T. Wheeler, J. von Braun, Climate change impacts on global food security. *Science* **341**, 508–513 (2013). doi: 10.1126/science.1239402; pmid: 23908229

143. R. P. Kelly, M. R. Caldwell, Ten ways states can combat ocean acidification (and why they should). *Harvard Environ. Law Rev.* **37**, 57–103 (2013). doi: 10.2139/ssrn.2020520

144. E. Mcleod *et al.*, Preparing to manage coral reefs for ocean acidification: Lessons from coral bleaching. *Front. Ecol. Environ* **11**, 20–27 (2013). doi: 10.1890/110240

145. A. L. Strong, K. J. Kroeker, L. T. Teneva, L. A. Mease, R. P. Kelly, Ocean acidification 2.0: Managing our changing coastal ocean chemistry. *Bioscience* **64**, 581–592 (2014). doi: 10.1093/biosci/biu072

146. R. Billé *et al.*, Taking action against ocean acidification: A review of management and policy options. *Environ. Manage.* **52**, 761–779 (2013). doi: 10.1007/s00267-013-0132-7; pmid: 23897413

147. Committee on Geoengineering Climate, "Technical evaluation and discussion of impacts," in *Climate Intervention: Reflecting Sunlight to Cool Earth* (National Academy of Sciences, Washington, DC, 2015).

148. M. D. Eisaman *et al.*, CO_2 extraction from seawater using bipolar membrane electrodialysis. *Energy Environ. Sci.* **5**, 7346 (2012). doi: 10.1039/c2ee03393c

149. M. J. H. van Oppen, J. K. Oliver, H. M. Putnam, R. D. Gates, Building coral reef resilience through assisted evolution. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2307–2313 (2015). doi: 10.1073/pnas.1422301112; pmid: 25646461

150. C. N. Young, S. A. Schopmeyer, D. Lirman, A review of reef restoration and coral propagation using the threatened genus *Acropora* in the Caribbean and Western Atlantic. *Bull. Mar. Sci.* **88**, 1075–1098 (2012). doi: 10.5343/bms.2011.1143

151. H. Govan *et al.*, *Status and Potential of Locally-Managed Marine Areas in the South Pacific: Meeting Nature Conservation and Sustainable Livelihood Targets Through Wide-Spread Implementation of LMMAs* (South Pacific Regional Environmental Program/WWF/WorldFish-Reefbase/Coral Reefs Initiative for the Pacific, Noumea, New Caledonia, 2009).

152. R. P. Kelly, S. R. Cooley, T. Klinger, Narratives can motivate environmental action: The Whiskey Creek ocean acidification story. *Ambio* **43**, 592–599 (2014). doi: 10.1007/s13280-013-0442-2; pmid: 24081705

153. L. Weatherdon, A. Rogers, R. Sumaila, A. Magnan, W. W. L. Cheung, *The Oceans 2015 Initiative, Part II: An Updated Understanding of the Observed and Projected Impacts of Ocean Warming and Acidification on Marine and Coastal Socioeconomic Activities/Sectors* (Institut du Développement Durable et des Relations Internationales, Paris, 2015).

154. R. Murti, C. Buyck, Eds., *Safe Havens: Protected Areas for Disaster Risk Reduction and Climate Change Adaptation* (IUCN, Gland, Switzerland, 2014).

155. G. Keppel *et al.*, The capacity of refugia for conservation planning under climate change. *Front. Ecol. Environ* **13**, 106–112 (2015). doi: 10.1890/140055

156. C. Cacciapaglia, R. van Woesik, Reef-coral refugia in a rapidly changing ocean. *Glob. Change Biol.* **21**, 2272–2282 (2015). doi: 10.1111/gcb.12851; pmid: 25646684

157. M. D. Mastrandrea *et al.*, *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties* (IPCC, New York, 2010).

RESEARCH ARTICLE SUMMARY

BIOMECHANICS

Why the seahorse tail is square

Michael M. Porter,* Dominique Adriaens,
Ross L. Hatton, Marc A. Meyers, Joanna McKittrick

INTRODUCTION: Although the predominant shapes of most animal tails are cylindrical, seahorse tails are square prisms. The skeleton of their tails consists of a bony armor arranged into several ringlike segments composed of four L-shaped plates that surround a central vertebra. These plates articulate with specialized joints that facilitate bending and twisting, as well as resist vertebral fracture from crushing. Muscles attached to the vertebral column transmit forces to the bony plates to provide motion for grasping and holding on to objects such as sea grasses, mangrove roots, and coral reefs, which allows them to hide and rely on camouflage when evading predators and capturing prey.

RATIONALE: We hypothesize that the square cross-sectional architecture of a seahorse tail improves mechanical performance in prehension (grasping ability) and armored functions (crushing resistance), relative to a cylindrical one. To test this hypothesis, we evaluated the mechanics of two three-dimensional (3D)–printed prototypes composed of articulating plates and vertebrae that mimic the natural (square prism) and a hypothetical (cylindrical) tail structure. We compared the bending, twisting, and compressive behavior of the biomimetic prototypes to show that the square profile is better than the circular one for two integrated functions: grasping ability and crushing resistance.

RESULTS: Seahorse tails (and the prototypes) have three primary joints that enable motion: ball-and-socket, peg-and-socket, and gliding. The ball-and-socket joints connect adjacent

vertebrae and constrain bending in both the square and cylindrical prototypes to the same degree, exhibiting a behavior similar to that of a natural seahorse tail. The peg-and-socket joints connect the plates of adjacent segments and substantially restrict twisting in the prototype with a square profile, as compared with the circular one. The square geometry limits
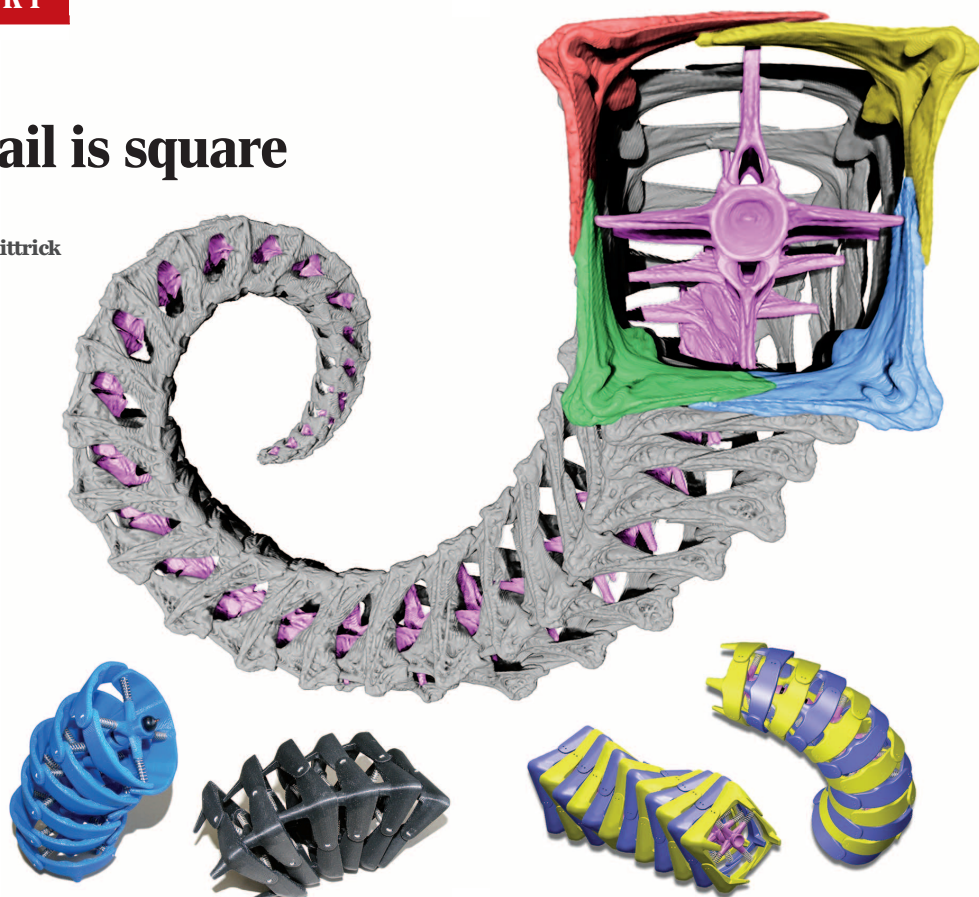
Engineering designs answer biological questions. 3D-printed models that mimic a seahorse tail were designed not only for potential engineering applications but also to answer the biological question, why might tails organized into square prisms be better than cylinders? A mechanical comparison of the prototypes shows that articulated square prisms perform better than do cylinders for grasping and resistance to crushing.

excessive torsion and preserves articulatory organization, which could provide seahorses a natural safety factor against torsion-induced damage and assist in tail relaxation. Further, the square architecture is flat (increasing surface contact) and undergoes an exterior shape change when twisted, which could allow seahorses to grasp objects with more control. Gliding joints are present at the plate overlaps along all four sides of both prototypes. Under transverse compression and impact (with a rubber mallet), the plates of the square prototype slide past one another with one degree of translation freedom (analogous to the crushing behavior of a natural seahorse tail), exhibiting a response that is stiffer, stronger, and more resilient than its cylindrical counterpart, whose plates translate and rotate on impact.

CONCLUSION: Exploration of these biologically inspired designs provides insight into

the mechanical benefits for seahorses to have evolved prehensile tails composed of armored plates organized into square prisms. Beyond their intended practical applications, engineering designs are convenient means to answer elusive biological questions when live animal data are unavailable (for example, seahorses do not have cylindrical tails). Understanding the role of mechanics in these prototypes may help engineers to develop future seahorse-inspired technologies that mimic the prehensile and armored functions of the natural appendage for a variety of applications in robotics, defense systems, or biomedicine. ∎

BIOMECHANICS

Why the seahorse tail is square

Michael M. Porter,[1]* Dominique Adriaens,[2] Ross L. Hatton,[3]
Marc A. Meyers,[4,5,6] Joanna McKittrick[4,5]

Whereas the predominant shapes of most animal tails are cylindrical, seahorse tails are square prisms. Seahorses use their tails as flexible grasping appendages, in spite of a rigid bony armor that fully encases their bodies. We explore the mechanics of two three-dimensional–printed models that mimic either the natural (square prism) or hypothetical (cylindrical) architecture of a seahorse tail to uncover whether or not the square geometry provides any functional advantages. Our results show that the square prism is more resilient when crushed and provides a mechanism for preserving articulatory organization upon extensive bending and twisting, as compared with its cylindrical counterpart. Thus, the square architecture is better than the circular one in the context of two integrated functions: grasping ability and crushing resistance.

Seahorses do not use their tails for swimming because they lack a caudal fin (*1*), but instead as flexible prehensile appendages that allow them to hide and rely on body crypsis to evade predators (*2*) and capture prey (*3*) by holding on to objects such as sea grasses, mangrove roots, and coral reefs (*4*). The skeleton of a seahorse consists of a rigid, subdermal armor of articulating bony plates, formed within the skin, that fully encase its body (Fig. 1A) (*5–13*). The tail armor is arranged into several ring-like segments composed of four L-shaped plates that surround a central vertebra (Fig. 1B). Muscles attached to the vertebrae transmit forces to the bony plates to control motion (*10–12*). Recent studies on the musculoskeletal structure, biomechanics, and kinematics of seahorse tails suggest that this bony-plated armor actively facilitates tail bending and twisting (Fig. 1B) (*10–12*) as well as resists vertebral fracture due to impact and crushing (Fig. 1C) (*13*). Selective pressures related to the defensive strategies [bony plates make eating seahorses difficult (*2*)] and feeding behavior [tail-attached, sit-and-wait suction feeding (*3*)] of seahorses may have driven the evolution of an ancestral plated system (square in cross section, but rather rigid and nonprehensile) into a more flexible system of similar plate organization for an added grasping functionality. Accordingly, we hypothesize that the cross-sectional square architecture of a seahorse tail possesses several mechanical advantages that

could have helped promote this adaptation, especially when compared with other grasping appendages in unrelated animal species and other plated (nonsquare) skeletons in unrelated fish species.

Whereas several animals have developed grasping appendages that are cylindrical and composed of soft tissues surrounding a central axis of articulating skeletal elements (for example, tails of New World monkeys and some lizards, sala-

manders, marsupials, and rodents) (*14–17*), a peripherally plated skeleton is not exclusive to fishes and is even an ancestral condition of the vertebral lineages (*18*, *19*), with different patterns having arisen independently (such as ostracoderms, polypterids, gasterosteids, loricariids, and pegasids) (*20*). One recurrent pattern is where consecutive plates articulate through a peg-and-socket connection, combining some level of flexibility with strength. Such a joint is also present in seahorses, as well as closely related species from the same family of Syngnathidae (pipehorses, pipefishes, and seadragons). What differs among them, however, is that there is notable variation in the size of pegs and shape of sockets, as has recently been observed (*12*). The ancestral syngnathid configuration is a rather rigid nonprehensile tail that is completely enclosed by strongly abutting but articulating (peg-and-socket) bony plates, such as those found in pipefishes (*12*). Seahorses and pipehorses are the only lineages that use their tails as prehensile organs. Yet, they share with other syngnathids, and closely related seamoths (Pegasidae), a similar tail cross section that is not ovally shaped (as would be expected for fishes), but approximately square (*5–13*). Because the latter (pipefishes, seadragons, and seamoths) do not have flexible tails, this suggests that the square morphology may not reflect a structural adaptation directly related to tail prehensility but could provide some form of protection. Although it is likely that larger fish may swallow seahorses and other syngnathids whole, the peripherally plated

[1]Department of Mechanical Engineering, Clemson University, Clemson, SC 29634, USA. [2]Evolutionary Morphology of Vertebrates, Ghent University, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium. [3]School of Mechanical, Industrial, and Manufacturing Engineering, Oregon State University, Corvallis, OR 97330, USA. [4]Materials Science and Engineering Program, University of California, San Diego, La Jolla, CA 92093, USA. [5]Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA 92093, USA. [6]Department of NanoEngineering, University of California, San Diego, La Jolla, CA 92093, USA.
*Corresponding author. E-mail: mmporte@clemson.edu

Fig. 1. Seahorse skeletons are composed of highly articulated bony plates that surround a central vertebral column. (**A** and **B**) μCT images of (A) a seahorse skeleton (*H. reidi*) and (B) its tail, illustrating its square cross section and particular flexibility. (**C**) μCT image of a seahorse tail segment (*H. kuda*) after crushing up to ~60% of its original height. For visual clarity, the vertebrae are color-masked magenta, and the bony plates are color-masked red, yellow, blue, and green. Scale bars, (A) 10 mm; (B) and (C) 2 mm.

skeleton can protect their soft tissues against bites from smaller predators (*13*). In the case of pipefishes, which closely resemble the syngnathid ancestral state, fortifying the body as a poorly bendable "stick" makes them difficult prey to swallow and digest (*2*). Also, their second largest group of predators (next to teleosts) are water-birds, with waders and seabirds accounting for 32% of all syngnathid predators (*2*). Because many birds use beaks to clamp down and capture prey before feeding, it is likely that the skeletal armor could protect seahorses from crushing (*13*). When a seahorse tail is crushed, for instance, the bony plates slide past one another, acting as an energy absorption mechanism, which resists fracture of the vertebral column (Fig. 1C) (*13*). This is different from the mechanisms observed in the skin and scales of other fish that provide an optimal flexural stiffness for both swimming and protection against predatory attacks (*21–23*). Hence, this complex skeletal system may provide seahorses an advantageous combination of joint flexibility for prehension with shielding from external forces for protection.

Similar to recent works that used robots to investigate animal behavior and locomotion (*24, 25*), we used three-dimensional (3D)–printed models to explore the mechanics of seahorse tails. Because cylindrical seahorse tails do not exist in nature, 3D printing allows us to build functional models of this hypothetical system so as to investigate what mechanical advantages an articulated square prism may have over a cylinder, which is the common shape of most other animal tails. This falsifiable methodology is particularly useful to examine the mechanics of musculoskeletal systems in vertebrates that are difficult or impossible to study in nature (for example, seahorses do not have cylindrical tails). Thus, via 3D printing we replicated comparative models that mimic the specialized joints and complex morphologies of both the natural (square prism) and hypothetical (cylindrical) systems.

In a quest to better understand armored prehensility in seahorses and develop useful tools inspired by their tails, we designed this new class of highly articulated structures—not only for biological research, but also for a wide range of potential robotics, armored systems, and biomedical applications. Such applications for the seahorse-inspired technology would be able to take advantage of a large body of kinematics and actuation work that has been developed for similar long, slender "continuum" structures inspired by organisms such as snakes and worms (*26–30*), squids and octopuses (tentacles) (*31, 32*), elephants (trunks) (*33*), and plants (tendrils and roots) (*34, 35*). The combination of articulated rigid plates and elastic deformability observed in a seahorse tail (Fig. 1) could strike a promising balance between the two dominant paradigms in robotics design: being lighter and more compliant than traditional "hard" robots built from servos and metal, but more robust and resistant to external tractions than the emerging class of "soft" robots with silicone-membrane bodies. However, akin to nature, the vast majority of these serially articulated structures (*26–35*) have cylindrical profiles. Therefore, we ask the question, why might square prisms be better than cylinders?

Inspired by seahorses, designed by engineers

Seahorse tails predominantly bend ventrally (*7*), an adaptation that allows seahorses to grasp objects in front of their line of sight. However, seahorses can also hyperextend and twist their tails to achieve moderate degrees of bending in the dorsal and lateral directions. Measuring the angle of tail rotation of a seahorse (*Hippocampus kuda*) shows that the tail can bend ~850° ventrally, ~290° dorsally, and ~570° laterally, whereas concurrent bending and twisting allows the tail to bend in the dorsal or lateral directions to nearly the same degree as in pure ventral bending (Fig. 2A). Although tail movement is primarily controlled by the epaxial, hypaxial, and ventral intervertebral

muscles, and constrained by the vertebral connections and segment interlocks (*7, 10, 12*), three additional morphological features present in the skeletal armor also influence tail bending (Fig. 2, B and C): (i) plate size, (ii) plate skew, and (iii) plate overlapping.

Recently, Praet *et al.* (*10*) measured the size and skewing of the bony plates along the length of a seahorse tail. The plates, and consequently the segments, decrease in size from the proximal to distal end of the tail; the resulting finer articulation enables a seahorse to smoothly bend its tail into a logarithmic spiral, with increasing curvature toward the tip (*10*). The plate skew [measured as the angle θ between the midline of the lateral plates and a normal vector to the central spinal column (Fig. 2B)] increases linearly along the length of the tail, so that the lateral plates and the spinal column are nearly orthogonal (θ = ~0°) at the proximal base, while highly skewed (θ = ~45°) at the distal tip (*10*). On the basis of a 2D geometric model of the tail kinematics that reduces the segment joints to their key salient features [two peg-and-socket joints that rotate about a single ball-and-socket joint (fig. S2C)], we identified two primary mechanisms by which the skew of the tail segments produces a bias toward ventral bending over dorsal bending (*36*): (i) The ventral plates are shifted proximally (toward the trunk), whereas the dorsal plates are shifted distally (away from the trunk). This proximal-distal shift changes the position of the pegs within the sockets when the tail is in a straightened position, so that the pegs are centered in the sockets on the ventral side but positioned increasingly closer to the distal-most end of the sockets on the dorsal side (fig. S2). This offset placement limits the range of motion of these joints in dorsal bending while increasing their range in ventral bending. (ii) Distally shifting the dorsal peg-and-socket joints angles them toward the ventral surface, so that they remain tangent to the rotation of the ball-and-socket joint (fig. S2). During ventral bending, this offset inclination

Fig. 2. Seahorse tails predominantly bend ventrally into a logarithmic spiral. (**A**) Plot of the angle of rotation of a seahorse tail (*H. kuda*) from the proximal base (0%) to the distal tip (100%) of the tail in different bending orientations. (**B** and **C**) μCT images of a seahorse tail (*H. reidi*) bent in the ventral direction, illustrating (B) the segment size, plate skew, and ventral-over-dorsal overlapping features, and (C) the ball-and-socket, peg-and-socket, and gliding joints. For visual clarity, the vertebral column is color-masked magenta, and the dorsal and ventral plates are color-masked green and red, respectively. Scale bars, (B) 6 mm; (C) 1.5 mm.
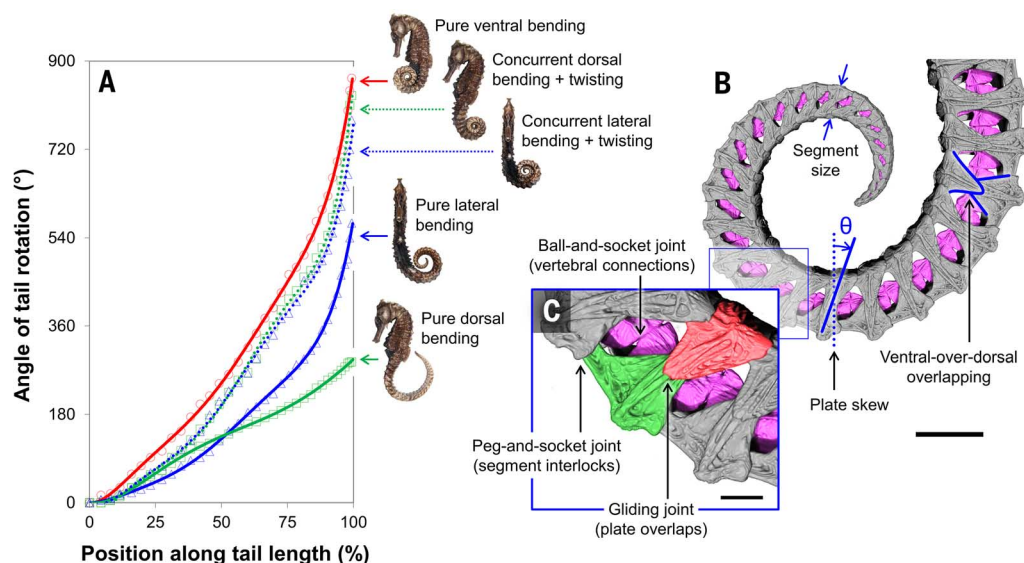
allows the sockets of the distal plates to lie flush with the pegs of the proximal plates on the dorsal side of the tail, instead of "peeling" away at these joints. These two mechanisms both increase the range of motion of the joints in the ventral direction, producing the different sagittal (ventro-dorsal) bend angles observed in Fig. 2A, and improve the alignment of adjacent plates when the tail is in a curled position (Fig. 2B). Further, it was previously noted that the ventral plates always overlap the dorsal plates along the lateral sides of the tail (Fig. 2B) (*13*). This bisymmetric ventral-over-dorsal overlapping arrangement likely promotes ventral tail bending as well (fig. S3), so that the ventral plates have more freedom to slide and rotate, and hence more articulation space, because they are not rigidly attached to the vertebral column [as in the case of the dorsal plates (*7*, *10*)]. Consequently, we hypothesize that eliminating all of these morphological features (plate size, skew, and ventral-over-

dorsal overlaps) from the skeletal armor of a seahorse would allow its tail to bend equally in all directions (ignoring the effect of muscular orientation and connective tissue composition).

To test this hypothesis, we designed and built two artificial skeletons composed of articulating plates and vertebrae arranged into either the natural square prism or a hypothetical cylindrical architecture (*36*). In Fig. 3, microcomputed tomography (µCT) images of three segments of a natural seahorse tail skeleton are juxtaposed with computer-aided design (CAD) models generated to fabricate the 3D-printed prototypes. Pictures of the two prototypes are shown in Fig. 4, highlighting the key structural elements incorporated into the designs. The primary components of the prototypes were scaled up and modified via CAD software (SolidWorks, Dassault Systemes, Waltham, MA) by mimicking the dominant geometries observed from µCT scans of a natural bony plate

and vertebra extracted from the proximal base of a seahorse tail. The computer models were then transferred to a 3D printer (Dimension 1200es, Stratasys, Eden Prairie, MN) to build the artificial plates and vertebrae from a rigid thermoplastic [ABS*plus* (Stratasys), where both the black and blue materials have nearly equivalent properties, with a compressive modulus of ~700 MPa], which were assembled together with compression springs (with a spring constant of 3.36 N/mm) and elastic bands (with a tensile modulus of 15 MPa) (Fig. 4 and table S1). The materials selected for the different components of the prototypes were chosen for convenience (ABS*plus* is available in-house, and the elastic bands are available at local retailers) and showed a similar order-of-magnitude difference in stiffness to that observed in natural fish bones (3.5 to 19.4 GPa) (*37*) and fibrous tissues (1.2 to 10 MPa) (*37*), respectively (table S1). The compression springs were selected to mimic the highly deformable compressive behavior observed in the vertebral struts of a natural seahorse tail [more detail is available in (*13*)]. Because the 3D-printed plates and vertebrae are nearly two orders of magnitude stiffer than the other load-bearing components (springs and elastic bands), they are approximated as fully rigid bodies throughout this study. Thus, the small volumetric differences between the "square" and "circular" plates can be considered negligible because the artificial plates and vertebrae are assumed to experience deformations much smaller than those observed in the overall structures (*36*).

In addition, both prototypes have three primary joints: ball-and-socket, peg-and-socket, and gliding (Fig. 4). These specialized joints provide the prototypes several internal degrees of rotational and translational freedom, allowing the structures to bend, twist, and compress with multiple modes of deformation. In building the prototypes, we excluded two features of the natural tail morphology: the tapering of the segments toward the distal end of the tail, and the

Fig. 3. Computer models designed to mimic the structural organization of a seahorse tail. (**A**) µCT images of a seahorse tail skeleton (*H. kuda*) composed of three segments. (**B**) CAD models of three segments from both the square prism (top) and cylindrical (bottom) artificial skeletons. For visual clarity, the vertebrae are color-masked magenta, and the plates are color-masked yellow, red, green, and blue.
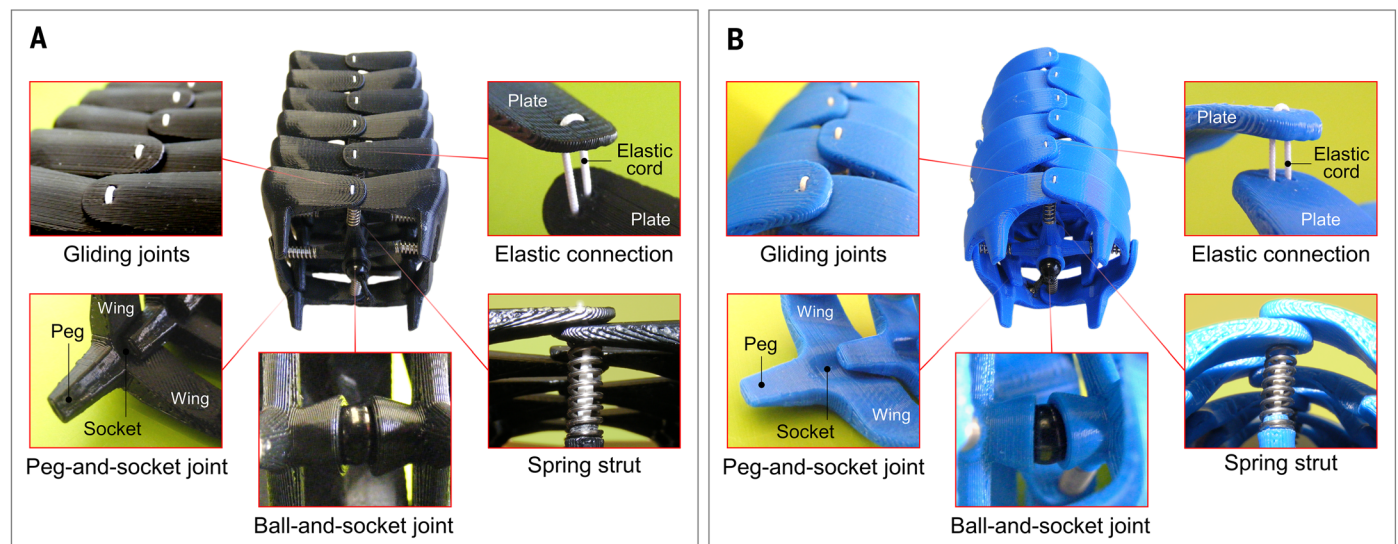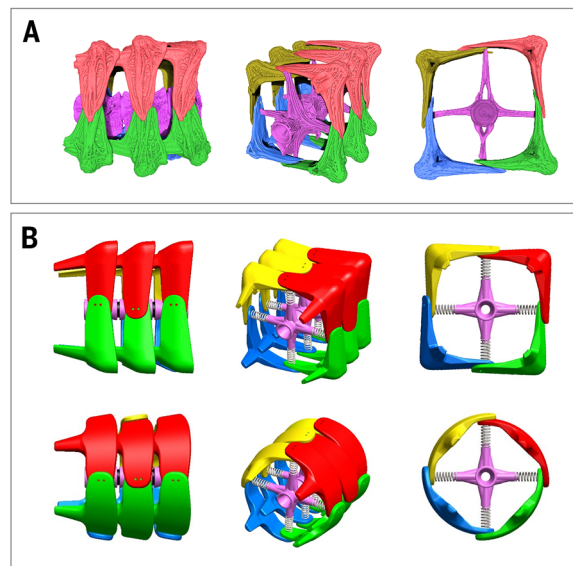

Fig. 4. 3D-printed prototypes designed to mimic the specialized joints of a seahorse tail. (**A** and **B**) Images of the seahorse-inspired prototypes with (A) a square cross section and (B) a circular cross section, depicting the primary structural features and joints incorporated into the designs.

skewing of the segments. This simplification focuses attention on the cross-sectional structure of the plates. The tapering is relatively small [<5% size difference between adjacent links near the proximal base (10)] and allows the tail to grasp small objects (10, 11). Similarly, the change in segment skew is relatively small (~2° to 3° angular difference between adjacent links) and can be accounted for in the tail's bias toward ventral bending (fig. S2). If we had incorporated the segment skew in our model, we anticipate that the resulting offset between the ventral and dorsal plates would have provided a small increase to the crushing resistance of the prototypes (by spreading loads applied to a ventral plate across two dorsal plates, and vice versa). The sequence of intrasegmental connections (plate overlaps) in the prototypes was asymmetrically alternated, so that each alternating segment contains right-over-left overlaps, followed by left-over-right overlaps, along the central midlines, eliminating the regular ventral-over-dorsal overlapping arrangement observed in the natural tail. These design modifications produced a quadraxial symmetry in both prototypes, which as hypothesized facilitates equal bending in all three directions (ventral, dorsal, and lateral).

Bending, twisting, and prehension

The images and CAD models of the square prismatic and cylindrical prototypes shown in Fig. 5 illustrate the limits of bending and twisting, which would likely change by altering the size and skewing of the plates (36). If the plates are bisymmetrically overlapped to more closely mimic the natural configuration, the prototypes favor bending in the "ventral" direction, which was confirmed via experimental observations (fig. S3). Regardless, for comparative purposes,

bending is primarily constrained by the ball-and-socket joints, which are designed to have a maximum rotation angle of ~12° in all directions (ignoring contact between the plates of adjacent segments). According to this constraint and as seen in Fig. 5, C and D, both artificial skeletons can bend ~90° over a span of approximately eight segments, matching that of the proximal base of a seahorse tail (12). When considering contact between the plates of adjacent segments, however, bending is slightly less restricted in the cylindrical prototype than in the square one. As shown in movie S1 and Fig. 5B, the cylindrical prototype can bend to nearly the same degree in any plane off-axis from the vertebral struts (covering all 360° of rotation), whereas the square prototype can bend to a higher degree in the planes oriented parallel and perpendicular to the vertebral struts (in the ventral-dorsal and lateral directions) than those off-axis because contact at the peg-and-socket joints restricts substantial out-of-plane bending.

Twisting, on the other hand, is fully constrained by the physical boundaries imposed at each peg-and-socket joint. As seen in Fig. 5, C and D, the overall twist of the cylindrical prototype is approximately twice that of the square prototype, with an angle of twist of ~30°—as opposed to only ~15°—between adjacent segments. This difference is due to the type of interference that occurs at the peg-and-socket joints. In the cylindrical prototype, interference between the pegs and vertebral struts limits twisting (Fig. 5D). However, in the square prototype, interference between the pegs and sockets occurs well before contact with the vertebral struts, restricting excessive torsion between segments (Fig. 5C).

This natural geometric constraint imposed by the peg-and-socket joints of the square architecture could help preserve musculoskeletal or-

ganization in a seahorse tail because it provides a natural safety factor against possible torsion-induced damage or misalignment from over-twisting. It also provides a passive mechanism for articulatory reorganization upon extensive bending and twisting, which could be advantageous for seahorses to expend less energy when returning their tails to a neutral resting position—slightly bent in the ventral direction (a result of skewing), but straight in the lateral directions [based on observations of a sedated seahorse (11)]. This mechanism observed in the square prototype is demonstrated in movie S1. After deformation, the square prototype always returns to the same linearly aligned resting position, whereas the cylindrical one does not (movie S1, dashed yellow lines). Instead, the lack of interference between its pegs and sockets allows the cylindrical prototype to rest in nearly any bent or twisted position.

In addition, the flat exterior surfaces of the square cross section increase contact area and stability when wrapped around an object, as compared with a circular cross section. When twisted, the square segments transform its exterior shape from a square prism to a solid helix, whereas the exterior profile of the cylindrical prototype remains unchanged under torsion (Fig. 5). The corner edges of this twisted configuration could provide the square architecture more prehensile control—for instance, when a seahorse adjusts its tail-grip on objects. Although not yet observed in nature, this type of twisted shape transformation could also be used to generate a sequence of repetitive forces allowing the articulated square prism to move or wriggle along a variety of substrates in a traveling-coil motion, akin to that demonstrated in modular snake robots (38). Thus, the flattened topology of the square architecture has several functional advantages related to bending
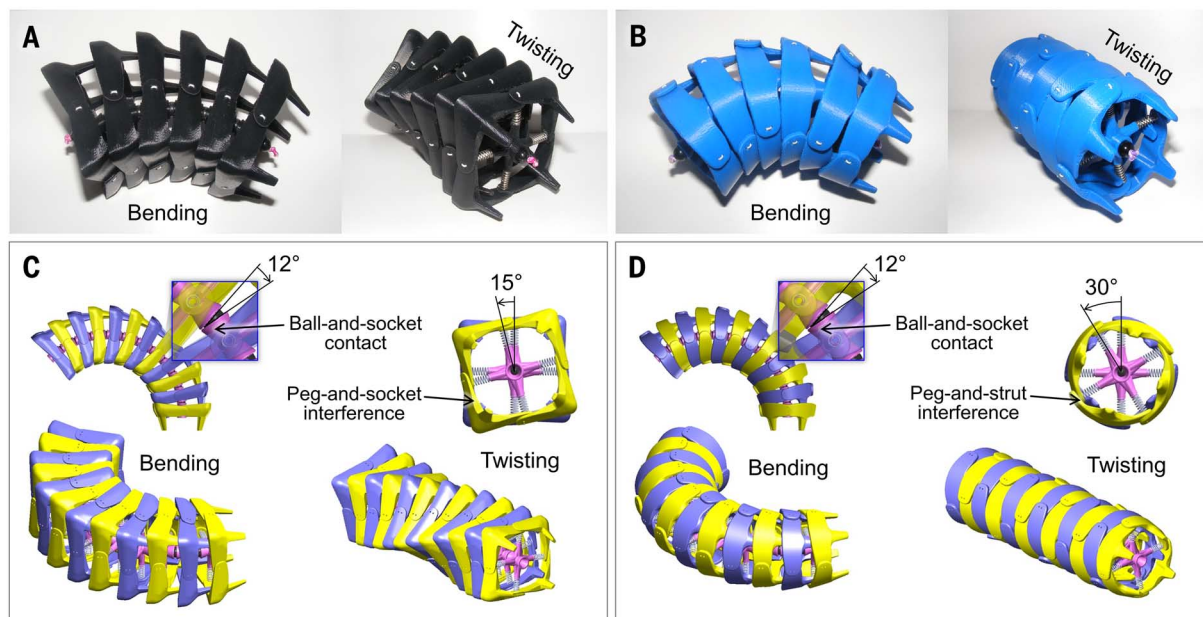

Fig. 5. Bending and twisting performance of the prototypes. (**A** and **B**) Photographs and (**C** and **D**) CAD models of the square (left) and cylindrical (right) prototypes, illustrating the limits of bending and twisting. A more complete comparison of bending and twisting is provided in movie S1.

and twisting that could allow seahorses to grasp objects with more control.

Impact, crushing, and protection

Beyond simple bending and twisting, the artificial skeletons also mimic the protective energy absorption mechanism observed in the highly deformable skeleton of a seahorse tail (Fig. 1C) (13). When subject to impact, such as the compressive force of a rubber mallet (Fig. 6, A and B, and movie S2), the gliding joints and vertebral struts of the prototypes allow the plates to slide past one another (or rotate, in the case of the cylindrical prototype), absorb the impact when compressed, and subsequently protect the internal vertebral column from brittle fracture. As seen in movie S2, the square prototype returns to its linearly aligned resting position after impact, whereas the cylindrical one remains partially deformed and misaligned (even when the first and last segments are taped to the supporting plane). Although there is some plate rotation observed in segments of the square prototype adjacent to the primary impact point, the plates directly impacted mainly show linear sliding. In contrast, the cylindrical prototype experiences extreme reverberations transferred through its vertebral column upon impact.

Under ideal uniaxial compression, the artificial square skeleton deforms linearly by sliding its plates past one another with approximately one degree of translational freedom (Fig. 6C). The cylindrical skeleton, on the other hand, deforms with approximately two degrees of freedom (one translational and one rotational), resulting in linear compressive and expansive deformations parallel and perpendicular to the applied force, respectively, leading to a cross-sectional shape change from circular to elliptical (Fig. 6D). When subject to biaxial compression, both the square and cylindrical skeletons deform bisymmetrically. However, the square architecture can contract its outer profile by ~30% (apparent reduction of cross-sectional area) (Fig. 6E), whereas the con-

tracted profile of the circular one is reduced by only ~10% (Fig. 6F). The difference in the compressive behavior of the two cross sections is a result of the different plate geometries. The flattened morphology of the plates in the square skeleton allows linear sliding to occur without disrupting its outer profile, whereas the rounded plates in the circular one interfere and create large gaps or rotation between adjacent plates. Although seahorses are unlikely to experience biaxial compression in nature, this type of contractive deformation is of interest for comparison to similar behaviors that occur in other organisms and mimetic technologies (such as peristalsis and burrowing) (29, 39).

In contrast, the mechanical behavior of square and circular solid rings (without gliding joints) subject to transverse uniaxial and transverse biaxial compression is shown to the right of each image in Fig. 6, C to F (where the resulting deformations are exaggerated for clarity). Derivations of analytical expressions describing these shapes, shown in figs. S4 and S5, are included in (36). Under transverse uniaxial compression, thin solid rings buckle at their midsections when the internal stresses accumulated in the rings exceed the yield stress of the ring material, forming plastic yield hinges where permanent deformation begins (Fig. 6, C and D) (40–42). By analogy, rings subject to biaxial compression also form plastic hinges at the locations of maximum stress where yielding begins (Fig. 6, E and F, red dots). The onset of these plastic hinges occurs when the absolute internal moment ($|M'|$) accumulated in the rings exceeds the plastic yield moment of the ring material (M_0), which are both approximated as a moments per unit length (b) for square and circular rings (40–42)

$$|M'| > M_0 = \frac{\sigma_0 t^2}{4} \qquad (1)$$

where σ_0 is the yield strength of the ring material and t is the ring thickness, assuming plane strain

so that the length of the ring (or tube) is much greater than its thickness ($b \gg t$). The locations of the plastic hinge zones are shown in fig. S6, determined analytically in (36), and verified experimentally for unilateral compression in Fig. 7, F and G. Accordingly, the plastic hinge zones are assumed to be controlled by the tensile properties of the ring material (36), as noted by the observation of crazing that occurs because of tension at these locations (Fig. 7, F and G).

In seahorses, the problem of plastic hinge formation and subsequent buckling of their ring-like armor is solved by having gliding joints at the locations of these plastic yield zones. The highly deformable vertebral struts and connective tissues—which exhibit local buckling and tensile stretching rather than brittle fracture and tearing, respectively—let the bony plates slide past one another to absorb energy when compressed (Fig. 1C) (13). This mechanism allows the tail (of a deceased seahorse) to deform to nearly 50% before the onset of permanent failure in the vertebral column (13). Exploiting this mechanism, we mimicked the vertebral struts and connective collagen with compression springs and elastic bands, respectively. When crushed, the springs compress—and bend or buckle, if necessary—to allow the artificial plates to slide and/or rotate (Fig. 7, B and C, and movie S2). As the plates move, the elastic bands connecting adjacent plates stretch, a response observed to be independent of the displacement rate (fig. S1). These two mechanisms bear the load in the two prototypes when subjected to transverse compression, exhibiting up to ~17% linear elastic deformation in the square prototype and ~12% in the circular one (Fig. 7A).

Upon further analysis, it is apparent that the square prototype under compression is stiffer, stronger, and more resilient than the circular one (Fig. 7A). The load carried by the square skeleton composed of six segments ($P_\square$) can be expressed as a function of its linear displacement (δ), treating
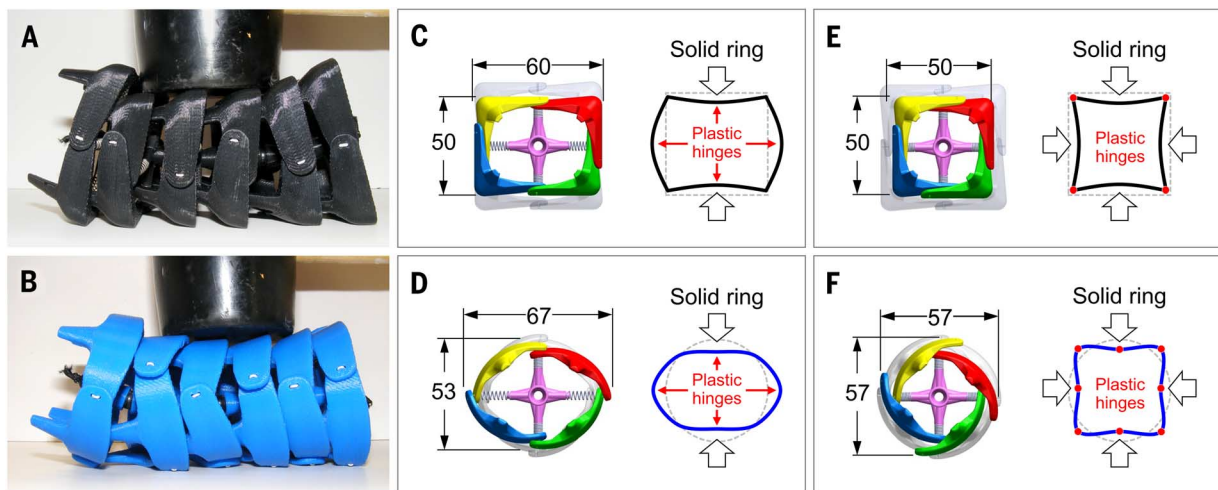
Fig. 6. Impact and crushing performance of the prototypes. (**A** and **B**) Images of the square (top) and cylindrical (bottom) prototypes subject to impact by a rubber mallet. A high-speed comparison of the prototypes subject to impact is provided in movie S2. (**C** to **F**) CAD models of the artificial skeletons (left) and representative schematics of solid rings (right) subject to transverse uniaxial and transverse biaxial compression, showing the locations of primary and secondary plastic hinges with red arrows in (C) and (D) and red dots in (E) and (F). All units are in millimeters.

the 3D-printed plates and vertebrae as rigid bodies, so that

$$P_{\square} = 6 \times \left(\frac{1}{2} k_s + 6 \frac{\beta_s}{L_s^3} + 4 k_e \right) \delta \qquad (2)$$

where $k_s = 3.36$ N/mm is the compressive spring constant, $\beta_s = 2.0 \times 10^{-4}$ N · m² is the flexural rigidity of the springs, $L_s = 10$ mm is the active length of the springs in bending, and $k_e = 0.20$ N/mm is the effective spring constant of the elastic bands (fig. S7) (36). Similarly, the load carried by the cylindrical skeleton composed of six segments ($P_{\circ}$) can also be expressed as a function of its linear displacement (δ)

$$P_{\circ} = 6 \times \left(\frac{1}{4} k_s (\delta + \gamma) \frac{a}{c} + k_e \zeta \frac{b}{c} \right) \qquad (3)$$

where a, b, and c are geometric relations dependent on the radius of plate curvature ($R = 30$ mm),

and ($\delta + \gamma$) and ζ are the vertical and lateral displacements of the cylindrical skeleton, respectively (36). To account for the six segments in each prototype compressed in parallel (Fig. 7, A to C), Eqs. 2 and 3 are both multiplied by a factor of six.

A plot of these equations (dashed lines) is shown in Fig. 7A, compared with experimental data obtained via compression tests of the prototypes (Fig. 7A, solid lines), each composed of six segments. The analytical expressions ($P_{\square}$ and $P_{\circ}$) both underpredict the actual load carried by the two prototypes. This discrepancy is because for simplicity, the frictional forces between adjacent plates and the prototypes with the compression platforms were neglected throughout the analyses. Regardless, the force-displacement curves in Fig. 7A illustrate the fact that the square prototype is ~3× stiffer, ~4× stronger, and ~1.5× more elastic than its circular counterpart. At the point

at which the load drops off because of a minor collapse of the structures (Fig. 7A, red arrows), the springs are almost fully compressed and begin to buckle before disjoining from the plates. At this point, the prototypes no longer exhibit a purely elastic behavior. Although recoverable, the structural integrity of the internal struts (compression springs) becomes compromised, making these points the elastic limits of the prototypes.

In juxtaposition, the compressive behavior of square and circular solid rings composed of the same 3D-print material (ABS*plus*) used to fabricate the seahorse-inspired prototypes is shown in Fig. 7, D to G. As seen in the plot of Fig. 7D, the solid rings are seemingly much stronger and stiffer than the articulated prototypes (Fig. 7A). Also, the square ring is an order of magnitude stiffer and stronger than the circular one of equal height but undergoes much less transverse
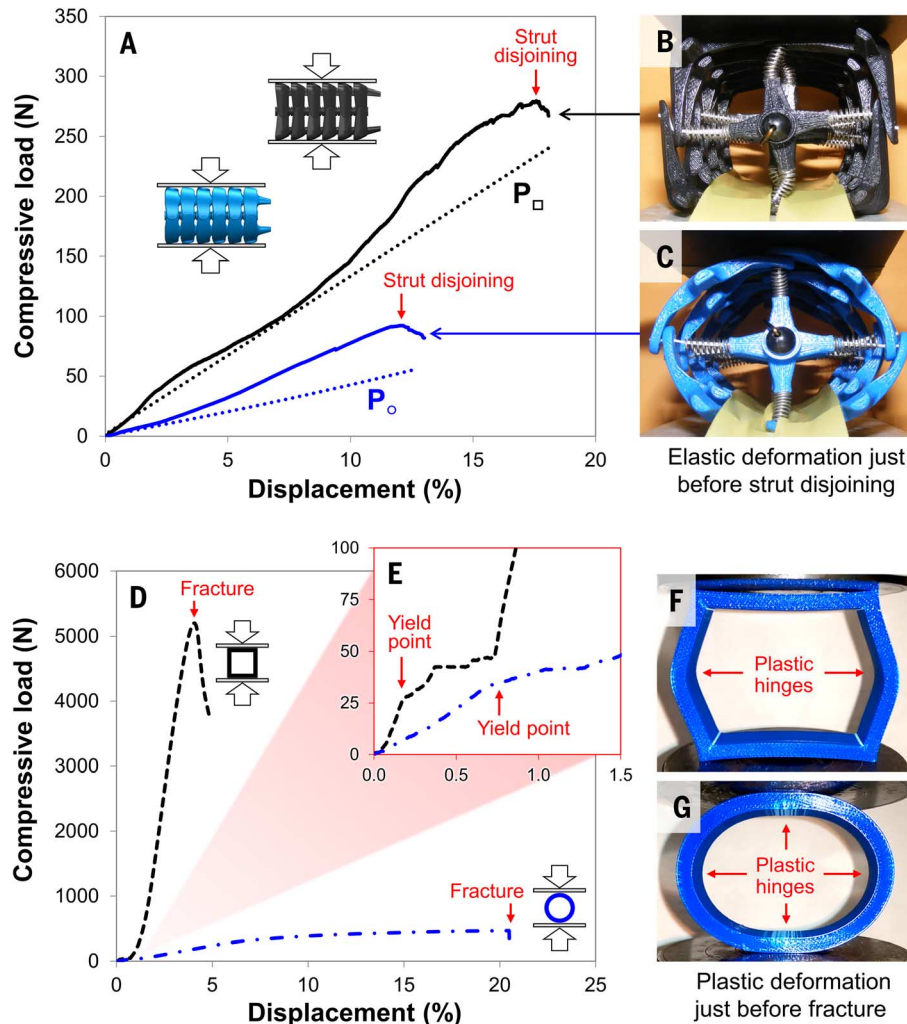
Fig. 7. Unilateral compression of the prototypes and solid rings. (**A**) Plot of the compressive load versus normalized displacement for the square (black) and cylindrical (blue) prototypes subject to unilateral compression between two rigid plates. The solid lines correspond to direct experimental measurements, and the dashed lines correspond to theoretical predictions of load and displacement from Eqs. 2 and 3. (**B** and **C**) Images of the square (top) and cylindrical (bottom) prototypes subject to transverse unilateral compression just before strut disjoining, correspond-ing to the red arrows in (A). (**D**) Plot of the compressive load versus normalized displacement for a square (black dashed line) and circular (blue dashed-dotted line) solid ring subject to transverse unilateral compression between two rigid plates. (**E**) Magnified plot illustrating the elastic region and yield points of the load-displacement curves in (D). (**F** and **G**) Images of the square (top) and circular (bottom) solid rings subject to transverse unilateral compression just before fracture, corresponding to the red ar-rows in (D).

displacement before failure. Nevertheless, these behaviors (Fig. 7D) occur when the solid rings are loaded within the plastic regime, beyond the initial formation of plastic hinges [which begin at the yield point, the first sign of permanent deformation (Fig. 7E)]. Considering this, it is more useful to compare the elastic behavior of the articulated prototypes (Fig. 7A) with that of the solid rings (Fig. 7E). As indicated in Fig. 7E, both solid rings begin to yield at ~30 N, whereas the square ring elastically deforms only ~0.1% and the circular ring ~0.8%. When crushed, the square prototype outperforms the cylindrical one, exhibiting a higher strength and stiffness, analogous to the mechanical performance of the solid rings. The square prototype is also more resilient than the cylindrical one, absorbing more energy before failure (strut disjoining), a trend not observed in the solid rings (solid rings with circular cross sections are more resilient than square ones). Thus, the specialized joints present in the prototypes (and likewise in a seahorse tail) provide the structures an enhanced range of motion, not only in bending and twisting but also in resistance to crushing.

Concluding remarks

The highly articulated bony plates that surround the central vertebral axis of a seahorse tail actively facilitate bending and twisting as well as resist vertebral fracture from impact and crushing. To explore why the bony plates are arranged into cross-sectional squares rather than circles, we analyzed the mechanics of 3D-printed models that mimic the natural (square prism) and hypothetical (cylindrical) architectures of a seahorse tail skeleton. Physical manipulation of the two prototypes revealed that the square architecture possesses several mechanical advantages over its circular counterpart in bending, twisting, and resistance to crushing. The enhanced performance realized in the square architecture provides insight into the way in which seahorses may benefit from having prehensile tails composed of armored plates organized into square prisms, rather than cylinders. This study demonstrates that engineering designs are convenient means to answer elusive biological questions when biological data are nonexistent or difficult to obtain. In addition, understanding the role of mechanics in these biologically inspired designs may help engineers to develop seahorse-inspired technologies for a variety of applications in robotics, defense systems, or biomedicine.

REFERENCES AND NOTES

1. M. A. Ashley-Ross, Mechanical properties of the dorsal fin muscle of seahorse (*Hippocampus*) and pipefish (*Syngnathus*). *J. Exp. Zool.* **293**, 561–577 (2002). doi: 10.1002/jez.10183; pmid: 12410605
2. D. Kleiber, L. K. Blight, I. R. Caldwell, A. C. J. Vincent, The importance of seahorses and pipefishes in the diet of marine animals. *Rev. Fish Biol. Fish.* **21**, 205–223 (2011). doi: 10.1007/s11160-010-9167-5
3. S. Van Wassenbergh, G. Roos, L. Ferry, An adaptive explanation for the horse-like shape of seahorses. *Nat. Commun.* **2**, 164 (2011). doi: 10.1038/ncomms1168; pmid: 21266964
4. S. A. Lourie, S. J. Foster, E. W. T. Cooper, A. C. J. Vincent, *A Guide to the Identification of Seahorses* (Project Seahorse, Washington, DC, 2004).
5. R. Anthony, L. Chevroton, Considérations sur les attitudes et la locomotion de l'hippocampe, étude chronophotographique. *Arch. Zool. Exp. Gen.* **51**, 11–22 (1913).
6. H. Peters, [On the ecologic physiology of the sea horse (*Hippocampus brevirostris*)]. *Z. Vgl. Physiol.* **33**, 207–265 (1951). doi: 10.1007/BF00395584; pmid: 24541372
7. M. E. Hale, Functional morphology of ventral tail bending and prehensile abilities of the seahorse, *Hippocampus kuda*. *J. Morphol.* **227**, 51–65 (1996). doi: 10.1002/(SICI)1097-4687(199601)227:1<51::AID-JMOR4>3.0.CO;2-S
8. E. Bruner, V. Bartolino, Morphological variation in the seahorse vertebral system. *Int. J. Morphol.* **26**, 247–262 (2008). doi: 10.4067/S0717-95022008000200002
9. S. Van Cauter *et al.*, Virtual design from nature: Kinematic modeling of the seahorse tail. *SIMULIA Customer Conference*, 770–783 (2010).
10. T. Praet *et al.*, Inspiration from nature: Dynamic modelling of the musculoskeletal structure of the seahorse tail. *Int. J. Numer. Method Biomed. Eng.* **28**, 1028–1042 (2012). doi: 10.1002/cnm.2499; pmid: 23027633
11. T. Praet, thesis, Ghent University, Ghent, Belgium (2013).
12. C. Neutens *et al.*, Grasping convergent evolution in syngnathids: A unique tale of tails. *J. Anat.* **224**, 710–723 (2014). doi: 10.1111/joa.12181; pmid: 24697519
13. M. M. Porter, E. Novitskaya, A. B. Castro-Ceseña, M. A. Meyers, J. McKittrick, Highly deformable bones: Unusual deformation mechanisms of seahorse armor. *Acta Biomater.* **9**, 6763–6770 (2013). doi: 10.1016/j.actbio.2013.02.045; pmid: 23470547
14. G. C. Hickman, The mammalian tail: A review of functions. *Mammal Rev.* **9**, 143–157 (1979). doi: 10.1111/j.1365-2907.1979.tb00252.x
15. L. Emmons, A. H. Gentry, Tropical forest structure and the distribution of gliding and prehensile-tailed vertebrates. *Am. Nat.* **121**, 513–524 (1983). doi: 10.1086/284079
16. K. C. Zippel, R. E. Glor, J. E. Bertram, On caudal prehensility and phylogenetic constraint in lizards: The influence of ancestral anatomy on function in Corucia and Furcifer. *J. Morphol.* **239**, 143–155 (1999). doi: 10.1002/(SICI)1097-4687(199902)239:2<143::AID-JMOR3>3.0.CO;2-O
17. J. M. Organ, V. B. Deleon, Q. Wang, T. D. Smith, From head to tail: New models and approaches in primate functional anatomy and biomechanics. *Anat. Rec.* **293**, 544–548 (2010). doi: 10.1002/ar.21132; pmid: 20235310
18. J.-Y. Sire, A. Huysseune, Formation of dermal skeletal and dental tissues in fish: A comparative and evolutionary approach. *Biol. Rev. Camb. Philos. Soc.* **78**, 219–249 (2003). doi: 10.1017/S1464793102006073; pmid: 12803422
19. P. C. J. Donoghue, I. J. Sansom, J. P. Downs, Early evolution of vertebrate skeletal tissues and cellular interactions, and the canalization of skeletal development. *J. Exp. Zoolog. B Mol. Dev. Evol.* **306**, 278–294 (2006). doi: 10.1002/jez.b.21090; pmid: 16555304
20. J. S. Nelson, *Fishes of the World* (Wiley, New York, 2006).
21. J. Long, M. Hale, M. Mchenry, M. Westneat, Functions of fish skin: Flexural stiffness and steady swimming of longnose gar, Lepisosteus osseus. *J. Exp. Biol.* **199**, 2139–2151 (1996). pmid: 9320050
22. B. J. F. Bruet, J. Song, M. C. Boyce, C. Ortiz, Materials design principles of ancient fish armour. *Nat. Mater.* **7**, 748–756 (2008). doi: 10.1038/nmat2231; pmid: 18660814
23. F. J. Vernerey, F. Barthelat, Skin and scales of teleost fish: Simple structure but high performance and multiple functions. *J. Mech. Phys. Solids* **68**, 66–76 (2014). doi: 10.1016/j.jmps.2014.01.005
24. T. Libby *et al.*, Tail-assisted pitch control in lizards, robots and dinosaurs. *Nature* **481**, 181–184 (2012). doi: 10.1038/nature10710; pmid: 22217942
25. A. J. Ijspeert, Biorobotics: Using robots to emulate and investigate agile locomotion. *Science* **346**, 196–203 (2014). doi: 10.1126/science.1254486; pmid: 25301621
26. K. Xu, N. Simaan, Actuation compensation for flexible surgical snake-like robots with redundant remote actuation. *IEEE International Conference on Robotics and Automation*, 4148–4154 (2006).
27. C. Wright *et al.*, Design of a modular snake robot. *IEEE/RSJ International Conference on Intelligent Robots and Systems* (2007).
28. J. K. Hopkins, B. W. Spranklin, S. K. Gupta, A survey of snake-inspired robot designs. *Bioinspir. Biomim.* **4**, 021001 (2009). doi: 10.1088/1748-3182/4/2/021001; pmid: 19158415
29. K. A. Daltorio *et al.*, Efficient worm-like locomotion: Slip and control of soft-bodied peristaltic robots. *Bioinspir. Biomim.* **8**, 035003 (2013). doi: 10.1088/1748-3182/8/3/035003; pmid: 23981561
30. C. D. Onal, R. J. Wood, D. Rus, An origami-inspired approach to worm robots. *IEEE/ASME Trans. Mechatron.* **18**, 430–438 (2013). doi: 10.1109/TMECH.2012.2210239
31. I. D. Walker *et al.*, Continuum robot arms inspired by cephalopods. *Conference on Unmanned Ground Vehicle Technology VII* **5804**, 303–314 (2005).
32. R. F. Shepherd *et al.*, Multigait soft robot. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20400–20403 (2011). doi: 10.1073/pnas.1116564108; pmid: 22123978
33. I. D. Walker, M. W. Hannan, A novel 'elephant's trunk' robot. *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 410–415 (1999).
34. J. S. Mehling, M. A. Diftler, M. Chu, M. Valvo, A minimally invasive tendril robot for in-space inspection. *IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics* **1–3**, 423–428 (2006).
35. A. Sadeghi, A. Tonazzini, L. Popova, B. Mazzolai, Robotic mechanism for soil penetration inspired by plant root. *IEEE International Conference on Robotics and Automation*, 3457-3462 (2013).
36. Materials and methods are available as supplementary materials on Science Online.
37. A. P. Summers, J. H. Long, Skin and bones, sinew and gristle: The mechanical behavior of fish skeletal tissues. *Fish Physiol.* **23**, 141 (2006).
38. M. Tesch *et al.*, Parameterized and scripted gaits for modular snake robots. *Adv. Robot.* **23**, 1131–1158 (2009). doi: 10.1163/156855309X452566
39. K. J. Quillin, Kinematic scaling of locomotion by hydrostatic animals: Ontogeny of peristaltic crawling by the earthworm lumbricus terrestris. *J. Exp. Biol.* **202**, 661–674 (1999). pmid: 10021320
40. C. Hwang, Plastic collapse of thin rings. *J. Aeronaut. Sci.* **20**, 819–826 (1953). doi: 10.2514/8.2868
41. J. A. DeRuntz, P. Hodge, Crushing of a tube between rigid plates. *J. Appl. Mech.* **30**, 391–395 (1963). doi: 10.1115/1.3636567
42. D. K. Sinha, N. R. Chitkara, Plastic collapse of square rings. *Int. J. Solids Struct.* **18**, 819–826 (1982). doi: 10.1016/0020-7683(82)90038-5

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6243/aaa6683/suppl/DC1
Materials and Methods
Figs. S1 to S8
Table S1
References (*43–52*)
Movies S1 and S2

REPORTS

PHYSICS

Observation of phononic helical edge states in a mechanical topological insulator

Roman Süsstrunk and Sebastian D. Huber*

A topological insulator, as originally proposed for electrons governed by quantum mechanics, is characterized by a dichotomy between the interior and the edge of a finite system: The bulk has an energy gap, and the edges sustain excitations traversing this gap. However, it has remained an open question whether the same physics can be observed for systems obeying Newton's equations of motion. We conducted experiments to characterize the collective behavior of mechanical oscillators exhibiting the phenomenology of the quantum spin Hall effect. The phononic edge modes are shown to be helical, and we demonstrate their topological protection via the stability of the edge states against imperfections. Our results may enable the design of topological acoustic metamaterials that can capitalize on the stability of the surface phonons as reliable wave guides.

The experimental hallmarks of the quantum spin Hall effect (QSHE) in semiconductor quantum wells (*1–5*) are two counterpropagating edge modes that differ by their spin degree of freedom. As long as time reversal symmetry is preserved, these two modes are independent and do not scatter into each other (*6, 7*). Much of the interest in condensed matter research involving topological states is driven by the use of these protected edge modes for technological applications such as spintronics (*8, 9*), magnetic devices (*10*), or quantum information processing (*11*). The transfer of the phenomenology of the QSHE from the quan-

tum mechanical realm to classical mechanical systems is therefore of fundamental interest, and its accomplishment would offer a gateway to new design principles in mechanical metamaterials.

Several key problems in the engineering of acoustic materials can potentially be addressed by capitalizing on the physics of the QSHE. The edge channels are robust counterparts to the well-known whispering gallery modes (*12, 13*). Any application that requires energy to be confined to the surfaces of some device—for example, vibration insulators—can potentially make use of such edge states. In contrast to the whispering gallery modes, which are extremely sensitive to the shape of the surface (*14*), the topological edge modes are stable under a variety of perturbations. Moreover, because of the stability of these modes, scattering-free phonon waveguides of almost arbitrary shape can be realized. This in turn enables the engineering of robust acoustic delay lines (*15–17*), which are useful for purposes such as acoustic lensing (*18*).

How a mechanical system described by Newton's equations can reproduce the phenomenology of a quantum mechanical model such as the QSHE has remained an open question (*19–23*). We sought to derive a mapping of the physics of the QSHE to a general mechanical system and to provide a specific experimental verification of our proposal. With a view to potential applications, it is essential to demonstrate that the physics of the

Institute for Theoretical Physics, ETH Zürich, 8093 Zürich, Switzerland.
*Corresponding author. E-mail: sebastian.huber@phys.ethz.ch
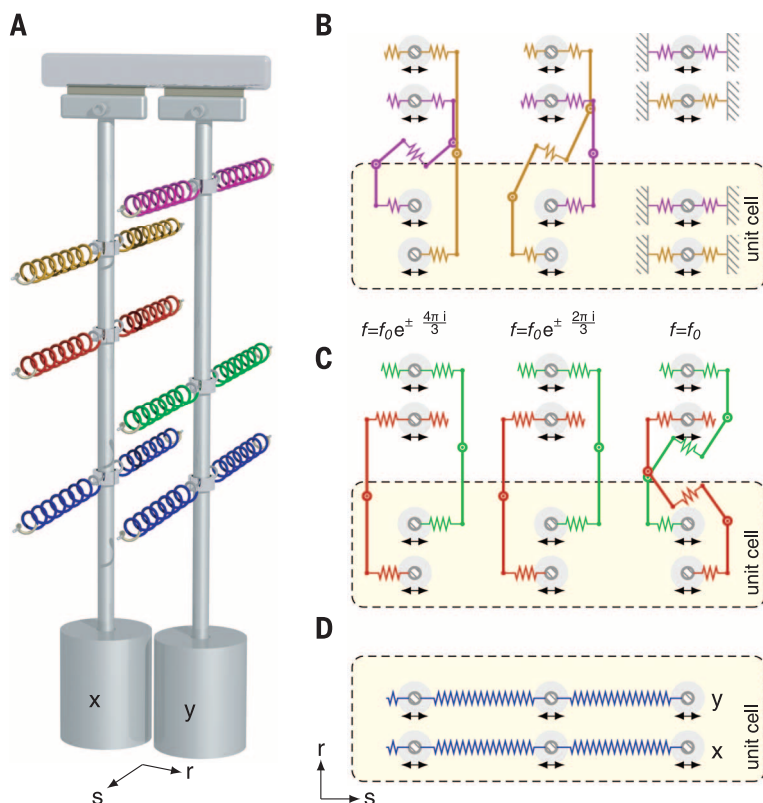
Fig. 1. Setup. (**A**) Illustration of two one-dimensional pendula, *x* and *y*, making up one effective site of our lattice model. (**B**) Schematic top view of the couplings perpendicular to the direction of motion of the pendula. The top two layers of springs (magenta and brown) implement the cross-coupling between *x* and *y* pendula. One lever arm yields a negative coupling, whereas two lever arms give rise to a positive coupling. The spring constants are chosen to give rise to the desired effective coupling strength Im(*f*). Note that there are three sites in one unit cell owing to the three different phases on the transverse couplings. (**C**) The next two layers of springs (green and red) implement the *x-x* and *y-y* couplings Re(*f*) in the transverse direction. (**D**) The bottom springs (blue) couple *x-x* or *y-y* springs with strength f_0 in the longitudinal direction.

In figure: $f=f_0 e^{\pm \frac{4\pi i}{3}}$ $f=f_0 e^{\pm \frac{2\pi i}{3}}$ $f=f_0$
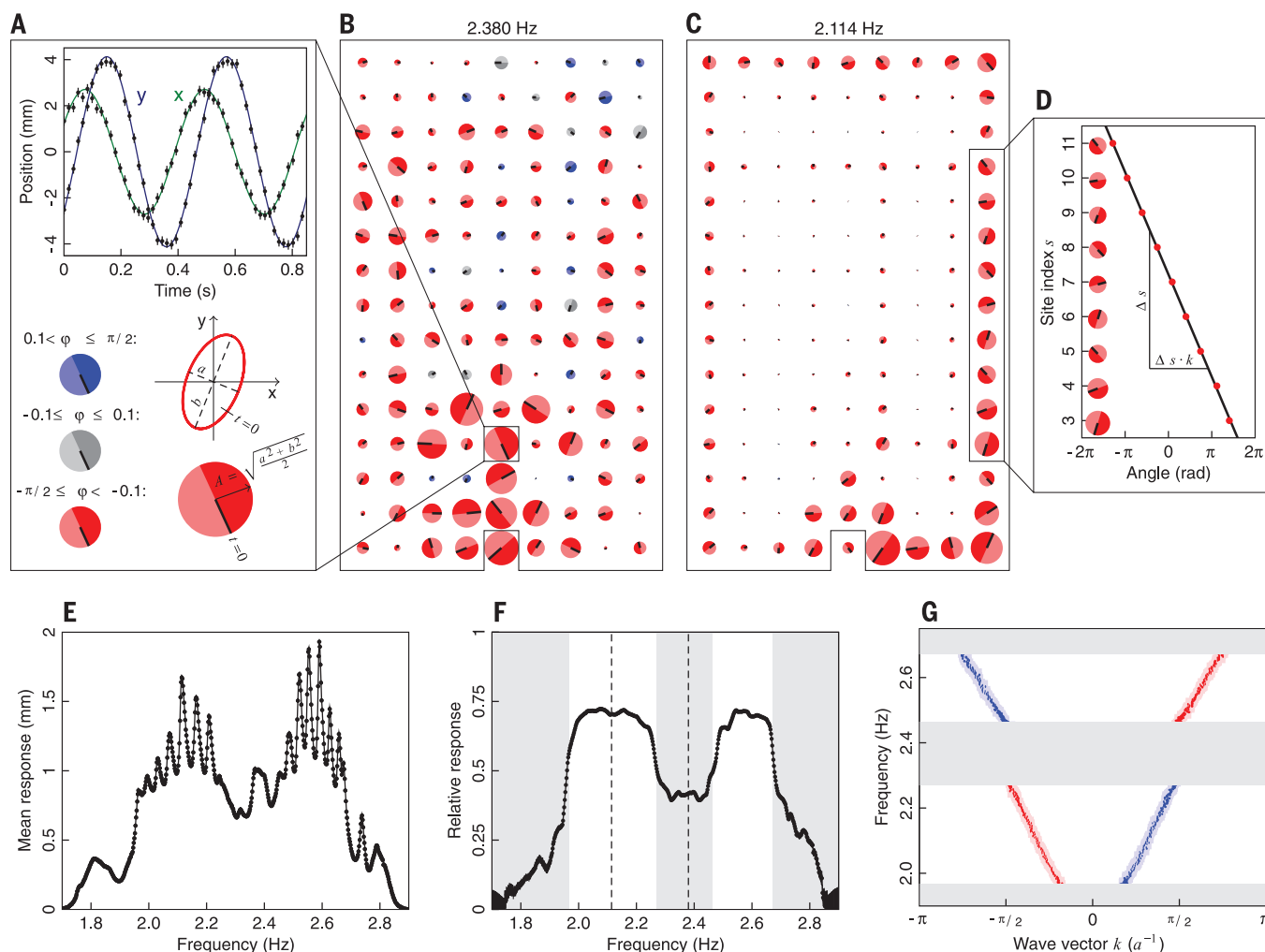
Fig. 2. Helical edge mode dispersion. (**A**) Time traces of two pendula. These can be interpreted as a two-dimensional trajectory, as shown. Steady states are displayed by colored disks representing their polarization φ. The radius of the circle corresponds to the mean deflection *A*; the black line indicates the position of the pendulum at a given fixed time. (**B** and **C**) Measured steady states at two different frequencies, 2.380 Hz and 2.114 Hz. The circles are normalized to the strongest deflection. The excluded site at the center of the bottom row is excited with left circular polarization. (**D**) For edge-dominated modes, the evolution of the angle indicated by the black line defines a wave vector *k*. (**E**) Mean response of the system (average *A*) as a function of the excitation frequency showing an overall bandwidth of ~1.2 Hz. Error bars (smaller than the symbol size) are explained in (*28*). (**F**) The mean response of the edge relative to the bulk. In the frequency ranges shaded in gray, the bulk response dominates. The frequencies marked with dashed lines correspond to (B) and (C). (**G**) The frequency of the edge modes as a function of the wave vector. The bulk bands are shaded in gray. The color labels the polarization as before, establishing the helical nature of the edge excitations. The shaded region marks a 0.04-Hz band corresponding to the loss-induced broadening of the eigenfrequencies.

Fig. 3. Beam splitter. (**A**) Geometry of the beam splitter. An arbitrary polarization is injected at the excluded edge site at 2.123 Hz. The relative weights *A* on the two shaded boxes L and R define a splitting ratio. (**B**) Splitting ratio on the Poincaré sphere. The north and south poles of the sphere respectively correspond to right and left circular polarizations of the drive. The color map is an interpolation based on 7000 data points. The maximal splitting is not reached at the poles, indicating the presence of disorder. The dots mark measurements along a great circle through the points of maximal splitting. (**C**) Splitting ratio along the great circle shown in (B). The black line marks a cosine expected for an optimal beam splitter. The maximally reached splitting is 99.80 ± 0.04%.
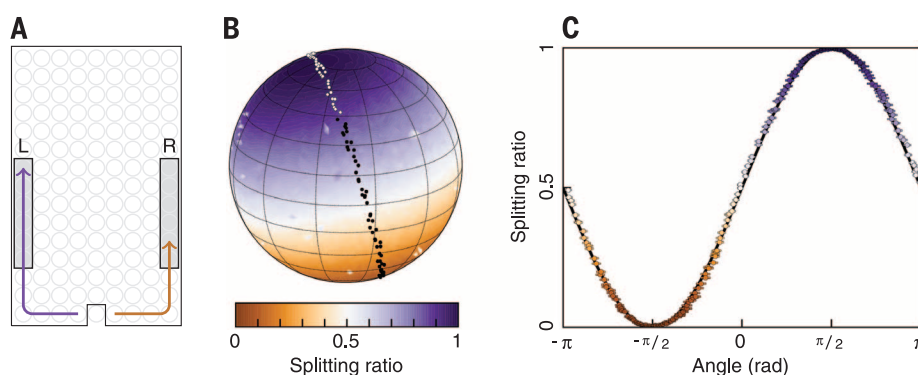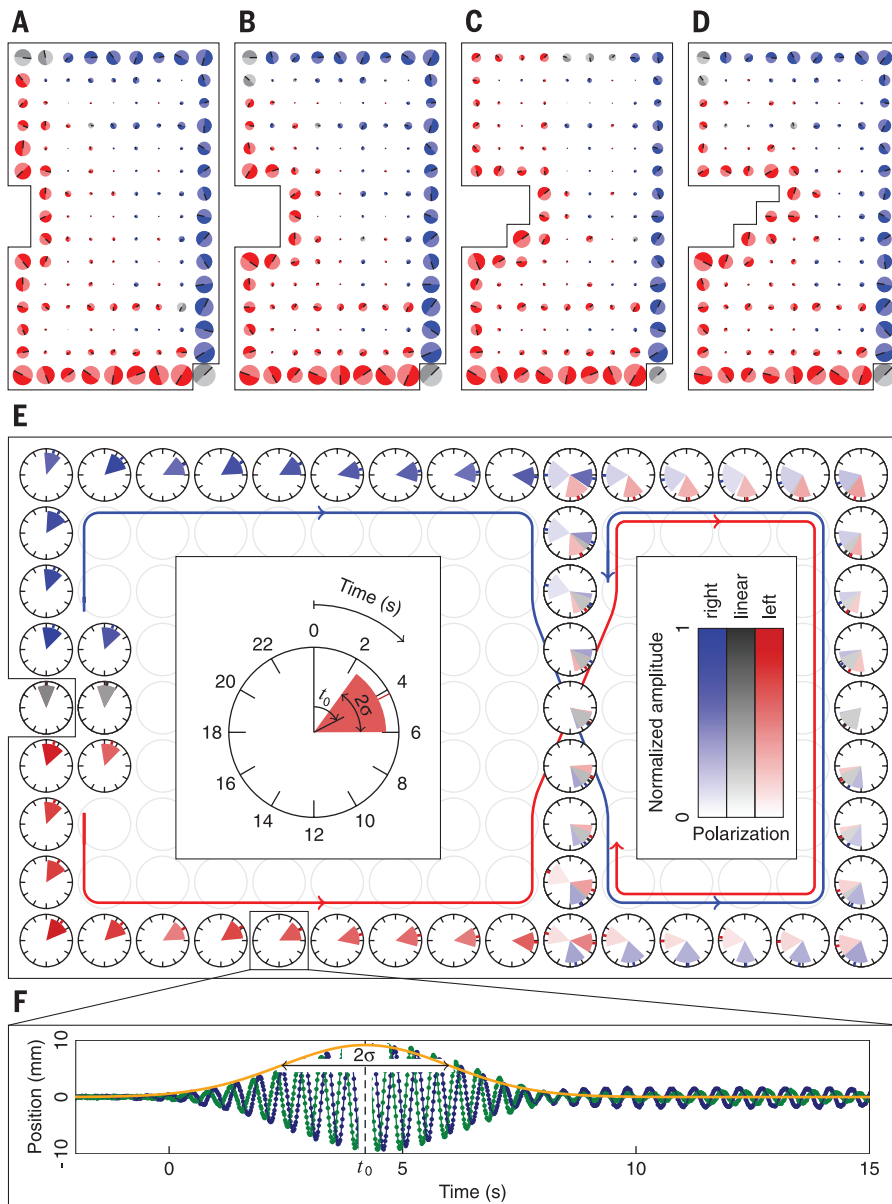
Fig. 4. Topological protection. (**A** to **D**) Steady states with a sequence of removed sites illustrating the stability of the edge states against boundary roughness. The bottom right site is driven with linear polarization at a frequency in the upper gap. (**E**) Wave packets launched at the excluded edge site with linear polarization at a frequency in the lower gap. Each clock represents an edge site on which the colored wedges are centered at the time t_0 when the wave packet traverses the site. The angle of the wedge indicates the width σ of the wave packet as illustrated in (F). The color represents its polarization. From these clocks, one can read off the propagation of the wave packets throughout the system. (**F**) Analysis of the wave packet at a given site from which the passing time t_0 and the width of the wave packet σ are extracted. The green and dark blue trajectories represent the measured motion of the x pendulum and y pendulum, respectively.

noninteracting QSHE can be observed in a real mechanical system, which necessarily suffers from losses, disorder, and nonlinearities inherent to coupled mechanical oscillators.

A quantum mechanical lattice problem is described by a Schrödinger equation of the form

$$i\hbar \dot\psi_i^\alpha = \mathcal{H}_{ij}^{\alpha\beta}\psi_j^\beta \qquad (1)$$

where ψ_i^α are the wave function amplitudes for an electron with spin α on lattice site i, and $\mathcal{H}_{ij}^{\alpha\beta}$ is

the hermitian Hamiltonian matrix describing the QSHE (24, 25). On the other hand, the dynamics of a collection of undamped classical harmonic oscillators is described by Newton's equation of motion

$$\ddot x_i = -\mathcal{D}_{ij} x_j \qquad (2)$$

where x_i are the coordinates of N pendula and $\mathcal{D}_{ij}$ is the real, symmetric, and positive semidefinite dynamical matrix containing the couplings between them. For either system, the existence

and properties of edge modes are features of the eigenstates of $\mathcal{H}$ or $\mathcal{D}$ alone and depend neither on the interpretation of ψ_i^α versus x_i nor on the nature of the dynamics (i.e., $i\partial_t$ versus ∂_t^2). Hence, we mean to design a dynamical matrix $\mathcal{D}$ incorporating the properties of the QSHE.

We start from two independent copies of the Hofstadter model (26) on the square lattice with flux $\Phi = \pm 2\pi/3$ per plaquette,

$$\mathcal{H} = \sum_{\alpha=\pm} \mathcal{H}_{\alpha,\Phi}$$

$$\mathcal{H}_{\alpha,\Phi} = f_0 \sum_{r,s} |r, s \pm 1, \alpha\rangle\langle r, s, \alpha| + |r, s, \alpha\rangle$$

$$\times \langle r \pm 1, s, \alpha| \exp(\pm i\alpha\phi_s) \qquad (3)$$

or, in matrix notation,

$$\mathcal{H} = \begin{pmatrix} \mathcal{H}_\Phi & 0 \\ 0 & \mathcal{H}_\Phi^* \end{pmatrix} \qquad (4)$$

Here, (r, s) denotes the location on the lattice of size $L_r \times L_s$, $\phi_s = \Phi s$, and f_0 represents the hopping amplitude. The two copies are labeled by a pseudo-spin index $\alpha = \pm$, which also represents the two blocks $\mathcal{H}_\Phi = \mathcal{H}_{+,\Phi}$ and $\mathcal{H}_\Phi^* = \mathcal{H}_{-,\Phi}$ in Eq. 4. H is symmetric under time reversal T and has three doubly degenerate bands, which are separated by nonzero gaps with a nontrivial topological $\mathbf{Z}_2$ index (27). Therefore, in a finite system, we expect one chiral edge state per pseudo-spin in both gaps.

The above matrix $\mathcal{H}$ is complex and hence does not comply with our design goals of a real matrix. However, owing to the time reversal symmetry, we can transform to a real basis by combining the local Kramers pairs $\alpha = \pm$,

$$\begin{pmatrix} x_{r,s} \\ y_{r,s} \end{pmatrix} = \sqrt{2} \underbrace{\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}^{-1}}_{=u^{-1}} \begin{pmatrix} \psi_{r,s}^+ \\ \psi_{r,s}^- \end{pmatrix}, U = u \otimes \mathbb{1}_{\text{lattice}}$$

$$(5)$$

This transformation yields

$$U^\dagger \mathcal{H} U = \begin{pmatrix} \text{Re } \mathcal{H}_\Phi & \text{Im } \mathcal{H}_\Phi \\ \text{Im } \mathcal{H}_\Phi^* & \text{Re } \mathcal{H}_\Phi \end{pmatrix} \equiv \mathcal{D} \qquad (6)$$

which is real, symmetric, and can easily be made positive definite. The four blocks of the matrix $\mathcal{D}$ encode x-x and y-y couplings through Re $\mathcal{H}_\Phi$ and the couplings between x and y via Im $\mathcal{H}_\Phi$. Both $x_{r,s}$ and $y_{r,s}$ describe a one-dimensional oscillator at lattice site (r, s). In the following, we interpret the combined coordinate $(x_{r,s}, y_{r,s})$ as an effective two-dimensional oscillator. It is then evident from the structure of u that the eigenmodes of the system characterized by $\alpha = \pm$ correspond to left and right circularly polarized motions. Hence, these two polarizations replace the notion of pseudo-spin up and down of the quantum mechanical problem.

We implement $\mathcal{D}$ using pendula, as shown in Fig. 1. $\mathcal{D}$ inherits its structure from the explicit Landau gauge in Eq. 3: In the s-direction, there is no cross-coupling between $x_{r,s}$ and $y_{r,s\pm1}$, as there is no imaginary part to the coupling $f = f_0$. Along the r-direction, $\mathcal{D}$ mixes the x- with x-, the y- with y-, as well as the x- with y-degrees of freedom (28) (Fig. 1). The total system consists of 270 pendula in a lattice of $L_r \times L_s = 9 \times 15$ sites. The bare

eigenfrequency of each pendulum due to gravity is $\omega_0/2\pi \approx 0.75$ Hz, which changes with the restoring forces of the spring couplings to $\omega/2\pi \approx 2.34$ Hz (28).

To analyze the system, we harmonically excite one lattice site with a well-defined polarization by forcing the position of the two pendula (see also movies S1 and S2). By tracking the position of all pendula, one can extract the two-dimensional traces $[x_{r,s}(t), y_{r,s}(t)]$. From these we obtain the mean deflection $A_{r,s}$ as well as the polarization (left-handed, linear, or right-handed) related to the lag between the x and y pendula (28) (Fig. 2A).

To establish the existence of the edge states, we scan the frequency at which one edge site is excited. For every frequency, we wait until the system reaches a steady state before we extract the total response $\chi = \sum_{r,s} A_{r,s}/N$, where N is the number of sites. Figure 2E shows that the system responds appreciably between 1.7 Hz and 2.9 Hz, as expected (28). Moreover, there are two regions with sequences of pronounced peaks. The width of these peaks is $\Gamma \sim 0.04$ Hz, indicating the damping of the oscillators and a quality factor $Q = \omega/2\pi\Gamma \approx 60$. To make the connection to edge modes, we separate the system into the outermost line of lattice sites $\chi_e = \sum_{\text{edge}} A_{r,s}/N_{\text{edge}}$ and the rest $\chi_b = \sum_{\text{bulk}} A_{r,s}/N_{\text{bulk}}$. The relative weight $\chi_e/(\chi_b + \chi_e)$ is shown in Fig. 2F. There are three bands where the response lies mainly in the bulk; these are separated by two frequency regions where the response is dominated by the edge. To illustrate this further, we show in Fig. 2, B and C, the recorded mode structure at a bulk and an edge frequency, respectively.

Given the analogy to the QSHE effect, we expect the edge states to be helical. The edge spectrum $\omega(k)$, where k is the wave vector along the edge, can be extracted from the steady states shown in Fig. 2C: Beyond $A_{r,s}$ and the polarization, we determine the position of each two-dimensional oscillator at a given time $[x_{r,s}(t_0), y_{r,s}(t_0)]$. The angle of this vector with respect to the positive x-direction defines a local phase $\phi_{r,s}$ (Fig. 2A). The change

$$k = \phi_{r,s+1} - \phi_{r,s} \qquad (7)$$

defines the wave vector in units of the inverse lattice constant a. The resulting edge dispersion $\omega(k)$ is shown in Fig. 2G. For each polarization, there is a chiral (that is, unidirectional) mode per gap, as expected.

The helicity of the edge dispersion indicates that the classical system faithfully implements the QSHE. Hence, for a clean system, a linearly polarized excitation on a boundary site is split into two counterpropagating circularly polarized modes. Consequently, these edge states act as a polarizing beam splitter.

The QSHE is protected by the quantum mechanical time reversal symmetry. In the classical case, the corresponding symmetry is the combination of time reversal together with the local exchange $(x_{r,s}, y_{r,s}) \rightarrow (y_{r,s}, -x_{r,s})$ (28). Disorder on the local couplings typically breaks this symmetry. To assess the effect of the symmetry breaking due to disordered spring constants, we characterized our edge states in terms of their efficiency as a beam splitter.

We measured the ratio of excitation $\sum_L A_{r,s}/\sum_R A_{r,s}$ between the left and the right long edge after exciting in the middle of a short edge (Fig. 3A). By changing the relative amplitude and time lag between the two local pendula, we scanned all possible polarizations. The resulting splitting ratios are represented in Fig. 3B as a color map on a Poincaré sphere, where the north and south poles correspond to right and left circular polarization of the drive, respectively. The maximal imbalance between the left and right edge is not reached on the north or south pole, but on two approximately antipodal points rotated by ~15° away from the poles. However, as shown in Fig. 3C, we reach a splitting fidelity of 99.80 ± 0.04% at the optimal points. From this we conclude that on the length scale of our system, only symmetry preserving disorder is relevant (28).

To demonstrate that our edge states are not mere whispering gallery modes (12, 13) and to highlight the robustness of the phononic edge states described here, we removed a sequence of sites from the dynamical problem, effectively creating a convex boundary. In Fig. 4, A to D, we show the resulting mode structure for a frequency in the lower gap. The results demonstrate that the exact shape of the boundary has no influence on the stability of the edge states.

To further strengthen the point that the edge states are topological rather than imposed by the finite-size geometry, we created a domain wall between two different topological sectors. We inverted the effective flux seen by the two polarizations on six rows of the system. At the boundary between the two sectors, the spin Chern numbers changed their values, which requires the presence of topological in-gap modes (28). We illustrate these modes along the sector boundaries by exciting a linearly polarized wave packet on the short edge of the larger sector. As shown in Fig. 4E, the wave packet splits into two circularly polarized packets, each traveling along one edge. At the boundary to the second sector, they are deflected into the interior of the lattice before they each travel independently along the physical edge of the second sector in reversed directions.

Our measured edge spectrum (Fig. 2G), the efficiency of the edge states as a beam splitter (Fig. 3), and their immunity to surface roughness (Fig. 4, A to D) corroborate that the QSHE phenomenology can be implemented in an imperfect mechanical system. Our findings indicate the robustness of our proposal against interactions, in particular against cubic (Duffing) nonlinearities (29) inherent to pendula. We indeed find that these nonlinear effects are present in our setup (28); however, they do not obstruct our results. In addition, unavoidable damping leads to the dissipation of the mechanical energy injected into the edge channels. However, their chirality would lead one to expect a linear scaling of decay length $\xi \sim Q$ with the quality factor of the local oscillators. For applications dependent on the controlled transmission of phonons in one-dimensional channels, this is an improvement over a standard bidirectional wave guide, where one expects $\xi \sim \sqrt{Q}$ in analogy to a random walk (28).

Although our experiment was performed with coupled pendula, the mapping from the quan-tum mechanical problem to a classical dynamical matrix is much more general and can be used in the design of acoustic metamaterials. The key property obtained from a mechanical QSHE is the combination of a phononic band gap with stable helical states on the surface of the material. The guaranteed presence of these states allows the efficiency of a band gap for isolation together with a reduction of reflection, as the surface state can absorb and channel energy around the material (30). Moreover, the stability of the surface states allows the specification of reliable phonon wave guides (31) or stable acoustic delay lines (18), which would be useful for applications such as acoustic lensing. Finally, the helical nature of the surface states provides one-way channels similar to sound and heat diodes (32).

REFERENCES AND NOTES

1. F. D. M. Haldane, *Phys. Rev. Lett.* **61**, 2015–2018 (1988).
2. C. L. Kane, E. J. Mele, *Phys. Rev. Lett.* **95**, 226801 (2005).
3. B. A. Bernevig, S.-C. Zhang, *Phys. Rev. Lett.* **96**, 106802 (2006).
4. M. König et al., *Science* **318**, 766–770 (2007).
5. J. E. Moore, *Nature* **464**, 194–198 (2010).
6. N. Jia, A. Sommer, D. Schuster, J. Simon, http://arxiv.org/abs/1309.0878 (2013).
7. V. V. Albert, L. I. Glazman, L. Jiang, *Phys. Rev. Lett.* **114**, 173902 (2015).
8. X. L. Qi, S.-C. Zhang, *Rev. Mod. Phys.* **83**, 1057–1110 (2011).
9. M. Z. Hasan, C. L. Kane, *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
10. A. R. Mellnik et al., *Nature* **511**, 449–451 (2014).
11. R. S. K. Mong et al., *Phys. Rev. X* **4**, 011036 (2014).
12. L. Rayleigh, *The Theory of Sound* (Macmillan, New York, 1889).
13. L. Rayleigh, *Philos. Mag.* **20**, 1001–1004 (1910).
14. D. K. Armani, T. J. Kippenberg, S. M. Spillane, K. J. Vahala, *Nature* **421**, 925–928 (2003).
15. M. Hafezi, E. A. Demler, M. D. Lukin, J. M. Taylor, *Nat. Phys.* **7**, 907–912 (2011).
16. M. Hafezi, S. Mittal, J. Fan, A. Migdall, J. M. Taylor, *Nat. Photonics* **7**, 1001–1005 (2013).
17. M. C. Rechtsman et al., *Nature* **496**, 196–200 (2013).
18. A. Spadoni, C. Daraio, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 7230–7234 (2010).
19. E. Prodan, C. Prodan, *Phys. Rev. Lett.* **103**, 248101 (2009).
20. N. Berg, K. Joel, M. Koolyk, E. Prodan, *Phys. Rev. E* **83**, 021913 (2011).
21. C. L. Kane, T. C. Lubensky, *Nat. Phys.* **10**, 39–45 (2013).
22. J. Paulose, B. G. Chen, V. Vitelli, *Nat. Phys.* **11**, 153–156 (2015).
23. B. G. Chen, N. Upadhyaya, V. Vitelli, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13004–13009 (2014).
24. B. A. Bernevig, T. L. Hughes, S.-C. Zhang, *Science* **314**, 1757–1761 (2006).
25. B. A. Bernevig, T. L. Hughes, *Topological Insulators and Superconductors* (Princeton Univ. Press, Princeton, NJ, 2013).
26. D. R. Hofstadter, *Phys. Rev. B* **14**, 2239–2249 (1976).
27. C. L. Kane, E. J. Mele, *Phys. Rev. Lett.* **95**, 146802 (2005).
28. See supplementary materials on *Science* Online.
29. I. Kovacic, M. J. Brennan, *The Duffing Equation: Nonlinear Oscillators and Their Behaviour* (Wiley, New York, 2011).
30. Z. Liu et al., *Science* **289**, 1734–1736 (2000).
31. M. J. A. Schuetz et al., http://arxiv.org/abs/1504.05127 (2015).
32. B. Li, L. Wang, G. Casati, *Phys. Rev. Lett.* **93**, 184301 (2004).

GEOMORPHOLOGY

Experimental evidence for hillslope control of landscape scale

K. E. Sweeney,[1]* J. J. Roering,[1] C. Ellis[2]

Landscape evolution theory suggests that climate sets the scale of landscape dissection by modulating the competition between diffusive processes that sculpt convex hillslopes and advective processes that carve concave valleys. However, the link between the relative dominance of hillslope and valley transport processes and landscape scale is difficult to demonstrate in natural landscapes due to the episodic nature of erosion. Here, we report results from laboratory experiments combining diffusive and advective processes in an eroding landscape. We demonstrate that rainsplash-driven disturbances in our experiments are a robust proxy for hillslope transport, such that increasing hillslope transport efficiency decreases drainage density. Our experimental results demonstrate how the coupling of climate-driven hillslope- and valley-forming processes, such as bioturbation and runoff, dictates the scale of eroding landscapes.

onvex hillslopes and concave valleys are ubiquitous in eroding, soil-mantled landscapes (1–3) (Fig. 1, A and B). These distinct landforms are produced by equally distinct sediment transport processes: On hillslopes, abiotic (4, 5) and biotic (6, 7) disturbance agents disperse sediment downslope, whereas in valleys, sediment is transported by debris flows (8) or flowing water (9). The transition between hillslopes and valleys has long been considered a fundamental landscape scale (3, 10, 11), but there is debate over what controls its style and position. Numerical results suggest that the hillslope-valley transition may be dictated by the minimum runoff necessary for sediment detachment or landslide initiation (11–13) or by the competition between diffusive transport on hillslopes and advective transport in channels (14, 15).

These geomorphic models predict expansion or contraction of the valley network from changes in climatic variables such as precipitation and vegetation (3, 12, 16). Hence, rigorous testing of controls on the hillslope-valley transition is central to our understanding of landscape response to environmental perturbations. Due to the slow and episodic nature of erosion, however, field tests are limited to comparisons of steady-state model predictions with natural topography [e.g., (2)]. Such approaches rely on the assumption that topography reflects long-term average fluxes, ignoring the potentially important effects of initial conditions (17) and temporal lags between landscape response and climatic and tectonic forcing (18, 19).

We conducted a series of laboratory experiments to determine whether the competition between hillslope transport and valley incision

sets the spatial scale of landscapes (15). The theoretical underpinnings of the process control on the hillslope-valley transition derive from a statement of mass conservation, where the rate of elevation change (dz/dt) is equal to

uplift rate (U), minus erosion due to disturbance-driven hillslope diffusion and channel advection by surface runoff

$$\frac{\partial z}{\partial t} = U + D\nabla^2 z - K(PA)^m S^n \qquad (1)$$

where D is hillslope diffusivity, K is the stream power constant, A is drainage area, S is slope, P is precipitation rate (assuming steady, uniform rainfall), and m and n are positive constants (9). In this framework, the strength of hillslope transport relative to channel processes can be quantified by the landscape Péclet number (Pe), assuming $n = 1$ (15)

$$Pe = \frac{KL^{2m+1}}{D} \qquad (2)$$

where L is a horizontal length scale and the hillslope-valley transition occurs at the critical length scale L_c where $Pe = 1$ (15). In plots of slope versus drainage area, L_c corresponds to a local maximum separating convex hillslope and concave valley topography (20) (Fig. 1, D and E). If this framework holds for field-scale and experimental landscapes, increasing the vigor of hillslope transport relative to valley incision (decreasing Pe) should result in longer hillslopes (higher L_c)
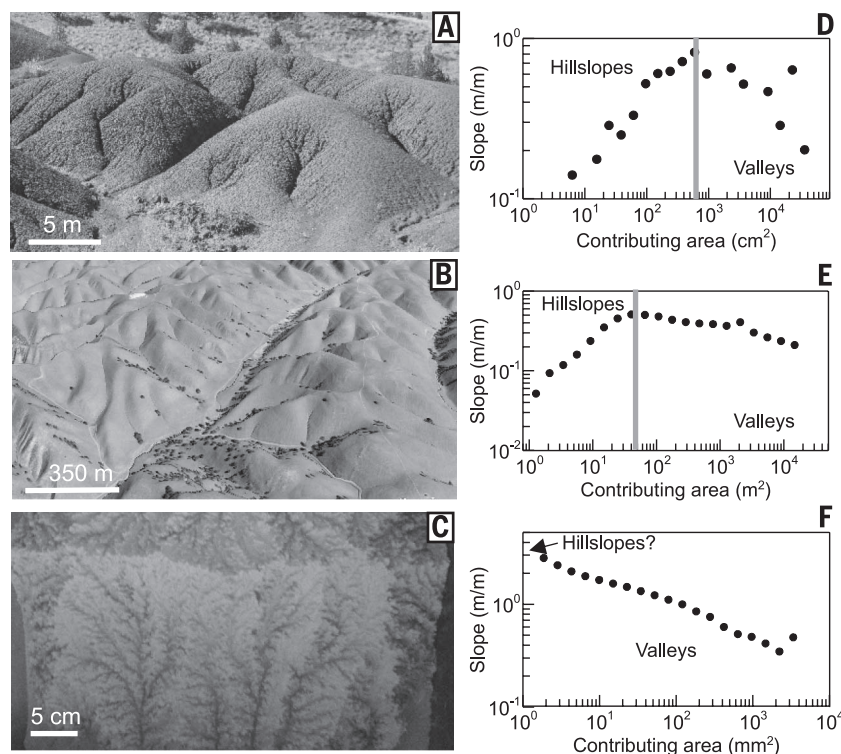
Fig. 1. Characteristic morphology of eroding landscapes. Photographs of eroding landscapes. (**A**) Painted Hills unit of John Day Fossil Beds National Monument, Oregon. (**B**) Gabilan Mesa, California. (**C**) Laboratory landscape from this study with no hillslope diffusion, and associated plots of local slope versus drainage area, calculated with steepest descent algorithm (**D** to **F**). Pictures in (A) and (C) taken by the author (K.E.S.); (B) from Google Earth. Topographic data to generate slope-area plots from (D) field surveys; (E) Lidar data from National Center for Airborne Laser Mapping; and (F) this study. Gray vertical bars in (D) to (F) demarcate the inferred hillslope-valley transition (34).

[1]Department of Geological Sciences, University of Oregon, 1272 University of Oregon, Eugene, OR 97403-1272, USA. [2]St. Anthony Falls Laboratory and National Center for Earth-Surface Dynamics, College of Science and Engineering, University of Minnesota, 2 Third Avenue SE, Minneapolis, MN 55414-2125, USA.
*Corresponding author. E-mail: kristin.e.sweeney@gmail.com

and a contraction of the valley network (i.e., a decrease in drainage density).

Experimental landscapes bridge the gap in complexity between numerical models and natural landscapes (*21*) by enabling us to control the confounding influences of tectonics, climate, and lithology and observe surface evolution through time. As previously noted [e.g., (*21*)], a complete dynamical scaling of erosional landscapes in the laboratory is typically intractable due to shallow water depths, large grain sizes relative to the size of the experiment, and other considerations. Nonetheless, past landscape experiments have successfully demonstrated the topographic manifestation of changing uplift rate (*22*, *23*), precipitation rate (*24*), and precipitation patterns (*25*). In these experiments, however, erosion was exclusively driven by surface runoff (e.g., Fig. 1, C and F), intentionally excluding the representation of diffusive hillslope processes (*22*) and hence precluding tests for the role of hillslope transport in setting landscape scale.

Following Eq. 1, we created an experimental system that distilled landscape evolution into three essential ingredients: base-level fall (uplift), surface runoff (channel advection), and sediment disturbance via rainsplash (hillslope diffusion) (Fig. 2). Our experiments in the eXperimental Landscape Model (XLM) at the St. Anthony Falls Laboratory systematically varied the strength of disturbance-driven transport relative to surface runoff (changing Pe) for steady, uniform uplift. The XLM consists of a 0.5 by 0.5 by 0.3 m³ flume with two parallel sliding walls, each connected to a voltage-controlled dc motor to simulate relative uplift (Fig. 2A). The experimental substrate was crystalline silica (median grain size = 30 μm) mixed with 33% water to increase cohesion and reduce infiltration (Fig. 2C). We began each experiment by filling the XLM with sediment and allowing it to settle for ~24 hours

to homogenize moisture content. Topographic data at 0.5 mm vertical accuracy were collected at regular time intervals on a 0.5 by 0.5 mm grid using a laser scanner.

Sediment transport in our experiments was driven by two distinct rainfall systems: the mister, a rotating ring fitted with 42 misting nozzles, and the drip box, a polyvinyl chloride constant head tank fitted with 625 blunt needles of 0.3 mm interior diameter arranged in a 2 by 2 cm grid (Fig. 2B). The fine drops produced by the mister lack sufficient energy to disturb sediment on impact and instead result only in surface runoff. By contrast, the 2.8 mm diameter drops from the drip box are sufficiently energetic to create rainsplash and craters on the experiment surface that result in sediment transport. We used four fans mounted on the corners of the experiment to generate turbulence and randomize drop location during drip box rainfall. Importantly, sediment transport due to drip box rainfall consists of both hillslope diffusion from the cumulative effect of drop impacts and nonnegligible advective transport due to the subsequent runoff of the drops. Thus, we expect that changing the relative contribution of rainsplash results in a change in both hillslope and channel transport efficiency (*D* and *K*, respectively).

We ran five experiments, varying the time of drip box rainfall (i.e., predominantly diffusive transport) from 0 to 100% of total experiment runtime (Fig. 3 and table S1) (*26*) and holding base-level fall and water flux from the mister and the drip box constant. During the experiment, we alternated between drip box rainfall and mister rainfall over 10-min periods (table S1); the fans used for drip box rainfall prevented simultaneous operation. We continued each experiment until we reached flux steady state such that the spatially averaged erosion rate was equal to the rate of base-level fall (figs S1 and S2). Each ex-

periment ran for 10 to 15 hours, encompassing 60 to 90 intervals of drip box and/or mister rainfall.

The steady-state topography of our experiments (Fig. 3, A to F) demonstrates how increasing the relative dominance of rainsplash disturbance affects landscape morphology. Qualitatively, landscapes formed by a higher percentage of drip box rainfall (Fig. 3E) appear smoother, with wider, more broadly spaced valleys and extensive unchannelized areas. In contrast, landscapes with more surface runoff transport (Fig. 3A), equivalent to past experimental landscapes (*24*, *27*), are

Fig. 3. Steady-state topography and hillslope morphology. (**A** to **E**) Hillshades of experimental topography for (A) 0% drip, (B) 18% drip, (C) 33% drip, (D) 66% drip, and (E) 100% drip overlain with channel networks (blue) and locations of hillslope profiles (red). Topography is 475.5 mm wide in plan-view. (**F**) Elevation profiles of hillslopes marked by red lines in (A) to (E). Vertical and horizontal length scales are equal.
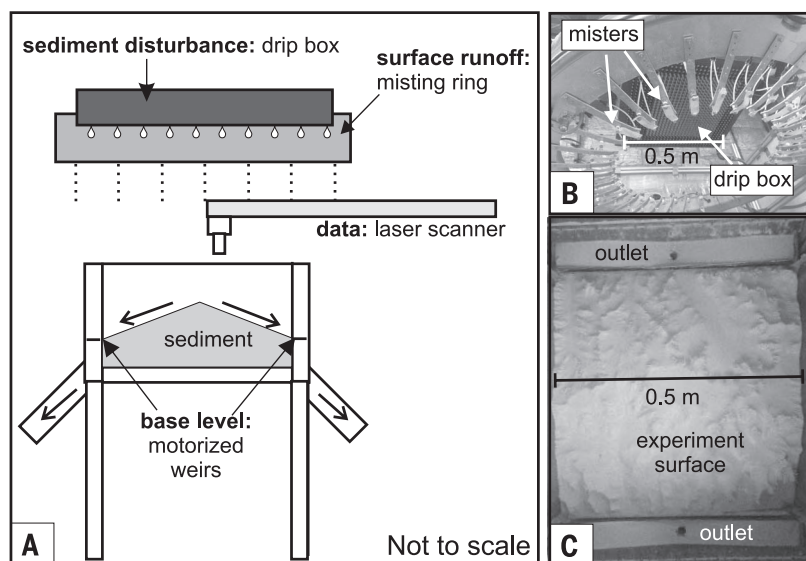

Fig. 2. Experimental setup. (**A**) Schematic profile of experimental apparatus (XLM). Line arrows show direction of sediment movement. (**B**) Photograph of misting ring and drip box looking from below. (**C**) Photograph of sediment surface during 100% drip run.
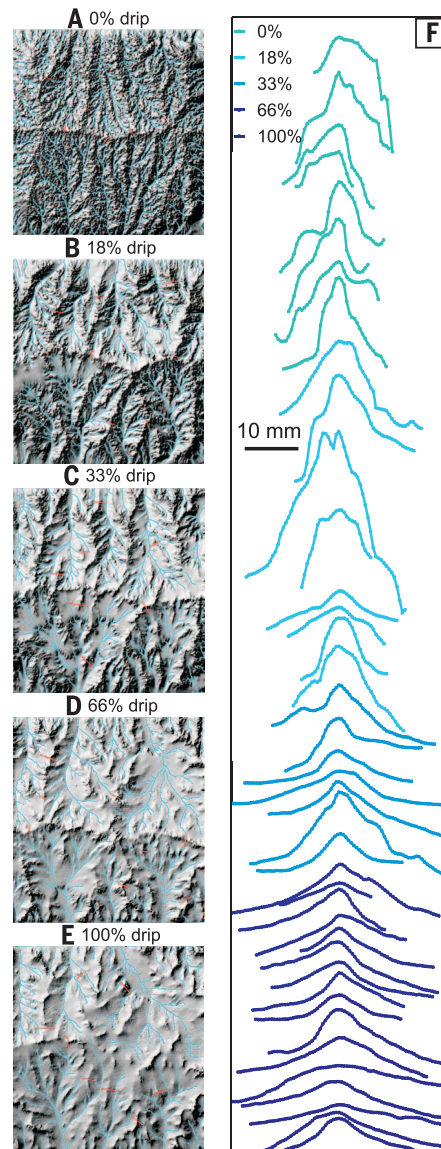
densely dissected. As the relative percentage of rainsplash increases, hillslope transects increase in both length and topographic curvature (Fig. 3F), confirming that our experimental procedure can be used to adjust hillslope transport efficiency. Hillslope relief in our experiments is approximately 10 to 20% of total landscape relief, a similar value to natural landscapes (28).

To test the expected relationship between Péclet number and landscape scale (15) (Eq. 2), we used steady-state relationships between landscape morphology and sediment transport laws to independently calculate D and K. This approach is often not possible in natural landscapes and thus extends our theoretical understanding beyond the slope-area plots shown in Fig. 1, D to F. Specifically, we used the approach of (28) to calculate D, which uses average hillslope length and gradient, thereby avoiding the stochastic imprint of individual raindrop impacts that can obscure local metrics of hillslope form, such as curvature. The following relationship relates mean hillslope gradient ($\overline{S}$) to hillslope length (L_h), critical slope (S_c), erosion rate (E, equal to U at steady state), and hillslope diffusivity (D)

$$\frac{\overline{S}}{S_c} = \frac{1}{E^*}\left(\sqrt{1+(E^*)^2} - \ln\left(\frac{1}{2}\left[1+\sqrt{1+(E^*)^2}\right]\right)-1\right) \quad (3)$$

where $E^* = E(2L_h/KS)$ (28). To calculate K and m for the stream power model, we used the prediction of Eq. 1 that local steady-state channel slope is a power-law function of drainage area (29):

$$S = \left(\frac{U}{KP^m}\right)^{\frac{1}{n}} A^{-m/n} \quad (4)$$

To quantify hillslope length and gradient, we mapped the channel network by explicitly identifying channel forms (28–31) (fig. S3), then traced hillslopes beginning at hilltop pixels by following paths of steepest descent to the nearest channel (28) (fig. S3). We take S_c to be sufficiently large (table S1) that Eq. 3 approximates linear diffusion. As the proportion of rainsplash transport increases, D calculated from Eq. 3 also increases (table S1), confirming that the morphologic trend of individual hillslope transects (Fig. 3F) reflects increasing hillslope transport efficiency.

To calculate the advective process parameters (Eq. 4), we extracted slope and steepest-descent drainage area data along networks defined by a minimum drainage area of 250 mm² (larger than the drainage area of channel initiation) and fit power-law relationships to plots of slope versus drainage area (29). Following (15, 22), we assume that $n = 1$ and use the intercept and slope of the power-law fits to calculate m and K for each experiment. Whereas m is relatively invariant for all our experiments, K tends to increase with the fraction of drip box transport, indicating that post-impact rainfall runoff contributes to advective as well as diffusive transport in our experiments.

Given that both D and K change in our experiments, we calculated Pe values (Eq. 2) for each of our experiments to quantify how diffusive and advective processes contribute to the observed transition from smooth and weakly channelized landscapes (100% drip box, Fig. 3A) to highly dissected terrain (mist only, 0% drip box, Fig. 3E). We calculated Pe for each experiment (Eq. 2) by assuming that $n = 1$ and taking the length scale L to be the horizontal distance from the main divide to the outlet (256 mm). Our results show that landscape Pe value increases with the fraction of drip box transport, demonstrating that an increase in hillslope transport efficiency, D, is the dominant result of increasing rainsplash frequency. Figure 4 reveals a positive relationship between Pe and drainage density, which is inversely related to hillslope length or L_c, such that increasing Pe in our experiments results in higher drainage density (i.e., shorter hillslopes). This finding is consistent with theoretical predictions for coupled hillslope-channel process controls on the scale of landscape dissection (14, 15).

In our experiments, hillslope transport imparts a first-order control on landscape scale, emphasizing the need to establish functional relationships between climate variables and hillslope process rates and mechanisms for real landscapes. Although climate change scenarios typically focus on the influence of vegetation and rainfall on overland flow and channel hydraulics (3, 12), climate controls on the vigor of hillslope transport (e.g., 32, 33) can drive changes in landscape form. Robust linkages between transport processes and topography, as discussed here, are an important component of interpreting planetary surfaces as well as decoding paleolandscapes and sedimentary deposits.

Fig. 4. Effect of landscape Péclet number on landscape scale. Landscape Péclet number for each experiment (circle, 0% drip; square, 18% drip; diamond, 33% drip; triangle, 66% drip; plus sign, 100% drip) versus drainage density of GeoNet-derived drainage networks.
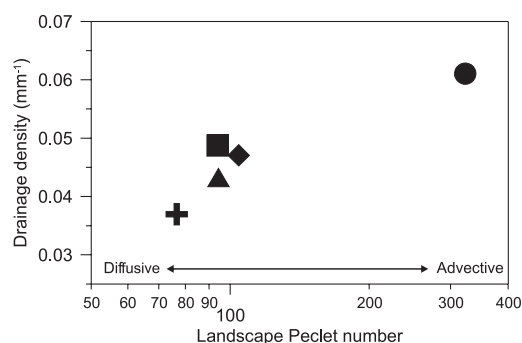
REFERENCES AND NOTES

1. W. M. Davis, *Science* **20**, 245 (1892).
2. G. K. Gilbert, *J. Geol.* **17**, 344–350 (1909).
3. G. Moglen, E. Eltahir, R. Bras, *Water Resour. Res.* **34**, 855–862 (1998).
4. N. Matsuoka, *Permafr. Periglac. Process.* **9**, 121–133 (1998).
5. D. J. Furbish, K. K. Hamner, M. W. Schmeeckle, M. N. Borosund, S. M. Mudd, *J. Geophys. Res.* **112** (F1), F01001 (2007).
6. E. J. Gabet, *Earth Surf. Process. Landf.* **25**, 1419–1428 (2000).
7. J. J. Roering, J. Marshall, A. M. Booth, M. Mort, Q. Jin, *Earth Planet. Sci. Lett.* **298**, 183–190 (2010).
8. J. Stock, W. E. Dietrich, *Water Resour. Res.* **39**, n/a (2003).
9. A. D. Howard, *Water Resour. Res.* **30**, 2261–2285 (1994).
10. M. J. Kirkby, *Spec. Publ. Inst. Br. Geogr.* **3**, 15–30 (1971).
11. D. Montgomery, W. Dietrich, *Water Resour. Res.* **25**, 1907–1918 (1989).
12. A. Rinaldo, W. E. Dietrich, R. Rigon, G. K. Vogel, I. Rodríguez-Iturbe, *Nature* **374**, 632–635 (1995).
13. R. Horton, *Geol. Soc. Am. Bull.* **56**, 275–370 (1945).
14. T. R. Smith, F. P. Bretherton, *Water Resour. Res.* **8**, 1506–1529 (1972).
15. J. T. Perron, J. W. Kirchner, W. E. Dietrich, *Nature* **460**, 502–505 (2009).
16. G. E. Tucker, R. Slingerland, *Water Resour. Res.* **33**, 2031–2047 (1997).
17. J. Taylor Perron, S. Fagherazzi, *Earth Surf. Process. Landf.* **37**, 52–63 (2012).
18. K. Whipple, G. E. Tucker, *J. Geophys. Res.* **104** (B8), 17661–17674 (1999).
19. J. J. Roering, J. W. Kirchner, W. E. Dietrich, *J. Geophys. Res.* **106** (B8), 16499–16513 (2001).
20. J. T. Perron, W. E. Dietrich, J. W. Kirchner, *J. Geophys. Res.* **113** (F4), F04016 (2008).
21. C. Paola, K. Straub, D. Mohrig, L. Reinhardt, *Earth Sci. Rev.* **97**, 1–43 (2009).
22. D. Lague, A. Crave, P. Davy, *J. Geophys. Res.* **108** (B1), 2008 (2003).
23. J. M. Turowski, D. Lague, A. Crave, N. Hovius, *J. Geophys. Res.* **111** (F3), F03008 (2006).
24. S. Bonnet, A. Crave, *Geology* **31**, 123 (2003).
25. S. Bonnet, *Nat. Geosci.* **2**, 766–771 (2009).
26. Materials and methods are available as supplementary material on *Science* Online.
27. L. Hasbargen, C. Paola, *Geology* **28**, 1067 (2000).
28. J. J. Roering, J. T. Perron, J. W. Kirchner, *Earth Planet. Sci. Lett.* **264**, 245–258 (2007).
29. C. Wobus *et al.*, in *Tectonics, Climate, and Landscape Evolution: Geological Society of America Special Paper 398*, S. D. Willett, N. Hovius, M. T. Brandon, D. M. Fisher, Eds. (Geological Society of America, Boulder, CO, 2006), pp. 55–74.
30. M. D. Hurst, S. M. Mudd, R. Walcott, M. Attal, K. Yoo, *J. Geophys. Res.* **117** (F2), F02017 (2012).
31. P. Passalacqua, T. Do Trung, E. Foufoula-Georgiou, G. Sapiro, W. E. Dietrich, *J. Geophys. Res.* **115** (F1), F01002 (2010).
32. J. L. Dixon, A. M. Heimsath, J. Kaste, R. Amundson, *Geology* **37**, 975–978 (2009).
33. O. A. Chadwick *et al.*, *Geology* **41**, 1171–1174 (2013).
34. W. Dietrich, C. Wilson, D. Montgomery, J. McKean, R. Bauer, *Geology* **20**, 675–679 (1992).

POROUS MATERIALS

Transient laser heating induced hierarchical porous structures from block copolymer–directed self-assembly

Kwan Wee Tan,[1] Byungki Jung,[1]* Jörg G. Werner,[1,2] Elizabeth R. Rhoades,[3] Michael O. Thompson,[1] Ulrich Wiesner[1]†

Development of rapid processes combining hierarchical self-assembly with mesoscopic shape control has remained a challenge. This is particularly true for high-surface-area porous materials essential for applications including separation and detection, catalysis, and energy conversion and storage. We introduce a simple and rapid laser writing method compatible with semiconductor processing technology to control three-dimensionally continuous hierarchically porous polymer network structures and shapes. Combining self-assembly of mixtures of block copolymers and resols with spatially localized transient laser heating enables pore size and pore size distribution control in all-organic and highly conducting inorganic carbon films with variable thickness. The method provides all-laser-controlled pathways to complex high-surface-area structures, including fabrication of microfluidic devices with high-surface-area channels and complex porous crystalline semiconductor nanostructures.

Hierarchically porous materials that bridge nano- to macroscopic length scales are expected to find use in a wide variety of applications including catalysis, energy conversion and storage, and membrane filtration (1, 2). Well-ordered nanoporous materials derived from small-molecule surfactant and block copolymer (BCP) self-assembly have been explored in the form of amorphous (3–7), polycrystalline (8, 9) and single-crystal (10) solids. Multiple time-consuming processing steps are typically required to generate the final structures. For example, removal of organic components by conventional thermal processing to create porosity typically takes several hours (6–8). Methods to fabricate hierarchical porous polymer scaffolds using BCPs have recently been devised that circumvent some of these complexities (11, 12). Further developments are still desirable, however, as hierarchical structures provide high surface area combined with high flux beneficial for a number of applications.

Laser-induced transient heating of organic materials has been explored as an alternative thermal processing approach to alter morphology, as well as materials properties, and to enable direct pattern transfer. For example, by laser heating in the millisecond to submillisecond time frames, the temperature stability of a methacrylate-based photoresist was enhanced by ~400°C relative to conventional hot-plate heating (13). CO_2 laser ablation of polyimide and polyetherimide resulted in the direct formation of three-dimensional (3D)

porous graphitic carbon network structures, but work did not involve a bottom-up self-assembly process, resulting in little control over mesostructure and pore size (14).

Three-dimensional porous structure fabrication by optical lithography has also been demonstrated using multiphoton absorption polymerization (MAP) and stimulated emission depletion (STED)–inspired direct laser writing techniques (15–19). However, direct laser writing of such 3D continuous structures with sub-100-nm feature size and periodicity remains challenging due to the inherent serial mode of operation and sophisticated instrumentation, diffraction limitations of MAP excitation lasers, and stringent materials criteria—e.g., for STED resists (16–19).

We demonstrate a simple and rapid method for direct generation of 3D continuous hierarchically porous structures with meso- to macroscopic shapes by combining BCP self-assembly directed resols structure formation with CO_2 laser-induced transient heating. A schematic of the BCP-based writing induced by transient heating experiments (B-WRITE) is shown in Fig. 1A (20). Lab-synthesized ABC triblock terpolymers [poly(isoprene)-block-poly(styrene)-block-poly(ethylene oxide), ISO-38/69] as well as commercially available ABA di-BCPs (Pluronic F127) were used as examples (20, 21). Oligomeric resols were synthesized by the polymerization of resorcinol/phenol and formaldehyde under basic conditions (22). The structure-directing BCP was first mixed with resols to form an all-organic hybrid thin film by spin-coating on a silicon substrate. Samples are designated as BCP-M-R, where M represents BCP molar mass and R denotes the resols additives. As-deposited or thermally cured films were then heated by a continuous wave CO_2 laser ($\lambda = 10.6$ μm), focused into a ~90 μm by 600 μm line beam, and scanned to generate a 0.5-ms dwell in air. Although the

organic layer does not interact strongly with the far-infrared photons, the boron-doped Si substrate absorbs most of the laser energy, thereby heating the film. As the laser passes, cooling occurs by thermal conduction on submillisecond time frames. Laser heating temperatures were determined using a procedure described in (13). Well-defined porous resin structures were formed by the simultaneous thermopolymerization of resols (negative-tone) and decomposition of BCPs (positive-tone) during transient heating by sequential CO_2 laser irradiations at locations predefined by the beam. Material in nonirradiated regions could be removed by rinsing with solvent, leaving resulting porous resin structures on the substrate.

Thermal stability was compared for organic films prepared from 5 weight percent (wt %) ISO-69, 5 wt % resols, and 6 wt % ISO-38/69-R solutions in tetrahydrofuran (THF) spin-coated on Si (Fig. 1B). The 6 wt % ISO-38/69-R hybrid thin films were prepared by mixing 2 to 3 parts ISO to 1 part resols in THF by mass (see table S1 for hybrid compositions). After thermal curing in vacuum at 100°C overnight (>12 hours), films were heated in air by either a single CO_2 laser irradiation for 0.5-ms dwell or furnace heating for 1 hour. Curves in Fig. 1B show film thickness as a function of temperature (and corresponding laser power). Under CO_2 laser heating for 0.5-ms dwell, ISO-69 (black) remained stable up to 550°C (35 W) and thermally decomposed entirely above 670°C (>40 W). Resols films (orange) extensively cross-linked, retaining more than 50% of the original film thickness at 670 to 805°C (40–45 W). For ISO-69-R hybrid films (blue), mixed behavior was observed. This suggested a laser heating process window to decouple BCP decomposition but retain the cross-linked resols (resin).

Initial experiments to generate nanoporous resin structures in ISO-69-R films by a single CO_2 laser irradiation with ISO decomposition at 670°C (40 W) for 0.5-ms dwell confirmed these concepts, albeit with poor uniformity and coverage (fig. S1). Thermal stability of ISO-69-R films was increased dramatically by more than 400°C when heating duration was reduced from 1 hour in a furnace to 0.5 ms with CO_2 laser-induced heating (compare dashed and solid blue lines in Fig. 1B, respectively). Submillisecond time frame CO_2 laser heating suppresses oxidation reactions in ambient environments, enabling organic systems to reach temperatures inaccessible by conventional heating.

For more controlled polymerization of resols and to inhibit thermocapillary dewetting effects (23), we introduced sequential CO_2 laser irradiations of increasing laser powers yielding thicker resin structures with similar enhanced thermal stability (Fig. 1C), but high film uniformity and coverage (Fig. 1, D and E). Scanning electron microscopy (SEM) (Fig. 2 and fig. S2) and Fourier transform infrared (FTIR) spectroscopy (fig. S2) were employed to investigate film morphology and chemical composition, respectively. Upon sequential irradiations, cured ISO-resols hybrid films

[1]Department of Materials Science and Engineering, Cornell University, Ithaca, NY 14853, USA. [2]Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, USA. [3]Cornell NanoScale Science and Technology Facility, Cornell University, Ithaca, NY 14853, USA.
*Present address: Intel Corporation, Hillsboro, OR 97124, USA.
†Corresponding author. E-mail: ubw1@cornell.edu

(fig. S2, A and C) densified due to cross-linking of resols with increasing transient temperatures (22), whereas above 550°C, BCP decomposition occurred (fig. S2, B and D) and was completed at about 670°C (Fig. 2, A and C). This is corroborated by decreasing IR signal intensities of hydroxyl stretching vibrations in the 3500 to 3200 cm^{-1} band and a sharp reduction of IR signal intensities of alkyl stretching vibrations in the 3000 to 2800 cm^{-1} band, respectively (22, 24) (fig. S2, E and F). Resulting resin structures retained ~30 to 55% of the original film thickness (Fig. 1C).

Optical images in Fig. 1D (and fig. S3) show macroscopic linear trenches in ISO-38/69-R hybrid thin films generated by sequential CO_2 laser irradiations for 0.5 ms (B-WRITE method). Each ISO-38-R–based trench was ~2 cm long, requiring ~4.5 min to fabricate in air (Fig. 1D). SEM (Fig. 1E) and profilometry (fig. S3B) show that the ~150-μm-wide B-WRITE trench indeed consists of a ~200-nm-thick nanoporous resin thin film. Laser-heated macroscopic shape dimensions are tunable, as demonstrated for ISO-69-R hybrid films with ~350-nm-thick structures in ~550-μm-wide B-WRITE trenches (fig. S3F). Cross-sectional SEM at lower magnifications in Fig. 2, E and F, shows highly uniform structure formation under these conditions over large areas, as defined by the CO_2 laser.

Resulting pore size and pore size distribution could be tailored by tuning hybrid film composition. Sequential CO_2 laser-heated resin structures formed in the submillisecond time frames were compared with mesoporous structures formed by furnace heating under N_2 for ≥1 hour (fig. S4). The average pore size of the N_2-furnace-heated ISO-69-R structures at 450°C is 39 ± 9 nm with an in-plane lattice spacing of ~47 nm (fig. S5). Laser-heated structures obtained from ISO-69-R hybrid films with the same ISO-to-resols mass ratio of 2.4:1 show a wide pore size distribution with values ranging from ~30 to 200 nm (Fig. 2C). The increase in pore size relative to the N_2-furnace-heated sample is ascribed to rapid evolution and release of gaseous decomposition products during transient heating, resulting in local pore expansion (14, 25). Hybrid film composition provides a convenient lever to tune pore size of B-WRITE structures. For example, Fig. 2D shows that the pore diameter of laser-heated ISO-69-R films was reduced to ~20 to 50 nm by decreasing the ISO-to-resols mass ratio to 1.5:1 (see Fig. 2B for ISO-38-R). From fast Fourier transform analysis, the in-plane lattice spacing was ~52 nm—i.e., close to the value obtained after furnace heating under N_2 (fig. S5).

Terpolymers are synthetically challenging, so BCPs were varied to also include the commercially available Pluronics family (22). Smaller molar mass Pluronic F127 copolymer mixed with resols at 1:1 mass ratio in ethanol led to the formation of resin structures with pore sizes of 20 to 50 nm (Fig. 2G). Pore size and pore size distribution were further reduced to 10 to 30 nm by decreasing the F127-to-resols mass ratio to 1:2 (Fig. 2H), thereby approaching the 5- to 10-nm limit of size control in conventional BCP phase separation (26). Film thickness could be tuned via solution concentration. For example, maintaining ISO-38-to-resols mass ratio at 3:1, sample thickness was increased with the mixed solution concentration (table S1) from ~500 nm (6 wt %) to 1.0 μm (10 wt %) and 1.6 μm (12.5 wt %). Figure 2, I and J, show thicker laser-heated structures with hierarchical pore size distributions of 50 to 400 nm and 50 to 600 nm for 10 wt % and 12.5 wt % solution–derived ISO-38-R samples, respectively. B-WRITE–based porous structures may be accessible for films thicker than 1.6 μm by further optimizing laser heating protocols.

Skipping the 100°C thermal curing step and directly laser-heating as-deposited ISO-resols samples achieved further process simplification. Figure 3A displays macroscopic porous resin lines on a clean Si substrate prepared by subjecting as-deposited ISO-69-R samples to an increased number of sequential CO_2 laser irradiations and removal of nonirradiated components by rinsing with THF (see stage III in Fig. 1A and SEM in
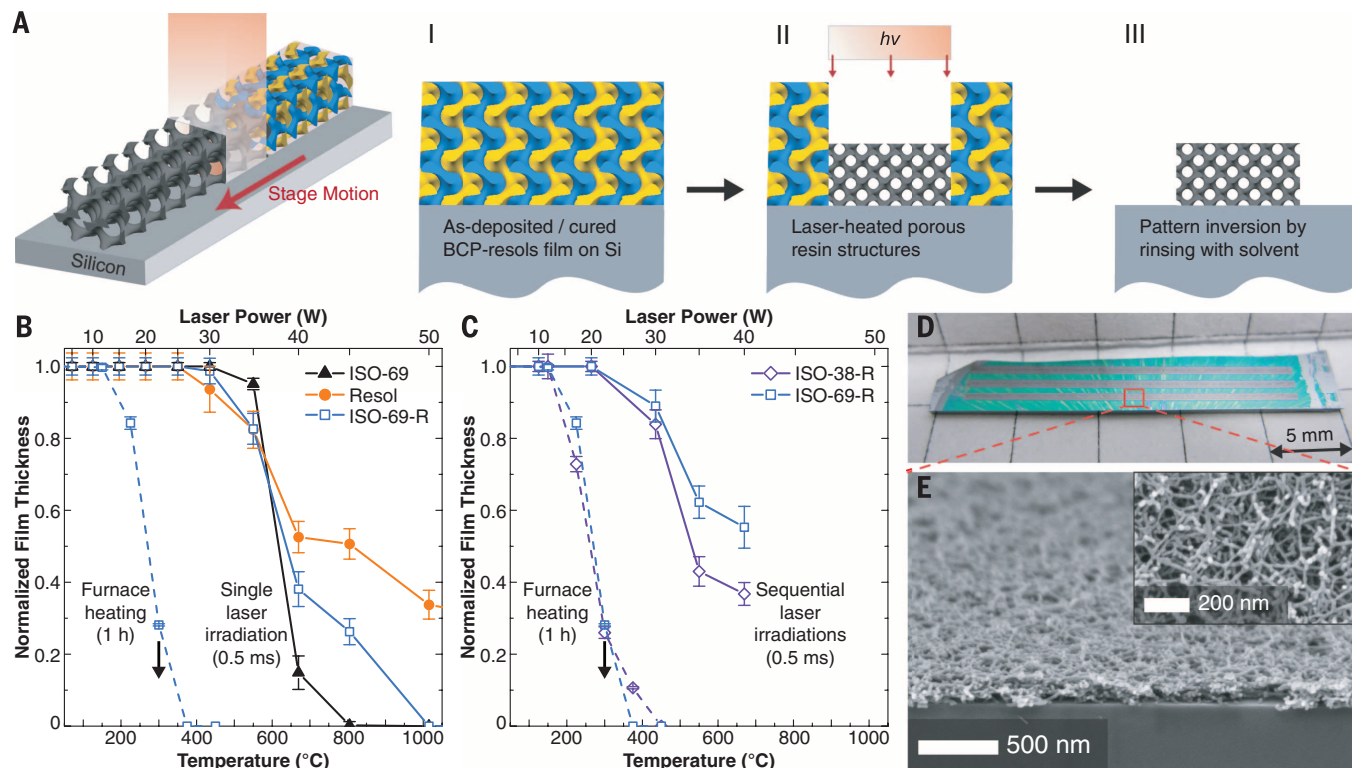

Fig. 1. Transient laser heating of BCP-directed resols hybrid samples. (**A**) Schematic of the B-WRITE method. (**B**) Thickness plots of 5 wt % ISO-69, 5 wt % resols, and 6 wt % ISO-69-R hybrid thin films heated by a single CO_2 laser irradiation of 0.5-ms dwell (solid lines) compared with furnace heating for 1 hour (dashed line), all in air. (**C**) Thickness plots of 6 wt % ISO-38/69-R hybrid thin films heated by sequential CO_2 laser irradiations of 0.5-ms dwell (solid lines) compared with furnace heating for 1 hour (dashed lines), all in air. (**D** and **E**) Optical image of macroscopic trenches in sample ISO-38-R after sequential CO_2 laser irradiations (D) with corresponding SEM cross section and plan view (inset) of the resulting structures in a trench (E). Each B-WRITE trench was generated in less than 5 min.

fig. S6A). Selective removal of resin material with a single CO_2 laser irradiation at 30 W of 25-ms dwell enabled macroscopic line-width adjustments and sidewall smoothening (see SEM in Fig. 2F and depth profile in Fig. 3B). More complex shapes were generated by removal of resin using laser scans in other directions, as shown for orthogonal scans in Fig. 3A.

A final furnace heating step under N_2 at temperatures up to 800°C resulted in carbonized and electrically conductive nanoporous carbon films without structure collapse. Raman spectroscopy shown in Fig. 3C indicated negligible graphitic carbon in laser-heated ISO-69-R samples, corroborating the organic nature of the B-WRITE porous resin structures. After furnace heating at 600°C, D and G bands centered at ~1325 and 1588 cm^{-1}, respectively, appeared in the Raman spectrum, indicating the conversion of phenolic resins into disordered carbon (22). At 800°C, D and G bands narrowed and shifted to ~1314 and 1598 cm^{-1}, respectively, sug-

gesting a higher degree of graphitization (22). Cross-sectional SEM (fig. S6B) revealed that B-WRITE structures contracted by 50 to 60% in the out-of-plane direction during carbonization at 800°C, but otherwise stayed intact. The electrical conductivity of B-WRITE films carbonized at 600°C and 800°C under N_2 was ~8 S/cm (fig. S6C) and 270 to 750 S/cm (Fig. 3D), respectively. This electrical conductivity is higher than that of dense carbon samples heated under identical conditions at 800°C (~52 S/cm) (fig. S6D) but lower than polycrystalline graphite (1250 S/cm) (27). The high electrical conductivity of B-WRITE carbon structures after furnace heating is attributed to the connectivity of the 3D continuous network. We speculate that porosity facilitates carbonization and release of oxygen/hydrogen species to form graphitic-like clusters (fig. S6E).

We constructed proof-of-concept microfluidic devices with highly porous organic channel floors for potential sensing or catalysis applications (Fig. 4, A to C, and fig. S7). A dense resols

film was spin-coated onto Si and cured at 100°C. A trench was then formed with a single CO_2 laser irradiation at 30 W of 25-ms dwell, completely decomposing the film, followed by spin-coating a 6 wt % ISO-38-R film as an overlayer. B-WRITE application formed nanoporous resin structures in the trench. The device was finally sealed with a polydimethylsiloxane film (Fig. 4A). From optical and electron microscopy, as well as profilometry, the microfluidic device channel was ~13 mm long (port-to-port), 630 μm wide, and 1.4 μm tall, containing ~200-nm-thick nanoporous resin structures at the channel bottom, increasing surface area (fig. S7). Bright field and fluorescence optical micrographs in Fig. 4, B and C, confirmed that tetramethylrhodamine (TRITC) dye dissolved in dimethyl sulfoxide (DMSO) was flowing through the microfluidic channel using a pressure-controlled pump.

We further demonstrated compatibility of the B-WRITE method with complex substrates obtained from established semiconductor processing
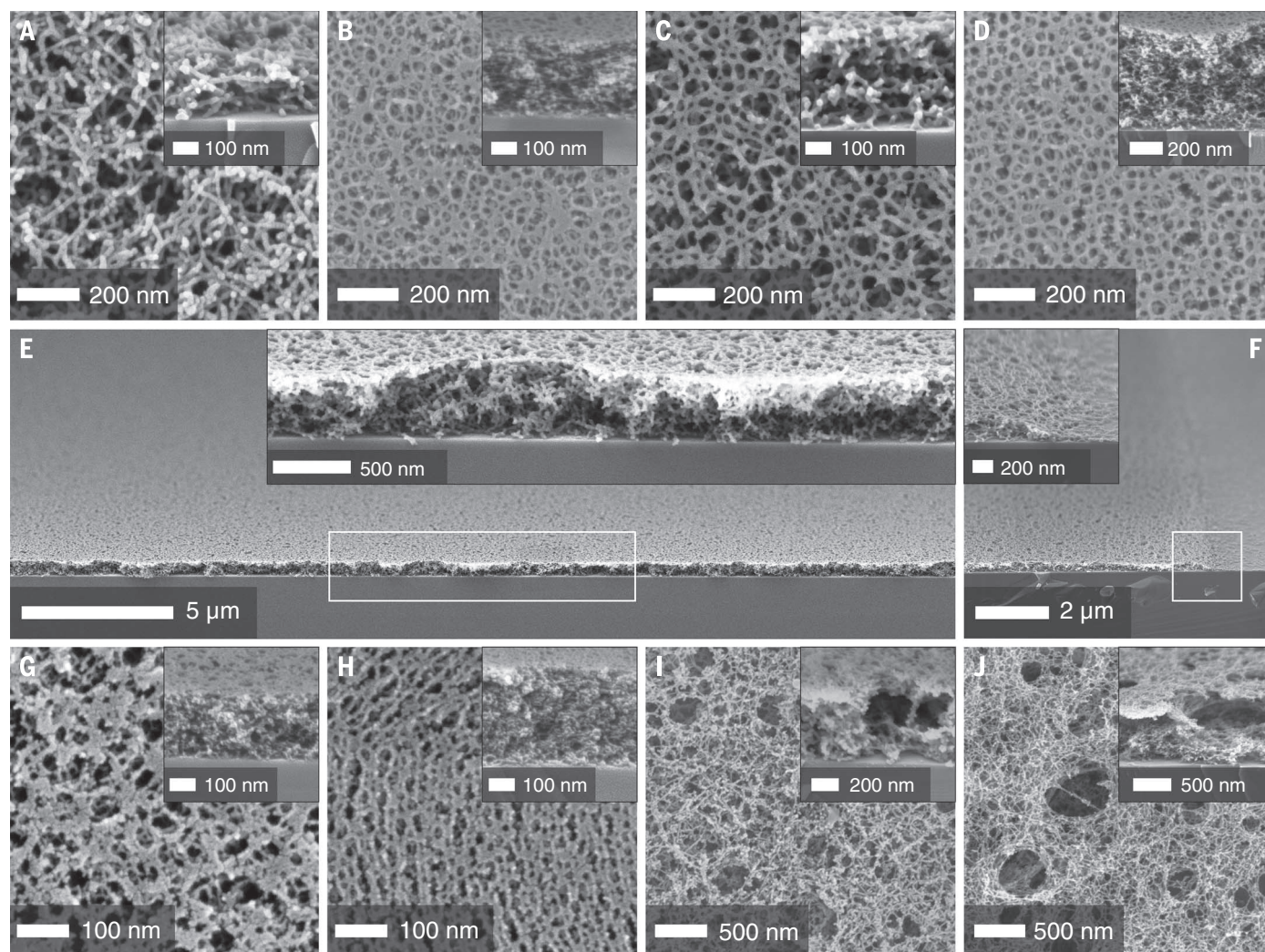

Fig. 2. SEM of B-WRITE resin structures. (**A** to **D**) Plan views and cross sections (insets) of laser-heated 6 wt % ISO-38-R [(A) and (B)] and 6 wt % ISO-69-R [(C) and (D)] solution–derived structures of 2 to 3:1 [(A) and (C)] and 1.5:1 [(B) and (D)] ISO/resols mass ratios, respectively. (**E** and **F**) Lower-magnification images of same sample as in (C). (**G** and **H**) Plan views and cross sections (insets) of laser-heated 20 wt % F127-R solution–derived structures of 1:1 (G) and 1:2 (H) F127/resols mass ratios, respectively. (**I** and **J**) Plan views and cross sections (insets) of laser-heated 10 wt % (I) and 12.5 wt % (J) ISO-38-R solution–derived structures.
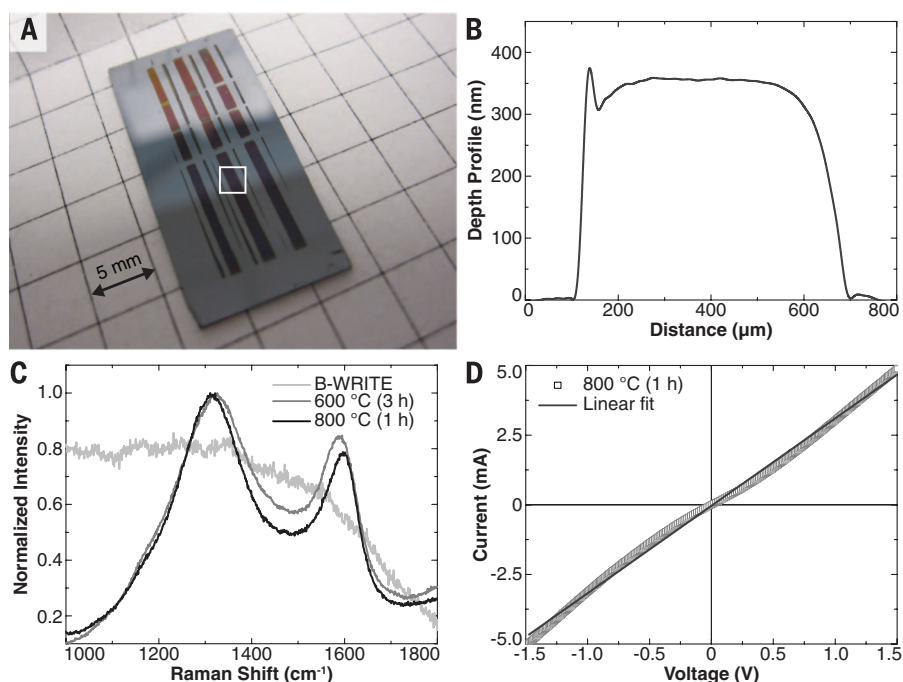
Fig. 3. Macroscopic pattern inversion and carbonization of B-WRITE structures. (**A** and **B**) Optical image of complex shapes (A) and corresponding depth profile (B) of laser-heated ISO-69-R after THF rinsing. Depth profile was acquired in the region denoted in (A). (**C**) Raman spectra (785-nm excitation) of laser-heated ISO-69-R formed by B-WRITE (light gray) and after carbonization at 600°C (gray) and 800°C (black). (**D**) Current/voltage plot of laser-heated porous carbon sample carbonized at 800°C.
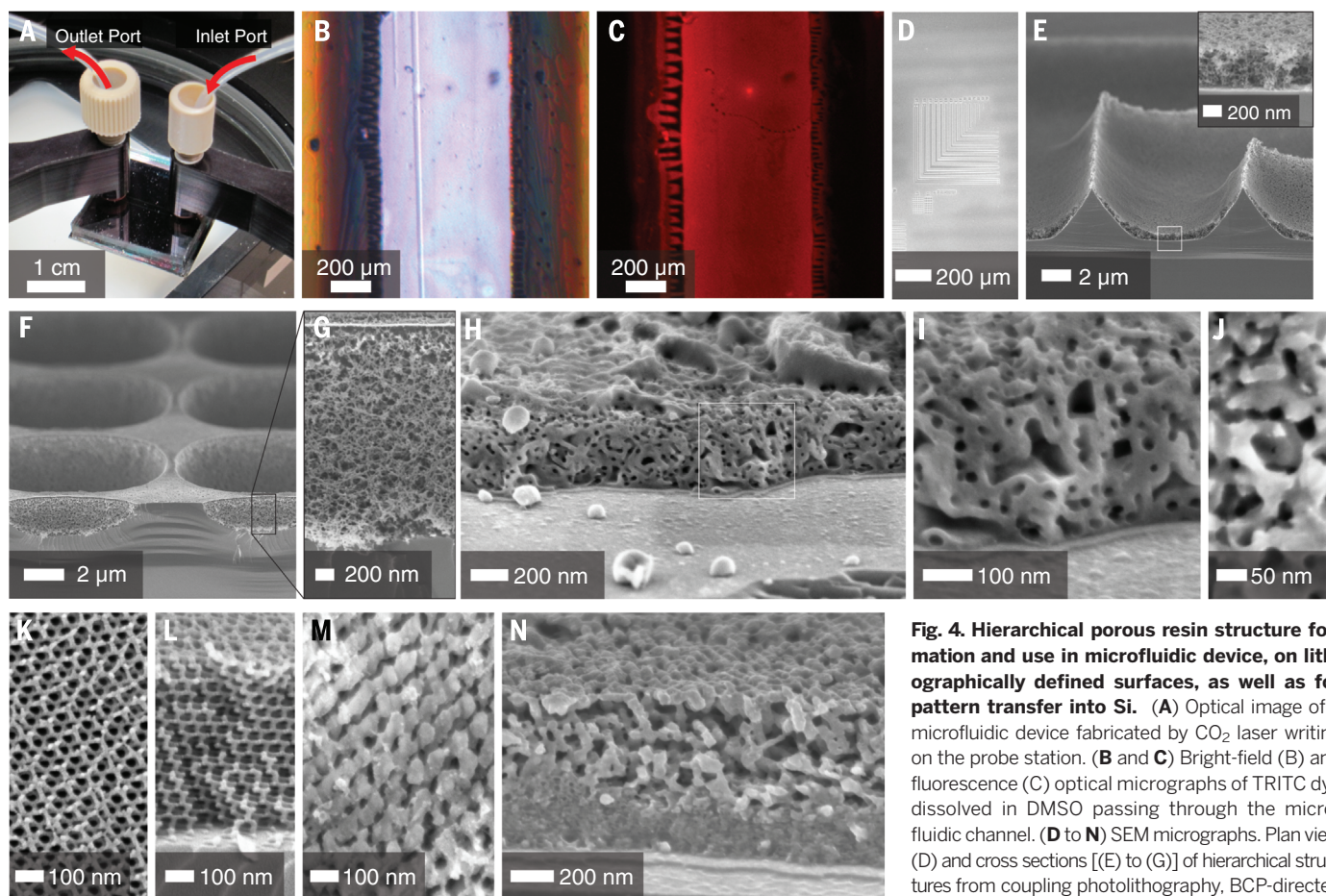

Fig. 4. Hierarchical porous resin structure formation and use in microfluidic device, on lithographically defined surfaces, as well as for pattern transfer into Si. (**A**) Optical image of a microfluidic device fabricated by CO_2 laser writing on the probe station. (**B** and **C**) Bright-field (B) and fluorescence (C) optical micrographs of TRITC dye dissolved in DMSO passing through the microfluidic channel. (**D** to **N**) SEM micrographs. Plan view (D) and cross sections [(E) to (G)] of hierarchical structures from coupling photolithography, BCP-directed self-assembly, and CO_2 laser heating. Inset in (E) shows the identified area at higher magnification. Cross section of excimer laser-induced crystalline Si nanostructures employing B-WRITE–derived organic template (H). Higher resolution cross section (I) and plan view (J) of area as indicated in (H). Plan views [(K) and (M)] and cross sections [(L) and (N)] of periodically ordered mesoporous gyroidal resin template [(K) and (L)] and resulting excimer laser-induced crystalline Si nanostructures after template removal [(M) and (N)].

technologies. Figure 4D displays a SEM plan view micrograph of an ISO-69-R hybrid sample deposited on a photolithographically patterned Si substrate and transformed into porous resin structures by the B-WRITE method. Cross-sectional SEM in Fig. 4, E to G, shows excellent conformal adhesion of the laser-derived structures despite the complex Si surface topography.

Finally, we demonstrated an all-laser-induced pattern transfer to create several-hundred-nanometer-thick 3D porous crystalline Si nanostructures using an excimer laser-induced Si melt-crystallization process described previously (10). To that end, B-WRITE–derived ISO-69-R structures were further heated for 1 hour at 400°C under N_2 to enhance adhesion to the substrate. The resin pore network was then backfilled with amorphous Si (a-Si) by chemical vapor deposition (CVD) (28) and irradiated with a single 40-ns pulsed XeCl excimer laser (λ = 308 nm) at 550 mJ/cm^2 to induce a transient 40- to 50-ns Si melt, which subsequently solidified into polycrystalline Si (fig. S8). SEM in Fig. 4, H to J, shows nanoporous polycrystalline Si nanostructures after template removal, using CF_4 reactive ion and wet acid etching treatments. What is notable about this experiment is the use of an organic BCP-directed resin template [as opposed to BCP-directed inorganic templates employed in (10)] that is not only stable for a-Si CVD at 350°C but also survives the transient melt of a-Si at temperatures ~1250°C, enabling high pattern-transfer fidelity (10, 29). All-organic BCP-directed structures open pathways to templates with high degrees of periodic order obtained—e.g., via thermal or solvent annealing before porous structure formation (30, 31). In a proof-of-concept experiment, SEM micrographs in Fig. 4, M and N, demonstrate that periodically ordered and networked mesoscopic crystalline Si nanostructures can be obtained after a-Si CVD into mesoporous gyroidal resin structures prepared by evaporation-induced self-assembly and pyrolysis (Fig. 4, K and L, and fig. S9) (21), excimer laser annealing, and template removal. Plan view (Fig. 4, K and M) and cross-sectional (Fig. 4, L and N) SEM images suggest that the interconnected and periodic pores of the mesoporous gyroidal resin template were conformally backfilled by the network struts of the resulting laser-induced crystalline Si nanostructures. Furthermore, the mesoporous all-organic resin template is inert to hydrofluoric (HF) acid, facilitating removal of native SiO_2 layers on the Si substrate using diluted HF (or HF vapor) before a-Si deposition and enabling formation of periodically ordered porous single-crystal epitaxial nanostructures (10).

REFERENCES AND NOTES

1. M. E. Davis, *Nature* **417**, 813–821 (2002).
2. U.S. Department of Energy Basic Energy Sciences Advisory Committee Mesoscale Science Committee W, *From Quanta to the Continuum: Opportunities for Mesoscale Science* (Government Printing Office, Washington, DC, 2012).
3. T. Yanagisawa, T. Shimizu, K. Kuroda, C. Kato, *Bull. Chem. Soc. Jpn.* **63**, 988–992 (1990).
4. C. T. Kresge, M. E. Leonowicz, W. J. Roth, J. C. Vartuli, J. S. Beck, *Nature* **359**, 710–712 (1992).
5. M. Templin et al., *Science* **278**, 1795–1798 (1997).
6. D. Zhao et al., *Science* **279**, 548–552 (1998).
7. S. H. Joo et al., *Nature* **412**, 169–172 (2001).
8. P. Yang, D. Zhao, D. I. Margolese, B. F. Chmelka, G. D. Stucky, *Nature* **396**, 152–155 (1998).
9. S. C. Warren et al., *Science* **320**, 1748–1752 (2008).
10. H. Arora et al., *Science* **330**, 214–219 (2010).
11. K.-V. Peinemann, V. Abetz, P. F. W. Simon, *Nat. Mater.* **6**, 992–996 (2007).
12. H. Sai et al., *Science* **341**, 530–534 (2013).
13. B. Jung et al., *ACS Nano* **6**, 5830–5836 (2012).
14. J. Lin et al., *Nat. Commun.* **5**, 5714 (2014).
15. S. Kawata, H.-B. Sun, T. Tanaka, K. Takada, *Nature* **412**, 697–698 (2001).
16. S. Maruo, J. T. Fourkas, *Laser & Photon. Rev.* **2**, 100–111 (2008).
17. L. Li, R. R. Gattass, E. Gershgoren, H. Hwang, J. T. Fourkas, *Science* **324**, 910–913 (2009).
18. J. Fischer, M. Wegener, *Laser & Photon. Rev.* **7**, 22–44 (2013).
19. Z. Gan, Y. Cao, R. A. Evans, M. Gu, *Nat. Commun.* **4**, 2061 (2013).
20. Materials and methods are available as supplementary materials on *Science* Online.
21. J. G. Werner, T. N. Hoheisel, U. Wiesner, *ACS Nano* **8**, 731–743 (2014).
22. Y. Meng et al., *Chem. Mater.* **18**, 4447–4464 (2006).
23. J. P. Singer et al., *Adv. Mater.* **25**, 6100–6105 (2013).
24. R. M. Silverstein, G. C. Bassler, T. C. Morrill, *Spectrometric Identification of Organic Compounds* (Wiley, New York, 1991).
25. M. F. El-Kady, V. Strong, S. Dubin, R. B. Kaner, *Science* **335**, 1326–1330 (2012).
26. F. S. Bates, *Science* **251**, 898–905 (1991).
27. R. O. Grisdale, *J. Appl. Phys.* **24**, 1288–1296 (1953).
28. S. A. Rinne, F. García-Santamaría, P. V. Braun, *Nat. Photonics* **2**, 52–56 (2008).
29. M. O. Thompson et al., *Phys. Rev. Lett.* **52**, 2360–2363 (1984).
30. S. Park et al., *Science* **323**, 1030–1033 (2009).
31. G. Deng et al., *Chem. Commun. (Camb.)* **50**, 12684–12687 (2014).

SULFUR CHEMISTRY

Gas phase observation and microwave spectroscopic characterization of formic sulfuric anhydride

Rebecca B. Mackenzie, Christopher T. Dewberry, Kenneth R. Leopold*

We report the observation of a covalently bound species, formic sulfuric anhydride (FSA), that is produced from formic acid and sulfur trioxide under supersonic jet conditions. FSA has been structurally characterized by means of microwave spectroscopy and further investigated by using density functional theory and ab initio calculations. Theory indicates that a $\pi_2 + \pi_2 + \sigma_2$ cycloaddition reaction between SO_3 and HCOOH is a plausible pathway to FSA formation and that such a mechanism would be effectively barrierless. We speculate on the possible role that FSA may play in the Earth's atmosphere.

There is an extensive literature on the chemistry of sulfur oxides and their derivatives. The area is rich in fundamental science and finds applications ranging from industrial chemistry to laboratory synthesis. Sulfur compounds are also active species in the atmosphere (*1*), and in particular, the oxides and oxyacids are important players in the formation of atmospheric aerosol (*2*). Here, we present a microwave spectroscopic study of the anhydride derived from formic and sulfuric acids, produced in a supersonic jet containing HCOOH and SO_3.

Department of Chemistry, University of Minnesota, 207 Pleasant Street SE, Minneapolis, MN 55455, USA.
*Corresponding author. E-mail: kleopold@umn.edu

The present work was stimulated by a series of studies concerned with the formation of sulfuric acid in the atmosphere. The acid, which can form via both gas-phase and aqueous-phase processes, is generated in the gas phase by the oxidation of SO_2 to SO_3, which is subsequently hydrated to give H_2SO_4

$$SO_3 + H_2O \rightarrow H_2SO_4 \qquad (1)$$

Both theoretical and experimental studies of this reaction indicate that viable mechanisms in the gas phase involve a facilitator molecule. For example, using ab initio theory, Morokuma and Muguruma (*3*) showed that the addition of
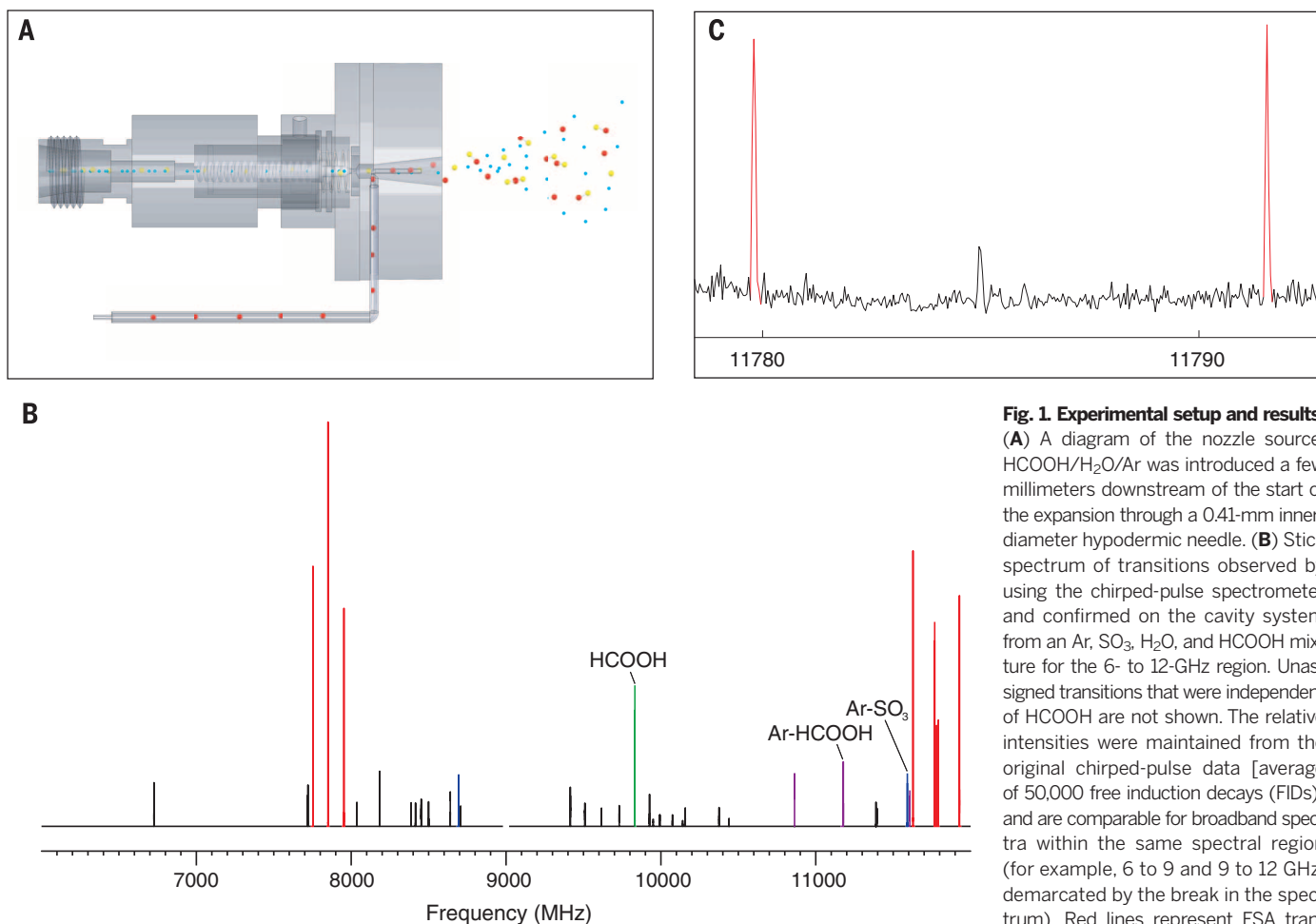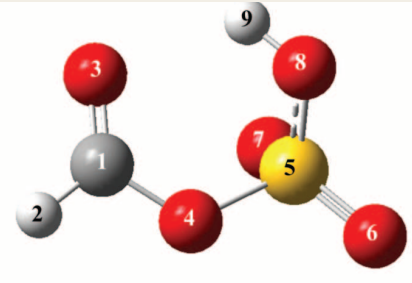
Fig. 1. Experimental setup and results. (**A**) A diagram of the nozzle source. HCOOH/H$_2$O/Ar was introduced a few millimeters downstream of the start of the expansion through a 0.41-mm inner-diameter hypodermic needle. (**B**) Stick spectrum of transitions observed by using the chirped-pulse spectrometer and confirmed on the cavity system from an Ar, SO$_3$, H$_2$O, and HCOOH mixture for the 6- to 12-GHz region. Unassigned transitions that were independent of HCOOH are not shown. The relative intensities were maintained from the original chirped-pulse data [average of 50,000 free induction decays (FIDs)] and are comparable for broadband spectra within the same spectral region (for example, 6 to 9 and 9 to 12 GHz, demarcated by the break in the spectrum). Red lines represent FSA transitions. Transitions assigned to Ar-SO$_3$, Ar-HCOOH, and HCOOH were left in the spectrum for comparison with FSA transitions and are marked in blue, pink, and green, respectively. The strongest transition observed for each is labeled on the figure. Black lines are unassigned. (**C**) Portion of the 9- to 12-GHz spectrum showing two FSA transitions highlighted in red.

a second water molecule to H$_2$O⋯SO$_3$ substantially lowers the activation barrier for Eq. 1, and indeed, kinetic data show a second-order dependence on water concentration (*4*). More recently, computational work by Hazra and Sinha has indicated that the conversion is essentially barrierless within the complex HCOOH⋯H$_2$O⋯SO$_3$ (*5*). Upon considering the concentration of formic acid, one of the most common atmospheric volatile organic compounds (VOCs), these authors argued that this pathway may be important for the formation of sulfuric acid. Furthermore, this mechanism terminates in the formation of the hydrogen-bonded complex H$_2$SO$_4$⋯HCOOH, which may be a preliminary step in a nucleation process. Indeed, the involvement of organic compounds in new particle formation has become a central topic in atmospheric particle research, and it now has been shown that organics participate not only in particle growth (*6–8*) but in nucleation as well (*9–15*). Carboxylic acids are abundant in the atmosphere (*16*), and thus, their primary interactions with sulfur-containing atmospheric species are of great interest.

Previous work in our laboratory has characterized a variety of atmospheric molecular complexes—including SO$_3$⋯H$_2$O (*17*), H$_2$SO$_4$⋯H$_2$O (*18*), and HNO$_3$⋯(H$_2$O)$_{n=1–3}$ (*19*)—in a supersonic jet by using microwave spectroscopy. Such systems, in general, are important to study because of the roles they play as intermediates in chemical reactions and precursors to atmospheric aerosol. Although conditions in the jet do not mimic those in the atmosphere, they can produce the same species, albeit under conditions amenable to microwave spectroscopy (in a collisionless

Table 1. Comparison of experimental and theoretical intermolecular distances.

Distance*	M06-2X†	MP2†	Experiment‡
C1-S5	2.589	2.593	2.594(5)
C1-H2	1.092	1.091	1.11(1)
C1-H9	2.634	2.572	2.62(1)
S5-H9	2.085	2.069	2.025(3)
S5-H2	3.514	3.520	3.513(2)

*All distances are in angstroms.   †M06-2X and MP2 calculations were done with 6-311++G(3df,3pd) basis set.   ‡Determined from Kraitchman's equations.

environment that is cold enough to ensure population of only the lowest rotational energy levels). Spectral analysis yields accurate information about gas-phase molecular and electronic structure, which in turn provides an important touchstone for computational studies. Therefore, inspired by Hazra and Sinha, and by the recent research suggesting the involvement of organic acids in aerosol formation, we set out to investigate the complexes such as $H_2SO_4 \cdots HCOOH$ and $SO_3 \cdots HCOOH$ by means of microwave spectroscopy in a supersonically expanded mixture of SO_3, H_2O, and HCOOH seeded in argon (Ar). What we found was entirely unexpected.

Rotational spectra were observed by means of Fourier transform microwave (FTMW) spectroscopy. Two methods were used: conventional cavity-type FTMW spectroscopy and the newer broadband, chirped-pulse FTMW technique. The cavity spectrometer in our laboratory has been described elsewhere (20). Briefly, molecules enter a tuned microwave cavity and are coherently excited by a 1- to 2-μs pulse of radiation. The resulting free induction decay is heterodyne-detected, digitized, and Fourier transformed so as to produce a signal in the frequency domain. Uncertainties in spectral transition frequencies are typically on the order of 2 to 3 kHz. In our broadband spectrometer (21), which applies the methods developed by Pate and coworkers (22), the cavity is eliminated, and the radiation for sample excitation is generated by up-conversion of a 0.2- to 3.2-GHz chirped pulse to the microwave spectral region of interest. A 20-W amplifier boosts the power before irradiation of the molecular sample, which is accomplished via a microwave horn. The resulting free induction decay is received by an identical companion horn, down-converted for digitization, and Fourier transformed so as to produce a 3-GHz-wide segment of the microwave spectrum. Linewidths are typically ~90 kHz, and the apparatus is less sensitive than that of the cavity spectrometer but it allows for rapid location and identification of spectra. In this work, spectra were initially recorded with the broadband spectrometer and subsequently remeasured at high resolution by using the cavity system.

A key feature of this experiment was the molecular source (21), which consisted of a supersonic expansion of SO_3 in Ar, to which a flowing stream of formic acid and water vapor was added. The Ar carrier gas was seeded with SO_3 by means of entrainment upon passage over solid, polymerized SO_3. A gaseous mixture of HCOOH and H_2O was added by bubbling Ar through 88% aqueous formic acid and injecting it into the expansion (Fig. 1A). This configuration provides "on-the-fly" mixing of reactive reagents during the first few tens of microseconds of the supersonic expansion.

Transitions from a variety of known species—including $Ar \cdots SO_3$, $H_2O \cdots SO_3$, H_2SO_4 (formed from SO_3 + H_2O), $H_2SO_4 \cdots H_2O$, $Ar \cdots HCOOH$, and $H_2O \cdots HCOOH$—were observable, as were numerous lines of unknown origin (Fig. 1B). Among these was a set of strong transitions whose pattern made them readily identifiable as the a-type spectrum of an asymmetric rotor, and a preliminary least-squares fit to a Watson A-reduced Hamiltonian gave residuals of <5 kHz. The appearance of these transitions required the presence of HCOOH, and our initial hypothesis was that they were due either to the $H_2SO_4 \cdots HCOOH$ or $SO_3 \cdots HCOOH$ weakly bound complex. However, theoretical predictions of the rotational constants for both species by using the density functional M06-2X (table S1) clearly indicated that neither is the carrier of the observed spectrum. Indeed, the computed rotational constants of $H_2SO_4 \cdots HCOOH$ as well as those of several weakly bound isomers of $SO_3 \cdots HCOOH$ (table S1 and fig. S3) differ by up to ~1 GHz and ~200 MHz, respectively, from those fitted from experimental data.

We performed additional calculations near the global minimum energy configuration of $SO_3 \cdots HCOOH$ with starting configurations at progressively shorter intermolecular separations. These eventually revealed a new chemically bonded species in which the acidic proton of the formic acid transfers to the SO_3 and a new sulfur-oxygen bond is formed (tables S2 and S3). We performed frequency calculations in order to verify the authenticity of the potential energy minimum, and the density functional theory (DFT) results were confirmed with MP2 calculations. The structure of this molecule, formic sulfuric anhydride (FSA), is shown in Table 1. This structure has substantial dipole moment components along each of its three inertial axes (μ_a, μ_b, and μ_c = 3.19, 0.48, and 0.97 D, respectively; μ_{Total} = 3.37 D), and with this information, b-type and c-type transitions were easily predicted and located. With a, b, and c-type transitions recorded, we performed a final fit, which yielded

Table 2. Comparison of experimental and calculated rotational constants for FSA.

Constant	Observed*	Theoretical*†	Difference*	Percent difference
A	4510.1155(12)	4542	−32	−0.7%
B	2013.28751(13)	2027	−14	−0.7%
C	1913.33629(13)	1927	−14	−0.7%

*All values are in megahertz. †Calculations were done with the M06-2X functional and 6-311++G (3df,3pd) basis set.
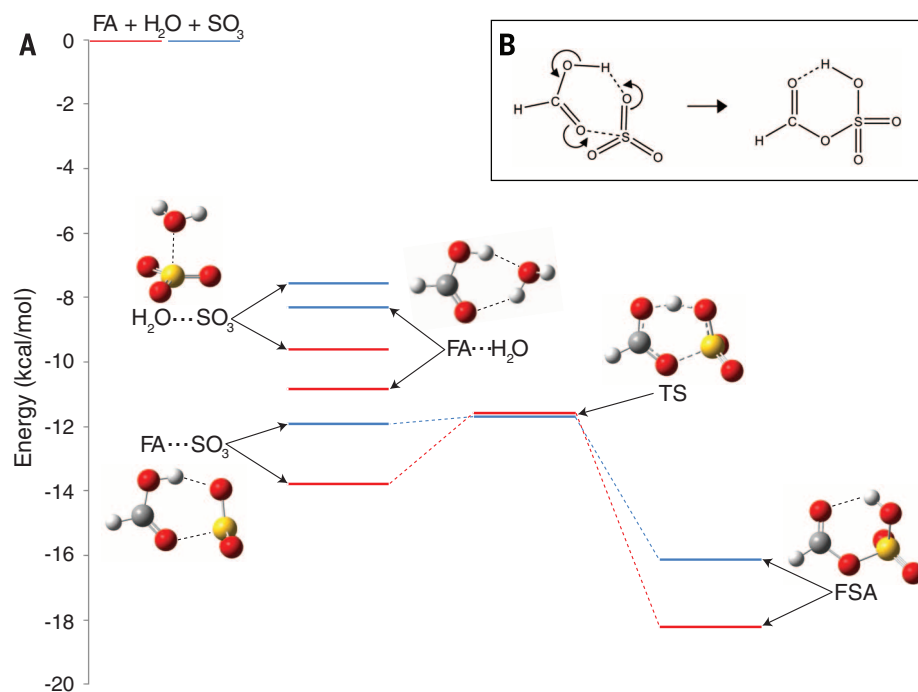

Fig. 2. FSA energetics. (**A**) Potential-energy landscape of the bimolecular complexes formed between H_2O, SO_3, and HCOOH (FA). Geometries were optimized by using MP2/6-311++G(3df,3pd), and single-point electronic energies were computed by using CCSD(T)/complete basis set with double and triple extrapolation [CBS(D-T)]. The red trace is uncorrected for zero-point energy (ZPE). The blue trace is ZPE-corrected with frequencies from the MP2/6-311++G(3df,3pd) calculations. Zero is defined as the sum of the monomer energies with and without ZPE corrections for the blue and red traces, respectively. The barrier to conversion from FA⋯SO_3 to FSA is 2.2 kcal/mol without ZPE corrections and 0.2 kcal/mol with ZPE corrections. (**B**) Possible formation mechanism of FSA from SO_3 and HCOOH.

the rotational constants listed in Table 2. As indicated in the table, the predicted rotational constants of FSA match the experimental values to within 0.7% (32 MHz for A and 14 MHz for B and C).
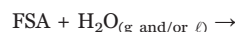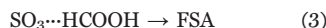
As a final test, we predicted and observed spectra of the ^{34}S-, ^{13}C-, and both monodeuterated isotopologues. We recorded ^{34}S- and ^{13}C-isotopologue spectra in natural abundance and used isotopically enriched HCOOD or DCOOH in experiments on the deuterated species. Transition frequencies and fitted spectroscopic constants for the observed isotopologues are provided in tables S4 to S9. All observed isotope shifts were in excellent agreement with those predicted from the theoretical structure. When we used HCOOD, the deuterium was found in the H9 position, confirming the occurrence (direct or indirect) of proton transfer (23); isotopic substitutions allowed determination of several interatomic distances by using Kraitchman's equations (24). These results are given and compared with the theoretical values in Table 1, where agreement is again seen to be excellent. The agreement between experimental and theoretical rotational constants, their isotope shifts, and the computed interatomic distances unambiguously establishes that the observed species is FSA.

The literature on compounds related to FSA appears sparse, although a few prior condensed phase studies are noteworthy. The sodium salt of acetic sulfuric anhydride has been described (25), but the parent acid (CH$_3$COOSO$_3$H) is reportedly unstable with respect to rearrangement or decomposition (26, 27). Dissolved salts of form [M^{n+}][SO$_3$OCHO$^-$]$_n$ have also been described in a patent concerning the preparation of isoflavones in a variety of nonaqueous solvents (28). We are unaware, however, of any previous gas-phase observations of FSA or its analogs. To better understand the pathway for the formation of FSA in the gas phase, we performed calculations (21) to locate the transition state connecting it with the SO$_3$···HCOOH van der Waals complex (table S10). For these calculations, using the optimized MP2 geometries, single-point coupled-cluster with single, double, and perturbative triple excitations [CCSD(T)] calculations were done by using the complete basis set extrapolation scheme of Neese and Valeev with the ANO-pVDZ to ANO-pVTZ basis sets (29). The transition state (TS in Fig. 2A) was found and corresponds to a concerted $\pi_2 + \pi_2 + \sigma_2$ cycloaddition (Fig. 2B), in which the acidic proton is transferred as the new S–O bond is formed. A similar reaction has been proposed for SO$_2$···HCOOH, but the formation of the resulting monomer, formic sulfurous anhydride, from SO$_2$ and HCOOH is endothermic (30). In the current case, however, FSA is 4.4 (4.2) kcal/mol lower in energy than the SO$_3$···HCOOH van der Waals complex, where the value in parentheses is zero-point corrected. The energy of the transition state is 2.2 (0.2) kcal/mol higher than that of SO$_3$···HCOOH, indicating that its conversion to FSA, with zero-point corrections, is essentially barrierless. In this light, it is not surprising that transitions of FSA were among the most prominent features in the observed spectrum. A second conformer of FSA, with the OH bond rotated 180°, was also identified from the M06-2X calculations (fig. S5). However, this local minimum lies 4.0 (3.6) kcal/mol above the global minimum of Table 1, and no spectra corresponding to this structure were identified, presumably because of insufficient population in the supersonic jet.

We also sought to establish the stability of the putative FSA precursor, SO$_3$···HCOOH, relative to that of other bimolecular complexes with potential roles as sulfuric acid and/or aerosol precursors. Computed energies of the complexes formed from SO$_3$, H$_2$O, and HCOOH are included in Fig. 2A, fig. S4, and table S10. With zero-point corrections, the SO$_3$···HCOOH complex is 3.6 and 4.4 kcal/mol more stable than H$_2$O···HCOOH and H$_2$O···SO$_3$, respectively. Similar calculations for H$_2$O···HCOOH and H$_2$O···SO$_3$ have been given elsewhere (3, 5) but are reproduced here in order to ensure comparisons at a uniform level of theory.

Both carboxylic and sulfonic acid anhydrides hydrolyze in water, and the hydrolysis of acetic sulfuric anhydride has been studied (27, 31). Thus, it is reasonable to hypothesize that the formation of SO$_3$···HCOOH with a barrierless conversion to FSA, followed by reaction with water and/or uptake into liquid droplets, may constitute an alternate pathway for H$_2$SO$_4$ production in the atmosphere

$$SO_3 + HCOOH \rightarrow SO_3 \cdots HCOOH \quad (2)$$

$$SO_3 \cdots HCOOH \rightarrow FSA \quad (3)$$

FSA + H$_2$O$_{(g \text{ and/or } \ell)}$ →
H$_2$SO$_{4(g \text{ and/or aq})}$ + HCOOH$_{(g \text{ and/or aq})}$ (4)

Although such a mechanism is not expected to dominate over established pathways involving direct reaction with water, it requires only the formation of a bimolecular—not a trimolecular—complex and may contribute to H$_2$SO$_4$ production, especially in areas with elevated concentrations of carboxylic acids. Additionally, the process described in Eq. 4 occurring in small water-containing clusters and/or liquid droplets may provide a pathway for the incorporation of volatile organic compounds into atmospheric aerosol. The mechanism may also be extendable to compounds of low volatility, which are more likely to contribute to prenucleation clusters (32). Indeed, we have performed additional calculations with larger carboxylic acids, which confirm the viability of their reaction to form FSA analogs (fig. S6). Moreover, laboratory studies that generated H$_2$SO$_4$ in situ have suggested that sulfur-containing species other than H$_2$SO$_4$ could be the initial nucleating agent (33). Under such a conclusion, FSA or its analogs could be nucleating agents themselves. Thus, although the atmospheric importance of Eqs. 2 to 4 is by no means certain, we suggest that scenarios involving FSA and larger sulfuric–carboxylic anhydrides should be explored in conjunction with models that involve trimolecular complexes such as (H$_2$O)$_2$···SO$_3$ and H$_2$O···HCOOH···SO$_3$ and in mechanisms for the early incorporation of organics in nucleation schemes. This could include, but is not limited to, mechanistic studies of the hydrolysis of sulfuric-carboxylic anhydrides in both water-containing clusters and bulk phase.

REFERENCES AND NOTES

1. R. Bentley, T. G. Chasteen, *Chemosphere* **55**, 291–317 (2004).
2. M. Kulmala et al., *Annu. Rev. Phys. Chem.* **65**, 21–37 (2014).
3. K. Morokuma, C. Muguruma, *J. Am. Chem. Soc.* **116**, 10316–10317 (1994).
4. C. E. Kolb et al., *J. Am. Chem. Soc.* **116**, 10314–10315 (1994).
5. M. K. Hazra, A. Sinha, *J. Am. Chem. Soc.* **133**, 17444–17453 (2011).
6. C. D. O'Dowd, P. Aalto, K. Hämeri, M. Kulmala, T. Hoffmann, *Nature* **416**, 497–498 (2002).
7. J. N. Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6634–6639 (2010).
8. J. Elm, T. Kurtén, M. Bilde, K. V. Mikkelsen, *J. Phys. Chem. A* **118**, 7892–7900 (2014).
9. R. Zhang et al., *Science* **304**, 1487–1490 (2004).
10. A. B. Nadykto, F. Yu, *Chem. Phys. Lett.* **435**, 14–18 (2007).
11. R. Zhang et al., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17650–17654 (2009).
12. A. Metzger et al., *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6646–6651 (2010).
13. P. Paasonen et al., *Atmos. Chem. Phys.* **10**, 11223–11242 (2010).
14. I. Riipinen et al., *Nat. Geosci.* **5**, 453–458 (2012).
15. S. Schobesberger et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17223–17228 (2013).
16. B. J. Finlayson-Pitts, J. N. Pitts, *Chemistry of the Upper and Lower Atmosphere* (Academic Press, San Diego, CA, 2000).
17. J. A. Phillips, M. Canagaratna, H. Goodfriend, K. R. Leopold, *J. Phys. Chem.* **99**, 501–504 (1995).
18. D. L. Fiacco, S. W. Hunt, K. R. Leopold, *J. Am. Chem. Soc.* **124**, 4504–4511 (2002).
19. G. Sedo, J. L. Doran, K. R. Leopold, *J. Phys. Chem. A* **113**, 11301–11310 (2009), and references therein.
20. J. A. Phillips et al., *J. Am. Chem. Soc.* **117**, 12549–12556 (1995).
21. Methods and materials are available as supplementary materials on *Science* Online.
22. G. G. Brown et al., *Rev. Sci. Instrum.* **79**, 053103 (2008).
23. Because formic acid inevitably contains water, experiments do not definitively establish whether the transformation of SO$_3$ + HCOOH occurs directly or whether an additional water molecule participates. However, the theoretical results of this work suggest that the reaction is possible without assistance.
24. W. Gordy, R. L. Cook, *Microwave Molecular Spectra* (John Wiley & Sons, New York, 1984).
25. A. J. van Peski, *Rec. Trav. Chim.* **40**, 103–118 (1921).
26. J. Russell, A. E. Cameron, *J. Am. Chem. Soc.* **60**, 1345–1348 (1938).
27. L. J. Tanghe, R. J. Brewer, *Anal. Chem.* **40**, 350–353 (1968).
28. D. C. Burdick, "Process for the Preparation of Isoflavones," International Publication Number WO 02/085881 A1, World Intellectual Property Organization, International Bureau (2002).
29. F. Neese, E. F. Valeev, *J. Chem. Theory Comput.* **7**, 33–43 (2011).
30. J. W. Keller, B. L. Harrod, S. A. Chowdhury, *J. Phys. Chem. A* **114**, 13182–13188 (2010).
31. S. J. Benkovic, R. C. Hevey, *J. Am. Chem. Soc.* **92**, 4971–4977 (1970).
32. M. Ehn et al., *Nature* **506**, 476–479 (2014).
33. T. Berndt et al., *Atmos. Chem. Phys.* **8**, 6365–6374 (2008).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6243/58/suppl/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S10
References (34–36)

ORGANIC CHEMISTRY

Catalytic asymmetric hydroamination of unactivated internal olefins to aliphatic amines

Yang Yang,[1]* Shi-Liang Shi,[1]* Dawen Niu,[1] Peng Liu,[2] Stephen L. Buchwald[1]†

Catalytic assembly of enantiopure aliphatic amines from abundant and readily available precursors has long been recognized as a paramount challenge in synthetic chemistry. Here, we describe a mild and general copper-catalyzed hydroamination that effectively converts unactivated internal olefins—an important yet unexploited class of abundant feedstock chemicals—into highly enantioenriched α-branched amines (≥96% enantiomeric excess) featuring two minimally differentiated aliphatic substituents. This method provides a powerful means to access a broad range of advanced, highly functionalized enantioenriched amines of interest in pharmaceutical research and other areas.

Unactivated internal olefins are manufactured on a very large scale during petroleum processing and thus represent excellent prospective building blocks for chemical synthesis. For example, 2×10^5 metric tons of 2-butene, the simplest member of this family, are produced annually through petroleum cracking and ethylene dimerization (*1*). In addition, the orthogonal reactivity of internal olefins with respect to carbonyls and other polar functional groups provides an exceptional opportunity for the late-stage diversification of complex molecules. In spite of these attractive attributes, few transition metal–catalyzed enantioselective transformations have been developed to convert unactivated internal olefins into broadly useful products with high levels of enantiocontrol (*2*, *3*). In particular, the catalytic asymmetric addition of a hydrogen atom and a functional group across an unactivated internal olefin (hydrofunctionalization) constitutes a class of chemical transformations of broad utility yet poses formidable challenges for synthetic chemists (*4*).

Several factors have impeded the development of efficient asymmetric hydrofunctionalization of unactivated internal olefins by using transition-metal hydride catalysts. The low binding affinity of internal olefins toward the metal center and the sluggish migratory insertion usually lead to intrinsically demanding hydrometalation; for a handful of well-tailored Co- (*5*) and Rh-based (*6*, *7*) catalyst systems that are capable of undergoing hydrometalation, the chain-walking mechanism consisting of iterative β-hydride elimination/migratory insertion steps rapidly convert the initially formed unstabilized secondary alkylmetal intermediate into the thermodynamically more stable, yet achiral, primary alkylmetal species. As a result, these processes are typically not amenable to enantioselective catalysis (*8*, *9*).

Enantioenriched amines represent key structural elements in a variety of natural products, pharmaceutical agents, agrochemicals, and functional materials. They also constitute an important family of chiral molecular scaffolds that are used stoichiometrically or catalytically to effect a diverse array of asymmetric transformations (*10*). Thus, the synthesis of optically pure amines has long been recognized as a preeminent goal for organic synthesis. Although enzymatic resolution (*11*) and stoichiometric chiral auxiliary-based methods (*12*) can be used to access α-branched aliphatic amines, these methods often require tedious multistep transformations from commercially available materials in order to arrive at the desired enantiopure amine products. To date, the general and modular catalytic assembly of α-branched aliphatic amines in a highly enantioselective fashion still presents a considerable challenge (*10*). Moreover, catalytic access to such compounds has proven elusive when the two aliphatic substituents are minimally differentiated by their steric and electronic properties (such as methyl- versus ethyl-).

The enantioselective hydroamination of alkenes has been successfully accomplished several times, albeit with a narrow substrate scope and/or moderate levels of enantioselectivity (*13*, *14*). Most notable are intramolecular processes (*15*, *16*) as well as intermolecular reactions of styrenes (*17*), terminal olefins (*18*, *19*), and norbornene (*20*, *21*). We have recently become interested in stereoselective amine synthesis by using copper-catalyzed olefin hydroamination (Fig. 1) (*22–27*). In this context, we envisioned the copper-catalyzed enantioselective hydroamination of unactivated internal olefins, if successful, would provide a general and powerful means of accessing enantioenriched α-branched aliphatic amines. The proposed catalytic cycle (Fig. 1C) commences with the enantioselective addition of a copper hydride species **I** across the double bond of the internal olefin **II** to provide the secondary alkylcopper intermediate **IIIa**. Electrophilic inter-

ception of the transient alkylcopper species with a hydroxylamine ester **IV** (*28*) in turn furnishes the enantioenriched aliphatic amine **Va** and releases the copper benzoate **VI**, which is reconverted to **I** upon reacting with hydrosilane **VII** (*29*). However, despite recent advances made in the field of copper(I) hydride–catalyzed asymmetric reduction (*30*, *31*), enantioselective transformations of unactivated internal olefins catalyzed by a copper hydride species have previously not been reported.

The hydrocupration of unactivated internal olefins is a very challenging process, and the impediments to this hydrocupration were evidenced by computational studies. As described in Fig. 1B, density functional theory (DFT) calculations with a 4,4′-Bi-1,3-benzodioxole-5,5′-diylbis(diphenyl-phosphane) (SEGPHOS) (**L1**)–based catalyst (**2a**) indicate that the hydrocupration of *trans*-2-butene (**1d**) is 7.9 and 2.8 kcal/mol more difficult than that of styrene (**1a**) and 1-propene (**1b**), respectively. These findings implied that unactivated internal olefin **1d** undergoes hydrocupration 6.6×10^5 times slower than that of styrene **1a** and 1.2×10^2 times slower than that of terminal olefin **1b**. Because competitive and unproductive reduction of the hydroxylamine ester **IV** is catalyzed by the same copper hydride species **I**, the sluggish hydrocupration of unactivated internal olefins could be detrimental to the hydroamination process (Fig. 1C). In further computations, we were encouraged to find that the barrier of hydrocupration of **1d** is 2.7 kcal/mol lower when **2a** is replaced by the DTBM-SEGPHOS (**L2**)–based catalyst (**2b**) that has previously shown enhanced capability for asymmetric induction in copper(I) hydride–mediated transformations (DTBM, 3,5-di-*tert*-butyl-4-methoxyphenyl) (*22*, *23*, *32*). On the basis of these results and our previous work (*22*, *23*), we posited that the use of this catalyst, which is highly active for hydrocupration yet still capable of effectively discerning the small steric bias between the prochiral faces of the internal olefin, would represent the key to the successful implementation of our proposed enantioselective hydroamination.

We began our investigation by exploring our previously developed catalyst systems for the enantioselective hydroamination of terminal olefins using *trans*-4-octene as the model substrate (Fig. 1D). Although the desired hydroamination product was not observed when (S)-SEGPHOS (**L1**) was used, we found that in the presence of 5 mole percent (mol %) (S)-DTBM-SEGPHOS–based catalyst, the desired amine product was formed in 34% yield with excellent enantioselectivity (98% enantiomeric excess). None of the isomeric amine products (**6b**, **6c**, and **6d**) originating from the chain-walking isomerization of the initially formed secondary alkylcopper species was observed (*33*). In an effort to further improve the catalytic efficiency and suppress the competitive reduction of **5**, we reasoned that the use of an electrophilic aminating reagent bearing a more electron-rich benzoyl group that is less prone to reduction yet exhibits comparable or even enhanced reactivity toward the nucleophilic

[1]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [2]Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, USA.
*These authors contributed equally to this work. †Corresponding author. E-mail: sbuchwal@mit.edu

alkylcopper intermediate could serve to address the problem. In addition, we postulated that the copper benzoate **VI** bearing an electron-rich benzoate would undergo faster σ-bond metathesis with the hydrosilane **VII** to regenerate the CuH species, accelerating catalyst turnover. Ultimately, fine-tuning of the electronic nature of the benzoyl group led to the identification of **5d** as the optimal electrophilic aminating agent. It is practically advantageous that **5d**-type electrophilic aminating reagents feature enhanced stability upon storage and could be conveniently prepared from the parent hydroxylamines and commercially available 4-(*N*,*N*-diethylamino)benzoic acid on a large scale in a single operation.

Having established the optimal reaction conditions for this copper-catalyzed hydroamination, we focused our effort on the substrate scope of this transformation (Fig. 2). We were particularly interested in the use of 2-butene as an

alkylating reagent to access branched aliphatic amines of high enantiopurity (Fig. 2A). The creation of "methyl-ethyl" stereocenters in a highly enantioselective fashion has widely been considered as a premier challenge in asymmetric catalysis in view of the minimal steric differentiation between the two substituents (*34*). In particular, catalytic asymmetric synthesis of methyl-ethyl α-branched tertiary amines is highly challenging. We found that in the presence of 2-butene, a wide range of hydroxylamine esters could be effectively converted into the corresponding chiral tertiary amines in a highly enantioselective manner. Electron-rich (**9a** and **9b**) and electron-deficient (**9c**) hydroxylamine esters represented compatible coupling partners. Furthermore, a broad range of sensitive functional groups such as an aniline (**9b**), an ester (**9d**), a phenol (**9d**), an alcohol (**9h**), an amide (**9j**), an acetal (**9l**), and a cyclic trisub-

stituted olefin (**9m**) were tolerated under these conditions. Additionally, the successful incorporation of functional group handles such as a boronate (**9e**) and a chloride (**9q** and **9r**) into the amine fragment provided opportunities for the further transformations of these valuable products by using modern cross-coupling techniques (*35*). Moreover, substrates bearing a variety of privileged heterocyclic motifs that are frequently found in medicinal agents—including indole (**9f**), thiophene (**9g**), furan (**9h**), pyridine (**9i**), and pyrimidine (**9n**)—could be successfully transformed into the corresponding tertiary amines with high levels of enantiocontrol. In addition to acyclic substrates, cyclic aminating reagents could also be accommodated, providing the desired product in excellent enantiomeric excess (**9n**). Last, the current catalyst system demonstrated excellent ability to control the diastereoselectivity when chiral electrophilic aminating
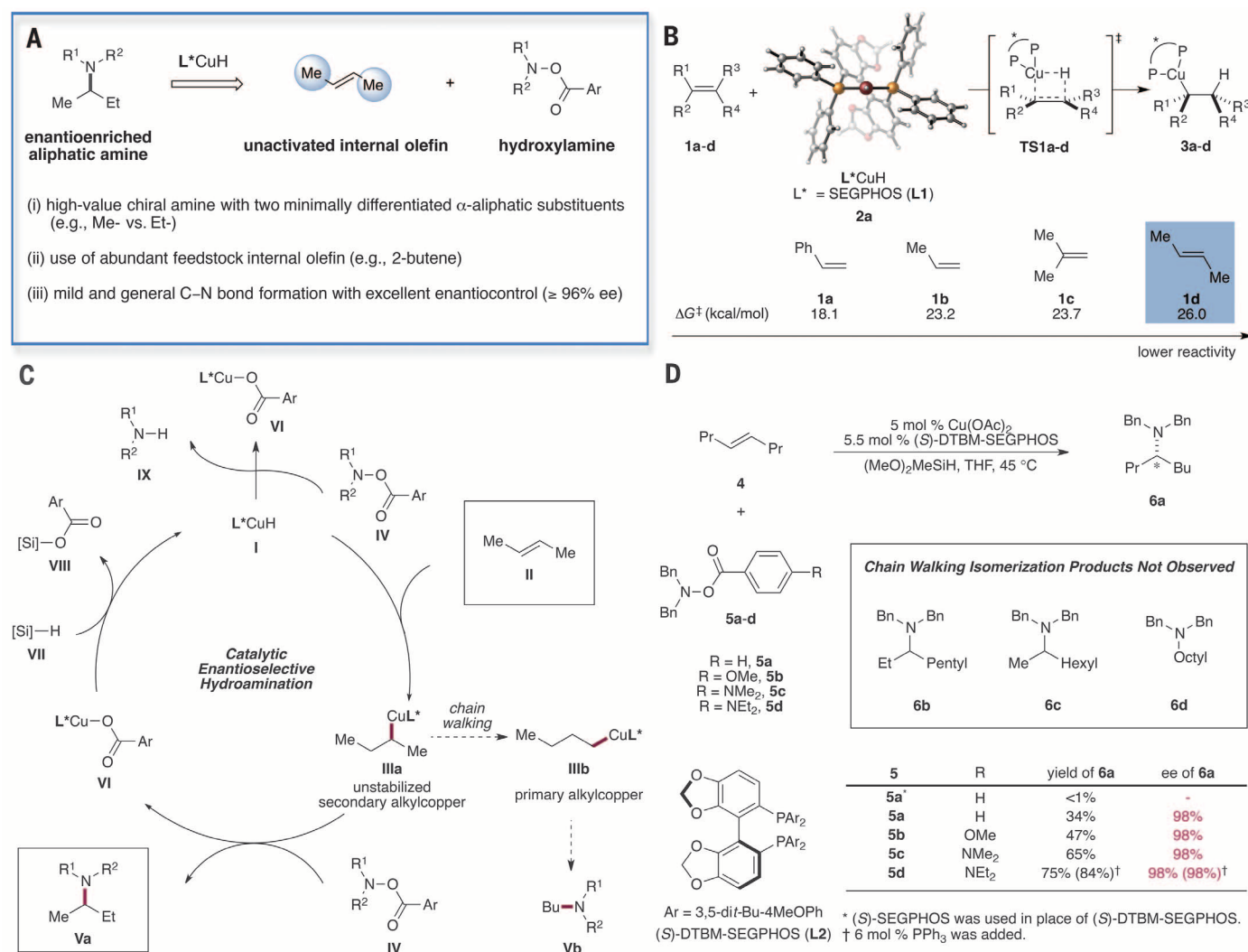

Fig. 1. Proposed asymmetric hydroamination of unactivated internal olefins to access enantioenriched branched aliphatic amines. (**A**) Advantageous properties of the reaction profile. (**B**) DFT-calculated activation barriers for the hydrocupration of olefins **1a** to **1d**. Energies are computed at the M06/SDD-6-311+G(d,p)/SMD (THF) level, with geometries optimized at the B3LYP/SDD-6-31G(d) level. (**C**) Proposed catalytic cycle. (**D**) Optimization studies. Reactions were performed by using **4** (0.60 mmol), **5** (0.20 mmol), (MeO)$_2$MeSiH (0.60 mmol), Cu(OAc)$_2$ (5 mol %), and **L** (5.5 mol %) in tetrahydrofuran (THF) (1.0 M) at 45°C for 36 hours. Yields were determined by means of gas chromatography analysis using dodecane as the internal standard.
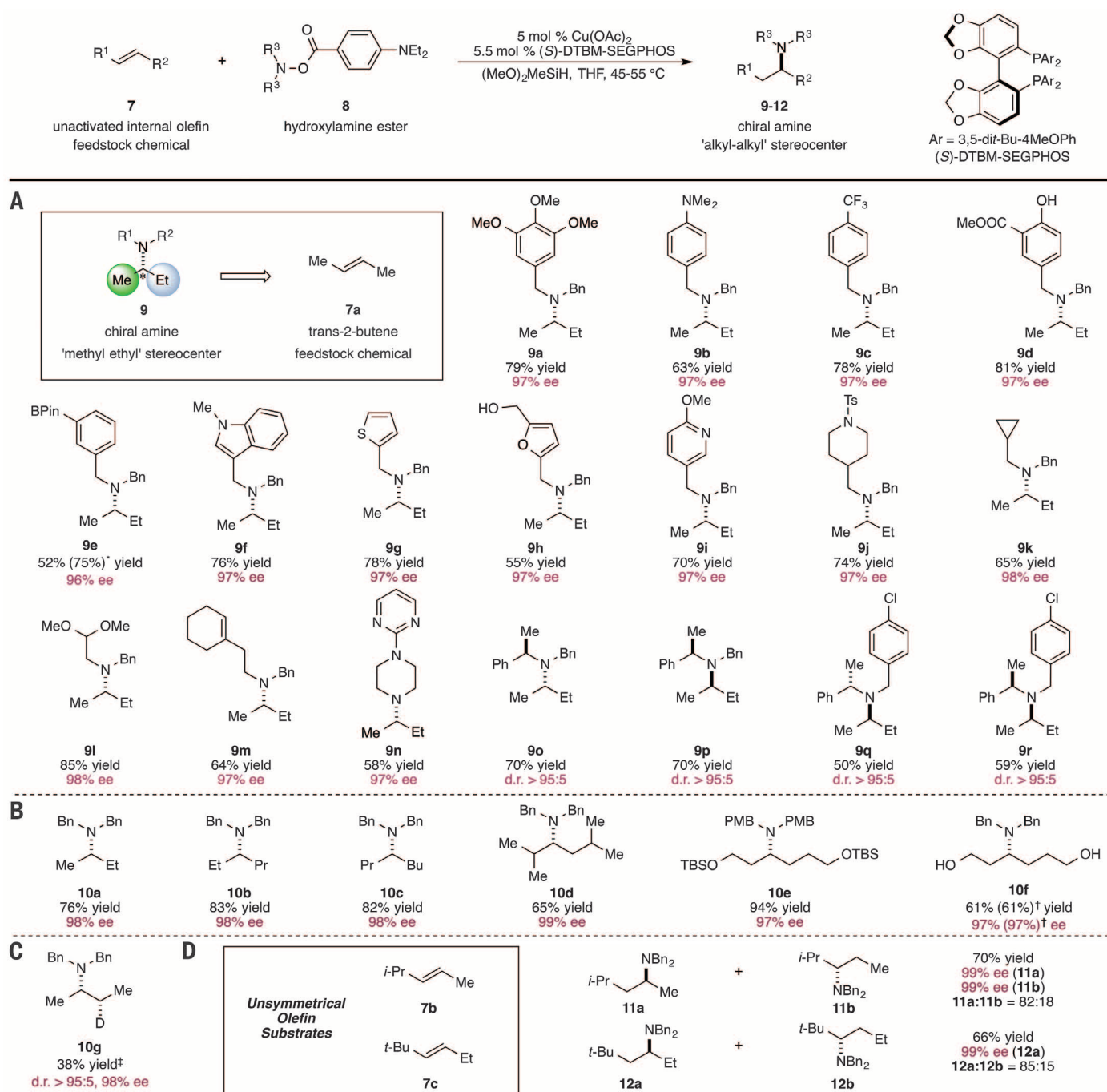
Fig. 2. Substrate scope of the copper-catalyzed enantioselective hydroamination of internal olefins. (**A**) Asymmetric hydroamination of 2-butene with a variety of electrophilic amines. (**B**) Scope of symmetrical internal olefins. (**C**) Deuterium incorporation. (**D**) Regioselectivity in the hydroamination of unsymmetrical internal olefins. Yields refer to isolated yields on the average of two runs. Enantiomeric excesses were determined by means of chiral high-performance liquid chromatography analysis or by using Swager's protocol (38). Asterisk indicates that the yield in parentheses was determined by means of ^{1}H nuclear magnetic resonance analysis. Dagger symbol (†) indicates that the reaction was performed on a 5-mmol scale with 1 mol % Cu(OAc)$_2$, 1.1 mol % (S)-DTBM-SEGPHOS, and 2 mol % PPh$_3$. Double-dagger symbol (‡) indicates that Ph$_2$SiD$_2$ was used in lieu of (MeO)$_2$SiMeH. d.r., diastereomeric ratio.

reagents bearing stereocenters adjacent to the nitrogen atom were applied (**9o** to **9r**).

We next surveyed the scope of unactivated internal olefins that could be used for the current transformation (Fig. 2B). A broad range of abundant and commercially available internal olefins was found to be suitable substrates, affording the corresponding chiral amines featuring two highly similar α-substituents in excellent enantioselectivities (**10a** to **10d**). In addition, internal olefins bearing branching secondary alkyl substituents (**10d**) readily participated in this transformation. Functional groups such as a silyl ether (**10e**) and a free alcohol (**10f**) were also compatible under these conditions. Moreover, we were able to perform this transformation on a 5-mmol scale while simultaneously lowering the catalyst loading to 1 mol % without any detrimental effect on the yield and the enantioselectivity (**10f**), demonstrating the scalability and practicality of this method. Furthermore, by capitalizing on the intrinsic flexibility of the hydrofunctionalization process, pharmacologically relevant β-deuteroamines (**10g**) (36) could be conveniently accessed in a highly
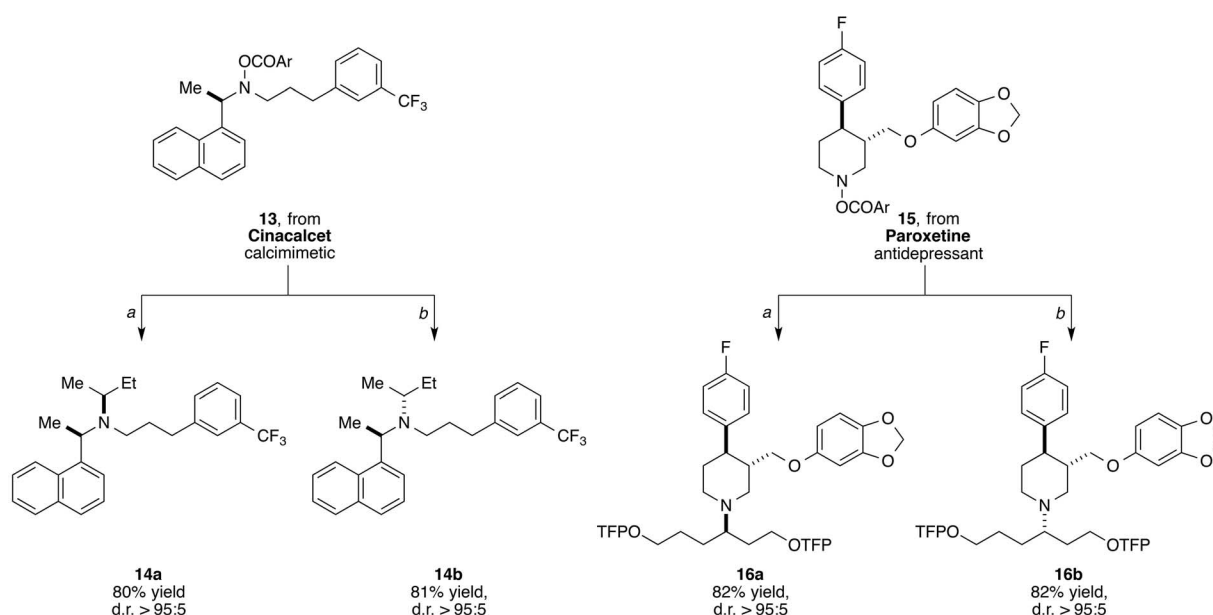
Fig. 3. Diversification of pharmaceutical agents. Conditions are **a**, 5 mol % Cu(OAc)$_2$, 5.5 mol % (*S*)-DTBM-SEGPHOS, (MeO)$_2$MeSiH, internal olefin, THF, 45°C; **b**, (*R*)-DTBM-SEGPHOS was used instead of (*S*)-DTBM-SEGPHOS. TFP, 2-(5-trifluoromethyl)pyridyl.
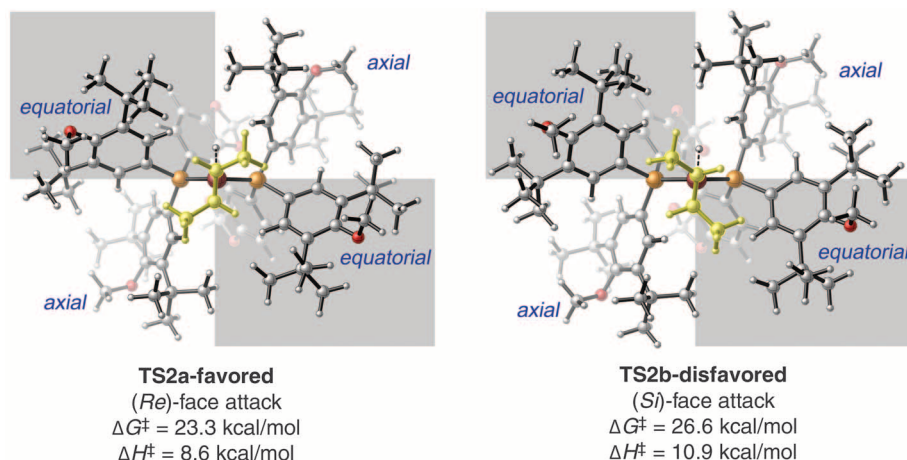

Fig. 4. Transition-state structures of the enantioselectivity-determining hydrocupration step with (*R*)-DTBM-SEGPHOS. Energies are computed at the M06/SDD-6-311+G(d,p)/SMD(THF) level, with geometries optimized at the B3LYP/SDD-6-31G(d) level.

diastereo- and enantioselective manner through the use of a commercially available source of deuteride (Ph$_2$SiD$_2$). Last, when unsymmetrical internal olefins (**7b** and **7c**) were used, the hydroamination proceeded in a regioselective manner, affording tertiary amines with uniformly high levels of enantiocontrol (**11** and **12**).

A central advantage of the current catalytic enantioselective hydroamination is its ability for the late-stage modification of advanced, highly functionalized synthetic intermediates. To demonstrate this strategy, we selected two widely prescribed pharmaceutical drugs, cinacalcet and paroxetine, and subjected the electrophilic amines (**13** and **15**) derived from these drugs to our reaction conditions (Fig. 3). All of these amine-

containing complex molecular scaffolds underwent effective hydroamination, with unactivated internal olefins to furnish either of the diastereomeric tertiary amine products in a catalyst-controlled fashion, exemplifying the potential of the current transformation in the rapid and efficient diversification of medicinally relevant molecules.

To gain insight into the origin of the high levels of enantiocontrol observed with our system, we performed DFT calculations on the hydrocupration of *trans*-2-butene that is likely enantiodetermining (Fig. 4). The hydrocupration with (*R*)-DTBM-SEGPHOS involves a four-membered transition state in which the hydride is delivered to the olefinic carbon while the Cu–C bond is

formed simultaneously (*37*). In this pseudo-four-coordinate transition state, the two P atoms on the ligand are perpendicular to the C=C double bond in the substrate. The C_2-symmetric bidentate phosphine ligand adopts the axial-equatorial coordination. The equatorial phosphine substituents are in closer contact with the olefin substrate and thus play a dominant role in enantiodiscrimination. In transition state **TS2a** involving the (*Re*)-face attack of the internal olefin, the unfavorable steric repulsion of both of the olefin's aliphatic substituents with the ligand's equatorial P substituents is minimized, leading to the formation of the preferred (*R*)-alkylcopper intermediate. The sterically encumbered *tert*-butyl groups present in the backbone of DTBM-SEGPHOS accentuate the energy differences between the (*Re*)- and the (*Si*)-face attack transition states, giving rise to an activation barrier difference ($\Delta\Delta G^{\ddagger}$) of 3.3 kcal/mol in favor of the (*Re*)-face attack pathway (**TS2a**). We expect this enantioselective hydroamination process to find broad application among practitioners of synthetic and pharmaceutical sciences. In addition to the utility of this protocol, we anticipate that the distinct reactivity of unactivated internal olefins revealed in the current reaction manifold will inspire the further advances in a range of enantioselective hydrofunctionalization processes using internal olefins.

REFERENCES AND NOTES

1. F. M. A. Geilen, G. Stochniol, S. Peitz, E. Schulte-Koerne, in *Butenes. Ullmann's Encyclopedia of Industrial Chemistry*, F. Obenaus, W. Droste, J. Neumeister, Eds. (Wiley, New York, 2014), pp. 1–13.
2. S. Bell *et al.*, *Science* **311**, 642–644 (2006).
3. T.-S. Mei, H. H. Patel, M. S. Sigman, *Nature* **508**, 340–344 (2014).
4. V. P. Anaikov, M. Tanaka, *Top. Organomet. Chem.* **43**, 1–20 (2013).
5. J. V. Obligacion, P. J. Chirik, *J. Am. Chem. Soc.* **135**, 19107–19110 (2013).

6. L. A. van der Veen, P. C. J. Kamer, P. W. N. M. van Leeuwen, *Angew. Chem. Int. Ed.* **38**, 336–338 (1999).
7. S. Pereira, M. Srebnik, *J. Am. Chem. Soc.* **118**, 909–910 (1996).
8. N. Sakai, K. Nozaki, H. Takaya, *J. Chem. Soc. Chem. Commun.* (4): 395–396 (1994).
9. T. Gadzikwa, R. Bellini, H. L. Dekker, J. N. H. Reek, *J. Am. Chem. Soc.* **134**, 2860–2863 (2012).
10. T. C. Nugent, *Chiral Amine Synthesis: Methods, Developments and Applications* (Wiley VCH, Weinheim, Germany, 2010).
11. U. T. Bornscheuer, R. J. Kazlauskas, *Hydrolases in Organic Synthesis* (Wiley VCH, Weinheim, Germany, 2006).
12. M. T. Robak, M. A. Herbage, J. A. Ellman, *Chem. Rev.* **110**, 3600–3740 (2010).
13. T. E. Müller, K. C. Hultzsch, M. Yus, F. Foubelo, M. Tada, *Chem. Rev.* **108**, 3795–3892 (2008).
14. P. W. Roesky, T. E. Müller, *Angew. Chem. Int. Ed.* **42**, 2708–2710 (2003).
15. M. R. Gagné *et al.*, *Organometallics* **11**, 2003–2005 (1992).
16. S. Hong, T. J. Marks, *Acc. Chem. Res.* **37**, 673–686 (2004).
17. M. Kawatsura, J. F. Hartwig, *J. Am. Chem. Soc.* **122**, 9546–9547 (2000).
18. Z. Zhang, S. D. Lee, R. A. Widenhoefer, *J. Am. Chem. Soc.* **131**, 5372–5373 (2009).
19. A. L. Reznichenko, H. N. Nguyen, K. C. Hultzsch, *Angew. Chem. Int. Ed.* **49**, 8984–8987 (2010).
20. R. Dorta, P. Egli, F. Zürcher, A. Togni, *J. Am. Chem. Soc.* **119**, 10857–10858 (1997).
21. J. S. Zhou, J. F. Hartwig, *J. Am. Chem. Soc.* **130**, 12220–12221 (2008).
22. S. Zhu, N. Niljianskul, S. L. Buchwald, *J. Am. Chem. Soc.* **135**, 15746–15749 (2013).
23. S. Zhu, S. L. Buchwald, *J. Am. Chem. Soc.* **136**, 15913–15916 (2014).
24. S.-L. Shi, S. L. Buchwald, *Nat. Chem.* **7**, 38–44 (2015).
25. N. Niljianskul, S. Zhu, S. L. Buchwald, *Angew. Chem. Int. Ed.* **54**, 1638–1641 (2015).
26. Y. Mikki, K. Hirano, T. Satoh, M. Miura, *Angew. Chem. Int. Ed.* **52**, 10830–10834 (2013).
27. R. P. Rucker, A. M. Whittaker, H. Dang, G. Lalic, *J. Am. Chem. Soc.* **134**, 6571–6574 (2012).
28. A. M. Berman, J. S. Johnson, *J. Am. Chem. Soc.* **126**, 5680–5681 (2004).
29. D. Lee, J. Yun, *Tetrahedron Lett.* **45**, 5415–5417 (2004).
30. D. H. Apella, Y. Moritani, R. Shintani, E. M. Ferreira, S. L. Buchwald, *J. Am. Chem. Soc.* **121**, 9473–9474 (1999).
31. C. Deutsch, N. Krause, *Chem. Rev.* **108**, 2916–2927 (2008).
32. B. H. Lipshutz, K. Noson, W. Chrisman, A. Lower, *J. Am. Chem. Soc.* **125**, 8779–8789 (2003).
33. G. M. Whitesides, E. R. Stedronsky, C. P. Casey, J. San Filippo Jr., *J. Am. Chem. Soc.* **92**, 1426–1427 (1970).
34. K. R. Fandrick *et al.*, *J. Am. Chem. Soc.* **133**, 10332–10335 (2011).
35. A. de Meijere, F. Diederich, *Metal-Catalyzed Cross-Coupling Reactions* (Wiley VCH, Weinheim, 2004).
36. T. G. Gant, *J. Med. Chem.* **57**, 3595–3611 (2014).
37. J. Won, D. Noh, J. Yun, J. Y. Lee, *J. Phys. Chem. A* **114**, 12112–12115 (2010).
38. Y. Zhao, T. M. Swager, *J. Am. Chem. Soc.* **137**, 3221–3224 (2015).

METALLOPROTEINS

A tethered niacin-derived pincer complex with a nickel-carbon bond in lactate racemase

Benoît Desguin,[1] Tuo Zhang,[2] Patrice Soumillion,[3] Pascal Hols,[3] Jian Hu,[2,4]* Robert P. Hausinger[1,2]*

Lactic acid racemization is involved in lactate metabolism and cell wall assembly of many microorganisms. Lactate racemase (Lar) requires nickel, but the nickel-binding site and the role of three accessory proteins required for its activation remain enigmatic. We combined mass spectrometry and x-ray crystallography to show that Lar from *Lactobacillus plantarum* possesses an organometallic nickel-containing prosthetic group. A nicotinic acid mononucleotide derivative is tethered to Lys[184] and forms a tridentate pincer complex that coordinates nickel through one metal-carbon and two metal-sulfur bonds, with His[200] as another ligand. Although similar complexes have been previously synthesized, there was no prior evidence for the existence of pincer cofactors in enzymes. The wide distribution of the accessory proteins without Lar suggests that it may play a role in other enzymes.

L actate dehydrogenases are nearly ubiquitous enzymes that reduce pyruvate to form either D- or L-lactic acid. The D-isomer is specifically incorporated into the cell wall of certain bacteria, where it can contribute to vancomycin resistance (*1*). Lactate racemase (Lar), which interconverts the D- and L-isomers of lactic acid, was recently shown to use nickel as a cofactor (*2*), bringing the number of known Ni-dependent enzymes to nine (*3*). The structure of the recombinant apoprotein from *Thermoanaerobacterium thermosaccharolyticum* (LarA$_{Tt}$) lacks the cofactor, leaving open questions of the metallocenter composition and coordinating ligands. Lar activity requires the products of the structural gene (*larA*) and three accessory genes (*larB*, *larC*, and *larE*). LarE is a putative adenosine 5′-triphosphate (ATP)–using enzyme of the PP-loop superfamily that, when isolated from cells also producing LarB and LarC, is able to activate Lar apoprotein in vitro (*2*). The requirement for these accessory proteins and the absence of an identified Ni-binding site in the reported LarA$_{Tt}$ structure were compatible with the presence of a Ni-binding cofactor in Lar.

To interrogate the nature of Ni-binding in Lar, we used liquid chromatography–electrospray ionization–mass spectrometry (LC-ESI-MS) to identify a prosthetic group on the enzyme. We also determined the x-ray structure of LarA containing the cofactor. Biochemical characterization of *Lactobacillus plantarum* lactate

racemase (LarA) was initially stymied by the loss of activity and Ni content in the first few hours after LarA protein purification from recombinant *Lactococcus lactis* cells (*2*). To circumvent this problem, we included sulfite in the purification. This compound was identified as a potent mixed inhibitor (fig. S1) and was shown both to stabilize LarA activity and delay Ni loss (fig. S2, A and B). Sulfite most probably protects LarA from oxidation because incubation of the enzyme in microaerobic environments had an equivalent effect (fig. S2C), whereas other reducing agents [dithiothreitol and *tris* (2-carboxyethyl)phosphine] did not protect Lar activity. LarA purified in the presence of sulfite showed greater Ni content and higher specific activity (table S1), and its absorbance spectrum revealed electronic transitions at 375, 440, and 550 nm (fig. S2D). These features were absent in LarA apoprotein obtained from cells not expressing the accessory proteins (fig. S2D), suggesting that the chromophore is associated with the metallocenter.

Protein molecular mass differences observed by using LC-ESI-MS further demonstrate the presence of an additional compound in LarA holoprotein. Whereas the mass of LarA apoprotein corresponded to its predicted 47,393 daltons (for protein missing its *N*-terminal Met) (Fig. 1A), the LarA holoprotein in the presence of 50 μM sulfite was 451 ± 1 daltons greater (Fig. 1B). Because of the LC purification prior to injection, the protein only retained strongly attached compounds, and hence, sulfite or other loosely bound molecules would not account for this mass increase. When incubated for 1 day at room temperature in buffer containing only 0.1 μM sulfite, a substantial portion of LarA holoprotein decreased in mass by 30 ± 1 and 56 ± 1 daltons (Fig. 1C), indicating

[1]Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824, USA. [2]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, USA. [3]Institute of Life Sciences, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium. [4]Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA.
*Corresponding author. E-mail: hujian1@msu.edu (JH); hausinge@msu.edu (RPH)

a degradation of the compound that is accompanied by loss of Lar activity (fig. S2A).

To further investigate whether the additional mass in LarA is associated with the enzyme activity and Ni content, we used mutagenesis to substitute eight of the most conserved and presumed active site residues (table S1). Four *larA* mutants (encoding the H108A, H172A, H200A, and K298A LarA variants) essentially abolished Lar activity in cell extracts, whereas another four mutants (encoding D72A, R75A, K184A, and Q295A LarA variants) showed residual activity in vivo, with the purified R75A and Q295A variants retaining in vitro activity (table S1). (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. In the variants, other amino acids were substituted at certain locations; for example, H108A indicates that histidine at position 108 was replaced by alanine.) All variants that could be purified (R75A, H172A, K184A, and Q295A) exhibited reduced Ni content, but only K184A LarA completely lacked both Ni and activity—which is equivalent to apoprotein (table S1). Moreover, the mass of K184A LarA corresponded to that of the variant apoprotein (Fig. 1D). This result implies that K184 is involved in the binding of the unknown compound that seems to be required for Lar activity and Ni coordination. The reduced Lar activity and Ni content in the other variants may be explained by the destabilization of the active site leading to rapid inactivation and Ni loss.
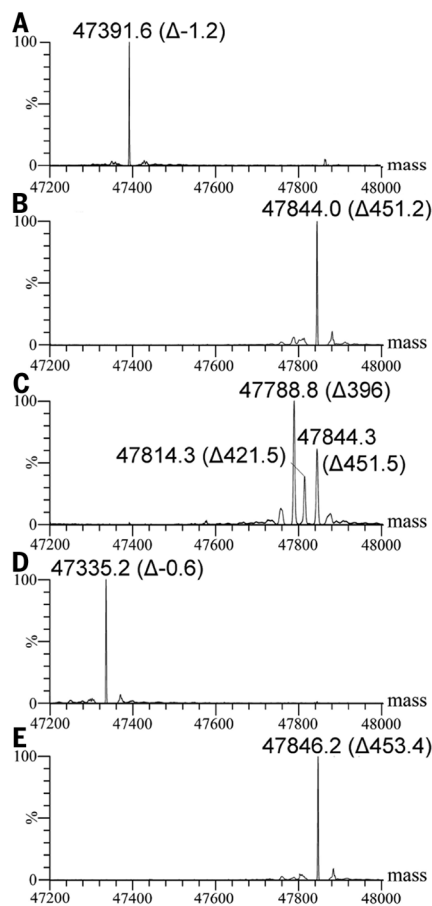

Fig. 1. Posttranslational modification of LarA by a nicotinic acid derivative. (**A** to **E**) The mass of (A) LarA apoprotein, (B) LarA holoprotein, (C) LarA holoprotein after incubating 1 day in buffer containing only 0.1 μM of sulfite, (D) K184A variant of LarA, and (E) LarA holoprotein purified from cells grown in the presence of d_4-nicotinic acid. The ordinate represents the abundance relative to the highest peak in percent. In parentheses are the mass differences compared with the theoretical mass of the LarA apoprotein lacking its *N*-terminal methionine.
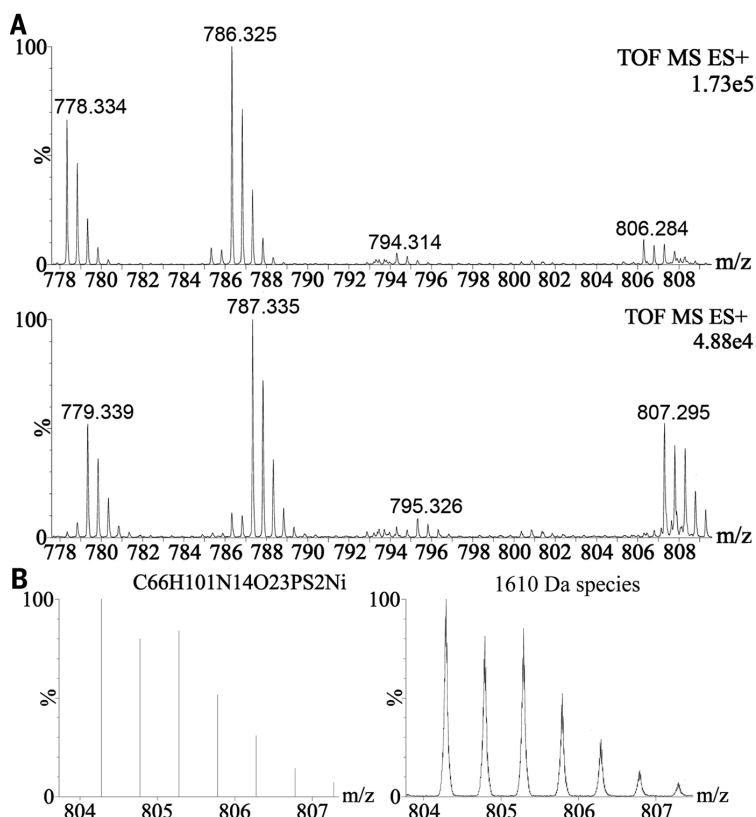

Fig. 2. LC-ESI-MS analysis of modified tryptic peptides of LarA in positive ionization mode and isotopic model of the 1610-dalton species. (**A**) Doubly positively charged tryptic peptides of LarA were analyzed by means of LC-ESI-MS with samples derived from cells grown in the presence of (top) nicotinic acid and (bottom) d_4-nicotinic acid. (**B**) Isotopic model of the doubly negatively charged 1610-dalton species (left) calculated with the Micromass MassLynx (Waters, Milford, MA) software package and (right) as observed in negative ionization mode.

Table 1. Mass spectrometric analysis of nicotinic acid–containing tryptic fragments of LarA.

Observed mass (OM)*	OM − peptide (−1161.676)†	OM − peptide − NAMN + H$_2$O (−1478.706)†	Δ‡	Theoretical mass of Δ
1610.570 ± 0.014	448.894 ± 0.014	131.864 ± 0.014	CS$_2$Ni^{2+} −2H$^+$	131.864
1586.637 ± 0.020	424.961 ± 0.020	107.931 ± 0.020	CO$_2$S$_2$	107.934
1570.646 ± 0.009	408.970 ± 0.009	91.940 ± 0.009	COS$_2$	91.939
1554.658 ± 0.010	392.982 ± 0.010	75.952 ± 0.010	CS$_2$/CO$_2$S	75.944/75.962

*Observed mass (OM, in daltons) is the average of masses observed by using positive and negative ionization modes (Fig. 2A and fig. S4). The "plus/minus" represents the 95% confidence interval (normal distribution, $n = 4$ m/z peaks).   †Mass of the tryptic peptide (residues 184 to 194) or tryptic peptide plus NAMN subtracted from OM (in daltons).   ‡Additional atoms that are likely to be present.

To investigate the nature of the cofactor, we grew recombinant *Lc. lactis* cells expressing the *lar* operon in chemically defined medium (*4*) that was selectively depleted of particular vitamins. Niacin (nicotinic acid or nicotinamide) was necessary for generating Lar activity (fig. S3), suggesting that niacin is a precursor of the LarA cofactor. To directly test this hypothesis, we grew *Lc. lactis* cells in the presence of tetra-deuterated nicotinic acid and analyzed the purified LarA by means of MS. The resulting mass of LarA showed a 2 ± 0.5–dalton mass increase compared with the holoenzyme isolated from control cultures (Fig. 1, B and E), indicating that the nicotinic acid was incorporated into the protein and that two deuterium atoms had undergone substitution.

To better define the structure of this nicotinic acid derivative, we digested LarA containing normal or deuterated cofactors with trypsin and analyzed the peptides by means of LC-ESI-MS. We detected four peptides eluting with similar retention times in both positive and negative ionization mode, all of which contained a 2.010 ± 0.003–dalton mass increase when d$_4$-nicotinic acid containing protein was digested, demonstrating the presence of a covalently bound nicotinic acid derivative (Fig. 2A, Table 1, and fig. S4A). The molecular mass of the 1610-dalton species corresponds to the addition of 448.894 ± 0.014 daltons to the peptide (R)K^{184}SVLPGIASYK(T) (1162 daltons), which is 2 daltons smaller than the 451 ± 1 dalton compound found in the holo-protein (Table 1 and Fig. 1B), which suggests that the same compound is present, probably in a more oxidized state. Moreover, the 1610-dalton species was 10 times more abundant after digestion in microaerobic conditions (fig. S4B). Together, these findings strongly suggest that the 1610-dalton species contains the original cofactor and that the other species represent degradation products arising, in part, from its oxidation. The 1610-dalton species and the two other major derivatized peptide species showed MS/MS fragment ions in positive ionization mode that matched fragments of the residue 184-197 peptide with modified Lys^{184} (fig. S5 and supplementary text), which is in agreement with our mutagenesis results. All four derivatized peptide species showed fragmentation patterns in negative ionization mode consistent with the loss of ribose phosphate (212 daltons), indicating the presence of a nicotinic acid mono-

nucleotide (NAMN) derivative in all species (fig. S6 and supplementary text). This result also suggests that the NAMN derivative is attached to the lysine through its carboxylic acid rather than by the sugar phosphate. After subtracting the mass of the NAMN, the additional mass (131.864 ± 0.014 daltons) corresponds to the addition of CS$_2$Ni^{2+} (131.864 daltons) because the closest alternative formula (CO$_2$SNi^{2+}: 131.882 daltons) falls outside the 95% confidence interval of the measurement (Table 1). The masses of the observed fragments also support this formula (fig. S5 and supplementary text). In addition, the isotopic distribution of this species is in agreement with the theoretical model. In particular, the lower abundance of the M+1 species versus the M and M+2 species is a fingerprint of Ni presence (Fig. 2B) because it reflects the isotopic distribution of Ni (^{58}Ni, 68%; ^{59}Ni, 0%; and ^{60}Ni, 26% abundance).

Assuming the 1610-dalton species is the native Ni-containing cofactor, the 1554-dalton species could directly arise from it by loss of Ni^{2+} (Table 1), which is consistent with the loss of 56 ± 1 daltons from the holoprotein depleted of sulfite (Fig. 1C), whereas the 1586- and 1570-dalton species could arise from gain of O$_2$ or half O$_2$, respectively, by the 1554-dalton species. An alternative 1554-dalton species could arise from loss of S or O from the 1586- or 1570-dalton species, respectively (Table 1, fig. S7, and supplementary text).

To establish the structure of the NAMN derivative and examine how Ni binds at the catalytic site, we purified LarA in the presence of sulfite and crystallized it within 3 days in order to minimize activity loss. We solved the structure at 3.0 Å resolution by means of molecular replacement using the structure of LarA$_{Tt}$ as a template (table S2) (*2*). The overall fold of LarA is similar to LarA$_{Tt}$, but the two LarA molecules in one asymmetric unit adopt distinct conformations (figs. S8A and S9): LarA with an open catalytic site is nearly superimposable to LarA$_{Tt}$, whereas LarA in the closed conformation has a much narrower cleft at the active site with the *N*- and *C*-domains approaching each other, leading to a sealed catalytic center (fig. S8B). The strong anomalous signal of Ni indicated that only the molecule in the closed conformation has a Ni atom in the active site (Fig. 3, A and B, and fig. S8C). Because Ni is absolutely required for LarA activity, this observation suggested that the closed conformation most likely

represents a snapshot of the active LarA. As shown in the Fo-Fc omit map (Fig. 3B), Ni is coordinated with the side-chain of His^{200} and a NAMN derivative that is covalently attached to Lys^{184} and interacts with several highly conserved residues. The electron density profile supports a linkage of C3 of the pyridinium ring to Lys^{184} with strong electron density, which is comparable with the density of the S of the spatially close Met^{224}, at the position expected for O of an amide bond (fig. S10), strongly suggesting S occupies this position. The electron density on C5 of the pyridinium ring has a shape and density (comparable with that of the S of Met^{224}) (fig. S10), suggesting a thiocarboxylic acid where the S faces Ni. Collectively, the crystallographic data suggest a cofactor with CS$_2$Ni added to the NAMN (Fig. 3B), which is consistent with the MS data (Table 1). As a result, Ni is coordinated by the S atoms of a thioamide and a thiocarboxylic acid along with His^{200} and C4 of the pyridine ring, forming a distorted square planar complex. A previous study suggested Ni coordination by two His (*2*), but this cofactor was not included when modeling those data and could substitute for one of the His side chains. The additional thiocarboxylic acid and the Ni-C bond explain the loss of two deuterium atoms from the d$_4$-nicotinic acid in the final cofactor (Figs. 1E and 2A).

The LarA metallocenter resembles pincer complexes that are widely applied in organic synthesis or organometallic catalysis (*5*, *6*), but such a complex has not been observed in enzymes. Following the chemical nomenclature, LarA has a (SCS)Ni(II) pincer complex that we name a pyridinium-3-thioamide-5-thiocarboxylic acid mononucleotide Ni pincer. Although organometallic (SCS)Ni(II) pincer complexes have been synthesized (*7*, *8*), the chemistry of these complexes remains underexplored (*9*). Pincer complexes are usually very stable (*5*); however, Ni is lost from the (SCS)Ni cofactor of LarA in aerobic environments in the absence of sulfite (fig. S2, B and C), correlating with the oxidation of the cofactor by molecular oxygen (Table 1). The basis of this oxidation remains unclear, but the S moieties of (SCS)Ni(II) complexes are presumably prone to oxidative decomposition (*9*). Oxidation of a sulfur in LarA (SCS)Ni might lead to the loss of coordination of Ni (fig. S11), accounting for the observed Ni loss and explaining the instability of enzyme activity.
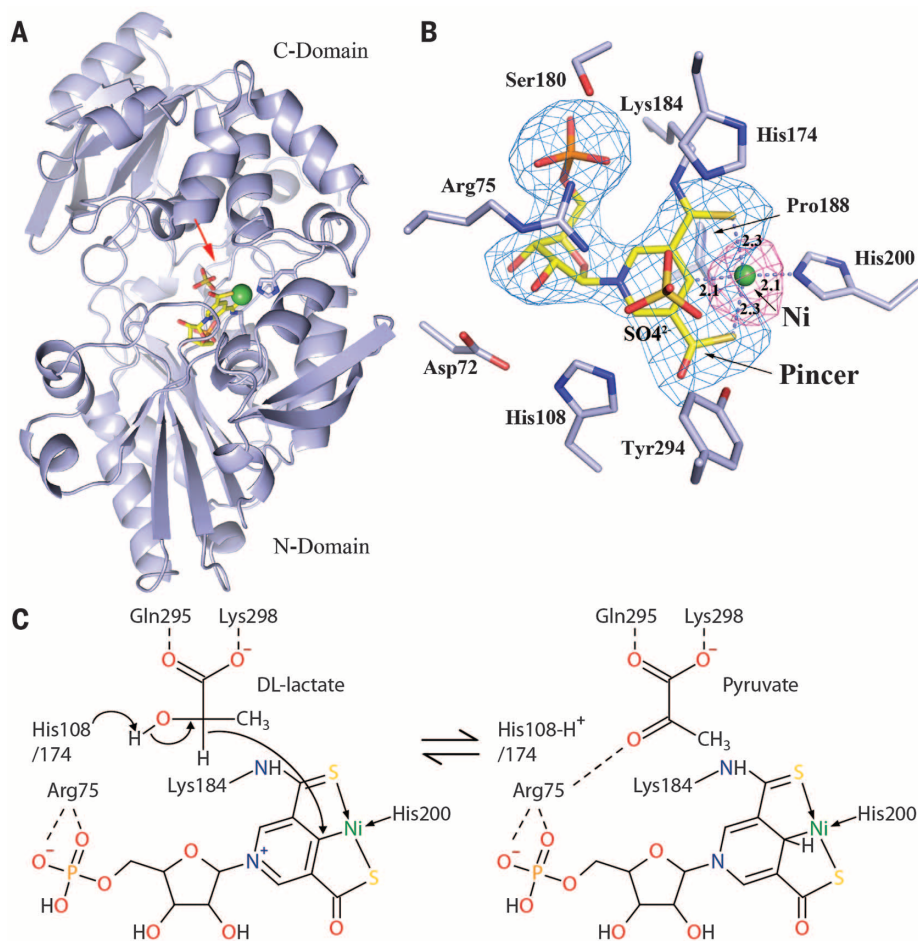
Fig. 3. Structure of LarA holoprotein and putative catalytic mechanism. (**A**) Structure of LarA in the active conformation. The protein is illustrated in cartoon mode, and the (SCS)Ni complex, His[200], and Lys[184] are shown as a stick model, with the cofactor C atoms in bright yellow, O in red, N in blue, S in dark yellow, and P in orange. Ni is depicted as a green sphere. (**B**) An expanded view of the LarA catalytic site as seen from the perspective of the arrow in (A). The color code for atoms is the same as in (A). The Fo-Fc omit map of the (SCS)Ni cofactor (σ = 4.0) is shown in blue, and the anomalous difference map of nickel (σ = 5.0) is shown in pink. The coordination bonds of nickel are shown as light blue dashed lines (length in angstroms). (**C**) Hypothetical catalytic mechanism of LarA. Straight arrows are coordination; curved arrows indicate the movement of electrons; dotted lines indicate the hydrogen bonds. An alternative site for hydride transfer is directly onto the Ni.

The presence of a metal-C bond is unusual in enzyme cofactors, but some well-known examples are adenosyl cobalamin (*10*), [Fe]-hydrogenase (*11*), and nitrogenase (*12*), and they are found as catalytic intermediates as in CO dehydrogenase (*13*). The reason for the presence of a Ni-C bond in Lar is not clear, but its presence suggests a role in catalysis. Past reports suggested a hydride transfer mechanism for lactate racemization (*14–17*). The participation of the (SCS)Ni cofactor in LarA catalysis would require that the substrate binds nearby. In the crystal structure, bound sulfate forms hydrogen bonds with six conserved residues at the presumed catalytic site (fig. S12), which strongly suggests a similar binding mode for lactate. If bound at this site, the C2 of lactate would be located directly on top of the C4 of the pyridinium ring and adjacent to the Ni atom, where hydride transfer is expected to take place (Fig. 3, B and C). Furthermore, two conserved histidine residues, His[108] and His[174], are situated on each side of the substrate and may participate in the racemization mechanism by acting as bases for deprotonating the hydroxyl group of one or the other lactate enantiomer and as acids for reprotonating the pyruvate intermediate (Fig. 3, B and C). Arg[75] forms hydrogen bonds with the (SCS)Ni cofactor and additionally could be involved in the stabilization of the hypothesized

pyruvate intermediate (Fig. 3C). Thus, the (SCS)Ni pincer complex may act to reversibly capture a hydride from the substrate, which is analogous to the role of nicotinamide adenine dinucleotide (NAD) in traditional hydride transfer enzymes (*18*), but by remaining tethered to the protein, it facilitates substrate racemization. Of additional interest, synthesis of the (SCS)Ni complex requires LarB, LarC, and LarE, which are widely distributed in microorganisms that lack LarA supporting a cofactor role in other enzymes (*2*).

REFERENCES AND NOTES

1. C. T. Walsh, *Science* **261**, 308–309 (1993).
2. B. Desguin *et al.*, *Nat. Commun.* **5**, 3615 (2014).
3. J. L. Boer, S. B. Mulrooney, R. P. Hausinger, *Arch. Biochem. Biophys.* **544**, 142–152 (2014).
4. G. Zhang, D. A. Mills, D. E. Block, *Appl. Environ. Microbiol.* **75**, 1080–1087 (2009).
5. J. T. Singleton, *Tetrahedron* **59**, 1837–1857 (2003).
6. G. van Koten, D. Milstein, Eds., *Organometallic Pincer Chemistry* (Springer, Germany, 2013), vol. 40.
7. S. M. Peterson, M. L. Helm, A. M. Appel, *Dalton Trans.* **44**, 747–752 (2015).
8. C. A. Kruithof *et al.*, *Organometallics* **27**, 4928–4937 (2008).
9. D. Zargarian, A. Castonguay, D. M. Spasyuk, in *Organometallic Pincer Chemistry*, G. van Koten, D. Milstein, Eds. (Springer-Verlag, Berlin, 2013), vol. 40, pp. 131–174.
10. L. Randaccio, S. Geremia, N. Demitri, J. Wuerges, *Molecules* **15**, 3228–3259 (2010).
11. S. Shima *et al.*, *Dalton Trans.* **41**, 767–771 (2012).
12. T. Spatzal *et al.*, *Science* **334**, 940 (2011).
13. S. W. Ragsdale, H. G. Wood, W. E. Antholine, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6811–6814 (1985).
14. T. Hiyama, S. Fukui, K. Kitahara, *J. Biochem.* **64**, 99–107 (1968).
15. H. Katagiri, T. Sugimori, K. Imai, *Agric. Biol. Chem.* **25**, 281–289 (1961).
16. D. Dennis, N. O. Kaplan, *Biochem. Z.* **338**, 485–495 (1963).
17. A. Cantwell, D. Dennis, *Biochemistry* **13**, 287–291 (1974).
18. P. Belenky, K. L. Bogan, C. Brenner, *Trends Biochem. Sci.* **32**, 12–19 (2007).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6243/66/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 to S3
References (*19–30*)

PRESYNAPTIC NETWORKS

Single-cell–initiated monosynaptic tracing reveals layer-specific cortical network modules

Adrian Wertz,[1]* Stuart Trenholm,[1]* Keisuke Yonehara,[1] Daniel Hillier,[1] Zoltan Raics,[1] Marcus Leinweber,[1] Gergely Szalay,[2] Alexander Ghanem,[3] Georg Keller,[1] Balázs Rózsa,[2] Karl-Klaus Conzelmann,[3] Botond Roska[1,4]†

Individual cortical neurons can selectively respond to specific environmental features, such as visual motion or faces. How this relates to the selectivity of the presynaptic network across cortical layers remains unclear. We used single-cell–initiated, monosynaptically restricted retrograde transsynaptic tracing with rabies viruses expressing GCaMP6s to image, in vivo, the visual motion–evoked activity of individual layer 2/3 pyramidal neurons and their presynaptic networks across layers in mouse primary visual cortex. Neurons within each layer exhibited similar motion direction preferences, forming layer-specific functional modules. In one-third of the networks, the layer modules were locked to the direction preference of the postsynaptic neuron, whereas for other networks the direction preference varied by layer. Thus, there exist feature-locked and feature-variant cortical networks.

In the cortex, many neurons selectively respond to distinct environmental features such as image motion in a specific direction or orientation (1, 2), the spatial position of the animal (3), or a specific part of a face (4). Each cortical neuron receives input from hundreds of nearby neighbors. Understanding the feature preference of the cortical neurons that provide input to a neuron with an identified feature preference could help us to understand how selectivity emerges and how cortical circuits are organized.

Within layer 2/3 of mouse primary visual cortex (V1), there is a close relation between orientation selectivity and synaptic connectivity (5). In contrast, in vivo single-cell recordings from layer 2/3 pyramidal cells in V1 have revealed different degrees of similarity between the preferred orientations at the dendritic input sites and at the cell body (6–8). The variability of orientation preferences at dendritic input sites could arise from differently tuned inputs from deeper cortical layers. The relation between the feature selectivity of the postsynaptic cell and the feature selectivity and functional organization of the deeper cortical neurons that provide synaptic input to individual layer 2/3 cells is still unknown.

We combined functionalized transsynaptic tracing with two-photon imaging and recorded the visually evoked responses of individual layer 2/3 pyramidal neurons together with their presynaptic neuronal networks across different cortical layers in vivo in mouse V1. Single-cell–initiated monosynaptic tracing (9) allows the expression of genetic tools in individual cortical neurons together with their monosynaptically connected presynaptic partners (10). This form of tracing is based on the delivery of three plasmids to a single neuron (one expressing a fluorophore, one expressing the avian receptor TVA, and one expressing the rabies virus envelope glycoprotein) and the local infection with the glycoprotein gene–deleted rabies virus coated with envelope-A [(EnvA), the ligand for the TVA receptor]. The EnvA-TVA ligand-receptor interaction restricts rabies virus infection to the TVA-expressing starter cell, and the glycoprotein allows rabies virus to move transsynaptically in the retrograde direction to only those neurons that monosynaptically connect to the starter cell (9). Rabies virus can be engineered to express genetically encoded calcium sensors, allowing the activity of the infected neurons to be recorded. The single-cell–initiated tracing system has been used for anatomical studies (9, 11, 12), but not yet for functional analyses.

Our modified version of single-cell–initiated monosynaptic tracing (9) differed in two aspects from previous approaches (Fig. 1A and supplementary materials and methods). First, we electroporated starter cells with four plasmids instead of three; the fourth plasmid, expressing a genetically encoded calcium sensor (GCaMP6s) (8), was necessary to record responses from the electroporated starter neuron. Second, we used a new rabies virus variant expressing GCaMP6s (materials and methods), which allowed for monitoring the activity from many presynaptic cells around the starter cell in a region spanning from layer 2/3 to layer 5. In each mouse, we labeled only a single layer 2/3

cell and its presynaptic network (Fig. 1B). We identified V1 using intrinsic in vivo imaging (13) (Fig. 1B and fig. S1) or post hoc confocal imaging in fixed brain slices (fig. S2). Soon after electroporation and rabies virus injection, the starter cell exhibited fluorescence (Fig. 1B). All starter cells were pyramidal cells. Next, presynaptic neurons expressing GCaMP6s appeared around the starter cell and increased in number as a function of time, whereas the responsiveness of the starter cell decreased as a function of time (Fig. 1, C and D, and fig. S3). Functional responses from presynaptic cells could be recorded up to ~2 weeks after electroporation, whereas starter cell responses could be recorded for ~1 week.

The total number of cells labeled in a presynaptic circuit, determined post hoc in immunostained brain slices, was 417 ± 74 (ranging from 70 to 846 cells, $n = 9$ presynaptic networks). The most abundant cluster of presynaptic cells surrounded the starter cell and was distributed across cortical layers (332 ± 64 presynaptic cells, ranging from 58 to 729 cells, $n = 9$) (Fig. 1, E to G, and fig. S4) (14). Within this local cluster, $82.5 \pm 2.4\%$ of cells were pyramidal cells (15). Outside of the local cluster of cortical neurons surrounding the starter cell, presynaptic cells were consistently labeled in several other brain regions that provide input to V1 (14) (fig. S5).

To determine the visual responses from the starter cell and its presynaptic network in V1, we presented animals with gratings that moved in eight directions and imaged GCaMP6s fluorescence with a two-photon laser scanning microscope from single optical planes with an area of 300 μm by 400 μm, from 40 to 600 μm below the brain surface, at 15- to 20-μm steps (Fig. 2A). Here we present functional data from 17 presynaptic networks connected to single pyramidal cells. From seven of these networks, in which the starter cell was electroporated with the four plasmids, we obtained visual motion responses from both the starter neuron and the presynaptic network. From 10 of the networks, electroporated with the three-plasmid approach (9), we obtained recordings from the presynaptic networks but not from the starter cells.

We imaged 98 ± 16 presynaptic cells in each presynaptic network. Nearly half ($43 \pm 4\%$) of these cells showed responses to image motion. Responses to motion were quantified using a direction-selective index (DSI) and an orientation-selective index (OSI) (calculated based on the vector sum of responses in all directions) (16) (fig. S6 and materials and methods). All presynaptic networks contained both direction- and orientation-selective neurons, and the degree of direction and orientation selectivity varied from neuron to neuron within a given network (Fig. 2, B to D, and fig. S6). Therefore, we analyzed all presynaptic networks for both direction and orientation selectivity.

Neurons within a single presynaptic network could be tuned to similar or different directions and orientations. We quantified the variability

[1]Neural Circuit Laboratories, Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. [2]Two-Photon Imaging Center, Institute of Experimental Medicine, Hungarian Academy of Sciences, Budapest, Hungary. [3]Max von Pettenkofer-Institute and Gene Center, Ludwig-Maximilians-University Munich, Munich, Germany. [4]Department of Ophthalmology, University of Basel, Basel, Switzerland.
*These authors contributed equally to this work. †Corresponding author. E-mail: botond.roska@fmi.ch

of the preferred directions and orientations of neurons within each of the 17 presynaptic networks. We determined the circular variance of the preferred directions and orientations in each of the presynaptic networks and estimated the probability that this variance could be obtained if equal numbers of neuronal responses were drawn randomly from a pool of responses measured using bulk rabies virus injection to V1 (Fig. 2, E to G; figs. S7 and S8; materials and methods). In 5 of the 17 presynaptic networks, the circular variance was significantly lower than the circular variances of the randomly chosen neuronal sets ($P < 0.01$) (Fig. 2G). We additionally performed two different statistical tests. First, the cumulative distributions of preferred directions and orientations in these five presynaptic networks were significantly different from the cumulative distribution of the pool of all responses (Kolmogorov-Smirnov test, $P < 0.05$) (fig. S9). Second, the hypothesis that the preferred directions and orientations of neurons within a presynaptic network were drawn randomly from a uniform distribution was
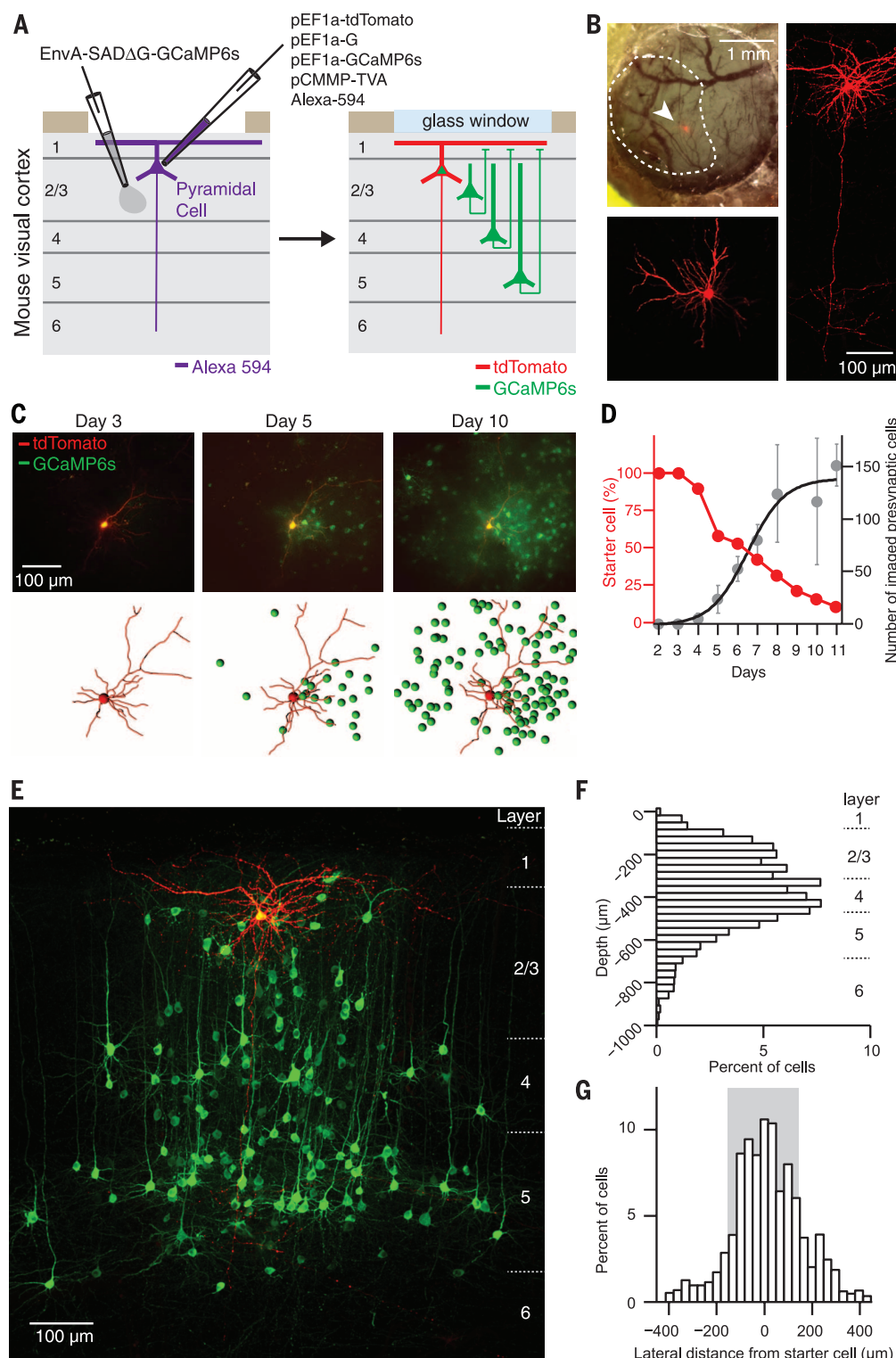
Fig. 1. GCaMP6s-functionalized, single-cell–initiated circuit tracing in mouse primary visual cortex. (**A**) Schematic of the experimental design. Individual layer 2/3 pyramidal cells in V1 were electroporated with a combination of four plasmids and Alexa 594 (purple) using the shadow-imaging technique (26, 27). Immediately after electroporation, EnvA-SADΔG-GCaMP6s rabies virus was injected near the electroporation site. After a few days, the electroporated cell expressed tdTomato (red) and GCaMP6s (green), whereas its presynaptic neurons expressed only GCaMP6s. (**B**) A tdTomato-expressing electroporated starter cell is shown in a cranial window using epifluorescence imaging (top left), two-photon imaging (bottom left), and post hoc confocal imaging (right). The dotted line (top left) represents the boundary of V1, as determined with intrinsic optical imaging (fig. S1 and materials and methods). (**C**) z projections of two-photon image stacks (top) and reconstructions (bottom) of an electroporated starter cell (red) and its local presynaptic neurons (green) over an 8-day period. (**D**) The presence of the electroporated starter cell (red) and the number of local presynaptic neurons (gray) imaged in vivo with a two-photon microscope at different days after electroporation ($n = 15$ networks; error bars indicate SEM). The black curve is a sigmoid fit to the mean number of presynaptic neurons over time. (**E**) Confocal image of an electroporated layer 2/3 pyramidal cell (red) and its local presynaptic network (green) after antibody staining (image is from a 300-μm-thick slice). (**F**) Average depth distribution of cell bodies of presynaptic neurons across cortical layers ($n = 9$). (**G**) Average distribution of the lateral distance between the cell bodies of the starter cell and its presynaptic cells in V1 ($n = 5$). The gray shaded area indicates the region from which we performed in vivo two-photon imaging. Both (F) and (G) are based on post hoc confocal analysis.

rejected (Hodges-Ajne test, $P < 0.05$) (fig. S9). Therefore, about one-third of all studied presynaptic networks (5/17) were "tuned": the responding neurons had similar preferred directions and orientations. However, the majority of presynaptic networks (12/17) were "untuned": the preferred directions and orientations of the responding neurons did not significantly differ from random draws from the pool of bulk responses (Fig. 2, E to G, and fig. S10).

To understand the relation between the visual responses of the starter cell and its presynaptic network, we compared the preferred direction and orientation of the starter cell to the average preferred direction and orientation of the presynaptic network in all seven networks from which we obtained visual responses from both the starter cell and the presynaptic network. In all cases in which the presynaptic network was tuned (3/7), the average preferred direction and orientation of the presynaptic network were within a small range of angles relative to the

preferred direction and orientation of the starter cell (direction: within $66 \pm 22°$; orientation: within $21 \pm 11°$; $n = 3$) (Fig. 3, A to C). Even in networks, which as a whole were untuned, layer 2/3 cells exhibited a bias such that they were tuned similarly to the starter neuron (direction: within $20 \pm 10°$; orientation: within $43 \pm 13°$; $n = 4$) (Fig. 3D and fig. S11).

Because layer 2/3 presynaptic subnetworks shared a similar preferred direction and orientation to the postsynaptic starter neuron, even in networks that were untuned, we analyzed the functional properties of presynaptic networks layer by layer. For each layer, we aligned the preferred direction and orientation of each neuron relative to the preferred direction and orientation of the layer subnetwork. We then averaged the aligned distributions layer by layer across all 12 untuned networks. Each layer subnetwork was significantly biased toward a single preferred direction and orientation [one-way analysis of variance (ANOVA), $P < 0.01$] (Fig. 3E and fig. S12).

Additionally, the circular variances of these relative direction and orientation distributions of layer 2/3 and layer 4 subnetworks were significantly lower than the circular variances of randomly chosen response sets from the pool of responses within the same layer obtained by bulk injection; this was also true when compared with the pool of all responses obtained by bulk injection ($P < 0.05$) (fig. S13). The Kolmogorov-Smirnov test and the Hodges-Ajne test were significant for layers 2/3, 4, and 5 ($P < 0.05$) (fig. S13). These results suggest the existence of functional layer modules within presynaptic networks.

We next investigated, in untuned networks, how the preferred direction of layer modules shifted from layer to layer. There was no significant bias in direction preference of layer 4 or 5 modules with respect to that of layer 2/3 modules (one-way ANOVA, $P > 0.05$) (Fig. 4, A and B, and fig. S14).

So far, the results we have presented were based on successively imaging individual optical

Fig. 2. Functional imaging reveals the existence of tuned and untuned presynaptic networks. (**A**) Schematic of the experimental protocol, in which mice were presented with moving gratings drifting in eight directions. GCaMP6s calcium responses were measured with a two-photon laser scanning microscope from different depths. (**B**) Example GCaMP6s responses from an electroporated starter cell (red) and three of its presynaptic neurons (black). Gray bars indicate when gratings were in motion, and arrows indicate the direction of stimulus motion. Polar plots of response amplitudes to each direction are shown to the right, as are the direction- and orientation-selective indices (DSI and OSI). Scale bars denote 100% $\Delta F/F$ (where F is fluorescence). (**C**) Raster plot for all cells with a DSI or OSI > 0.2 from an example network. Each horizontal line represents the normalized response amplitude of a single cell to eight directions of motion (shown below, scale bar at right). For each cortical layer (indicated on the right), cells are arranged based on the direction of motion driving the largest response. (**D**) Example 3D representation of the locations of a starter cell and its presynaptic neurons. Each filled circle represents a neuron and is colored according to the preferred motion direction (color code is shown at bottom right). Cells that did not respond to the visual stimuli are represented by small black circles. Cells that responded to motion equally in all directions are represented by small gray circles. (**E**) Distribution of the preferred motion directions in a representative tuned network. (**F**) Distribution of the preferred motion directions in a representative untuned network. (**G**) Plot of the P values of the circular variance of preferred directions (p-value$_{dir}$) and orientations (p-value$_{ori}$) of each network, calculated by comparing responses in each network to a randomly drawn set of responses from a pool (materials and methods). The dotted lines represent $P = 0.01$, with red dots indicating tuned networks.
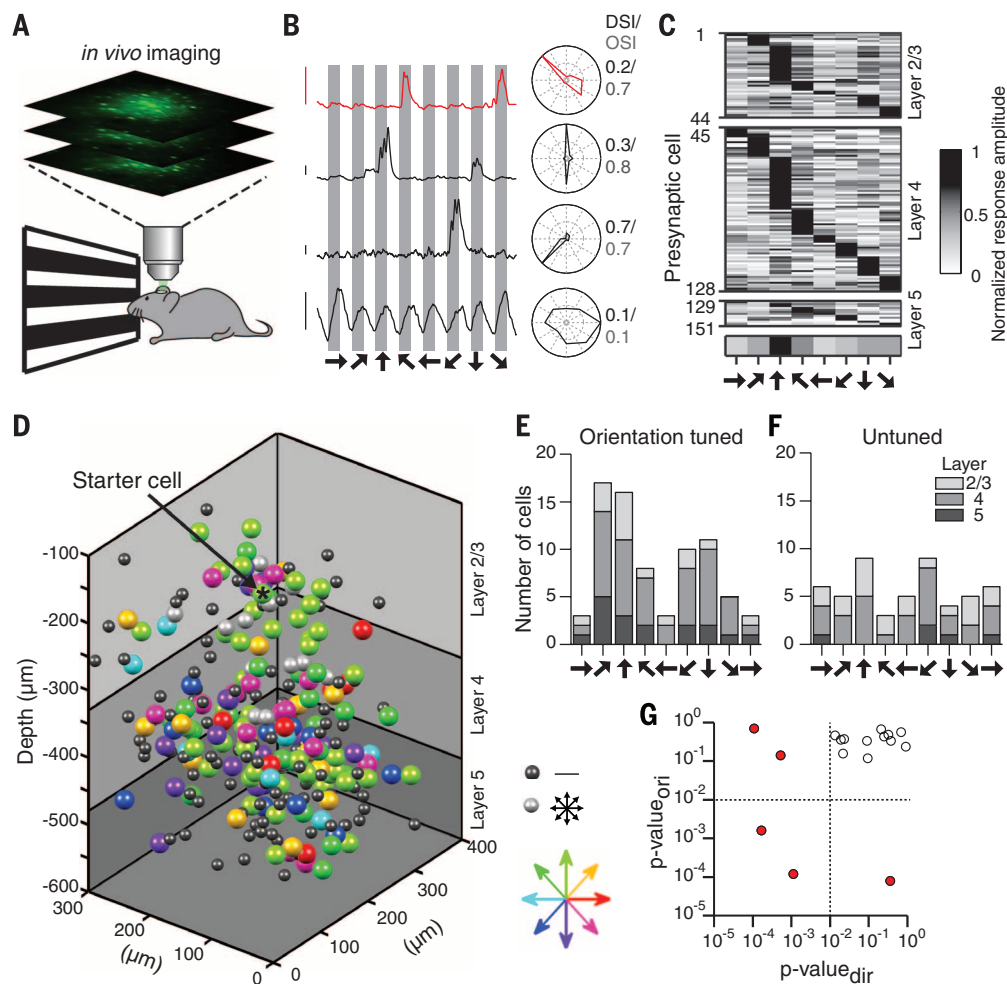
planes in V1. To test whether the tuning of a presynaptic network was conserved if the activity of the three-dimensional (3D) network was imaged nearly simultaneously, we used a 3D acousto-optic two-photon microscope to

measure network activity (*17*). We found no significant difference (Kolmogorov-Smirnov test, $P = 0.93$) in the overall motion direction preferences of networks when their activity was measured one optical plane at a time ver-

sus when the network was imaged in 3D in real time (fig. S15).

In summary, we used functionalized single-cell–initiated, monosynaptically restricted trans-synaptic tracing and two-photon imaging in V1

Fig. 3. Presynaptic networks are built from tuned layer modules. (**A**) Reconstruction of an electroporated layer 2/3 postsynaptic starter cell (red). Presynaptic cells are represented by filled circles. The gray scale of each circle refers to the cortical layer in which the presynaptic cell resides (gray scale code at bottom left). (**B**) The GCaMP6s response of an orientation-selective starter cell (top; the red curve is an interpolation) is compared to the distribution of preferred motion directions of its presynaptic neurons pooled across all recorded layers in a tuned presynaptic network (bottom). The direction of moving gratings is indicated below. (**C**) The average GCaMP6s responses of all starter cells belonging to tuned networks (top; red, $n = 3$), aligned to their preferred direction (the preferred direction was set to 0°), are compared to the relative preferred directions (relative to the preferred direction of the starter cell) of presynaptic neurons in these networks (bottom). (**D**) The average GCaMP6s responses of all starter cells belonging to untuned networks (top; red, $n = 4$), aligned to their preferred direction, are compared to the relative preferred directions of presynaptic neurons in layer 2/3 (bottom). (**E**) Distribution of relative preferred directions within layers (relative to the preferred direction of each layer subnetwork) in all untuned networks (materials and methods). The preferred direction of the layer subnetwork was set to 0°. **$P < 0.01$ (one-way ANOVA). Error bars represent SEM.

Fig. 4. Feature-locked and feature-variant networks. (**A**) In each of the 17 recorded networks, the preferred directions of layer modules were binned into eight directions (materials and methods) and are shown relative to the preferred direction of their layer 2/3 module. Solid lines link together the layer modules belonging to the same presynaptic network. Presynaptic networks that are tuned to the same direction across layers are indicated in red and highlighted by the gray shaded area. (**B**) Distributions of preferred directions of neurons in different cortical layers from networks in which each layer module was tuned to the same direction [left, red in (A)]. Preferred directions are shown relative to the preferred direction of the layer 2/3 module (materials and methods). Distributions of the preferred directions of neurons in different cortical layers from networks in which the preferred direction of each layer module was shifted compared with that of layer 2/3 [right, black in (A)]. **$P < 0.01$; *$P < 0.05$; one-way ANOVA. Error bars represent SEM. (**C**) Schematic representation of feature-locked (left) and feature-variant (right) networks. The starter cell is indicated with a red outline. The preferred direction of each neuron is indicated with an arrow.

to record the responses from individual layer 2/3 pyramidal cells together with their presynaptic partners in different cortical layers. Presynaptic cells within layers 2/3 and 4 (and, to some extent, layer 5) are tuned similarly for motion direction and orientation, forming layer-specific functional modules. The preferred direction and orientation of different layer modules can be aligned, resulting in presynaptic networks that are "feature-locked," or can be shifted relative to each other, giving rise to "feature-variant" networks (Fig. 4C).

The existence of feature-locked and feature-variant networks may explain why some studies found more variability than others in the tuning of dendritic input sites of layer 2/3 pyramidal cells (6–8) and may suggest that variability is likely due to inputs from deeper cortical layers. The combination of distinct layer modules in feature-variant networks is consistent with previous studies in brain slices showing cross-talk between different subnetworks in layer 2/3 and layer 5 (18, 19). In the visual cortex, the strength of connections among neurons correlates with similarity in visual responses (20), raising the possibility that feature-locked networks have a higher density of strong connections compared with feature-variant networks. Also, whether different subtypes of cortical interneurons (21, 22) are differentially represented in feature-locked and feature-variant networks remains an open question. Finally, it will be interesting to test whether postsynaptic cells in feature-locked and feature-variant networks exhibit different population coupling strengths (23).

What could be the role of feature-variant presynaptic networks? One possibility is that feature-variant networks are plastic. Top-down modulation or learning (24) could force the preferred direction and orientation of layer modules to align, resulting in a transition from a feature-variant to a feature-locked network. This recruitment of relevant circuits could allow more robust feature representations of behaviorally important stimuli. Another possibility is that variant layer modules enhance responses of the postsynaptic cell during object motion. Approaching and receding objects, for example, have edges moving in different directions. Some of these edges may stimulate inputs from deeper layers, which are not strong enough to drive responses of the postsynaptic cell alone but could boost responses of the postsynaptic cell to an edge moving in its preferred direction. Indeed, responses to combinations of orientations have been demonstrated in primate V2 (25).

REFERENCES AND NOTES

1. D. H. Hubel, T. N. Wiesel, *J. Physiol.* **148**, 574–591 (1959).
2. U. C. Dräger, *J. Comp. Neurol.* **160**, 269–290 (1975).
3. T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, E. I. Moser, *Nature* **436**, 801–806 (2005).
4. W. A. Freiwald, D. Y. Tsao, M. S. Livingstone, *Nat. Neurosci.* **12**, 1187–1196 (2009).
5. H. Ko et al., *Nature* **473**, 87–91 (2011).
6. S. L. Smith, I. T. Smith, T. Branco, M. Häusser, *Nature* **503**, 115–120 (2013).
7. H. Jia, N. L. Rochefort, X. Chen, A. Konnerth, *Nature* **464**, 1307–1312 (2010).
8. T.-W. Chen et al., *Nature* **499**, 295–300 (2013).
9. J. H. Marshel, T. Mori, K. J. Nielsen, E. M. Callaway, *Neuron* **67**, 562–574 (2010).
10. I. R. Wickersham et al., *Neuron* **53**, 639–647 (2007).
11. E. A. Rancz et al., *Nat. Neurosci.* **14**, 527–532 (2011).
12. M. Vélez-Fort et al., *Neuron* **83**, 1431–1443 (2014).
13. S. Schuett, T. Bonhoeffer, M. Hübener, *J. Neurosci.* **22**, 6549–6559 (2002).
14. Y.-J. Liu et al., *Curr. Biol.* **23**, 1746–1755 (2013).
15. K. D. Harris, G. M. G. Shepherd, *Nat. Neurosci.* **18**, 170–181 (2015).
16. M. Mazurek, M. Kager, S. D. Van Hooser, *Front. Neural Circuits* **8**, 92 (2014).
17. G. Katona et al., *Nat. Methods* **9**, 201–208 (2012).
18. B. M. Kampa, J. J. Letzkus, G. J. Stuart, *Nat. Neurosci.* **9**, 1472–1473 (2006).
19. Y. Yoshimura, J. L. M. Dantzker, E. M. Callaway, *Nature* **433**, 868–873 (2005).
20. L. Cossell et al., *Nature* **518**, 399–403 (2015).
21. C. A. Runyan et al., *Neuron* **67**, 847–857 (2010).
22. A. M. Kerlin, M. L. Andermann, V. K. Berezovskii, R. C. Reid, *Neuron* **67**, 858–871 (2010).
23. M. Okun et al., *Nature* **521**, 511–515 (2015).
24. J. P. Gavornik, M. F. Bear, *Nat. Neurosci.* **17**, 732–737 (2014).
25. A. Anzai, X. Peng, D. C. Van Essen, *Nat. Neurosci.* **10**, 1313–1321 (2007).
26. K. Kitamura, B. Judkewitz, M. Kano, W. Denk, M. Häusser, *Nat. Methods* **5**, 61–67 (2008).
27. B. Judkewitz, M. Rizzi, K. Kitamura, M. Häusser, *Nat. Protoc.* **4**, 862–869 (2009).

BRAIN STRUCTURE

Cortical folding scales universally with surface area and thickness, not number of neurons

Bruno Mota[1] and Suzana Herculano-Houzel[2,3]*

Larger brains tend to have more folded cortices, but what makes the cortex fold has remained unknown. We show that the degree of cortical folding scales uniformly across lissencephalic and gyrencephalic species, across individuals, and within individual cortices as a function of the product of cortical surface area and the square root of cortical thickness. This relation is derived from the minimization of the effective free energy associated with cortical shape according to a simple physical model, based on known mechanisms of axonal elongation. This model also explains the scaling of the folding index of crumpled paper balls. We discuss the implications of this finding for the evolutionary and developmental origin of folding, including the newfound continuum between lissencephaly and gyrencephaly, and for pathologies such as human lissencephaly.

The expansion of the cerebral cortex, the most obvious feature of mammalian brain evolution, is generally accompanied by increasing degrees of folding of the cortical surface into sulci and gyri (1). Cortical folding has been considered a means of allowing numbers of neurons in the cerebral cortex to expand beyond what would be possible in a lissencephalic cortex, presumably as the cortical sheet expands laterally with a constant number of neurons beneath the surface (2, 3). Although some models have shown cortical convolutions

[1]Instituto de Física, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. [2]Instituto de Ciências Biomédicas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. [3]Instituto Nacional de Neurociência Translacional, INCT/MCT, São Paulo, Brazil.
*Corresponding author. E-mail: suzanahh@gmail.com

to form as a result of cortical growth (*4*, *5*), the mechanisms that drive gyrification remain to be determined, and the field still lacks a mechanistic and predictive, quantitative explanation for how the degree of cortical folding scales across species. Moreover, recent systematic analyses of cortical folding have made clear that gyrification actually scales differently across mammalian orders, across clades within an order, and across individuals as a function of increasing brain volume (*6–9*). These apparent discrepancies have led to the view that different mechanisms must regulate cortical folding at the evolutionary, species-specific, and ontogenetic levels (*7*).

We undertook a systematic analysis of the variation in cortical folding across a large sample of mammalian species in search of a universal, unifying relationship between cortical folding and morphological properties of the cerebral cortex. We examined two data sets: our own, which includes numbers of cortical neurons and cortical surface areas (*10–21*), and another consisting of published data on cortical surface area, thickness, brain volume, and folding index, but not numbers of cortical neurons (*1, 22–24*) (table S1).

In the combined data set, there is a general correlation between total brain mass and the degree of cortical folding, and the two data sets overlap in their distribution (Fig. 1A, compare black and colored data points). However, the power function that relates the folding index of gyrencephalic species to brain mass has a fairly low r^2 and a 95% confidence interval that excludes many species (Fig. 1A). Striking and well-known outliers in this relationship are cetaceans (as a whole) and the manatee, but the capybara, the greater kudu, and humans also lie outside of the confidence interval (Fig. 1A). This indicates that cortical folding is not a homogeneous function of brain mass.

Although all cortical hemispheres with fewer than 30 million neurons are lissencephalic in our data set, and the correlation between folding index and number of neurons is significant across gyrencephalic species (Spearman correlation, ρ = 0.7741, $P < 0.0001$), the degree of gyrification is much larger in artiodactyls than in primates for similar numbers of cortical neurons (Fig. 1B). Additionally, the elephant cortex is about twice as folded as the human cortex, although the former has only about one-third the number of neurons found in the latter (Fig. 1B, "e" and "h"). The cortical surface area across species expands sublinearly with the number of cortical neurons in primates and supralinearly in other species (Fig. 1C). As a consequence, the average number of neurons per mm^2 of cortical surface is highly variable across species, ranging in our data set from 10,752 in the African elephant (*15*) to 138,606 in the squirrel monkey (*20*). Cortical expansion and folding are therefore neither a direct consequence of increasing numbers of neurons nor a requirement for increasing numbers of neurons in the cortex.

In comparison to the poor fit between folding index and total brain mass (Fig. 1A), a better fit is found for total surface area of the cerebral cortex in the two data sets (Fig. 1D). In this case, there is better overlap across afrotherians, glires, primates, and artiodactyls, although cetaceans, the manatee, and humans are still major outliers. Interestingly, all species with a cortical surface area below 400 mm^2 are lissencephalic in the two data sets. Similarly, all species with average cortical thickness below 1.2 mm are lissencephalic, but the folding index does not vary as a significant power function of cortical thickness across gyrencephalic species (Fig. 1E).

The folding index shows a sharp inflection between smooth and gyrated cortices, so it is unlikely that a universal model in terms of this variable alone could be derived. Because the folding index is the ratio of total surface area A_G to exposed surface area A_E, we next examined directly how A_E scales with A_G (Fig. 1F). In the combined data sets, for the species with small A_G (<400 mm^2) there is no folding, such that A_E equals A_G (Fig. 1F, green
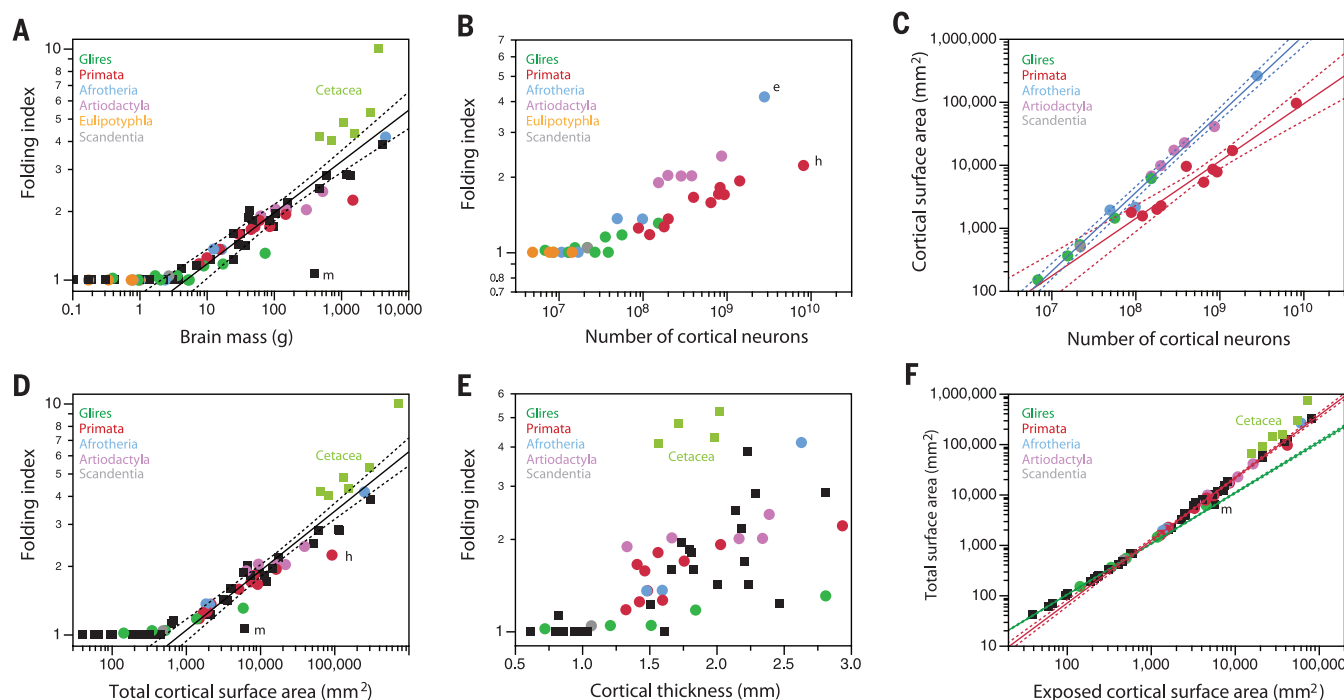

Fig. 1. Scaling of cortical folding index and total cortical surface area. Data points in black are taken from the literature; points in colors are from our own data set, except for cetaceans. (**A** to **E**) Folding index scales across all gyrencephalic species in the combined data sets as power functions of (A) brain mass, with exponent 0.221 ± 0.018 (r^2 = 0.751, $P < 0.0001$); (B) number of cortical neurons, with exponent 0.168 ± 0.032 (r^2 = 0.573, $P < 0.0001$; not plotted); (D) total cortical area, with exponent 0.257 ± 0.014 (r^2 = 0.872, $P < 0.0001$); and (E) average cortical thickness, with a nonsignificant exponent (r^2 = 0.054, $P = 0.1430$; not plotted). (C) Total cortical surface area of the cerebral cortex scales across primate species with an exponent of 0.911 ± 0.083 (r^2 = 0.938, $P < 0.0001$) and across nonprimate species with an exponent of 1.248 ± 0.037 (r^2 = 0.989, $P < 0.0001$). (**F**) Total cortical surface area varies across lissencephalic species as a linear function of the exposed surface area, but as a power function with an exponent of 1.242 ± 0.018 across noncetacean gyrencephalic species (r^2 = 0.992, $P < 0.0001$). Dashed lines are 95% confidence intervals for the fitted functions.

line). This linear relationship extends to the manatee cerebral cortex, even though its A_G is much larger than 1000 mm^2. In contrast, for all noncetacean gyrencephalic species, A_G increases with $A_E^{1.242 \pm 0.018}$ ($r^2 = 0.992$, $P < 0.0001$), significantly above linearity (Fig. 1F, red line), meaning that as total surface area increases, it becomes increasingly folded. Cetaceans fall above the 95% confidence interval of the function, which indicates that these cortices are more folded than similarly sized cortices in noncetaceans.

The finding that A_G scales as a power law of A_E means that gyrification is a property of a cortical surface that is self-similar down to a fundamental scale (the limit area between lissencephaly and gyrencephaly). This strongly suggests the existence of a single universal mechanism responsible for cortical folding (the alternative being some improbable multiscale fine-tuning) that over a range of scales generates self-similar, or fractal, surfaces.

Fractals can be characterized by the power-law scaling between intrinsic and extrinsic measures of an object's size, such as A_G and A_E. In this case, the fractal dimension d of the cortical surface is twice the value of the exponent relating A_G to A_E (given that A_E in turn scales with the square of the linear dimension of the cortex). Given that A_G scales with $A_E^{1.242 \pm 0.018}$ across noncetacean gyrencephalic brains, then $d = 2.484 \pm 0.036$. This value is remarkably close to the fractal dimension 2.5 of crumpled sheets of paper (25), which are fractal-like self-avoiding surfaces thin enough to fold under external compression while maintaining structural integrity.

Empirically, we conceive the fractal folding (or lack thereof) of the cortical surface as a consequence of the minimization of the effective free energy of a self-avoiding surface of average thickness T that bounds a volume composed of fibers connecting distal regions of said surface. Our model incorporates the known mechanics and organization of elongating axonal fibers (26, 27), as described in the supplementary materials. It predicts that from a purely physical perspective, A_G, A_E, and T are related by the power law $T^{1/2}A_G = kA_E^{5/4}$. (The exponent 5/4 is the only value for which the constant k is adimensional.)

We first tested whether our model based on the minimization of the effective free energy of a self-avoiding surface could explain the well-known fractal folding of a self-avoiding surface: paper. We examined how the exposed surface area of crumpled paper balls, A_E, scales with increasing total surface area, A_T, and thickness, T, of office paper (in this case, under forces applied externally by the experimenter's hands). As shown in Fig. 2A, $A_T \approx A_E^{1.234 \pm 0.033}$ for crumpled single sheets, a value similar to that for gyrencephalic cortices. Increasing T (by stacking sheets before crumpling) displaces the curves to the right (Fig. 2A) but leaves their slope largely unaltered, resulting in similar-looking but less folded paper balls (Fig. 2B). However, the product $T^{1/2}A_T$ varies proportionately to $A_E^{1.105 \pm 0.022}$ as a single, universal power function across all paper balls of different surface areas and thicknesses (Fig. 2C), as predicted by our model. This conformity indi-

cates that the coarse-grained folding of a sheet of paper subjected to external compression depends simply on a combination of its surface area and thickness.

We next examined whether our model predicts the folding of the mammalian cerebral cortex by plotting the product $T^{1/2}A_G$ as a function of A_E for the combined data sets. This yielded a power function with an exponent of 1.329 ± 0.014, with a very high r^2 of 0.996 for the noncetacean gyrencephalic species in the combined data sets (Fig. 3A, red line). Note that this function, although calculated for gyrencephalic species, overlaps with lissencephalic species. Including lissencephalic species (but still excluding cetaceans) actually improved the fit, with $r^2 = 0.998$, and yielded an exponent of 1.305 ± 0.010, which is close to the expected value of 1.25. Adding cetaceans to the analysis resulted in a small change of the fit (Fig. 3A, black line). Remarkably, the function fitted exclusively for lissencephalic species also predicted the relationship between $T^{1/2}A_G$ and A_E in gyrencephalic species (Fig. 3A, green line)—and species such as the manatee and other afrotherians are no longer outliers.

Given the theoretical relation $T^{1/2}A_G = kA_E^{1.25}$, it follows that lissencephalic species (for which $A_G = A_E$) are those that meet the condition $T = k^2A_G^{1/2}$. In contrast, all species for which $T < k^2A_G^{1/2}$ are predicted to be gyrencephalic, with $A_G > A_E$ (the alternative where $T > k^2A_G^{1/2}$ would result in $A_G < A_E$, which is geometrically impossible). Indeed, in the combined data set, we find that $T \approx A_G^{0.555 \pm 0.053}$ ($P < 0.0001$) across lissencephalic species (Fig. 3B, green line). All gyrencephalic species data points fall to the right of the lissencephalic distribution; that is, their A_G values are larger than predicted for a cortical thickness that would allow lissencephaly. The precise relationship between T and A_G across gyrencephalic species differs across orders, with a much smaller exponent for primates than for artiodactyls (Fig. 3B, red and pink lines). Thus, within the single universal relationship that describes cortical expansion, there is a transition point between smooth and folded cortices: Gyrencephaly ensues when A_G expands in area faster than T^2. For gyrencephalic species, the rate of expansion of cortical thickness relative to expansion of the cortical surface varies across orders, but the product $T^{1/2}A_G$ still varies as a universal function of $A_E^{1.26}$ to $A_E^{1.33}$.

We also found the same universality between the product $T^{1/2}A_G$ and A_E across coronal sections
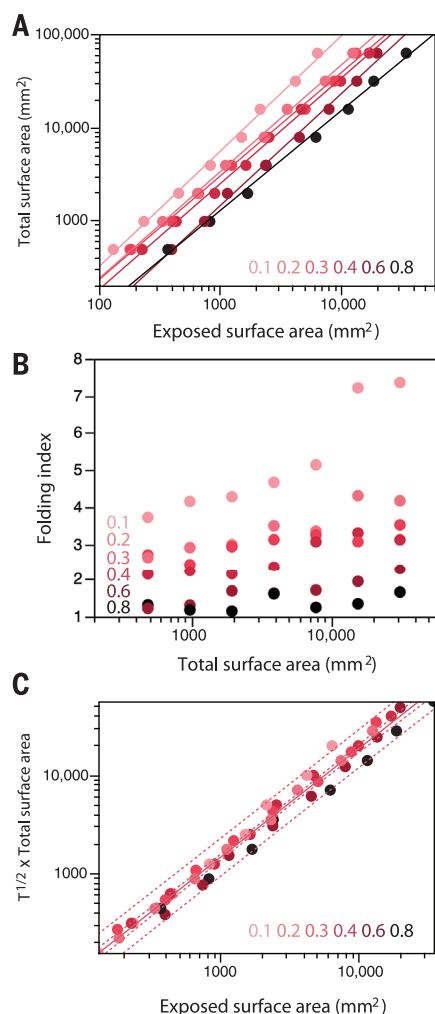

Fig. 2. The degree of folding of crumpled paper balls is a function of surface area and thickness as predicted by our model. (**A**) Relationship between total surface area of A4 to A11 sheets of office paper and the exposed surface area of the crumpled sheet of paper, with a power function of exponent 1.234 ± 0.033 for a single sheet of thickness 0.1 mm. (**B**) Increasing the thickness of the paper to be crumpled by stacking two to eight sheets displaces the curves to the right, that is, decreases the folding index of the resulting paper balls. (**C**) However, all crumpled paper balls of varying total surface area and thickness exhibit the same relationship, with the product $T^{1/2}A_T$ varying proportionately to $A_E^{1.105 \pm 0.022}$ ($r^2 = 0.983$, $P < 0.0001$). Color gradations correspond to thickness in millimeters, as shown in each panel.

along the anteroposterior axis of the cortical hemisphere of a single individual, of different individuals, and even different species ranging from small rodents to human and elephant (fig. S1).

The finding that A_E scales across all lissencephalic and gyrencephalic mammals (and even across species usually regarded as outliers such as the manatee and cetaceans) as a single power law of $T^{1/2}A_G$ indicates that gyrification is an intrinsic property of any mammalian cortex. Further, because the degree of folding can be described by the simple equation generated by our model (which also applies to crumpled sheets of paper), folding must occur as it minimizes the effective free energy of the cortical surface. Folding is therefore an intrinsic, fractal property of a self-avoiding surface, whether biological or not, subjected to crumpling forces. As such, this scaling of cortical folding does not depend on numbers of neurons or how they are distributed in the cortical sheet, but simply on the relative lateral expansion of this sheet relative to its thickness, regardless of how densely neurons are distributed within it.

The finding that cortical folding scales universally across clades, species, individuals, and parts of the same cortex implies that the single mechanism based on the physics of minimization of effective free energy of a growing surface subject to inhomogeneous bulk stresses applies across cortical development and evolution. This is in stark contrast to previous conclusions that different mechanisms regulated folding at different levels (7); such conclusions may reflect the traditional emphasis on the relationship between folding degree and brain volume (1, 8), which is indeed diverse across orders, across species, and across individuals of a same species (6, 8). Also, the dependence of cortical folding on a simple combination of A_G and T implies that any alterations, such as defects in cell migration, that lead to increased T or decreased A_G (or both) are expected to decrease cortical folding, exactly as found in human pathological lissencephaly (28). This might also be the case for the lissencephalic brain of birds, where a very thick telencephalon of small surface area surrounds the subpallial structures.

Finally, our findings indicate that cortical folding did not evolve, in the sense of a new property specific to some clades but not others. Similarly, there is no such thing as "secondary lissencephaly" (29), nor are there two clusters of gyrencephaly (9). Rather, what has evolved, we propose, is a faster increase in A_G relative to T^2 in development—and at different rates in different mammalian clades, which thus become gyrencephalic at different functions of A_G or different numbers of neurons.

Remarkably, there is no a priori reason for lissencephaly, considering that A_G and T ultimately result from different biological processes: lateral expansion of the progenitor cell population for A_G, radial neurogenesis and cell growth for T (30). Similarly, there is no a priori reason for the cortex to become gyrencephalic once past a certain surface area—unless the rate of (lateral) progenitor cell expansion inevitably outpaces the rate of (radial) neurogenesis at this point, which apparently occurs typically when A_G reaches 400 mm². We propose that, starting from the earliest and smallest (and smooth) mammalian brains (31), the cortical surface initially scaled isometrically, with $A_G \approx T^2$. Gyrencephaly ensued in each clade as soon as this lockstep growth changed, with A_G now increasing faster than T^2. Probable mechanisms involved are those that control the rate of neurogenesis and increases in cell size relative to the rate of progenitor and intermediate progenitor cell proliferation in early cortical development. Rapid increases in numbers of intermediate progenitor cells would lead to gyrencephaly, although not through the generation of larger numbers of neurons, as previously thought (7, 30, 32), but rather through the simple lateral expansion of the resulting cortical surface area at a rate faster than the cortical thickness squared.

REFERENCES AND NOTES

1. M. A. Hofman, *Brain Behav. Evol.* **27**, 28–40 (1985).
2. A. J. Rockel, R. W. Hiorns, T. P. Powell, *Brain* **103**, 221–244 (1980).
3. P. Rakic, *Trends Neurosci.* **18**, 383–388 (1995).
4. R. Toro, Y. Burnod, *Cereb. Cortex* **15**, 1900–1913 (2005).
5. T. Tallinen, J. Y. Chung, J. S. Biggins, L. Mahadevan, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 12667–12672 (2014).
6. K. Zilles, N. Palomero-Gallagher, K. Amunts, *Trends Neurosci.* **36**, 275–284 (2013).
7. E. Lewitus, I. Kelava, W. B. Huttner, *Front. Hum. Neurosci.* **7**, 424 (2013).
8. P. Pillay, P. R. Manger, *Eur. J. Neurosci.* **25**, 2705–2712 (2007).
9. E. Lewitus, I. Kelava, A. T. Kalinka, P. Tomancak, W. B. Huttner, *PLOS Biol.* **12**, e1002000 (2014).
10. F. A. C. Azevedo et al., *J. Comp. Neurol.* **513**, 532–541 (2009).
11. M. Gabi et al., *Brain Behav. Evol.* **76**, 32–44 (2010).
12. S. Herculano-Houzel, B. Mota, R. Lent, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12138–12143 (2006).
13. S. Herculano-Houzel, C. E. Collins, P. Wong, J. H. Kaas, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3562–3567 (2007).
14. S. Herculano-Houzel et al., *Brain Behav. Evol.* **78**, 302–314 (2011).
15. S. Herculano-Houzel et al., *Front. Neuroanat.* **8**, 46 (2014).
16. R. S. Kazu, J. Maldonado, B. Mota, P. R. Manger, S. Herculano-Houzel, *Front. Neuroanat.* **8**, 128 (2014).
17. K. Neves et al., *Front. Neuroanat.* **8**, 5 (2014).
18. D. K. Sarko, K. C. Catania, D. B. Leitch, J. H. Kaas, S. Herculano-Houzel, *Front. Neuroanat.* **3**, 8 (2009).
19. S. Herculano-Houzel, B. Mota, P. Wong, J. H. Kaas, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 19008–19013 (2010).
20. L. Ventura-Antunes, B. Mota, S. Herculano-Houzel, *Front. Neuroanat.* **7**, 3 (2013).
21. P. F. M. Ribeiro et al., *Front. Neuroanat.* **7**, 28 (2013).
22. H. Elias, D. Schwartz, *Z. Saugetierkd.* **36**, 147–163 (1971).
23. R. L. Reep, T. J. O'Shea, *Brain Behav. Evol.* **35**, 185–194 (1990).
24. T. M. Mayhew, G. L. Mwamengele, V. Dantzer, S. Williams, *J. Anat.* **188**, 53–58 (1996).
25. Y. Kantor, M. Kardar, D. R. Nelson, *Phys. Rev. A* **35**, 3056–3071 (1987).
26. D. H. Smith, *Prog. Neurobiol.* **89**, 231–239 (2009).
27. D. H. Smith, J. A. Wolf, D. F. Meaney, *Tissue Eng.* **7**, 131–139 (2001).
28. S. E. Hong et al., *Nat. Genet.* **26**, 93–96 (2000).
29. I. Kelava, E. Lewitus, W. B. Huttner, *Front. Neuroanat.* **7**, 16 (2013).
30. J. H. Lui, D. V. Hansen, A. R. Kriegstein, *Cell* **146**, 18–36 (2011).
31. T. B. Rowe, T. E. Macrini, Z. X. Luo, *Science* **332**, 955–957 (2011).
32. I. Reillo, C. de Juan Romero, M. A. García-Cabezas, V. Borrell, *Cereb. Cortex* **21**, 1674–1694 (2011).

Fig. 3. The degree of folding of the mammalian cerebral cortex is a single function of surface area and thickness across lissencephalic and gyrencephalic species alike, although thickness scales as order-specific functions of cortical surface area. (**A**) The product $T^{1/2}A_G$ varies with $A_E^{1.329\pm0.014}$ ($r^2 = 0.996$, $P < 0.0001$) across noncetacean gyrencephalic species in the combined data set (red line), with $A_E^{1.325\pm0.009}$ ($r^2 = 0.997$, $P < 0.0001$, $k = 0.157 \pm 0.012$) across all species (including cetaceans; black line), and with $A_E^{1.292\pm0.027}$ ($r^2 = 0.994$, $P < 0.0001$) across lissencephalic species alone (green line). Note that the function plotted for lissencephalic species predicts the product $T^{1/2}A_G$ for gyrencephalic species equally well as the functions plotted for gyrencephalic species themselves. (**B**) Cortical thickness varies with cortical surface area $A_G^{0.555\pm0.053}$ ($r^2 = 0.887$, $P < 0.0001$) across lissencephalic species in the combined data set (green line), but with $A_G^{0.160\pm0.025}$ ($r^2 = 0.703$, $P < 0.0001$) across primates (red line), and with $A_G^{0.334\pm0.072}$ ($r^2 = 0.879$, $P = 0.0185$) across artiodactyl species (pink line). All fits exclude cetaceans. Dashed lines indicate the 95% confidence intervals for the fitted functions.
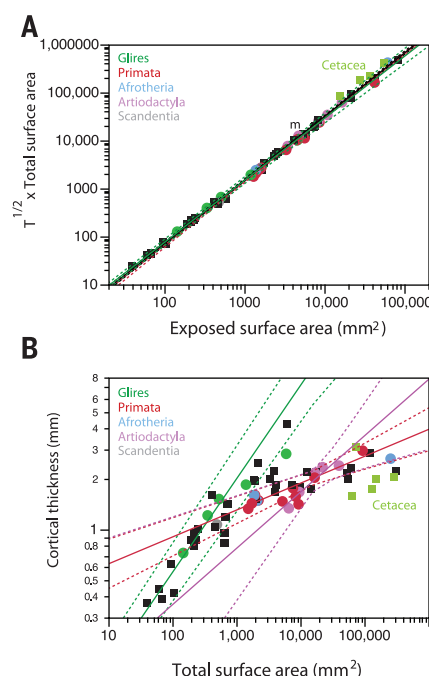
GREEN CHEMISTRY

Shape-selective zeolite catalysis for bioplastics production

Michiel Dusselier,*† Pieter Van Wouwe, Annelies Dewaele, Pierre A. Jacobs, Bert F. Sels*

Biodegradable and renewable polymers, such as polylactic acid, are benign alternatives for petrochemical-based plastics. Current production of polylactic acid via its key building block lactide, the cyclic dimer of lactic acid, is inefficient in terms of energy, time, and feedstock use. We present a direct zeolite-based catalytic process, which converts lactic acid into lactide. The shape-selective properties of zeolites are essential to attain record lactide yields, outperforming those of the current multistep process by avoiding both racemization and side-product formation. The highly productive process is strengthened by facile recovery and practical reactivation of the catalyst, which remains structurally fit during at least six consecutive reactions, and by the ease of solvent and side-product recycling.

Development of biodegradable plastics is part of the long-term solution of mitigating the negative environmental consequences of plastic pollution, exemplified by marine accumulation of inert microplastics (1, 2). Polylactic acid (PLA) is the major synthetic bio-based plastic on the market, and it serves in an impressive range of applications (3–5). Despite its environmentally benign performance in life-cycle assessments (6), a firm economic driver for megaton-scale PLA production is missing because of its high manufacturing and feedstock costs (2, 4). The current PLA production begins with the anaerobic fermentation of renewable sugars to aqueous (usually L-) lactic acid (LA) (Fig. 1). Though cost- and waste-efficient production of large volumes of LA is challenging by current fermentation processes, a slow but steady access to cheaper LA is forecasted because of novel emerging catalytic routes to LA directly from renewable (sugar, glycerol) feedstock (7–9) and concurrent progress in biotechnology (10). The industrial route to PLA from LA relies on the intermediate synthesis of lactide, the cyclic dimer of LA (Fig. 1). This dimer is converted to quality PLA through a controlled ring-opening polymerization (3). Unfortunately, current lactide production is very costly and therefore one of the great challenges in the large-scale production of PLA (4).

The current two-stage process for lactide synthesis from LA proceeds via polycondensation of L-LA with water removal to a prepolymer, followed by a reaction releasing lactide from the chain (Fig. 1, red frame). The second stage involves an endothermic transesterification, which is facilitated by the presence of nonrecoverable soluble metal salt (8). To avoid the thermodynamically favored reverse reaction in this stage, continuous removal of lactide (e.g., by distillation) is essential. Both stages are time- and energy-intensive owing to their reliance on high temperatures in vacuo. Moreover, moderate L,L-lactide selectivity (50 to 60%) is achieved, mainly due to racemization leading to unwanted D,L-lactide and formation of polymeric waste residue (11–13). The two-step process thus involves strenuous down-stream purification. An ideal one-stage alternative route for lactide formation should therefore involve LA dimerization to linear dimer (L_2A) and its direct cyclization by condensation, during water-removal (Fig. 1, blue frame). This route is only possible if the competitive, thermodynamically (and by higher reaction order) favored growth of oligomers can be stopped (14, 15). The absence of selectivity control in concentrated, process-relevant LA solutions during condensation explains the current use of the two-step procedure.

Steering the outcome of a reaction with heterogeneous catalysts is omnipresent in the petrochemical industry, and promising examples in the conversion of sugars have emerged as well (7, 16–19). Zeolite catalysts, in particular, are well known for product selectivity control due to pore confinement effects. This type of steering, known as shape selectivity since Weisz (20), is key in the synthesis of refinery products (21, 22), organic transformations (23), and biomass pyrolysis (24). The use of this concept (and of heterogeneous catalysts in general) in the valorization of bio-based chemicals to bio-based polymers is unexplored but could offer unrivaled advantages.

Therefore, we investigated the direct conversion of concentrated aqueous LA to lactide with Brønsted acidic zeolite catalysts. The rationale for the use of zeolites is their ability to facilitate condensation reactions and the large molecular

Center for Surface Chemistry and Catalysis, KU Leuven, Kasteelpark Arenberg 23, B-3001 Heverlee, Belgium.
*Corresponding author. E-mail: michiel.dusselier@gmail.com (M.D.); bert.sels@biw.kuleuven.be (B.F.S.) †Present address: Chemical Engineering, California Institute of Technology, 1200 East California Boulevard, MSC 210-4, Pasadena, CA 91125, USA.
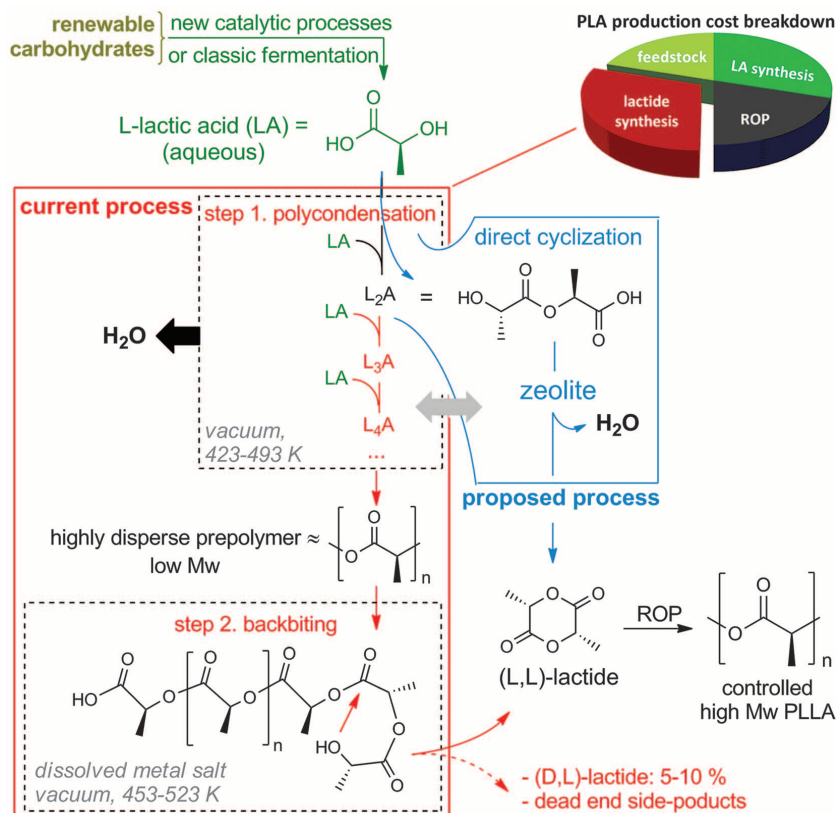
Fig. 1. Current and proposed chemical process for making lactide and PLA from lactic acid. Red: current industrial two-step process; blue: proposed direct lactide synthesis by selectivity control during condensation. L_nA: linear oligomer of n lactyl units; ROP: ring-opening polymerization. Pie chart shows a cost estimate for current PLA production from sugar to pellet.

differences, both in size and functionality, between (growing) oligomers and cyclic lactide. This contribution realizes fast and selective lactide synthesis directly from LA, by using microporous 12-membered ring H-zeolites under continuous water removal via solvent reflux. Record lactide space-time yields of >250 g liter^{-1} hour^{-1}, devoid of racemization, are attained with our one-step process, by steering away from lactyl oligomers in the reaction (Fig. 1). The combination of the zeolite's stability and its reuse potential with the ease of product separation, solvent recovery, and side-product reintegration allows us to propose a highly competitive lactide production process.

In particular, zeolite H-Beta is highly selective for lactide synthesis, yielding nearly 79% at full LA conversion, compared to 2% at only 35% conversion in the absence of a catalyst (for analysis, see Materials and Methods, figs. S1 to 3). Common soluble acid catalysts, such as sulfuric and sulfonic acids, accelerate LA condensation, mainly to lactyl oligomers instead of lactide (Fig. 2B and fig. S4), with the number-average oligomer length, n, exceeding a value of 10. Pronounced oligomerization is clearly the effect of an unselective acid-catalyzed condensation under water removal, in line with the well-known polydisperse LA polycondensation via solvent reflux (3). Macroporous resins like Nafion and Amberlyst show a product slate similar to that of the soluble acids and only yield about 20% of lactide at full LA conversion. High-performance liquid chromatography (HPLC) product profiles nicely illustrate the zeolite-induced lactide-selective condensation control compared

to an unselective resin catalyst (Fig. 2C and fig. S5). The open resin forms oligomers according to Flory-Schulz–like kinetics, whereas the zeolite converts LA almost quantitatively into lactide and only a minor portion of short oligomers, being mostly L$_{3-4}$A at full LA conversion. The catalytic results of an acidic mesoporous aluminosilicate, H-Al-MCM-41, with pore diameters about four times that of zeolite Beta (fig. S6), emphasize the zeolite's shape-selective control, as it is clearly less selective (and active) toward lactide (Fig. 2B). The presence of Brønsted acidity is essential, as a zeolite in its nonacidic form (NH$_4$-Beta, Si/Al = 12.5) and a low-aluminum H-Beta (Si/Al = 180) showed almost no conversion and very low lactide yields on top of an auto-catalyzed background reaction (table S1). H-USY zeolites, having larger mesopores next to micropores in the crystals, show intermediate activity and lactide selectivity.

Accumulation of linear dimer (L$_2$A) at short contact times points to its key role as primary product en route to lactide (Fig. 2, A and D). It results from the energetics of the discrete reaction pathways in the zeolite, ring-closing condensation of L$_2$A to lactide being more demanding than LA dimerization. We assessed the kinetics of the overall reaction at varying reflux temperatures with the same solvent by reducing pressures (figs. S7 to S9). The apparent activation energy (E_A) of the overall reaction with H-Beta roughly amounted to 75 kJ mol^{-1}, an expected value for condensation chemistry, pointing to the absence of pore-diffusion limitation in the rate-

limiting event (25). As the conversion of the linear dimer L$_2$A showed a very similar apparent E_A value, yielding 93% lactide at full conversion (table S1), the zeolite-catalyzed LA conversion to lactide runs in chemical regime, with ring-closing (and thus lactide formation) determining the reaction rate. Henceforth, the use of higher–boiling point solvents and thus reaction temperatures (figs. S10 and S11) will allow the formation rate of lactide to be increased. In o-xylene, 83% of lactide at full conversion was thus attained, at an initial rate of 19.9 mM hour^{-1} compared to 79% yield at only 3.2 mM hour^{-1} in toluene. Xylenes are benign toluene substitutes, and promising renewable routes for their production have been reported (24). Our zeolite-catalyzed reaction in o-xylene is both fast and selective. In a reactor holding 9 weight % (wt %) of LA and 5 wt % of zeolite with respect to the solvent, volumetric lactide productivities (space-time yields) around 170 g liter^{-1} hour^{-1} are accomplished, corresponding to a yield of 64% within 15 min. The lactide productivity is further enhanced to 267 g liter^{-1} hour^{-1} by doubling LA and catalyst loadings, while maintaining a high catalyst efficiency for lactide at 3.0 g g$_{zeolite}$$^{-1}$ hour^{-1}. Another important advantage of the zeolite-based over the current two-step process, besides the lactide yield, consists in the absence of racemization. Both chromatographic and nuclear magnetic resonance (NMR) analyses pointed to the presence of less than 1 mole percent (mol %) of D,L-lactide in the total lactide fraction. Values of D,L-lactide between 5 and 15 mol % are common in the current industrial

Fig. 2. Catalytic data and benchmarking of the shape-selective reaction. (**A**) Reaction pathways. (**B**) Product distribution after reaction with selected zeolites, soluble acids, and macro- and mesoporous heterogeneous catalysts (0.565 mmol H$^+$, e.g., 0.5 g for H-Beta zeolite Si/Al = 12.5) on 1.67 g of aqueous L-LA (50 wt %) in 10 ml of toluene refluxed for 3 hours with water removal. Product [oligomers (red) and lactide (blue)] and reagent [LA (green)] distribution based on ^{1}H-NMR yields. Labels indicate the number-average length of the oligomer fraction (L$_n$A, with $n \geq 2$). (**C**) HPLC product profiles with a typical zeolite- (shape-selective H-Beta) and resin- (non–shape-selective NAFION) catalyzed reaction at full LA conversion. (**D**) Typical time profile of a zeolite reaction: LA (♦), lactide (●), L$_2$A (▲), L$_n$A, $n \geq 3$ (■) (^{1}H-NMR with HPLC-based L$_2$A-deconvolution from oligomers).
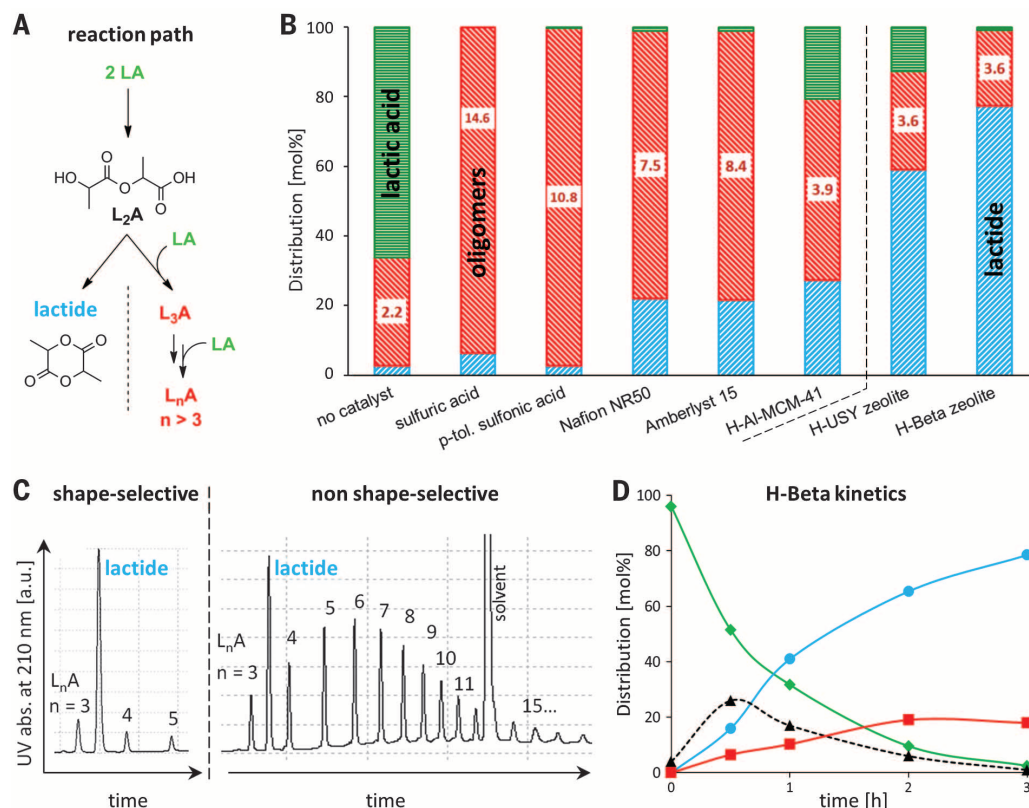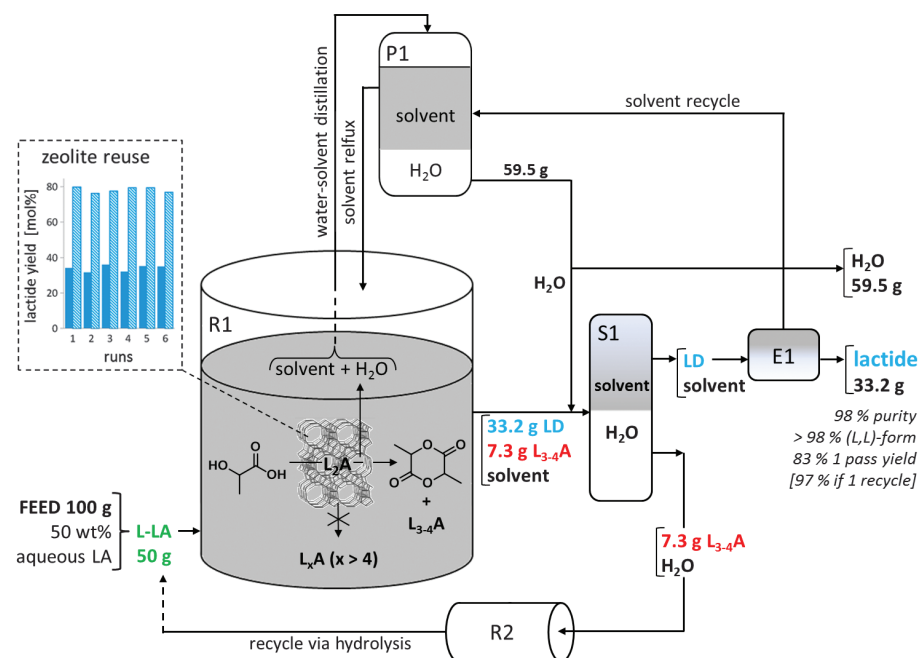
Fig. 3. H-Beta zeolite-based process scheme for making lactide. Feedstock, product, and solvent fluxes based on an exemplary 100-g input of aqueous LA, extrapolated from small-scale batch reactions and extractions. R1: main catalytic reactor; P1: phase settler. R1 and P1 guarantee reflux with water removal at constant solvent volume. S1: liquid-liquid extraction solvent-water; E1: solvent evaporator/lactide (LD) crystallization; R2: reactor for oligomer hydrolysis. Inset: zeolite reuse study in initial (dark blue) and productive conditions (light blue).

process, starting from enantiopure L-LA (11–13). D,L- from L,L-lactide separation is a decisive cost-determining factor.

The high lactide selectivity in the presence of microporous zeolites like H-Beta is a shape-selective effect, in accordance with Weisz's original definitions. A shape selectivity factor of 7.5 may be calculated, being clearly larger than unity (20). Because only short oligomers like di- and trimers were noticed in the zeolite crystal (extraction experiments of spent zeolite; see supplementary text), restricted growth of lactyl oligomers in the zeolite seems to be the basis of the shape selectivity. As a consequence, aside from the equilibrium nature of these reactions, zeolite pores are not blocked substantially, in contrast to the commonly reported deactivation of zeolites in biomass conversions involving irreversible tar and humin formation (26). Instead, lactic acid is converted to lactide in the absence of transport phenomena, resembling the characteristics of transition-state shape selectivity. This is corroborated by (i) the high observed E_A of lactic acid to lactide conversion; (ii) the intermediate accumulation of L_2A (Fig. 2D and fig. S11); and (iii) the identical lactide kinetics of a custom-made H-Beta zeolite with 30 to 50 times enhanced crystal sizes but similar Brønsted acid content (fig. S12). Whether the shape selectivity has a thermodynamic basis (27), or whether pure steric constraints (20) are involved, remains an open question.

The promising catalytic data prompted us to conceive an integrated process using practical reactor technology (Fig. 3). The zeolite-catalyzed reaction is carried out in a stirred tank reactor (R1) in an organic solvent under reflux, entraining condensation water, with a phase settler (P1) separating removed water from the water-immiscible solvent. When a dried solvent is returned from the settler, a lactide yield of 85% in toluene is obtained (fig. S13). Ideally, the outlet should continuously be subjected to liquid-liquid extraction (S1) with water, the polar small oligomeric acids preferring the aqueous phase, and the apolar lactide nearly exclusively residing in the solvent (98% pure, batch extraction fig. S14). An evaporator (E1) allows solvent recovery and convenient L,L-lactide crystallization (fig. S15). The aqueous effluent of S1 is easily recycled because lactyl oligomers are prone to spontaneous hydrolysis in excess of water (R2) (14). The resulting LA can be reinserted into P1, as the process is flexible in the amount of water initially present in the feed (fig. S16), an important trait when adapting this process to the dilute fermentative LA production. The process scheme in Fig. 3 essentially highlights the complementarity of the selectivities of both catalyst and extraction, as well as water removal via solvent distillation and lactide crystallization. A highly carbon-efficient process is thus obtained, with a minimal number of unit operations. The stability and reusability of the H-Beta zeolite catalyst were verified by six consecutive runs in o-xylene (in both initial and productive conditions) using intermediate calcination at 550°C. Full activity and selectivity regeneration were possible in each run (Fig. 3), and

characterization showed that the zeolite did not change structurally (figs. S17 and S18).

The applicability of this process to other α-hydroxy acids of interest in polyester chemistry such as α-hydroxy-butyric and α-hydroxy-butenoic acids is encouraging (8, 28). Highly selective ethylglycolide synthesis has been demonstrated (88% yield, fig. S19), as well as the synthesis of an asymmetric dimer by feeding a mixture of acids. These building blocks, for which traditional synthesis is cumbersome (29), grant access to a range of novel functional polyesters (tentatively drawn in fig. S20). The versatile zeolite technology may therefore facilitate the synthesis of a wide range of renewable, highly performant, and degradable bio-based plastics.

REFERENCES AND NOTES

1. K. L. Law, R. C. Thompson, *Science* **345**, 144–145 (2014).
2. R. A. Gross, B. Kalra, *Science* **297**, 803–807 (2002).
3. R. Auras, L.-T. Lim, S. E. M. Selke, H. Tsuji, Eds., *Poly(Lactic Acid): Synthesis, Structures, Properties, Processing, and Applications* (Wiley, New York, 2010).
4. L. Shen, E. Worrell, M. Patel, *Biofuels Bioprod. Biorefin.* **4**, 25–40 (2010).
5. C. O. Tuck, E. Pérez, I. T. Horváth, R. A. Sheldon, M. Poliakoff, *Science* **337**, 695–699 (2012).
6. E. T. H. Vink, K. R. Rábago, D. A. Glassner, P. R. Gruber, *Polym. Degrad. Stabil.* **80**, 403–419 (2003).
7. M. S. Holm, S. Saravanamurugan, E. Taarning, *Science* **328**, 602–605 (2010).
8. M. Dusselier, P. Van Wouwe, A. Dewaele, E. Makshina, B. F. Sels, *Energy Environ. Sci.* **6**, 1415 (2013).
9. L. S. Sharninghausen, J. Campos, M. G. Manas, R. H. Crabtree, *Nat. Commun.* **5**, 5084 (2014).
10. M. A. Abdel-Rahman, Y. Tashiro, K. Sonomoto, *Biotechnol. Adv.* **31**, 877–902 (2013).
11. M. Mueller, European patent 261572A1 (1988).
12. P. R. Gruber et al., U.S. patent 5247058 (1993).
13. W. G. O'Brien, L. A. Cariello, T. F. Wells, WO patent 9606092A1 (1996).
14. D. T. Vu et al., *Fluid Phase Equilib.* **236**, 125–135 (2005).
15. V. N. Emel'yanenko et al., *Russ. J. Phys. Chem. A* **84**, 1491–1497 (2010).
16. A. Corma, S. Iborra, A. Velty, *Chem. Rev.* **107**, 2411–2502 (2007).
17. G. W. Huber, R. D. Cortright, J. A. Dumesic, *Angew. Chem. Int. Ed.* **43**, 1549–1551 (2004).
18. Y. Román-Leshkov, C. J. Barrett, Z. Y. Liu, J. A. Dumesic, *Nature* **447**, 982–985 (2007).
19. T. P. Vispute, H. Zhang, A. Sanna, R. Xiao, G. W. Huber, *Science* **330**, 1222–1227 (2010).
20. P. B. Weisz, *Pure Appl. Chem.* **52**, 2091 (1980).
21. A. Corma, *Chem. Rev.* **95**, 559–614 (1995).
22. P. A. Jacobs, M. Dusselier, B. F. Sels, *Angew. Chem. Int. Ed.* **53**, 8621–8626 (2014).
23. C. W. Jones, K. Tsuji, M. E. Davis, *Nature* **393**, 52–54 (1998).
24. Y.-T. Cheng, Z. Wang, C. J. Gilbert, W. Fan, G. W. Huber, *Angew. Chem. Int. Ed.* **51**, 11097–11100 (2012).
25. G. Ertl, H. Knözinger, F. Schüth, J. Weitkamp, *Handbook of Heterogeneous Catalysis* (Wiley-VCH, Weinheim, Germany, ed. 2, 2008).
26. J.-P. Lange, W. D. van de Graaf, R. J. Haan, *ChemSusChem* **2**, 437–441 (2009).
27. B. Smit, T. L. M. Maesen, *Nature* **451**, 671–678 (2008).
28. M. Dusselier et al., *ACS Catal.* **3**, 1786–1800 (2013).
29. W. W. Gerhardt et al., *Biomacromolecules* **7**, 1735–1742 (2006).

PLANT VOLATILES

Biosynthesis of monoterpene scent compounds in roses

Jean-Louis Magnard,[1] Aymeric Roccia,[1,2] Jean-Claude Caissard,[1]
Philippe Vergne,[2] Pulu Sun,[1] Romain Hecquet,[1] Annick Dubois,[2]
Laurence Hibrand-Saint Oyant,[3] Frédéric Jullien,[1] Florence Nicolè,[1]
Olivier Raymond,[2] Stéphanie Huguet,[4] Raymonde Baltenweck,[5] Sophie Meyer,[5]
Patricia Claudel,[5] Julien Jeauffre,[3] Michel Rohmer,[6] Fabrice Foucher,[3]
Philippe Hugueney,[5]* Mohammed Bendahmane,[2]* Sylvie Baudino[1]*

The scent of roses (*Rosa* x *hybrida*) is composed of hundreds of volatile molecules. Monoterpenes represent up to 70% percent of the scent content in some cultivars, such as the Papa Meilland rose. Monoterpene biosynthesis in plants relies on plastid-localized terpene synthases. Combining transcriptomic and genetic approaches, we show that the Nudix hydrolase RhNUDX1, localized in the cytoplasm, is part of a pathway for the biosynthesis of free monoterpene alcohols that contribute to fragrance in roses. The RhNUDX1 protein shows geranyl diphosphate diphosphohydrolase activity in vitro and supports geraniol biosynthesis in planta.

Roses are used as ornamental plants in gardens, as cut flowers, and as sources of essential oils for perfume and cosmetics. Breeding with a focus on cut flowers and visual attributes can leave scent traits disadvantaged (*1*). The cause for the lack of fragrance in these flowers is unknown and does not seem to be linked to increased vase life (*2*). Monoterpene alcohols and 2-phenylethanol characterize typical rose scents; volatile phenolic compounds characterize tea-scented roses (*3*). Although genes involved in the biosynthesis of phenolic scent compounds, 2-phenylethanol, and sesquiterpenes have been characterized (*4*–*6*), the basis for monoterpene biosynthesis remains obscure.

In plants, geranyl diphosphate (GPP), precursor of monoterpenes, is synthesized in plastids from dimethylallyl diphosphate and isopentenyl diphosphate supplied by the methylerythritol 4-phosphate pathway (*7*). The monoterpenes in essential oils are produced through the activity of various monoterpene synthases (*8*). For example, geraniol synthase (GES) converts GPP into geraniol in basil (*9*). Here we investigate mono-

terpene biosynthesis in roses. We used cDNA amplification fragment length polymorphism–differential display (AFLP-DD) and DNA microarrays to compare the transcriptomes of two rose cultivars that have different scents (see supplementary materials and methods).

The Papa Meilland (PM) cultivar emits a heavy typical rose scent, mostly composed of monoterpene alcohols and 2-phenylethanol. The Rouge Meilland (RM) cultivar produces very little scent and only trace amounts of these compounds (table S1). With AFLP-DD we identified two amplicons favored in PM, one with homology to the Nudix hydrolase family (DIF1) and one with homology to a laccase protein (DIF38) (table S2). With microarrays we found 91 genes expressed more in PM than in RM (data S1). The gene with the highest differential expression (PM1, 7583-fold increase in PM relative to RM) (table S2 and data S1) also corresponded to the Nudix hydrolase. We have named this gene *RhNUDX1* (GenBank accession number JQ820249). *RhNUDX1* encodes a 150–amino acid protein that contains the characteristic Nudix domain (*10*) and is similar (59% identity) to AtNUDX1 from *Arabidopsis thaliana* (fig. S1). In a survey of 10 rose cultivars with contrasting scent profiles (tables S1 and S2 and data S2), *RhNUDX1* expression correlated with the presence of monoterpene alcohols (geraniol, nerol, citronellol) and sesquiterpenes (farnesol, farnesene, and farnesyl acetate) (Fig. 1). Nudix hydrolases remove nucleoside diphosphates linked to other moieties (*10*). Some may also accept non-nucleoside substrates (*11*). These enzymes have various functions and may act as diphosphoinositol polyphosphate phosphohydrolases, coenzyme A pyrophosphatases, adenosine diphosphate–ribose pyrophosphatases, diadenosine polyphosphate hydrolases, and mRNA decapping enzymes (fig. S1B). Nudix hydrolyses are found in animals, plants, and bacteria. The number of Nudix representatives in each species varies from one in *Myco-*

plasma to more than 50 in eukaryotes (*12*). In sequenced genomes of *Arabidopsis*, rice, and grapevine, the number of genes coding for putative NUDX proteins is 28, 33, and 30, respectively (*12, 13*). RhNUDX1 shows the closest similarity to AtNUDX1 (fig. S1). This protein was proposed to have a similar function to *Escherichia coli* mutator protein (MutT), which acts to eliminate harmful compounds, such as 8-oxo–deoxyguanosine triphosphate (8-oxo-dGTP), which may be misincorporated in DNA during replication (*14*). We have searched rose transcriptome database (*15*) and identified 55 expressed sequence tags (ESTs) corresponding to putative NUDX genes, indicating that, like in other species, NUDX1 belongs to a gene family. All ESTs corresponding to *RhNUDX1* showed high expression levels in fully opened flowers (data S3). The other ESTs show no or weak expression levels in blooming flowers.

In PM, *RhNUDX1* was expressed in petals, where scent is produced, with little to no expression in stamens, sepals, or young leaves (Fig. 2). Expression increased at later stages of flower development (Fig. 2B, stages 3 to 5) when scent
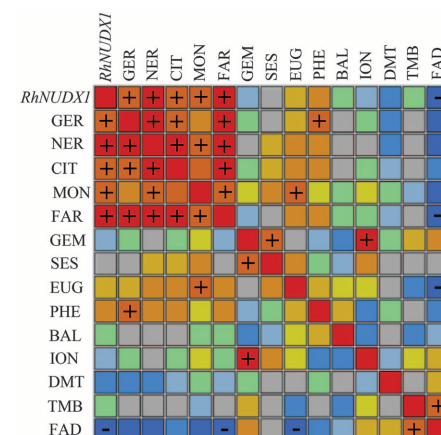
Fig. 1. Correlation map of the expression of *RhNUDX1* and quantity of scent compounds found in 10 rose cultivars. A nonparametric Spearman correlation test was used. Strengths of correlations are depicted by colors. Dark blue squares with minus signs indicate a significant negative correlation with correlation coefficient r close to −1 ($P < 0.05$). Dark red squares with plus signs indicate a significant positive correlation with r close to +1 ($P < 0.05$). BAL, benzyl alcohol and benzaldehyde; CIT, citronellol; DMT, 3,5-dimethoxytoluene; EUG, eugenol and methyleugenol; FAD, hexanal, *E*-2-hexenal, *Z*-3-hexenol, *E*-2-hexenol, 1-hexanol, *Z*-3-hexenyl acetate, hexyl acetate, and nonanal; FAR, *E*-α-farnesene, farnesol, farnesal, and farnesyl acetate; GEM, germacrene D, germacrene D-4-ol, and bicyclogermacrene; GER, geraniol, geranial, geranic acid, and geranyl acetate; ION, 3,4-dihydro-α-ionone and dihydro-α-ionol; MON, β-myrcene, *Z*-β-ocimene, and *E*-β-ocimene; NER, nerol and neral; PHE, 2-phenylethanol and phenylacetaldehyde; SES, δ-cadinene, elemol, α-cadinol, τ-cadinol, and τ-muurolol; TMB, 1,3,5-trimethoxybenzene.

[1]Laboratoire BVpam, EA3061, Université de Lyon/Saint-Etienne, 23 Rue du Dr Michelon, F-42000, Saint-Etienne, France. [2]Laboratoire Reproduction et Développement des Plantes UMR Institut National de la Recherche Agronomique (INRA)–CNRS, Université Lyon 1-ENSL, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France. [3]INRA, Institut de Recherche en Horticulture et Semences (INRA, AGROCAMPUS-OUEST, Université d'Angers), SFR 4207 QUASAV, BP 60057, 49071 Beaucouzé Cedex, France. [4]Génomiques Fonctionnelles d'Arabidopsis, Unité de Recherche en Génomique Végétale, UMR INRA 1165–Université d'Evry Val d'Essonne–ERL CNRS 8196, Evry, France. [5]INRA, Université de Strasbourg, UMR 1131 Santé de la Vigne et Qualité du Vin, 28 Rue de Herrlisheim, F-68000 Colmar, France. [6]Université de Strasbourg–CNRS, UMR 7177, Institut Le Bel, 4 Rue Blaise Pascal, 67070 Strasbourg Cedex, France.
*Corresponding author. E-mail: sylvie.baudino@univ-st-etienne.fr (S.B.); philippe.hugueney@colmar.inra.fr (P.H.); mohammed.bendahmane@ens-lyon.fr (M.B.)

emission is at its maximum (*16*). RhNUDX1 protein accumulated at stage 4 in PM petals at the predicted molecular mass of 16.8 kD, but not in petals of the scentless cultivar RM (Fig. 2C). RhNUDX1 seems to carry no transit peptide, as evidenced by its sequence (*17*) and by cytosolic localization after transient expression in *Nicotiana benthamiana* of RhNUDX1 protein fused to green fluorescent protein (GFP) at its N or C terminus (fig. S2).

Expression of RhNUDX1 correlated with terpene biosynthesis. When we used an RNAi construct to knock down expression of *RhNUDX1* in *Rosa chinensis* cv. Old Blush (OB), a rose that produces high levels of the monoterpene geraniol, three independent transgenic events were generated, one of which (line A) showed reduced *RhNUDX1* expression (Fig. 3A). Volatile monoterpenes geraniol and geranial from petals were positively correlated with *RhNUDX1* expression levels (Fig. 3B and table S3). Line A had the least monoterpene content. Amounts of other volatiles were unaffected (fig. S3). To confirm the consequences of *RhNUDX1* down-regulation, we used a transient transformation strategy, which has been shown to allow efficient gene silencing in floral tissues (*18*). *Agrobacterium*-mediated transient transformation of rose petals was found to be extremely genotype-dependent. The best transformation efficiencies were obtained using the heavily scented genotype known as The Mac Cartney Rose, as shown by the expression of GFP (fig. S4). These petals were infiltrated with *Agrobacterium* harboring the *RhNUDX1* RNA interference (RNAi) construct used for stable transformation. Compared with petals expressing GFP, petals expressing the *RhNUDX1* RNAi construct had fewer monoterpenes; other classes of scent compounds were not significantly affected (fig. S5 and table S4).

To study the genetic basis of monoterpene biosynthesis in roses, we analyzed an F1 progeny from crosses between genotypes with different scent profiles. This mapping population consists of a full-sib family of 116 hybrids derived from a diploid cross between OB and a hybrid *Rosa wichurana* (RW) originating from the Bagatelle garden (Paris, France). The female parent (OB) produced high amounts of geraniol, whereas the male parent (RW) did not. *RhNUDX1* was expressed ~4000 times more in OB than in RW (fig. S6). Gas chromatography–mass spectrometry (GC-MS) analyses of petal volatiles showed that the geraniol content segregated in the progeny (fig. S7C). We detected a major quantitative trait locus (QTL) for geraniol biosynthesis on the female linkage group 2 (LG2). This QTL explained 76% of the observed variation in geraniol content (LOD score 39) (fig. S7B). When mapped on the female LG2, *RhNUDX1* colocalized with this QTL for geraniol production (fig. S7A). Together with genetic colocalization, the transgenic down-regulation data supports a role of RhNUDX1 in geraniol biosynthesis in roses.

Our hypothesis is that the RhNUDX1 protein, located in the cytosol, is in the biosynthetic pathway for monoterpenes. Recombinant RhNUDX1

Fig. 2. Analysis of *RhNUDX1* expression in *Rosa x hybrida*. Flower developmental stages were defined as in (*16*). SE values are indicated by vertical bars (*n* = 6 replicates). Means with different letters (a, b, c) are significantly different (Tukey's HSD test, *P* < 0.05). (**A**) Expression in young leaves and floral organs at stage 4 (S4), analyzed by quantitative polymerase chain reaction (qPCR). (**B**) Expression in petals during development (from S1 to S5), analyzed by qPCR. (**C**) Western blot analysis of RhNUDX1 protein in *Rosa x hybrida* petals. M, Molecular weight marker; R, recombinant protein produced in *E. coli*.

Fig. 3. Characterization of *RNAi-RhNUDX1* transgenic rose lines. (**A**) Real-time quantitative reverse transcription PCR analyses of *RhNUDX1* expression in petals of nontransformed plants, in plants transformed with the Gus reporter gene under the control of the *35S* promoter (*35S::GUS*), and in three transgenic *RNAi-RhNUDX1* lines. SE values are indicated by vertical bars (*n* = 6 replicates). (**B**) GC-MS analyses of the petal volatile monoterpenes in *RNAi-RhNUDX1* transgenic rose lines. Amounts are expressed in micrograms per gram of fresh leaf weight (μg/g FW). SE values are indicated by vertical bars (*n* = 8 to 12 replicates). For both panels, means with different letters (a and b) are significantly different (Tukey's HSD test, *P* < 0.05).

failed to yield geraniol or farnesol after incubation with GPP and farnesyl diphosphate (FPP), suggesting that RhNUDX1 lacks terpene synthase activity (fig. S8). Recombinant RhNUDX1 (Fig. 4B) showed diphosphohydrolase activity when incubated with GPP and FPP. The products were geranyl monophosphate (GP) and farnesyl monophosphate (FP), respectively (Fig. 4A, fig. S9, and table S5). Optimal activity enzymatic activity occurred around pH 8, with very low activity below pH 6 (fig. S10). RhNUDX1 exhibited low Michaelis constant (K_m) values for GPP and FPP (140 and 480 nM, respectively) (table S5). The K_m for GPP was two orders of magnitude lower than that of GES (*9*) and in the same range as that of the MutT nudix protein of *E. coli* with its substrate 8-oxo-dGTP (*14*). Conversely, RhNUDX1 exhibited poor activity with 8-oxo-dGTP and dGTP.

To investigate RhNUDX1 involvement in geraniol biosynthesis, we compared its activity to that of GES from basil (*9*). We used *Agrobacterium*-mediated transient expression to express both in *N. benthamiana*. We verified that the RhNUDX1 protein was accumulated in tobacco leaves (fig. S11). Expression of GES in plastids resulted in the

Fig. 4. Functional characterization of RhNUDX1 in vitro and in planta. (**A**) Decrease in GPP and concomitant increase in GP after incubation with recombinant NusA-RhNUDX1 fusion protein. NusA-RhNUDX1 (250 ng) was incubated in the presence of 1 μM GPP at 30°C. At the indicated time points, GPP and GP were quantified by liquid chromatography–mass spectrometry (LC-MS). (**B**) SDS–polyacrylamide gel electrophoresis analysis of recombinant NusA-RhNUDX1 protein. 1, molecular weight marker; 2, soluble protein extract from NusA-RhNUDX1–expressing *E. coli*; 3 and 4, purified NusA-RhNUDX1 protein. Molecular weights (in kilodaltons) are indicated. (**C**) Accumulation of geranyl glycosides after transient expression

of *RhNUDX1* in *N. benthamiana*. GES from basil (*9*) was expressed as a full-length protein including its transit peptide (plastidic GES, 35*S*::plaGES) and as a truncated protein without transit peptide (cytosolic GES, 35*S*::cytGES). This latter construct was aimed at comparing GES and RhNUDX1 activity in a cytosolic context. GFP was used as a control. Geranyl glycosides were quantified by LC-MS 96 hours after transformation. Amounts are expressed in micrograms per gram of fresh leaf weight (μg/g FW) of geranyl glucoside equivalent, as means of triplicate assays. Bars indicate SE.
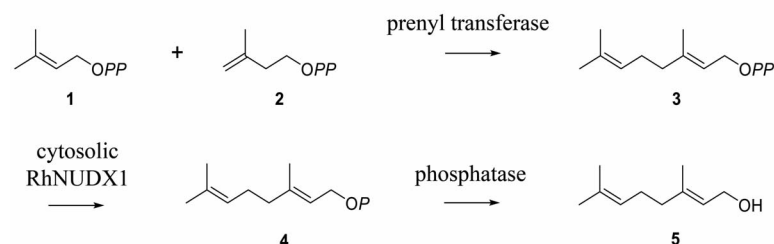
Fig. 5. Hypothetical biogenetic scheme for the formation of free geraniol in rose petals. 1, dimethylallyl diphosphate; **2**, isopentenyl diphosphate; **3**, GPP; **4**, GP; **5**, geraniol.

release of geraniol (fig. S12) and the accumulation of geranyl glycosides (Gglyc) (Fig. 4C). Expression of GES or RhNUDX1 in the cytosol both resulted in similar accumulations of Gglyc and geraniol (Fig. 4 and fig. S12). As RhNUDX1 did not dephosphorylate GP to geraniol in vitro (fig. S8), we assumed that enzymes (presumably phosphatases) carried out that conversion in planta. Indeed, we incubated rose petal proteins with GP and found that it was efficiently converted to geraniol (fig. S13). Although the RhNUDX1 enzyme was able to use FPP as a substrate in vitro, farnesol or farnesol glycoside did not accumulate in RhNUDX1-expressing tissues. Altogether, our data show that RhNUDX1 supports efficient geraniol and Gglyc biosynthesis in planta and that, in a cytosolic context, RhNUDX1 is equivalent to GES.

As GPP is a substrate for RhNUDX1, which is itself localized in the cytoplasm, we ask about the origin of GPP. GPP is a precursor for monoterpenes through pathways generally localized in plastids. Only a few studies suggest that GPP could be generated by the mevalonate pathway in some plants like rose and other Rosaceae (*19–22*). GPP for use in the cytosol for the biosynthesis of monoterpenes may be exported from plastids (*23*) or could also be a by-product of FPP synthase activity (*24*) or may be produced by a cytosolic GPP synthase.

We propose that RhNUDX1 is a cytosolic component of a terpene synthase–independent pathway for monoterpene biosynthesis that leads to scent production in roses (Fig. 5). The intensity of scent in roses may rely on RhNUDX1 and the

monoterpenes it produces, and our finding may provide leverage to restore scent in roses with low levels of RhNUDX1 expression.

REFERENCES AND NOTES

1. A. Vainstein, E. Lewinsohn, E. Pichersky, D. Weiss, *Plant Physiol.* **127**, 1383–1389 (2001).
2. A. M. Borda, D. G. Clark, D. J. Huber, B. A. Welt, T. A. Nell, *Postharvest Biol. Technol.* **59**, 245–252 (2011).
3. A. Joichi, K. Yomogida, K.-I. Awano, Y. Ueda, *Flavour Fragrance J.* **20**, 152–157 (2005).
4. G. Scalliet *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 5927–5932 (2008).
5. Y. Kaminaga *et al.*, *J. Biol. Chem.* **281**, 23357–23366 (2006).
6. I. Guterman *et al.*, *Plant Cell* **14**, 2325–2338 (2002).
7. M. Rodríguez-Concepción, A. Boronat, *Plant Physiol.* **130**, 1079–1089 (2002).
8. F. Chen, D. Tholl, J. Bohlmann, E. Pichersky, *Plant J.* **66**, 212–229 (2011).
9. Y. Iijima, D. R. Gang, E. Fridman, E. Lewinsohn, E. Pichersky, *Plant Physiol.* **134**, 370–379 (2004).
10. M. J. Bessman, D. N. Frick, S. F. O'Handley, *J. Biol. Chem.* **271**, 25059–25062 (1996).
11. A. G. McLennan, *Cell. Mol. Life Sci.* **63**, 123–143 (2006).
12. E. Kraszewska, *Acta Biochim. Pol.* **55**, 663–671 (2008).
13. K. Yoshimura, S. Shigeoka, *Biosci. Biotechnol. Biochem.* **79**, 354–366 (2015).
14. H. Maki, M. Sekiguchi, *Nature* **355**, 273–275 (1992).
15. A. Dubois *et al.*, *BMC Genomics* **13**, 638–648 (2012).
16. V. Bergougnoux *et al.*, *Planta* **226**, 853–866 (2007).
17. O. Emanuelsson, H. Nielsen, S. Brunak, G. von Heijne, *J. Mol. Biol.* **300**, 1005–1016 (2000).
18. Y. Shang *et al.*, *Plant Methods* **3**, 1–12 (2007).
19. P. J. Dunphy, *Phytochemistry* **67**, 1110–1119 (2006).
20. M. J. O. Francis, M. O'Connell, *Phytochemistry* **8**, 1705–1708 (1969).
21. D. Hampel, A. Mosandl, M. Wüst, *J. Agric. Food Chem.* **54**, 1473–1478 (2006).
22. D. Hampel, A. Swatski, A. Mosandl, M. Wüst, *J. Agric. Food Chem.* **55**, 9296–9304 (2007).
23. M. Gutensohn *et al.*, *Plant J.* **75**, 351–363 (2013).
24. S. Frick *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4194–4199 (2013).
25. S. Gagnot *et al.*, *Nucleic Acids Res.* **36** (suppl. 1), D986–D990 (2008).
26. T. Barrett *et al.*, *Nucleic Acids Res.* **35** (suppl. 1), D760–D765 (2007).

CONSERVATION

Genetic assignment of large seizures of elephant ivory reveals Africa's major poaching hotspots

S. K. Wasser,[1]* L. Brown,[2] C. Mailand,[1] S. Mondol,[1]† W. Clark,[3] C. Laurie,[2] B. S. Weir[2]

Poaching of elephants is now occurring at rates that threaten African populations with extinction. Identifying the number and location of Africa's major poaching hotspots may assist efforts to end poaching and facilitate recovery of elephant populations. We genetically assign origin to 28 large ivory seizures (≥0.5 metric tons) made between 1996 and 2014, also testing assignment accuracy. Results suggest that the major poaching hotspots in Africa may be currently concentrated in as few as two areas. Increasing law enforcement in these two hotspots could help curtail future elephant losses across Africa and disrupt this organized transnational crime.

The illegal wildlife trade has become the world's fourth largest transnational organized crime, and African elephant ivory is a major part of that trade (1). An estimated 40,000 African elephants were killed in 2011 (2), with 41 tons of illegal ivory seized (3). Given the 51 tons of ivory seized in 2013, the number of elephants killed that year could have exceeded 50,000, out of an estimated 434,000 elephants remaining (4). Such losses would exacerbate the already serious ecological, economic, and security impacts from rapid decline of this keystone species (2, 3, 5–8).

Here we use DNA-based methods to assign population of origin to African elephant ivory from 28 large seizures (≥0.5 metric tons) made across Africa and Asia from 1996 to 2014 (table S1). Large seizures of raw ivory constitute over 70% of the total weight of ivory shipped since 2006 (7). Our results reveal recurrent patterns that can help guide law enforcement effort to diminish organized trade in ivory and protect vulnerable elephant populations.

We assign geographic origin of large ivory seizures by statistically matching ivory genotypes from savanna or forest elephants to a geographic-specific allele frequency map of 16 microsatellite DNA loci (9–12), assembled separately for each of these two subspecies. We use a Markov chain Monte Carlo algorithm with spatial smoothing (13, 14) to compute separate continuous allele frequency maps across Africa from savanna or forest elephant reference samples. A total of 1350 reference samples, 1001 savanna and 349 forest, were collected at 71 locations across 29 African countries, with 1 to 95

samples per location (table S2). The respective allele frequency maps are then used to assign geographic origin to unknown savanna or forest ivory samples, one sample at a time with a uniform prior (13), or as part of a group with a Voronoi prior (15). We test the accuracy of these results by assigning each of the reference samples, treating them as samples of unknown origin, under several cross-validation scenarios (table S3).

We acquired 20% of all large seizures made between 1996 and 2005, 28% made between 2006 and 2011 (7), and 61% made between 2012 and 2014. All but one of the 28 large ivory seizures (96%) that we assigned were geographically concentrated in a total of four areas, and seizures made after 2007 were concentrated in just two areas (Figs. 1 to 4). Forest elephant ivory in the five consignments seized between 1996 and 2005 were all assigned to eastern Democratic Republic of Congo (DRC) (Fig. 1). Two seizures of savanna elephant ivory made in 2002 and 2007 were both largely assigned to Zambia. The first was concentrated in Kifue National Park, southwestern Zambia, and was the

largest seizure (6.5 tons) since the 1989 ivory ban. The second was concentrated in the Luangwa Valley (Fig. 2). Samples from DRC and Zambia were absent in origin assignments of seizures made after that time.

Six of the seven forest elephant ivory consignments seized between 2006 and 2014 were largely assigned to the area within or in close proximity to the Tridom (Tri-National Dja-Odzala-Minkébé) protected ecosystem that spans northeastern Gabon, northwestern Republic of Congo, and southeastern Cameroon, and the adjacent Dzanga Sangha Reserve in southwestern Central African Republic (CAR) (Fig. 3). The seventh seizure was made during the 2013 warehouse raid in Togo, allegedly belonging to a major ivory dealer in West Africa, and thus was not part of a transiting shipment. The 60 whole tusks from that raid were broadly assigned to West Africa (Ghana, the Ivory Coast, and possibly Togo) and the Tridom and Dzanga Sangha in Central Africa, with a few samples assigned to areas in between those two locales. Consistently, all but one (Fig. 3C) of the other seizures that transited or were seized in Togo included a smaller portion of samples from those same West Africa countries (primarily from Ghana and Côte d'Ivoire). The massive 6-ton seizure made in Malaysia during 2012 (Fig. 3B), which transited through Togo, also included a large portion (61%) of savanna ivory (see below).

Excluding the single 2007 seizure assigned to Zambia (Fig. 2), all of the 15 remaining savanna elephant seizures made between 2006 and 2014 were largely assigned to Tanzania, but spanning into northern Mozambique (Fig. 4). Seven out of the first 10 seizures made between 2006 and 2011 were almost entirely concentrated in the cross-border ecosystem of the Selous and Niassa Game Reserves, respectively spanning southeastern Tanzania and northern Mozambique. The three exceptional seizures (Fig. 4, F, H, and I) each had distinctly more diffuse geographic distributions, as well as atypical transit locations. The 2010 and 2011 Kenya seizures were both made in Jomo Kenyatta International Airport; port seizures of such large size are more

Philippines, 1996–2005
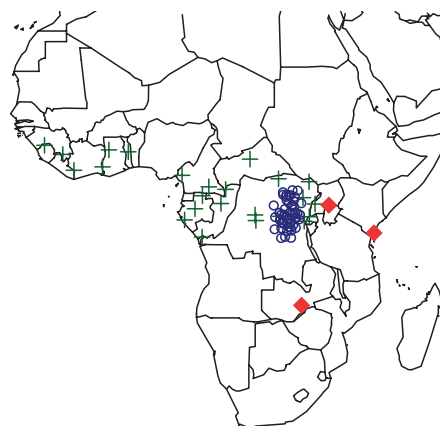
Fig. 1. Philippines seizures from 1996 to 2005 were assigned to eastern DRC. Blue circles represent ivory assignments. Green crosses represent locations of forest elephant reference samples used in the assignments. Red diamonds represent initial transit locations out of Africa: Uganda, Kenya, Zambia.

[1]Center for Conservation Biology, Department of Biology, University of Washington, Box 351800, Seattle, WA 98195-1800, USA. [2]Department of Biostatistics, University of Washington, Box 357232, Seattle, WA 98195-7232, USA. [3]INTERPOL, Environmental Security Sub-Directorate (ENS), Lyon, France.
*Corresponding author. E-mail: wassers@u.washington.edu
†Present address: Wildlife Institute of India, Chandrabani, Dehradun (Uttarakhand), India.

commonly made at shipping ports. The 2010 Kenya seizure also included a high concentration of tusks spanning north-south along eastern Tanzania, which is now devoid of elephants. The third exceptional seizure was the massive 6-ton seizure made in Malaysia (Fig. 4I). This seizure transited through Togo in West Africa, making it the only seizure of savanna ivory in our data set that did not transit eastern Africa. It was also the same seizure that contained a large portion (40%) of forest elephant tusks from the Tridom and surrounding area (Fig. 3B). Thus, large quantities of forest and savanna ivory from the two major hotspot areas identified in Figs. 3 and 4 were consolidated, prior to this consignment being smuggled out of Togo.

The next seizure in this sequence (Fig. 4J) was again concentrated in the Selous-Niassa ecosystem. However, all subsequent seizures showed a progressive shift several hundred kilometers northwest of the Selous, eventually centered around Ruaha National Park and the adjacent Rungwa Game Reserve (Fig. 4), where Tanzania's next-largest elephant population is located. Some of the samples from those latter seizures were also assigned to areas north of Ruaha, creeping toward southern Kenya.

Finally, we note that 23 of the 28 seizures that we examined were shipped (or about to be shipped) out of a different country from where they were assigned (table S1). The first five seizures were shipped out of Tanzania, with the

shift occurring just after Tanzania was identified as a potential hotspot (8, 16).

The results of this study have implications for law-enforcement efforts aimed at tackling transnational organized trade in ivory and the increasing poaching of elephants, provided that our assignments have sufficient accuracy. Accuracy was evaluated by assigning our 1001 savanna and 349 forest reference as unknowns. We either removed one sample, half the samples, or all reference samples for assignment from a sampling location, recalculating allele frequencies in their absence prior to their assignment. Accuracy is expressed as the distance in kilometers between inferred and actual site of geographic origin. Table S3 summarizes these results

Fig. 2. Singapore seizures from 2002 and 2007 were largely assigned to Zambia. Blue circles represent ivory assignments. Orange crosses represent locations of savanna elephant reference samples used in the assignments. Red diamonds represent initial transit locations out of Africa: seizure (**A**) = Malawi via South Africa; seizure (**B**) = Egypt.
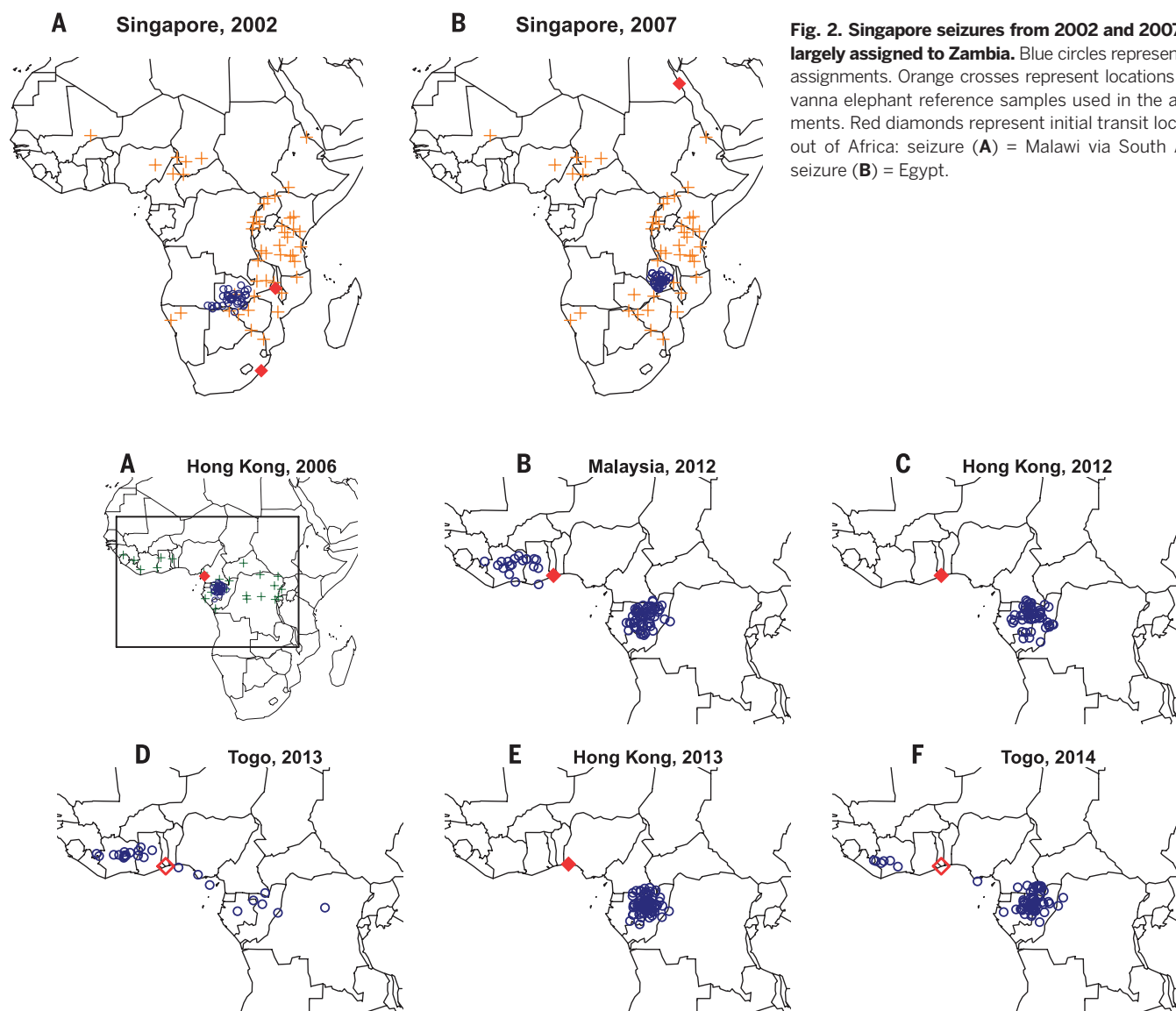
Fig. 3. Forest elephant seizures from 2006 to 2014 were largely assigned to the TRIDOM in northeastern Gabon, northwestern Republic of Congo, and southeastern Cameroon and neighboring Dzanga Sanga in southwestern CAR. Forest elephant reference sample locations (green crosses) are shown in the first map only, along with a box designating the magnified areas displayed in all subsequent maps. Blue circles represent ivory assignments. Initial transit locations out of Africa (solid diamonds) and areas where ivory was seized prior to transit out of Africa (open diamonds) include the following: seizure (**A**) = Cameroon; seizures (**B**) to (**D**) and (**F**) = Togo; seizure (**E**) = Nigeria.

for each of the 71 reference locations, subdivided into the three savanna elephant and two forest elephant geographic regions (*13*). Results also detail how accuracy varies with population structure (i.e., the extent to which populations can be genetically subdivided), the type of prior used [uniform (*13*), analyzing a single sample assignment at a time, or Voronoi (*15*), for group assignment], and sample size.

Half-location-out assignments using a Voronoi prior displayed higher accuracy and less variation than any other assignment combination (table S3). This is the type of assignment portrayed in Figs. 1 to 4 and is most applicable to cases where a moderate to large group of tusks are assigned to a location with multiple reference samples. We find that 50% of forest elephant samples were assigned to within 301 km of their actual origin for anywhere in Africa and the majority (75%) of samples within 416 km of their actual origin. For savanna elephants, 50% of samples were assigned to within 267 km of their actual origin for anywhere in Africa, and the majority within 430 km of their actual origin. The highest accuracy occurred in the most isolated populations, presumably because of population differentiation associated with
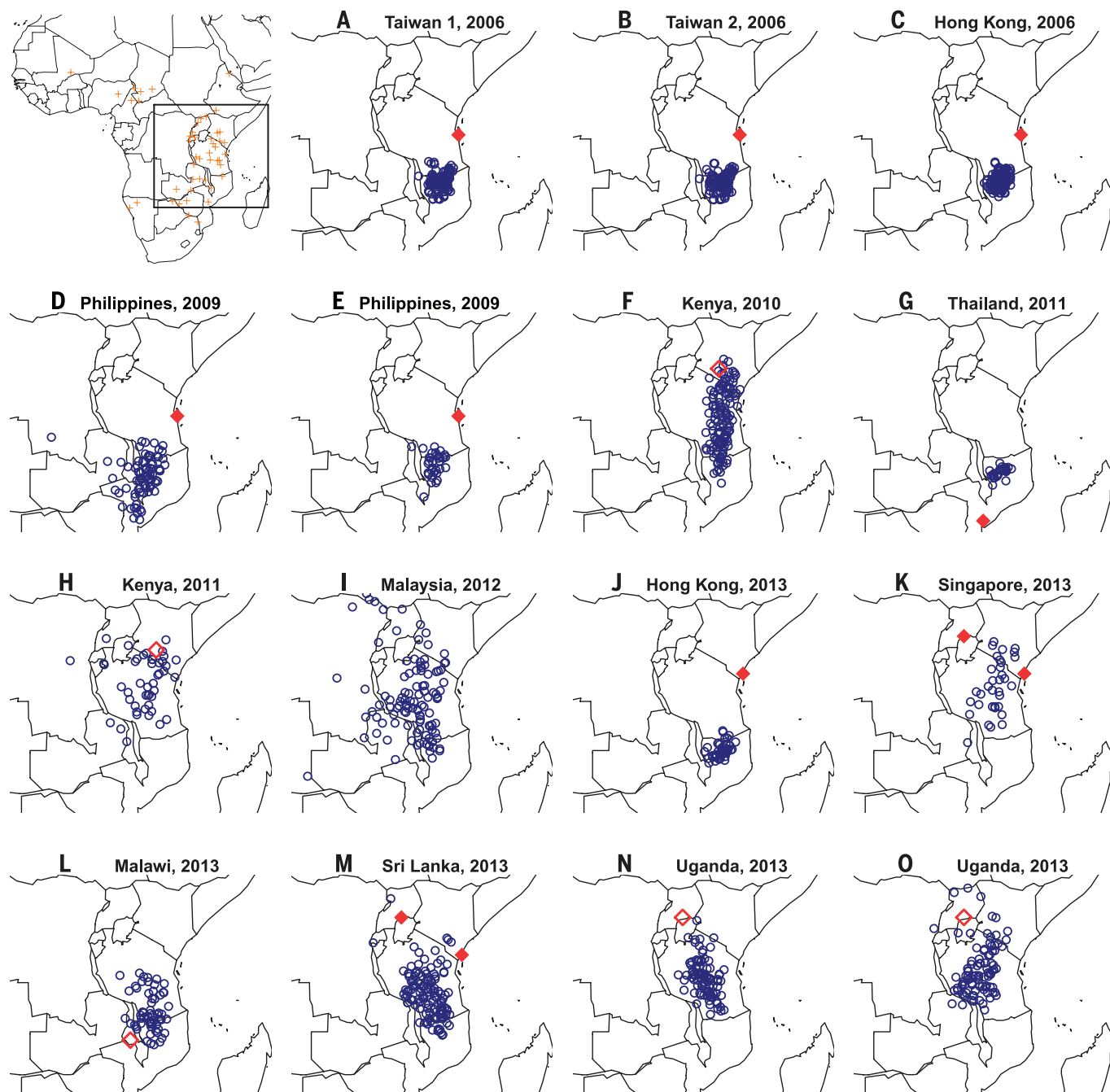
Fig. 4. Savanna elephant seizures from 2006 to 2014 were largely assigned to southeastern Tanzania and adjacent northern Mozambique, but eventually shifted northward within Tanzania. Savanna elephant reference sample locations (orange crosses) are shown in the first map only, along with a box designating the magnified areas displayed in all subsequent maps. Blue circles represent ivory assignments. Most savanna ivory transited (solid diamonds) or was seized (open diamonds) from East African countries: seizures (**A**) to (**E**) = Tanzania; seizures (**F**), (**H**), and (**J**) = Kenya; seizure (**G**) = Mozambique; seizures (**K**) and (**M**) = Uganda to Kenya, seizure (**L**) = Malawi, and seizures (**N**) and (**O**) = Uganda. Seizure (**I**) was exceptional, transiting Togo, West Africa.

isolation. These were the West African forest elephant samples (50% within 161 km, majority within 270 km) and the North African savanna elephant samples (50% within 180 km, majority within 396 km) (table S3). Figure S4, A and B, details assignment results for each of the 42 (out of 71) forest and savanna elephant reference locations with sample sizes ≥ 10.

All of the above precision estimates are conservative because the leave-half-out validations used only half of the reference samples from each area in order to assign the other half as unknowns. The large numbers of ivory samples typically assigned from a large seizure would further increase precision due to the benefits of group assignment using a Voronoi prior (supplementary materials, fig. S2, and table S3).

Our work suggests that the majority of ivory in large seizures since 2006 was poached in just two key areas. We did not obtain all seizures made during that period, but there was only one seizure that deviated from this pattern among the 22 seizures that we analyzed. The exceptional seizure was from Zambia. DRC and Zambia were hotspots prior to 2006 (see also supplementary materials), but seizures of poached ivory from these areas markedly dissipated thereafter. All other large seizures made between 2006 and 2014 consisted of ivory assigned to essentially two areas: Between 86 and 93% of the savanna elephant ivory from that period was predominantly assigned to southeastern Tanzania and adjacent northern Mozambique; 86 to 93% of the forest elephant ivory from that period was predominantly assigned to the Tridom in northern Gabon and Republic of Congo and the adjacent Dzanga Sangha of southwestern CAR.

Maisels *et al.* (*17*) report a 61% decline in Central African forest elephant populations over the past decade, with the last remaining elephant stronghold in Gabon and other areas surrounding the "Tridom." The high concentration of elephants restricted to this part of Central Africa helps explain why it constituted the only forest elephant hotspot in our data set since 2006. By contrast, the greater number of savanna elephants distributed across Africa raises questions about why Tanzania has so consistently remained the primary poaching hotspot for savanna elephants.

If the numbers of hotspots are as few and concentrated as our data suggest, a focus of major international law-enforcement efforts on these key areas could result in substantial reductions in elephant poaching. Targeting those areas would choke the biggest flow of contraband ivory entering the ballooning criminal networks that allow this transnational crime to operate (*1*). We acknowledge the possibility that poaching areas may shift on a shorter time scale than the seizure and processing of illegal ivory, and are taking measures to shorten that time lag.

Other data sets also complement these efforts. MIKE (monitoring illegal killing of elephants) collects country-specific data on change in elephant population sizes (*4*) and the propor-

tion of illegally killed elephants (PIKE) (*3*). ETIS (elephant trade information system) provides data on weights and sizes of all ivory seized along with details on transit locations (*7*). MIKE and ETIS data pertain to all elephant poaching, large and small. Our data are based solely on the sources of large ivory seizures that bear the signature of large-scale organized crime, and these sources are most often different from the country out of which the ivory transited.

Collectively, these combined data sets could provide a comprehensive picture of the illegal ivory trade over time. However, the PIKE data are limited by the small proportion of available carcasses counted (as low as 1% in some of the most heavily poached areas) (*2*) and marked variation in carcass detection probabilities between countries. Data are largely collected by rangers during patrols, and effort—as well as areas covered—varies by year, available funds, and ranger motivation (*2, 7*). Such unaccounted for variation in detection probabilities limits interpretation of carcass occurrence estimates (*18*). PIKE models also suffer from holding natural mortality constant between populations in order to pull out the impacts of poaching (*2*). Poaching causes tremendous variation in age structure, which in turn affects reproductive and mortality rates (*19*). Fortunately, all of these issues can be resolved by training and additional quality controls.

DNA assignment data can be similarly constrained by variation in a country's willingness to provide seizures for analysis in a timely manner, particularly since all seizures are assumed to represent only 10% of what is smuggled (*15*). If this were a serious problem in our study, we would have expected at least some seizures in our analyses to fall outside the small number of hotspots that we identified. Fortunately, our results found virtually no such exceptions. More importantly, in 2013, delegates to the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) unanimously passed Decision 16.83, urging countries to turn over their large seizures for origin analysis within 90 days of the seizure. That decision has greatly increased the rate at which large seizures are now arriving in our lab, substantially improving our ability to track major shifts in poaching pressure on a real-time basis.

Methods that determine geographic origin of seized wildlife contraband offer powerful law-enforcement tools for a variety of forensic applications. This approach requires development of species-specific DNA reference maps, and acquisition of reference DNA from readily available dung samples makes this possible for a variety of species. Only a small number (~25) of reference samples are needed to characterize the allele frequencies from any given protected area, and efficient techniques such as detection dogs can be used to readily locate samples across the target species range (*20*). This study illustrates the value of this relatively inexpensive technology for combating the illegal wildlife trade.

REFERENCES AND NOTES

1. United Nations, *The Globalization of Crime* (United Nations Office on Drugs and Crime, New York, 2010); http://unodc.org/unodc/en/data-and-analysis/tocta-2010.html.
2. G. Wittemyer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13117–13121 (2014).
3. Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), SC65 Doc. 42.1 (2014); http://cites.org/sites/default/files/eng/com/sc/65/E-SC65-42-01_2.pdf.
4. www.elephantdatabase.org/preview_report/2013_africa/Loxodonta_africana/2012/Africa
5. S. Blake, S. L. Deem, E. Mossimbo, F. Maisels, P. Walsh, *Biotropica* **41**, 459–468 (2009).
6. D. Western, *In the Dust of Kilimanjaro* (Island Press, Washington, DC, 2001).
7. F. M. Underwood, R. W. Burn, T. Milliken, *PLOS ONE* **8**, e76539 (2013).
8. S. Wasser *et al.*, *Science* **327**, 1331–1332 (2010).
9. A. L. Roca, N. Georgiadis, J. Pecon-Slattery, S. J. O'Brien, *Science* **293**, 1473–1477 (2001).
10. K. E. Comstock, S. K. Wasser, E. A. Ostrander, *Mol. Ecol.* **9**, 1004–1006 (2000).
11. K. E. Comstock *et al.*, *Mol. Ecol.* **11**, 2489–2498 (2002).
12. S. Nyakaana, P. Arctander, *Mol. Ecol.* **7**, 1436–1437 (1998).
13. S. K. Wasser *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 14847–14852 (2004).
14. P. Vounatsou, T. Smith, A. E. Gelfand, *Biostatistics* **1**, 177–189 (2000).
15. S. K. Wasser *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 4228–4233 (2007).
16. S. K. Wasser, B. Clark, C. Laurie, *Sci. Am.* **301**, 68–76 (July 2009).
17. F. Maisels *et al.*, *PLOS ONE* **8**, e59469 (2013).
18. D. I. MacKenzie *et al.*, *Ecology* **83**, 2248–2255 (2002).
19. K. S. Gobush, B. M. Mutayoba, S. K. Wasser, *Conserv. Biol.* **22**, 1590–1599 (2008).
20. S. K. Wasser *et al.*, *Can. J. Zool.* **82**, 475–492 (2004).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/349/6243/84/suppl/DC1
Materials and Methods
Tables S1 to S3
Figs. S1 to S5
References (*21–33*)

DENGUE VIRUS

Cryo-EM structure of an antibody that neutralizes dengue virus type 2 by locking E protein dimers

Guntur Fibriansah,[1,2] Kristie D. Ibarra,[3] Thiam-Seng Ng,[1,2] Scott A. Smith,[4,5] Joanne L. Tan,[1,2] Xin-Ni Lim,[1,2] Justin S. G. Ooi,[1,2] Victor A. Kostyuchenko,[1,2] Jiaqi Wang,[1,2] Aravinda M. de Silva,[6] Eva Harris,[3] James E. Crowe Jr.,[5,7]* Shee-Mei Lok[1,2]*

There are four closely-related dengue virus (DENV) serotypes. Infection with one serotype generates antibodies that may cross-react and enhance infection with other serotypes in a secondary infection. We demonstrated that DENV serotype 2 (DENV2)–specific human monoclonal antibody (HMAb) 2D22 is therapeutic in a mouse model of antibody-enhanced severe dengue disease. We determined the cryo–electron microscopy (cryo-EM) structures of HMAb 2D22 complexed with two different DENV2 strains. HMAb 2D22 binds across viral envelope (E) proteins in the dimeric structure, which probably blocks the E protein reorganization required for virus fusion. HMAb 2D22 "locks" two-thirds of or all dimers on the virus surface, depending on the strain, but neutralizes these DENV2 strains with equal potency. The epitope defined by HMAb 2D22 is a potential target for vaccines and therapeutics.

Dengue virus (DENV), consisting of four serotypes (DENV1 to 4), is a major human pathogen transmitted by mosquitoes (*1*, *2*). DENV causes disease ranging from mild dengue fever to the severe dengue hemorrhagic fever/dengue shock syndrome. Preexisting antibodies against one serotype can enhance infection by virus of a second serotype. This is probably due to the targeting of virus complexed with non-neutralizing antibodies to monocytic cells via interaction with the Fcγ-receptor, thereby increasing virus infection, a process called antibody-dependent enhancement (ADE) of infection (*3*). Therefore, a safe dengue vaccine should elicit equivalent levels of potently neutralizing responses against all four serotypes. Recent phase 3 clinical trials of a tetravalent vaccine showed poor efficacy, especially against DENV2 (*4*, *5*).

Previous in vitro studies showed that DENV2-specific human monoclonal antibody (HMAb 2D22) has potent neutralization capacity (*6*). We showed that HMAb 2D22 protects against DENV2 when the antibody is administered before (Fig. 1A) or after (Fig. 1B) DENV2 inoculation in an AG129 mouse model (supplementary text). This indicates the potential of using this HMAb as both a prophylactic and therapeutic agent. We also showed that therapeutic administration of the LALA mutant variant of HMAb 2D22 (which abolishes Fc receptor binding) to AG129 mice pretreated with polyclonal DENV1 serum and then inoculated with DENV2, prevents development of antibody-enhanced lethal vascular leak disease (supplementary text) (Fig. 1C).

DENV-neutralizing antibodies primarily target the viral envelope (E) protein. The E protein contains three domains: DI, DII, and DIII (fig. S1). The cryo–electron microscopy (cryo-EM) structure of DENV2 at 4°C (*7*) showed E proteins arranged in icosahedral symmetry, with three individual E proteins (A, B, and C molecules) in each asymmetric unit (fig. S1). The E proteins exist as dimers, and three of the dimers lie parallel to each other, forming a raft (*7*, *8*). The 30 rafts are arranged in a herringbone pattern on the virus surface. This structure represents DENV that was grown in mosquito cells (28°C) and then kept at 4°C. However, when exposed to 37°C, the surface proteins of three DENV2 laboratory-adapted

[1]Program in Emerging Infectious Diseases, Duke–National University of Singapore Graduate Medical School, Singapore. [2]Centre for BioImaging Sciences, National University of Singapore, Singapore. [3]Division of Infectious Diseases and Vaccinology, School of Public Health, University of California, Berkeley, CA, USA. [4]Department of Medicine, Vanderbilt University, Nashville, TN, USA. [5]The Vanderbilt Vaccine Center, Vanderbilt University, Nashville, TN, USA. [6]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC, USA. [7]Departments of Pediatrics and Pathology, Microbiology and Immunology, Vanderbilt University, Nashville, TN, USA.
*Corresponding author. E-mail: sheemei.lok@duke-nus.edu.sg (S.-M.L.); james.crowe@vanderbilt.edu (J.E.C.)

Fig. 1. Prophylactic and therapeutic efficacy of HMAb 2D22 and 2D22-LALA in DENV2-inoculated mice. (**A**) Prophylactic studies. AG129 mice injected with HMAb 2D22 24 hours before sublethal challenge with DENV2 strain D2S10 showed significant reduction in serum viremia and bone marrow viral load, compared to the control mice receiving IgG1 isotype control. The DENV2 limit of detection (LOD) is indicated by the dashed line. Significance was determined using a two-tailed Wilcoxon rank-sum test (*P < 0.05). Shown is one of two representative experiments, with n = 4 mice per group. (**B**) Therapeutic efficacy after high-dose-DENV2-lethal infection. AG129 mice were first inoculated with DENV2 D2S10, and then 20 μg of HMAb 2D22-LALA was administered 24 hours later. Chimeric human-mouse MAb E60-N297Q or IgG1 isotype controls were used to treat control animals. Shown are the combined results from two independent experiments, with n = 9, 8, or 4 mice for the IgG1, 2D22-LALA, or E60-N297Q groups, respectively. Significance was determined using the Mantel-Cox log-rank test (P < 0.001, 2D22-LALA versus IgG1). (**C**) Therapeutic efficacy after ADE-DENV2-lethal infection. AG129 mice were administered a lethal enhancing dose of anti-DENV1 serum 24 hours before inoculation with DENV2 D2S10. The indicated MAb was delivered intraperitoneally 24 hours after inoculation. The results from two independent experiments are shown, with n = 8 mice for the IgG1, 2D22-LALA (20 or 50 μg) and E60-N297Q (20 μg) groups, n = 4 for the group receiving 50 μg of E60-N297Q, and n = 3 for E87-N297Q. Mice receiving 20 or 50 μg of 2D22-LALA displayed a significant level of protection compared to IgG1 control mice (P < 0.001, Mantel-Cox log-rank test).

strains (NGC, WHO, and 16681) undergo structural rearrangement (*9*, *10*), resulting in bumpy-surfaced expanded virus particles (fig. S2). This was also observed in a mouse-adapted DENV2 strain (fig. S2). On the other hand, the virus particles of clinical isolate DENV2 PVP94/07 did not undergo structural changes at 37°C (fig. S2). Thus, we solved the cryo-EM structures of Fab 2D22 complexed with both DENV2(PVP94/07) and DENV2(NGC) strains (fig. S2).

The cryo-EM structures of Fab 2D22:DENV2 (PVP94/07) at 4° and 37°C were determined to a resolution of 6.5 Å (Fig. 2A and figs. S3 and S4) and 7 Å (figs. S3 and S5, A and B), respectively. As the maps were very similar, the 4°C-2D22-PVP94/07 structure was used to identify the Fab-E protein interactions. There are 180 copies of the

Fabs on the virus (Fig. 2A and table S1). The Fab binds across E proteins within a dimer (Fig. 2, B and C). The interactions of the Fabs with each of the three dimers (A-C′, B-B′, and C-A′) in a raft vary slightly (Fig. 2C and table S2). Some of the Fabs also bind to the E protein inter-dimer and inter-raft interface. However, these additional residues are unlikely to be important for antibody binding, because the studies described below show that HMAb 2D22 also binds to the expanded 37°C-DENV2(NGC) structure, which has an altered quaternary structure. In addition, increasing the contour of the 4°C-2D22-PVP94/07 cryo-EM density map showed that the three individual Fab molecule densities in an asymmetric unit are equally strong (fig. S6). This suggests that the interaction of a Fab with an E protein dimer, which

is common between these Fabs in the asymmetric unit, is sufficient for binding.

The light chain of Fab 2D22 bound to DIII and the glycan loop on DI of one E protein, whereas the heavy chain bound to DII, including the fusion loop, of another (Fig. 2C, fig. S7, and table S2). The Fab also caused the glycan-containing loop on DI on an E protein to change in position (Fig. 2D and fig. S4D). Examination of the possible binding of two arms of an immunoglobulin G (IgG) molecule to the DENV virion surface showed that they could bind to adjacent E protein dimers but not to those farther apart (fig. S8).

Previous studies showed that the DENV2 NGC strain can change in structure at 37°C (fig. S2) (*9*), suggesting that antibody binding may be affected. We solved the structure of the 2D22-NGC DENV2
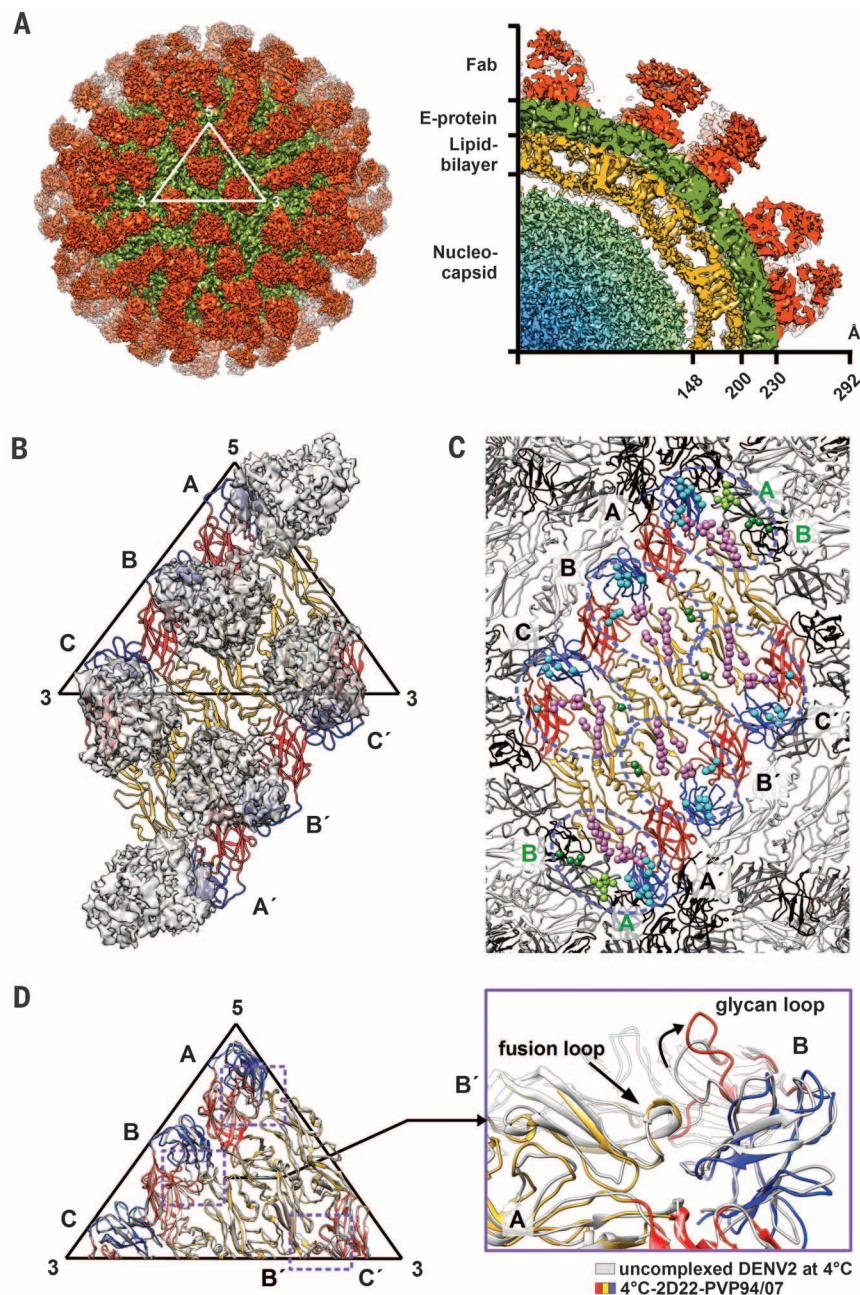
Fig. 2. The 6.5 Å resolution cryo-EM structure of 4°C-2D22-PVP94/07. (**A**) The surface (left) and cross section (right) of the cryo-EM map. One icosahedral asymmetric unit is indicated by a white triangle. (**B**) Three Fab 2D22 molecules bind per asymmetric unit. The three individual E proteins in an asymmetric unit are labeled as molecules A, B, and C. The same molecules from an adjacent asymmetric unit in a raft structure are labeled as A′, B′, and C′. DI, DII ,and DIII of E protein are colored in red, yellow, and blue, respectively. (**C**) The 2D22 epitopes on an E protein raft. The epitopes on molecules A, B, and C are largely similar; however, that on molecule A also has some interactions with adjacent E proteins. The DI, DII, and DIII of the surrounding E proteins are colored in light gray, gray, and black, respectively. The residues that interact with heavy and light chains of the Fab are shown as violet and cyan spheres, respectively. Additional residues from adjacent E protein dimers that formed part of the epitope are shown as green spheres. The boundary of each epitope is indicated with a light blue dotted circle. (**D**) The glycan loop (residues 144 to 157) on one E protein moves away from the fusion loop of the opposite E protein in the dimer when Fab is bound. The E protein of the uncomplexed DENV2 (gray) is superimposed onto the E protein of the complexed structure.

complex at 4º and 37ºC. The 4ºC cryo-EM structure of the NGC-2D22 complex (fig. S9A and table S1) was similar to the structure of the Fab 2D22: DENV2(PVP94/07) complex (Fig. 2A and fig. S5A).

There are two possible scenarios for how this antibody could bind to DENV2 (NGC) when in-fecting humans: It may bind to the virus before and/or after it has changed in structure at 37ºC. Therefore, we examined the virus-Fab complex structures when the antibody was added before (BF-37ºC-2D22-NGC) or after (AF-37ºC-2D22-NGC) shifting the temperature to 37ºC. In both BF-37ºC-

2D22-NGC and AF-37ºC-2D22-NGC samples, two classes of particles (class I and II) were observed. All class I structures were similar to the Fab complexed to the unexpanded virus structure (fig. S9, B and C). Therefore, only class II structures will be described.

The class II BF-37ºC-2D22-NGC structure (20 Å resolution, fig. S10) showed 120 Fab molecules on the virus surface (Fig. 3A, fig. S11, and table S1), with all E protein dimers on a higher radius as compared to the unexpanded 4ºC control complex structure (Fig. 3B, i). Fab molecules remained bound to all A-C dimers, whereas those on the B-B′ dimer had disassociated (Fig. 3A, right). In the BF-37ºC-2D22-NGC structure, the Fab that was previously bound to the B-B′ dimer at 4ºC (fig. S9A) must have dissociated when the temperature was increased. Movement of the A-C dimers from their position in the 4ºC control structure to that in the BF-37ºC-2D22-NGC structure would cause clashes of these E proteins with the Fab that was bound to the B-B′ dimer (Fig. 3C). This steric conflict may "knock off" the Fab and cause the B-B′ dimer to rotate, as observed in the BF-37ºC-2D22-NGC structure. Also, as suggested by the uncomplexed stage 3 DENV2(NGC) 37ºC structure, the B-B′ dimer may dissociate, contributing to the detachment of Fab as observed in the BF-37ºC-2D22-NGC structure.
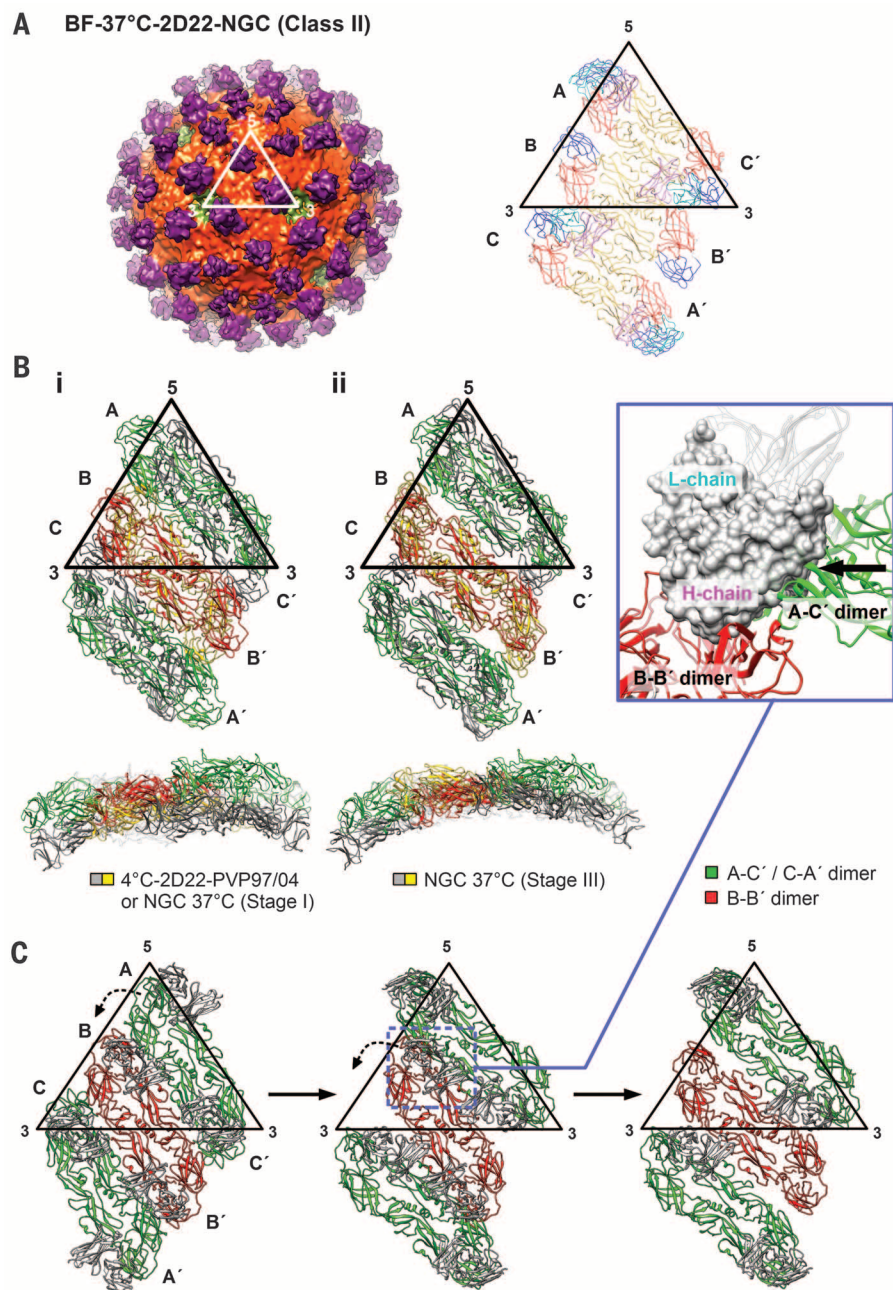
Fig. 3. Cryo-EM structure of BF-37ºC-2D22-NGC (Class II). (A) Cryo-EM map of BF-37ºC-2D22-NGC (Class II) (left) and the fitted structure of the variable region of the Fab complexed with E dimers on a raft (right). The Fab 2D22 heavy and light chains are colored in violet and cyan, respectively. (B) (i) The E proteins of BF-37ºC-2D22-NGC are on a higher radius than those of 4ºC-2D22-NGC or DENV2(NGC) stage 1. (ii) The E protein arrangement of BF-37ºC-2D22-NGC is more similar to that of the 37ºC-DENV2(NGC) (stage 3) structure (top). The A-C′ dimer of BF-37ºC-2D22-NGC is on a slightly higher radius, whereas the B-B′ are located lower when compared to the stage 3 NGC structure. (C) Postulated movement of the Fab-E protein complex from the 4ºC-2D22-NGC to the BF-37ºC-2D22-NGC structures. The partial lifting and rotation of the A-C dimers would knock off the bound Fab on the B-B′ dimer (zoomed-in panel). This may then be followed by the rotation of the B-B′ dimer.
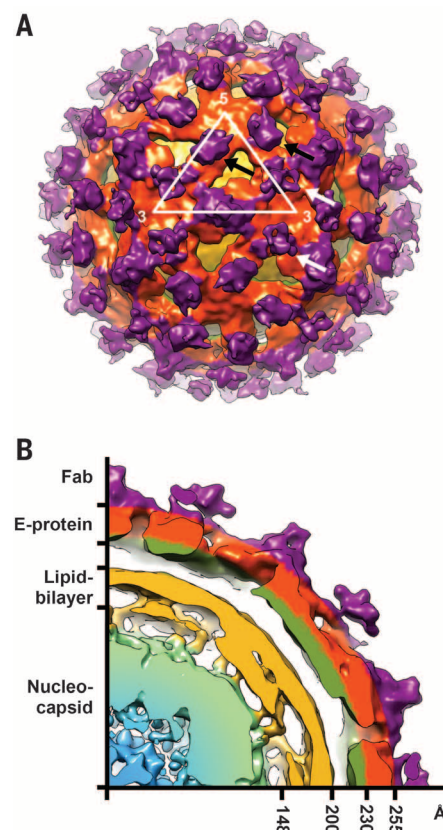
Fig. 4. The cryo-EM map of AF-37ºC-2D22-NGC. (A) Surface of the cryo-EM map and (B) its central cross section. Fabs bound near the five- and three-fold vertices are indicated by black and white arrows, respectively.

The class II BF-37°C-2D22-NGC structure may represent a structure trapped at an intermediate stage of expansion at 37°C. Previous cryo-EM studies of the uncomplexed 37°C DENV2(NGC) sample showed four stages of structural change (9), of which only stage 1 and 3 structures were interpreted. The first stage is similar to the un-expanded structure. The third stage showed that all dimers had moved to a higher radius. The A-C dimer rotated, whereas the B-B′ dimer dissociated from each other. Comparison of the class II BF-37°C-2D22-NGC structure with the DENV2 37°C stage 3 structure (Fig. 3B, ii) showed similar organization to that of the A-C dimers. The B-B′ dimer, on the other hand, seemed to be a transitional structure between the DENV2 NGC 37°C stage 1 and 3 structure, as it lay at a radius in between these two structures (Fig. 3B, i and ii).

The ~21 Å resolution class II AF-37°C-2D22-NGC cryo-EM map (Fig. 4, A and B, and fig. S10) showed clear Fab densities near the fivefold vertices and much weaker Fab densities near the threefold vertices. The E protein densities were sparse; therefore, the map was not interpreted. The positions of Fab density near the fivefold vertices was similar to those in the BF-37°C-2D22-NGC map, but not those near the threefold vertices. The Fab densities near the fivefold and threefold vertices probably represent those that are bound to each end of the A-C′ dimer. Neutralization profiles of BF-37°C-2D22-NGC and AF-37°C-2D22-NGC samples were similar (fig. S12B), suggesting that the poor Fab density near the threefold vertices may be due to local movement in the structure rather than low Fab occupancy.

DENV fusion in endosomes requires E protein dimers to disassociate and then reassociate into trimeric structures. Fab 2D22 locks both ends of all dimers on DENV2(PVP94/07), thereby preventing E protein reorganization. In the BF-37°C-2D22-NGC and AF-37°C-2D22-NGC samples, only two-thirds of the dimers on the virus surface are locked. The remaining free dimer in each raft is probably unable to form trimers. Indeed, HMAb 2D22 effectively neutralized DENV2 strains PVP94/07 and NGC, even though the latter bound one-third fewer antibody molecules (fig. S12A).

Many human antibodies that strongly neutralize dengue bind to quaternary epitopes (epitopes involving more than one E protein molecule) (6, 11, 12). Human antibodies that neutralize DENV serotypes 1 and 3 bound to quaternary epitopes, which require virion assembly (11, 12). In contrast, HMAb 2D22 binds to a simpler epitope that requires only the formation of E homodimers. Several DENV serotype cross-neutralizing human antibodies (13) were also shown to bind E protein dimer epitope (EDE) (fig. S13, B, C, and D). These EDEs are largely similar to 2D22 epitope (fig. S13, A, B and C). Compared to HMAb C8 and C10 (fig. S13, A, B, and C), HMAb 2D22 has more interactions on DIII that are unique to DENV2 (fig. S13E), leading to its serotype specificity.

The DENV2 surface is more dynamic than that of the other serotypes (9, 10). Thus, antibodies that bind across different dimers and rafts may lose potency, depending on the temperature and strain

of DENV2. Therefore, antibodies binding to "simpler" epitopes, such as monomers or dimers, may be more effective against this serotype.

In areas of high dengue endemicity, a potential therapeutic needs to be protective in the presence of preexisting antibodies (14). Certain MAb LALA variants protect therapeutically against an ADE infection, because they are neutralizing and at the same time suppress the enhancing potential of preexisting fusion loop antibodies by displacing their binding (15–17). In contrast, highly neutralizing DIII MAbs that do not block fusion loop enhancing antibodies protect in high-dose-DENV2-lethal, but not ADE-DENV2-lethal challenge (Fig. 1C) (17). The increased efficacy of HMAb 2D22 may be due to its ability to lock E proteins and also block the binding of low-affinity fusion-loop enhancing antibodies.

The molecular features of the 2D22 epitope and the ability of HMAb 2D22-LALA to prevent ADE will aid in the development of vaccines and therapeutics, respectively.

REFERENCES AND NOTES

1. S. Bhatt et al., Nature 496, 504–507 (2013).
2. S. J. Thomas, T. P. Endy, Curr. Opin. Infect. Dis. 24, 442–450 (2011).
3. S. B. Halstead, Adv. Virus Res. 60, 421–467 (2003).
4. L. Villar et al., N. Engl. J. Med. 372, 113–123 (2015).
5. M. R. Capeding et al., Lancet 384, 1358–1365 (2014).
6. R. de Alwis et al., Proc. Natl. Acad. Sci. U.S.A. 109, 7439–7444 (2012).
7. X. Zhang et al., Nat. Struct. Mol. Biol. 20, 105–110 (2013).
8. R. J. Kuhn et al., Cell 108, 717–725 (2002).
9. G. Fibriansah et al., J. Virol. 87, 7585–7592 (2013).
10. X. Zhang et al., Proc. Natl. Acad. Sci. U.S.A. 110, 6795–6799 (2013).
11. G. Fibriansah et al., EMBO Mol. Med. 6, 358–371 (2014).
12. E. P. Teoh et al., Sci. Transl. Med. 4, 139ra83 (2012).
13. A. Rouvinski et al., Nature 520, 109–113 (2015).
14. B. R. Murphy, S. S. Whitehead, Annu. Rev. Immunol. 29, 587–619 (2011).
15. S. J. Balsitis et al., PLOS Pathog. 6, e1000790 (2010).
16. J. D. Brien et al., J. Virol. 87, 7747–7753 (2013).
17. K. L. Williams et al., PLOS Pathog. 9, e1003157 (2013).

SELENOPROTEINS

CRL2 aids elimination of truncated selenoproteins produced by failed UGA/Sec decoding

Hsiu-Chuan Lin,[1,2] Szu-Chi Ho,[1] Yi-Yun Chen,[3] Kay-Hooi Khoo,[2,3] Pang-Hung Hsu,[4] Hsueh-Chi S. Yen[1,2]*

Selenocysteine (Sec) is translated from the codon UGA, typically a termination signal. Codon duality extends the genetic code; however, the coexistence of two competing UGA-decoding mechanisms immediately compromises proteome fidelity. Selenium availability tunes the reassignment of UGA to Sec. We report a CRL2 ubiquitin ligase–mediated protein quality-control system that specifically eliminates truncated proteins that result from reassignment failures. Exposing the peptide immediately N-terminal to Sec, a CRL2 recognition degron, promotes protein degradation. Sec incorporation destroys the degron, protecting read-through proteins from detection by CRL2. Our findings reveal a coupling between directed translation termination and proteolysis-assisted protein quality control, as well as a cellular strategy to cope with fluctuations in organismal selenium intake.

The canonical genetic code includes 20 amino acids. Additionally, selenocysteine (Sec/U) and pyrrolysine (Pyl/O) are the 21st and 22nd amino acids and are coded by the otherwise termination codons UGA and UAG, respectively (1, 2). Sec is cotranslationally incorporated into selenoproteins, a distinct set of proteins largely functioning as oxidoreductases, with Sec in the active sites (3–5). At least 25 selenoproteins have been identified in humans

(6). Translating UGA into Sec requires a Sec insertion sequence (SECIS) element in the 3′ untranslated region (3′UTR) of mRNA transcripts, Sec–transfer RNA (tRNA), Sec-specific elongation factor (eEFSec), and the SECIS-binding protein SBP2 (7–10). This renders UGA/Sec redefinition failure-prone, facing competition between Sec-tRNA and the release factor for UGA decoding (11). The reassignment efficiency is greatly influenced by dietary selenium (12). Stop codon reprograming expands the genetic code at the risk of introducing premature translational termination due to missed stop codon reassignment. Cells process potentially detrimental truncated proteins produced from failed UGA-to-Sec translation via previously undetermined mechanisms.

We developed GPS, a cell-based system for measuring global protein stability (13). In this system, the expression cassette contains a single promoter with an internal ribosome entry site, permitting the expression of two fluorescent proteins from one mRNA transcript. The

first fluorescent protein, RFP (red fluorescent protein), is the internal control, whereas the second fluorescent protein GFP (green fluorescent protein) is fused to the N terminus of the protein of interest. The GFP/RFP ratio is a surrogate for protein stability measurements reading the relative steady-state abundance of GFP-fusion protein over RFP (13, 14). Coupling GPS with functional ablation of ubiquitin ligase, we generated a generic platform to isolate ubiquitin ligase substrates (14, 15). This strategy identified 102 substrates for the CRL2 ubiquitin ligase, including the five selenoproteins SEPHS2, SELV, SEPX1/MSRB1, SELK, and SELS/VIMP from a GPS library containing 15,483 human open reading frames (ORFs) (fig. S1, A and B). We subcloned these selenoprotein genes into backgrounds resembling native transcripts (+UTR, Fig. 1D). Inhibition of CRL2 activity by either genetic perturbation or pharmacological treatment stabilized these selenoproteins, but not their paralogs without Sec (SEPHS1 and SEPW2) (Fig. 1A and fig. S1, C and D). The stability of selenoproteins was positively correlated with selenium availability (Fig. 1B); selenium supplementation attenuated CRL2-mediated selenoprotein degradation (Fig. 1C).

The five selenoproteins identified share no sequence similarity (fig. S1E). We generated

various selenoprotein mutants to uncover the determinants for their degradation (Fig. 1D). Analysis of selenoprotein constructs exclusively expressing truncated proteins (ΔUTR, Δ, UAA, and UAG) and those producing only full-length (FL) proteins via replacement of UGA to other codons, revealed that CRL2 selectively targeted truncated but spared FL selenoproteins (Fig. 1, E and F, and fig. S2, A and B).

We asked whether CRL2 is responsible for removing prematurely terminated selenoproteins arising from failures in UGA/Sec reprograming. FL selenoproteins were more stable than truncated ones (Fig. 2A); the stability of selenoproteins expressed from the UTR construct fell in between, as expected from a mixture of FL and truncated proteins. Indeed, two populations of proteins were translated from the UTR-containing mRNAs upon CRL2 suppression: a shorter product terminated at the UGA codon and a longer read-through product, with the former as the only CRL2 substrate (Fig. 2B). Similar to FL selenoproteins created by substituting the UGA codon (Fig. 1, E and F), Sec-containing FL selenoproteins were stable and exempted from CRL2 surveillance (Fig. 2C and fig. S2C). Selenium availability had no effect on the stability of truncated or FL selenoproteins (fig. S2D). Rather, it enhanced the efficiency of

[1]Institute of Molecular Biology, Academia Sinica, Taiwan. [2]Genome and Systems Biology Degree Program, National Taiwan University, Taiwan. [3]Institute of Biological Chemistry, Academia Sinica, Taiwan. [4]Department of Life Science, Institute of Bioscience and Biotechnology, National Taiwan Ocean University, Taiwan.
*Corresponding author, E-mail: hyen@imb.sinica.edu.tw

Fig. 1. Selenoprotein degradation by CRL2. (**A**) HEK293T GPS reporter cells expressing selenoproteins from the UTR construct were treated or not treated with dominant-negative Cul2 (DNCul2) and then analyzed. (**B**) GPS assay for cells cultured in serum-free medium supplemented with various concentrations of sodium selenite. (**C**) GPS assay for cells cultured in serum-free medium with or without sodium selenite supplement and DNCul2 treatment. (**D**) A schematic representation and nomenclature of each selenoprotein mutant construct. (**E**) GPS analysis of selenoprotein mutants in (D). FL and truncated selenoproteins are presented using different x-axis scales to avoid off-scaling. (**F**) Western blot analysis of SEPHS2 or SELV mutants. FL and truncated (Δ) selenoproteins are indicated by arrowheads. Asterisks mark degradation products from FL SELV. GAPDH and tubulin were loading controls.
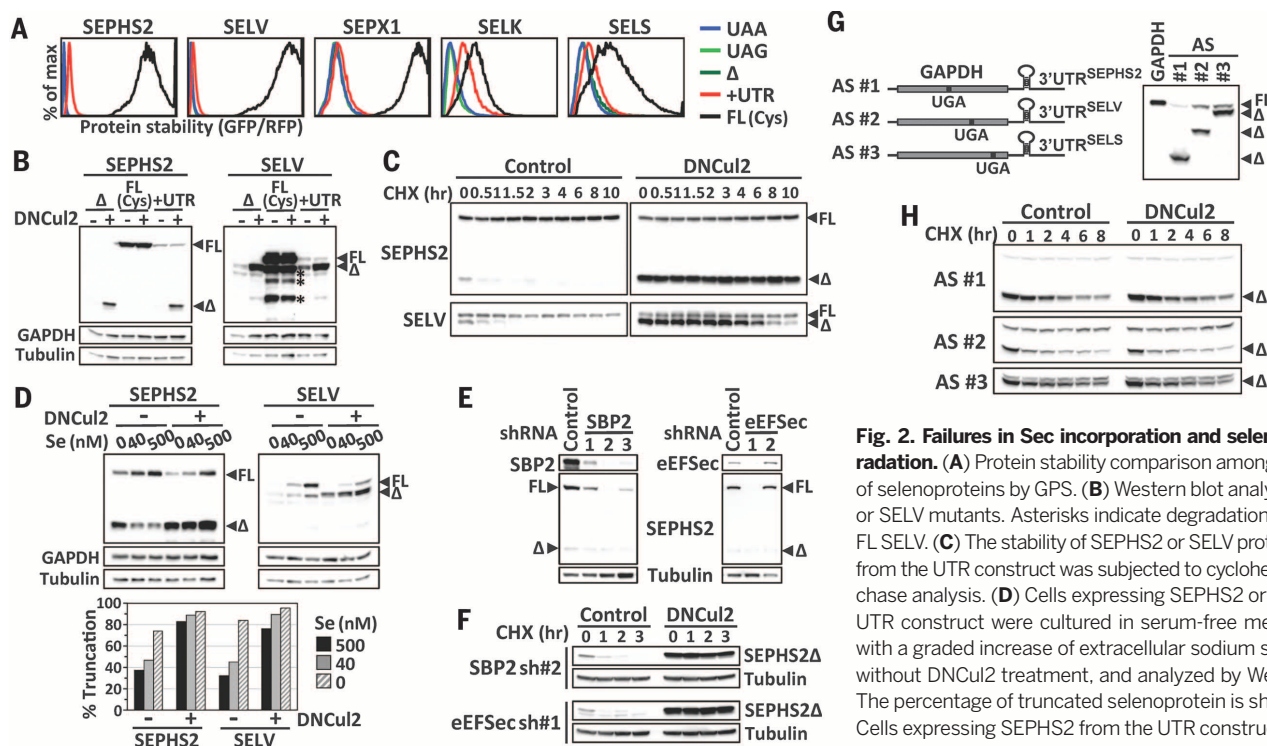
Fig. 2. Failures in Sec incorporation and selenoprotein degradation. (**A**) Protein stability comparison among various forms of selenoproteins by GPS. (**B**) Western blot analysis of SEPHS2 or SELV mutants. Asterisks indicate degradation products from FL SELV. (**C**) The stability of SEPHS2 or SELV proteins expressed from the UTR construct was subjected to cycloheximide (CHX)–chase analysis. (**D**) Cells expressing SEPHS2 or SELV from the UTR construct were cultured in serum-free medium supplied with a graded increase of extracellular sodium selenite, with or without DNCul2 treatment, and analyzed by Western blotting. The percentage of truncated selenoprotein is shown below. (**E**) Cells expressing SEPHS2 from the UTR construct were treated with short hairpin RNAs (shRNAs) for SBP2 or eEFSec and then analyzed by Western blotting. (**F**) The protein stability of SEPHS2 in cells treated with shRNAs for SBP2 or eEFSec was analyzed by CHX chase. (**G**) A schematic representation of GAPDH artificial selenoprotein (AS) transcripts. Cells expressing wild-type GAPDH or AS were analyzed by Western blotting. The introduced in-frame UGA codon terminated GAPDH at amino acid positions 152, 247, and 301. (**H**) CHX-chase analysis of GAPDH expressed from AS transcripts in (G).
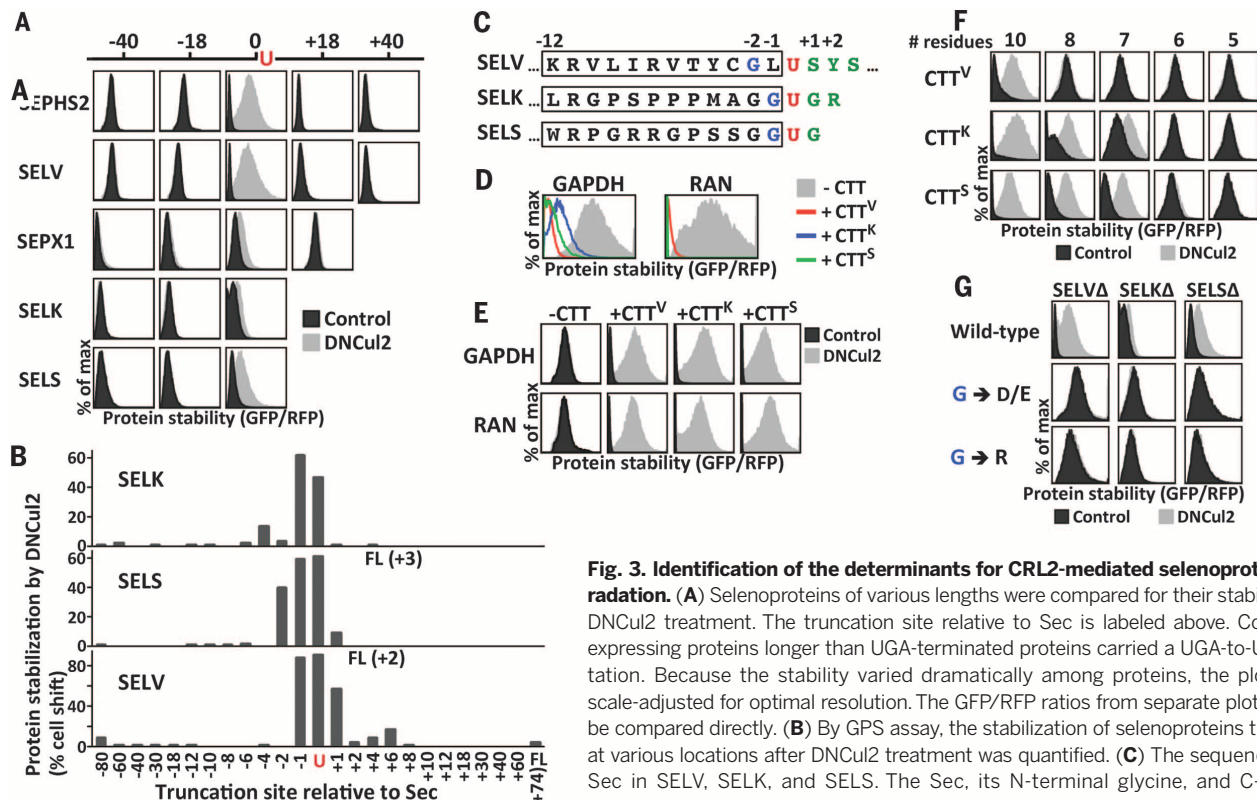
Fig. 3. Identification of the determinants for CRL2-mediated selenoprotein degradation. (**A**) Selenoproteins of various lengths were compared for their stability upon DNCul2 treatment. The truncation site relative to Sec is labeled above. Constructs expressing proteins longer than UGA-terminated proteins carried a UGA-to-UGU mutation. Because the stability varied dramatically among proteins, the plots were scale-adjusted for optimal resolution. The GFP/RFP ratios from separate plots cannot be compared directly. (**B**) By GPS assay, the stabilization of selenoproteins truncated at various locations after DNCul2 treatment was quantified. (**C**) The sequences near Sec in SELV, SELK, and SELS. The Sec, its N-terminal glycine, and C-terminal residues are labeled in red, blue, and green, respectively. The 12-residue C-terminal tail (CTT) is marked with a box. (**D** to **E**) The protein stability of GAPDH or RAN without or with CTT tags at the C terminus was analyzed. CTT^V, CTT^K, and CTT^S, represent the CTT of SELVΔ, SELKΔ, and SELSΔ, respectively. (**F**) The stability of GAPDH tagged with various lengths of CTTs. (**G**) The stability of UGA-terminated selenoproteins with mutations in the glycine N-terminal to Sec.

Sec insertion (Fig. 2D). The accumulation of truncated selenoproteins was heightened by CRL2 inhibition, regardless of the selenium supplies (Fig. 2D and fig. S2E). Endogenously made truncated selenoproteins were accumulated upon CRL2 inhibition (fig. S3). Taken together, these results reveal a role of CRL2 as a gatekeeper in selenoprotein quality control.

CRL2 can recognize truncated SELK and SELS, even though these two proteins differ from their FL counterparts by only three or two amino acids, respectively (fig. S1E). CRL2 may recognize truncated selenoproteins by sensing incomplete translation via scouting the UTR of mRNA transcripts. Alternatively, components in Sec incorporation machinery may assist CRL2 to distinguish UGA-terminated selenoproteins. However, selenoproteins expressed from constructs lacking 3′UTRs or sequences 3′ to UGA remained CRL2 substrates (Fig. 1, E and F). Pro-

tein half-lives of truncated selenoproteins from transcripts with or without 3′UTR were comparable (fig. S4A). Moreover, knocking down SBP2 or eEFSec, which are trans-acting elements essential for Sec incorporation, decreased Sec insertion efficiency but did not influence CRL2-mediated degradation (Fig. 2, E and F). To examine whether CRL2-mediated degradation exhibits substrate specificity, we created artificial selenoproteins; CRL2 could not target UGA-terminated glyceraldehyde phosphate dehydrogenase (GAPDH) made from GAPDH transcripts with in-frame UGA codons within the ORF and 3′UTR from authentic selenoproteins (Fig. 2, G and H). Collectively, our data support the notion of direct recognition of truncated protein products by CRL2.

We systematically shortened (−) or extended (+) the length of the UGA-terminated selenoprotein (Δ) to elucidate how CRL2 recognized

truncated selenoproteins. Truncations had to be made at the position within one to two amino acid residues originally translated into Sec so as to promote CRL2-mediated degradation (Fig. 3, A and B, and fig. S4B). A 12-residue C-terminal tail (CTT) of UGA-terminated selenoproteins was sufficient to promote CRL2-mediated degradation (Fig. 3C and fig. S5, A and B). Fusion of CTTs to GAPDH and RAN, which are not natural CRL2 substrates, resulted in their degradation by CRL2 (Fig. 3, D and E). Thus, CTTs comprise transferable CRL2 degrons (degradation signals). The minimal CRL2 degrons can be as small as 10 or 7 residues in length (Fig. 3F).

We identified a critical glycine at the −1 position of SELK and SELS, or at the −2 position of SELV (Fig. 3C). Replacing this glycine with other amino acids abolishes CRL2-dependent degradation (Fig. 3G and fig. S5B). Changing the leucine at the −1 position of SELV, next to

Fig. 4. The BC-box proteins contribute to CRL2 recognition of truncated selenoproteins. (**A**) GPS assay for cells carrying UGA-terminated (Δ) or FL selenoproteins infected with viruses expressing various BC-box proteins. (**B**) GPS assay for cells expressing Δ or FL selenoproteins treated with shRNAs against BC-box proteins. FL and Δ selenoproteins were presented using different *x*-axis scales. (**C**) GST pull-down using lysates expressing hemagglutinin (HA)–tagged BC-box proteins and GST or GST-tagged Δ or FL selenoproteins. (**D**) GPS assay for cells expressing various forms of SELV or SELK infected with viruses expressing VHL, APPBP2, or KLHDC2. The truncation sites relative to Sec are labeled above. (**E**) GST pull-down using lysates expressing HA-tagged KLHDC2 and GST or various forms of GST-tagged SELK. (**F**) GST pull-down using lysates expressing GFP-tagged GAPDH-CTT fusion and GST or GST-tagged BC-box proteins. (**G**) The protein stability of truncated selenoproteins terminated with various CTTs. (**H**) The stability of UGA-terminated selenoproteins in various cells.
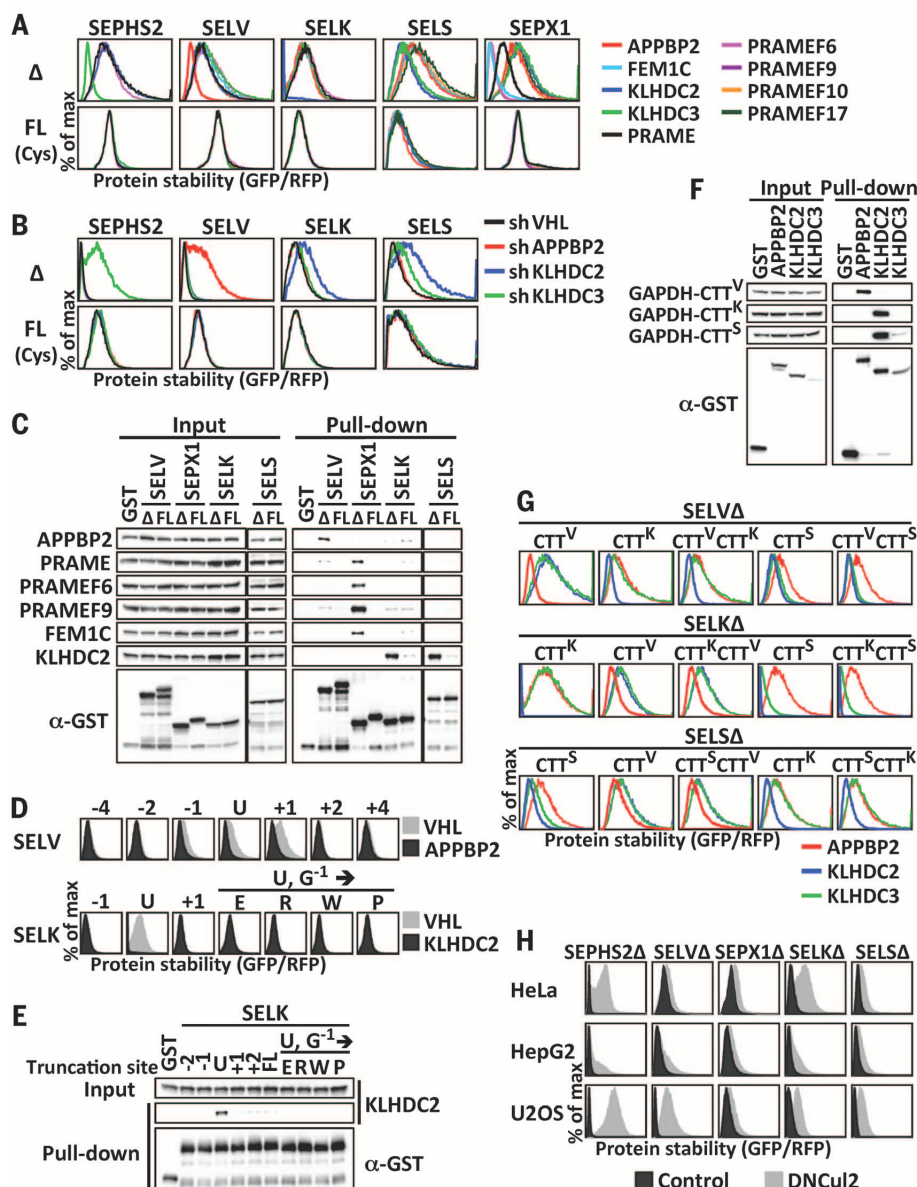
this glycine, did not affect SELV degradation (fig. S5C). The residues C-terminal to Sec, covering the degron, can tolerate more extreme replacement (fig. S5D). Fusing the C-terminal end of UGA-terminated selenoproteins to GFP fully abrogated selenoprotein degradation (fig. S5E). An illegitimate C terminus, once exposed, triggers CRL2-mediated selenoprotein quality control. Rather than serving as a general inspector to eliminate every abnormal selenoproteins, CRL2 mediates the degradation to clear truncated translational products of unsuccessful UGA-to-Sec decoding.

CRL2 is a modular ubiquitin ligase that uses an interchangeable set of BC-box proteins as substrate receptors when forming approximately 40 different CRL2 complexes with a host of substrate specificities ([16]–[19]). We found that CRL2 targeted various UGA-terminated selenoproteins via distinct BC-box proteins (Fig. 4, A and B, and fig. S6). KLHDC3, APPBP2, and KLHDC2 were used to target SEPHS2, SELV and SELK, respectively; SELS could be targeted by both KLHDC2 and KLHDC3. SEPX1 could be recognized by four BC-box proteins, namely PRAME, PRAMEF6, PRAMEF9, and FEM1C. Each BC-box protein was preferentially associated with the corresponding UGA-terminated selenoprotein over the FL one (Fig. 4C and fig. S6, D and E). Furthermore, the substrate selectivity of CRL2 was attributed to BC-box proteins (Fig. 4, D and E). Supporting the idea that CTTs comprise CRL2 degrons, GAPDH-CTT fusions confer binding to the acceptor BC-box proteins (Fig. 4F). Swapping respective BC-box proteins can be achieved by exchanging CTTs or adding an extra CTT C-terminal to UGA-terminated selenoprotein (Fig. 4G and fig. S7A).

In examining the prevalence of proteolysis-assisted selenoprotein quality control, we detected CRL2-mediated quality surveillance for SEPHS2, SELV, SEPX1, SELK, and SELS in all cell types tested (Fig. 4H and fig. S8A). We surveyed six additional selenoproteins and identified one more CRL2 substrate, SEPW1. CRL2 also selectively degraded the UGA-terminated version of SEPW1 while sparing the read-through version (fig. S8B). Regardless of the involvement of CRL2, FL selenoproteins were more stable than their truncated counterparts, and truncated proteins were degraded by the ubiquitin-proteasome pathway (fig. S8, C and D).

When selenium is limited, the transcripts of some, but not all, selenoproteins are driven toward degradation by the nonsense-mediated mRNA decay pathway ([20]–[22]). CRL2 provides an additional layer of defense against translational errors due to the duality in codon assignment. The tiered and complementary nature of these two safeguards grants robustness and fidelity to selenoprotein quality control. Here we report a mechanism by which CRL2 recognizes aberrant selenoproteins via various substrate receptors. We have allocated the peptide immediately N-terminal to Sec as the CRL2-targeting degron triggering degradation when placed at the C-terminal end of a protein (fig. S9A). The CRL2 degrons are highly conserved across species (fig. S9B). The BC-box proteins responsible for selenoprotein recognition do not share common substrate recognition motifs. Instead, they contain various structural motifs involved in general protein-protein interaction, such as Kelch, LRR, ANK, and TPR repeats (fig. S7B). APPBP2, KLHDC2, KLHDC3, PRAME, and FEM1C have all been implicated in human diseases, although selenoproteins are their only identified substrates to date. Beyond exclusively serving as selenoprotein-specific inspectors, these CRL2 substrate receptors may play a broader role in targeted protein degradation.

REFERENCES AND NOTES

1. A. Ambrogelly, S. Palioura, D. Söll, *Nat. Chem. Biol.* **3**, 29–35 (2007).
2. A. V. Lobanov, A. A. Turanov, D. L. Hatfield, V. N. Gladyshev, *Crit. Rev. Biochem. Mol. Biol.* **45**, 257–265 (2010).
3. S. R. Lee et al., *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2521–2526 (2000).
4. D. L. Hatfield, V. N. Gladyshev, *Mol. Cell. Biol.* **22**, 3565–3576 (2002).
5. L. Zhong, A. Holmgren, *J. Biol. Chem.* **275**, 18121–18128 (2000).
6. G. V. Kryukov et al., *Science* **300**, 1439–1443 (2003).
7. L. V. Papp, J. Lu, A. Holmgren, K. K. Khanna, *Antioxid. Redox Signal.* **9**, 775–806 (2007).
8. D. M. Driscoll, P. R. Copeland, *Annu. Rev. Nutr.* **23**, 17–40 (2003).
9. D. L. Hatfield, B. A. Carlson, X. M. Xu, H. Mix, V. N. Gladyshev, *Prog. Nucleic Acid Res. Mol. Biol.* **81**, 97–142 (2006).
10. J. Donovan, P. R. Copeland, *Antioxid. Redox Signal.* **12**, 881–892 (2010).
11. A. Mehta, C. M. Rebsch, S. A. Kinzy, J. E. Fletcher, P. R. Copeland, *J. Biol. Chem.* **279**, 37852–37859 (2004).
12. M. T. Howard, B. A. Carlson, C. B. Anderson, D. L. Hatfield, *J. Biol. Chem.* **288**, 19401–19413 (2013).
13. H. C. Yen, Q. Xu, D. M. Chou, Z. Zhao, S. J. Elledge, *Science* **322**, 918–923 (2008).
14. M. J. Emanuele et al., *Cell* **147**, 459–474 (2011).
15. H. C. Yen, S. J. Elledge, *Science* **322**, 923–929 (2008).
16. N. Mahrour et al., *J. Biol. Chem.* **283**, 8005–8013 (2008).
17. T. Kamura et al., *Genes Dev.* **18**, 3055–3065 (2004).
18. F. Okumura, M. Matsuzaki, K. Nakatsukasa, T. Kamura, *Front. Oncol.* **2**, 10 (2012).
19. M. D. Petroski, R. J. Deshaies, *Nat. Rev. Mol. Cell Biol.* **6**, 9–20 (2005).
20. R. A. Sunde, A. M. Raines, K. M. Barnes, J. K. Evenson, *Biosci. Rep.* **29**, 329–338 (2009).
21. A. Seyedali, M. J. Berry, *RNA* **20**, 1248–1256 (2014).
22. X. Sun et al., *Mol. Biol. Cell* **12**, 1009–1017 (2001).

STRUCTURAL VIROLOGY

Conformational plasticity of a native retroviral capsid revealed by x-ray crystallography

G. Obal,[1,2]* F. Trajtenberg,[3]* F. Carrión,[1] L. Tomé,[1,2]† N. Larrieux,[3] X. Zhang,[4] O. Pritsch,[1,2]‡ A. Buschiazzo[3,5]‡

Retroviruses depend on self-assembly of their capsid proteins (core particle) to yield infectious mature virions. Despite the essential role of the retroviral core, its high polymorphism has hindered high-resolution structural analyses. Here, we report the x-ray structure of the native capsid (CA) protein from bovine leukemia virus. CA is organized as hexamers that deviate substantially from sixfold symmetry, yet adjust to make two-dimensional pseudohexagonal arrays that mimic mature retroviral cores. Intra- and interhexameric quasi-equivalent contacts are uncovered, with flexible trimeric lateral contacts among hexamers, yet preserving very similar dimeric interfaces making the lattice. The conformation of each capsid subunit in the hexamer is therefore dictated by long-range interactions, revealing how the hexamers can also assemble into closed core particles, a relevant feature of retrovirus biology.

Retroviruses undergo an obligatory maturation step in the formation of infectious particles ([1]–[3]). The cleavage of Gag generates several mature proteins, including capsid (CA), which self-assembles into a fullerene-like core, enclosing the RNA genome. Revealing the molecular features of the retroviral mature core and its assembly mechanism is important for understanding retrovirus biology and developing novel antiretroviral drugs.

Retroviral cores display variable shapes and sizes (4, 5), sharing a similar organization into a fullerene-like array of hexameric CA building units, with 12 interspersed CA pentamers to allow for particle closure (6, 7). The CA protomer is composed of two α-helical domains, the N- and C-terminal domains (NTD and CTD) (8–11), connected by a flexible linker (2). The crystal structure of the intact CA molecule has been challenging because of its intrinsic flexibility (12–15) and tendency to self-assemble into heterogenous structures. Stabilization of HIV-1 CA hexamer and pentamer building units by cysteine engineering eventually lead to solving their crystal structures (14, 16), albeit including additional mutations that destabilize critical contacts. Together with state-of-the-art cryoelectron microscopy analyses of

[1]Institut Pasteur de Montevideo, Unit of Protein Biophysics, Mataojo 2020, 11400, Montevideo, Uruguay. [2]Departamento de Inmunobiología, Facultad de Medicina, Universidad de la República, Avenida General Flores 2125, 11800, Montevideo, Uruguay. [3]Institut Pasteur de Montevideo, Unit of Protein Crystallography, Mataojo 2020, 11400, Montevideo, Uruguay. [4]Institut Pasteur, Unité de Virologie Structurale, Département de Virologie and CNRS Unité Mixte de Recherche 3569, 28, Rue du Docteur Roux, 75015, Paris, France. [5]Institut Pasteur, Department of Structural Biology and Chemistry, 25, Rue du Dr Roux, 75015, Paris, France.
*These authors contributed equally to this work. †Present address: Institut Pasteur, Viruses and RNA Interference Unit and Centre National de la Recherche Scientifique, UMR 3569, 28, Rue du Docteur Roux, 75015, Paris, France. ‡Corresponding author. E-mail: pritsch@pasteur.edu.uy (O.P.); alebus@pasteur.edu.uy (A.B.)

tubular assemblies of native HIV-1 CA (17, 18), unprecedented insight into the architecture of the core particle has been provided, extending earlier lower-resolution studies (19): A central fivefold or sixfold symmetric, seemingly rigid ring of NTDs is surrounded by an outer belt of mobile CTDs involved in lateral CTD-CTD dimer contacts among hexamers to form the fullerene lattice (14). Trimeric CTD interfaces, observed only in tubular assemblies, have been proposed to play a key role for lattice curvature (17, 18).

We now report the crystal structure of the native, mature CA protein from bovine leukemia virus (BLV) at 2.75 Å resolution. A tumorigenic B-lymphotropic deltaretrovirus, BLV infects cattle worldwide (20–22) and is closely related to human T-lymphotropic viruses (HTLV) (23). BLV CA (Fig. 1A) crystallized in space group P6₅, with one hexamer in the asymmetric unit (Fig. 1B) and a similar overall architecture as HIV-1 CA (14). However, the BLV CA hexamer deviates substantially from sixfold symmetry, yet is able to make a flat pseudohexagonal lattice (Fig. 1C). The 6_5 screw axis relates successive layers along the c axis of the crystal (fig. S1), with the quasi–sixfold (Q6) axis of the hexamer off-centered from the crystallographic one and tilted by ~7 degrees (Fig. 1D and fig. S1). The structures of the individual NTD and CTD subdomains of BLV CA were determined, respectively, to 1.44 and 2.45 Å resolution (table S1 and fig. S2), which was instrumental in the structure determina-

tion process (24). The NTD and CTD structures are organized as their counterparts in other retroviruses (fig. S3), displaying highest similarity to HTLV-1 (table S2).

The BLV CA hexamer is stabilized through two types of interfaces (Fig. 2), NTD-NTD and NTD-CTD, as in other retroviral CA proteins (11, 14, 18, 25). The core of the inner NTD ring is formed by the first three α helices of each protomer, packed together through van der Waals interactions, with buried surfaces ranging from 360 to 460Å² for each protomer at the six NTD-NTD interfaces (Fig. 2B). Two hydrogen bonds were also consistently observed (Fig. 2B), with one or two additional hydrogen bonds present only in some pairs. Water-mediated hydrogen bonds could not be satisfactorily modeled at this resolution. On the other hand, NTD-CTD intrahexamer contacts were observed both within and between protomers. Intraprotomer contacts associate helices α2 and α7 of the NTD with CTD α8 and the α8-α9 loop, burying 280 to 300 Å². Interprotomer NTD-CTD interactions (Fig. 2C) involve NTD α4 and CTD α8 and α11, burying areas of 280 to 400 Å² per protomer. All interfaces are thus relatively small, consistent with the large variation found in the relative positioning of the various pairs of interacting protomers (Fig. 2, B and C), related to the imperfect Q6 symmetry of the hexamer (Fig. 1B).

Comparison of the crystallized hexamer with the closest sixfold symmetrized hexamer (24)

Fig. 1. Structure of BLV CA and crystal packing.
(**A**) Linear representation of BLV CA with numbered secondary structure elements. Domain boundaries and missing segments found in HIV-1 CA are indicated. (**B**) Schematic representation of the contents of the asymmetric unit of mature BLV CA crystals, corresponding to a hexameric unit. Each protomer is depicted in a different color. The ring of NTDs is outlined, and angles between centers of mass of neighbor NTD pairs are indicated. (**C**) Solvent accessible surface of seven interacting hexameric units, forming a pseudohexagonal planar array. The central hexamer is highlighted with darker tones, and the two colors distinguish NTDs (green) from CTDs (blue). In this orientation, the CTD rings are shown closer to the reader. Independent Q2 (yellow ovals) and Q3 (yellow triangles) axes are highlighted, and interacting CTDs at selected Q2 and Q3 interfaces are outlined (further details in fig. S5). (**D**) Tiled organization of successive hexameric building units in the crystal. The 2D lattice plane is horizontal in this view. The quasi-sixfold noncrystallographic axis is shown as a black stick on the left-most hexamer. The dashed lines (perpendicular to the Q6 axes) connect identical positions of different protomers within the hexamers, highlighting the ~10% vertical translational shift in certain adjacent interfaces.
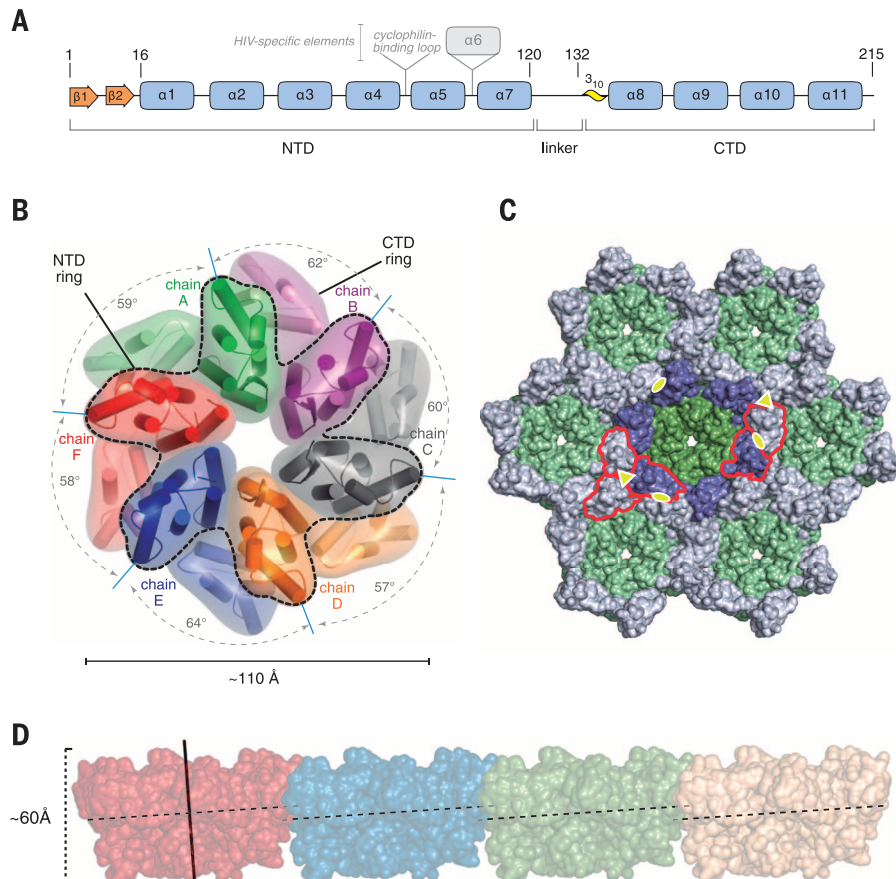
yielded a rather high root mean square deviation (RMSD) of 2.8 Å (for 1136 Cα superposed atoms). Moreover, symmetrizing the NTD and CTD rings separately resulted in RMSDs of 2.4 and 1.9 Å (732 and 420 Cα superposed atoms, respectively), relative to their counterparts in the crystal structure. These deviations from strict C_6 (proper sixfold rotational) symmetry indicate that the CA hexamer is plastic, giving rise to both intracapsomer (fig. S4) and intercapsomer (fig. S1C) quasi-equivalent protein contact interfaces.

In the planar lattice of the crystals, the tilt of the Q6 axis of the hexamer with respect to the crystal c axis results in a tiled pattern of juxtaposed hexamers in one direction (Fig. 1D). The packing is such that it gives rise to three independent quasi-twofold (Q2) and two quasi-threefold (Q3) lateral contacts (Fig. 1C), fitting adjacent hexamers tightly together along continuous planes (fig. S5). The CTD "dimers" related by the three independent Q2 axes are very similar to each other (fig. S6), involving

mainly hydrophobic packing of the 3_{10} helix, the α8-α9 loop, and helix α9. The side chains of residues Trp[133] (within the 3_{10} helix) and Ile[168] (in α9) play prominent roles, resembling the geometry observed in the crystal structures of the isolated HIV-1 CTD dimer (26). Buried interface areas at the BLV Q2 axes vary between 440 and 550Å^2 per protomer, comparable to the single dyad in HIV-1 structures and displaying an overall similar arrangement (14, 27). The Trp[133]/Ile[168] pair (fig. S6) plays an analogous
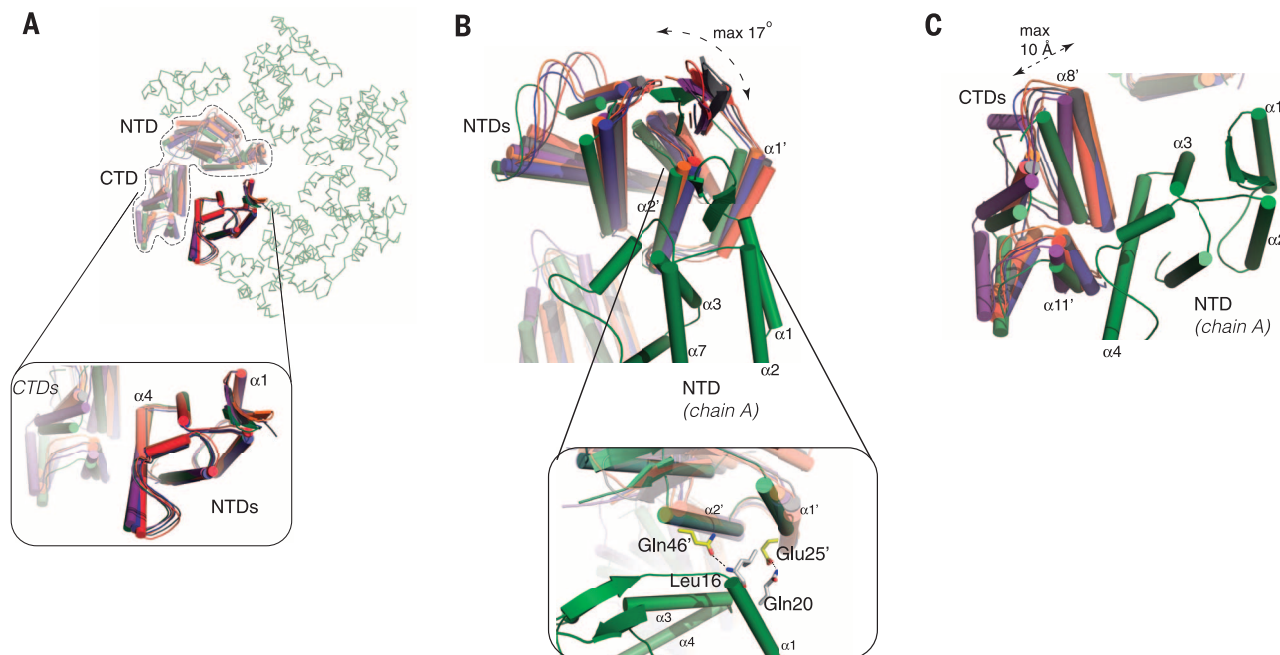
Fig. 2. Variations among intrahexamer contacts. (**A**) The hexamer is shown as a green transparent ribbon. One protomer is outlined for reference. The NTDs of the six chains (solid cartoons, colored as in Fig. 1B) were superposed onto protomer A. The resulting rotated positions of neighboring subunits are shown as semitransparent cartoons. The inset highlights the good superposition among all NTDs (0.5 to 0.6 Å RMSD). (**B**) The view in (A) is rotated by ~45° along a horizontal axis. Only one of the superposed NTDs is shown in green (chain A), with neighboring NTDs colored by chain and α-helix labels identified with a prime. Large variations among different interprotomer NTD-NTD interfaces are observed (maximum angular shift observed between chains A and F is labeled). The inset shows a close-up, with residues (in sticks) involved in H bonds in all the interprotomer pairs. (**C**) The same NTD is shown as a green cartoon (chain A), same orientation as (A). CTD-CTD interactions are highlighted, revealing important variation of CTD-CTD interfaces. The dashed arrow highlights the maximum shift observed between helices α8 of chains A and D.
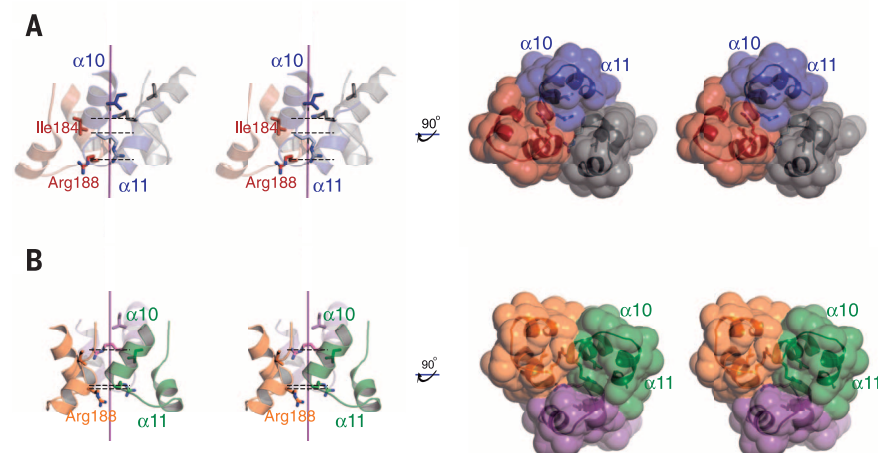
Fig. 3. Close-up of the two independent trimeric interfaces stabilizing the BLV CA lattice, as observed in the crystal asymmetric unit. Stereo views of two orthogonal orientations of the interface between (**A**) protomers B, D, and F and (**B**) protomers A, C, and E, from three neighbor hexamers, are shown. Chains are colored as in Fig. 1B. The panels on the left show side views with respect to the quasi-threefold axes (depicted as purple sticks). The positions of Arg[188] Cα atoms (dashed lines) highlight variable height shifts among the interacting protomers. Arg[188] and Ile[184] (on the same side of α10) are shown as sticks for reference. On the right, solvent-accessible surface representations show the close contacts among the three interacting protomers.

role as Trp[184]/Met[185] of HIV-1, suggesting a functional conservation of these hydrophobic residues. Similarly, the pair Val[163]/Pro[164] in BLV occupies an equivalent position to HIV-1 Val[181]. Residues located at different secondary structure elements direct their side chains to occupy matching volumes in space, docking the CTD dimer unit in a conserved way.

The two independent Q3 contacts, in contrast to the Q2 interfaces, are quite different from each other (Fig. 3). They engage equivalent interaction surfaces as those observed in HIV CA tubes (17), encompassing helices α10 and α11 of the three CTDs. Several side chains at the Q3 interfaces in BLV CA display weak electron density, indicating relative mobility of these residues. The surfaces contributed by chains A-C-E at one of the Q3 axes bury areas of 460, 360, and 325 Å², respectively, whereas the B-D-F interface buries 375, 200, and 350 Å². Residues at the C terminus of α10, mainly Gln[186]-Gly[187]-Arg[188], together with residues from α11, make loose interactions among the three contacting CTDs. The Q3 interfaces seem remarkably supple, consistent with the large variations among the equivalent contact areas present in HIV-1 CA tubes (17), as well as the shifts observed in dehydrated HIV-1 CA crystals (27). Moreover, the threefold interfaces are not restricted to curved lattices, in contrast to previous suggestions (17), because they are present in the planar lattice of BLV CA involving comparable interface areas as the dimeric ones. The two-dimensional (2D) hexagonal arrays in the P6 structures of HIV-1 CA are stabilized via dimer CTD contacts, with no protein interactions on the threefold axes (fig. S5), yet the latter do show key interactions in tubular HIV-1 CA assemblies (17, 18). The P6 crystal packing may thus be favoring a symmetric organization at the expense of lateral contacts. Altogether, a snug fit among hexamers with simultaneous twofold and threefold contacts appears to be linked to global asymmetry in the forms of capsomer deformation or 2D lattice curvature.

To gain further functional insights of CA plasticity, we studied in vitro the self-assembly of full-length BLV CA, comparing the mature form (used to crystallize) to a variant mimicking the immature state. The latter included one additional N-terminal residue coming from Gag, expected to disrupt the N-terminal β hairpin of mature CA. Whereas the mature form generated large 2D sheets, consistent with the planar layers observed in the crystal structure, the one lacking the β hairpin exclusively made tubular assemblies (fig. S7). To understand the striking difference, we attempted crystallization of the immature variant and obtained crystals only for the isolated NTD (termed NTD-nobeta) (table S1). The crystal structure revealed no electron density for the N-terminal segment (amino acids 1 to 11), indicating that these residues are disordered in the absence of β-hairpin formation. It has been postulated that β-hairpin formation upon the maturation cleavage of Gag could be linked to structural rearrangements of helix α1 (28, 29), eventually favoring formation of the mature CA hexamer. Comparing the NTD structures of BLV with and without β hairpin, a significant shift of helix α1 was indeed observed (fig. S8).

The range of different NTD-NTD and NTD-CTD orientations within the hexamer (Fig. 2) provides a connected network for the propagation of long-range forces. Modifications at the NTD correlate with lattice modulation (figs. S7 and S8), even if NTDs are not in direct contact among adjacent hexameric units. Similar observations with HIV-1 CA (30) are consistent with this hypothesis. By binding small molecules (27, 31) or protein partners (32) in the infected cell, NTDs could induce CTD-CTD interface rearrangements deeper within the core, ultimately modulating capsid stability.

In conclusion, we have now captured the natural plasticity of the retroviral capsid assembly unit by working with a CA protein with no engineered mutations. With one hexamer in the asymmetric unit, no crystallographic symmetry constraint is imposed onto individual protomers. Our structure provides a snapshot of one asymmetric conformation, probably among different ones accessible to the hexameric unit. The simplest packing of hexamers into a planar lattice would be a truly symmetric honeycomb array, as observed in the layers of the P6 crystals from HIV-1 CA (14). Such a packing is obviously not favored for BLV CA, likely due to nonoptimal interacting surfaces for hexamers to pile directly on top of each other. Yet the BLV CA hexamer finds a way to pack in 2D sheets by shifting laterally and tilting from the crystal c axis. The constraint of building layers to form a P6₅ crystal leads to the deformation of the ring such that nonidentical Q2 and Q3 lateral contacts of tilted hexamers result in a planar lattice. This is a demonstration of the striking ability of the CA protein to conform to long-range interactions and provides insight on how closure of the CA core may be linked to the actual conformation of the individual hexamers. It also sheds new light into the observed polymorphism of retrovirus mature cores in general. In contrast with the current paradigm, which states that the CTDs in the outer ring adapt to different packing environments by altering their relations to the inner, rigid NTD ring (14, 17), we now observe that the inner NTD ring is also plastic, deviating substantially from a symmetric configuration in order to conform to the dominant long-range interactions. This model of capsid plasticity contributes with a novel framework to develop more accurate hypotheses of capsid self-assembly and uncoating. Implications of this mechanism seem relevant for the discovery of allosteric effectors with novel antiretroviral properties.

REFERENCES AND NOTES

1. J. A. Briggs, H. G. Kräusslich, *J. Mol. Biol.* **410**, 491–500 (2011).
2. B. K. Ganser-Pornillos, M. Yeager, O. Pornillos, *Adv. Exp. Med. Biol.* **726**, 441–465 (2012).
3. F. K. Schur *et al.*, *Nature* **517**, 505–508 (2014).
4. B. K. Ganser-Pornillos, U. K. von Schwedler, K. M. Stray, C. Aiken, W. I. Sundquist, *J. Virol.* **78**, 2545–2552 (2004).
5. J. Benjamin, B. K. Ganser-Pornillos, W. F. Tivol, W. I. Sundquist, G. J. Jensen, *J. Mol. Biol.* **346**, 577–588 (2005).
6. S. Li, C. P. Hill, W. I. Sundquist, J. T. Finch, *Nature* **407**, 409–413 (2000).
7. B. K. Ganser, S. Li, V. Y. Klishko, J. T. Finch, W. I. Sundquist, *Science* **283**, 80–83 (1999).
8. R. K. Gitti *et al.*, *Science* **273**, 231–235 (1996).
9. T. R. Gamble *et al.*, *Science* **278**, 849–853 (1997).
10. C. C. Cornilescu, F. Bouamr, X. Yao, C. Carter, N. Tjandra, *J. Mol. Biol.* **306**, 783–797 (2001).
11. G. B. Mortuza *et al.*, *Nature* **431**, 481–485 (2004).
12. S. Khorasanizadeh, R. Campos-Olivas, M. F. Summers, *J. Mol. Biol.* **291**, 491–505 (1999).
13. R. Campos-Olivas, J. L. Newman, M. F. Summers, *J. Mol. Biol.* **296**, 633–649 (2000).
14. O. Pornillos *et al.*, *Cell* **137**, 1282–1292 (2009).
15. S. Du *et al.*, *J. Mol. Biol.* **406**, 371–386 (2011).
16. O. Pornillos, B. K. Ganser-Pornillos, M. Yeager, *Nature* **469**, 424–427 (2011).
17. G. Zhao *et al.*, *Nature* **497**, 643–646 (2013).
18. I. J. Byeon *et al.*, *Cell* **139**, 780–790 (2009).
19. B. K. Ganser-Pornillos, A. Cheng, M. Yeager, *Cell* **131**, 70–79 (2007).
20. Y. Aida, H. Murakami, M. Takahashi, S. N. Takeshima, *Front Microbiol* **4**, 328 (2013).
21. N. Gillet *et al.*, *Retrovirology* **4**, 18 (2007).
22. S. M. Rodríguez *et al.*, *Viruses* **3**, 1210–1248 (2011).
23. M. Fujii, M. Matsuoka, in *Fields Virology*, D. M. Knipe, P. M. Howley, Eds. (Wolters Kluwer/Lippincott Williams & Wilkins Health, Philadelphia, PA, 2013), vol. 2, chap. 48, pp. 1474–1501.
24. Materials and methods are available as supplementary materials on *Science* Online.
25. G. Cardone, J. G. Purdy, N. Cheng, R. C. Craven, A. C. Steven, *Nature* **457**, 694–698 (2009).
26. D. K. Worthylake, H. Wang, S. Yoo, W. I. Sundquist, C. P. Hill, *Acta Crystallogr. D Biol. Crystallogr.* **55**, 85–92 (1999).
27. A. T. Gres *et al.*, *Science* **349**, 99–103 (2015).
28. K. Chen, G. Piszczek, C. Carter, N. Tjandra, *J. Biol. Chem.* **288**, 1511–1520 (2013).
29. M. Obr *et al.*, *Retrovirology* **11**, 94 (2014).
30. I. Gross, H. Hohenberg, C. Huckhagel, H. G. Kräusslich, *J. Virol.* **72**, 4798–4810 (1998).
31. A. J. Price *et al.*, *PLOS Pathog.* **10**, e1004459 (2014).
32. H. Yang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18372–18377 (2012).

STRUCTURAL VIROLOGY

X-ray crystal structures of native HIV-1 capsid protein reveal conformational variability

Anna T. Gres,[1,2] Karen A. Kirby,[1,3] Vineet N. KewalRamani,[4] John J. Tanner,[2,5] Owen Pornillos,[6] Stefan G. Sarafianos[1,3,5]*

The detailed molecular interactions between native HIV-1 capsid protein (CA) hexamers that shield the viral genome and proteins have been elusive. We report crystal structures describing interactions between CA monomers related by sixfold symmetry within hexamers (intrahexamer) and threefold and twofold symmetry between neighboring hexamers (interhexamer). The structures describe how CA builds hexagonal lattices, the foundation of mature capsids. Lattice structure depends on an adaptable hydration layer modulating interactions among CA molecules. Disruption of this layer alters interhexamer interfaces, highlighting an inherent structural variability. A CA-targeting antiviral affects capsid stability by binding across CA molecules and subtly altering interhexamer interfaces remote to the ligand-binding site. Inherent structural plasticity, hydration layer rearrangement, and effector binding affect capsid stability and have functional implications for the retroviral life cycle.

T he mature capsid of HIV-1 is formed from a single capsid protein (CA) containing N-terminal (CA_{NTD}, residues 1 to 145) and C-terminal (CA_{CTD}, residues 150 to 231) domains connected by a flexible linker region (1–5). The capsid contains ~250 CA hexamers and 12 CA pentamers. CA hexamers comprise six CA_{NTD}s held together by CA_{NTD}-CA_{NTD} and CA_{NTD}-CA_{CTD} contacts between adjacent CAs, whereas the six CA_{CTD}s engage in interhexamer interactions (1–5). CA-CA interactions affect capsid structural integrity and infectivity (1–3, 6–11). Following viral entry, the capsid undergoes controlled disassembly (uncoating), which seems coordinated with reverse transcription (7, 9). Antivirals targeting CA (12–17) include PF-3450074 (PF74), which has a bimodal mechanism of action (18–21). At lower concentrations (nanomolar to ~2 μM), it competes with binding of host factors CPSF6 and NUP153, affecting nuclear entry. At higher concentrations, PF74 blocks uncoating and reverse transcription (18–23). Crystal structures of PF74 with CA_{NTD} (CA_{NTD}-PF74) (14) or cross-linked CA hexamers ($CA_{XL-PF74}$) (18, 19) have shown that PF74 binds at the same site as CPSF6 and NUP153. However, the structural mechanism by which therapeutically relevant high concentrations of PF74 affect uncoating remains incompletely defined.

Cryo–electron microscopy (cryo-EM) studies have helped build informative models of the mature HIV-1 capsid (2, 3, 24). X-ray and nuclear magnetic resonance (NMR) structures of CA domains or CA monomers or dimers have described some interactions at interfaces (1–5, 16, 25, 26). Structures of cross-linked CA (CA_{XL}) elucidated the CA interactions in hexamers and pentamers (1, 3). However, engineered mutations left interhexamer interactions that govern virus uncoating and assembly insufficiently described.

We crystallized native full-length CA (Fig. 1A) and solved its structure in space group P6 with one molecule per asymmetric unit (Fig. 1B and table S1; see also supplementary materials and methods). CA subunits from neighboring hexamers are related by two- and threefold crystallographic symmetry (Fig. 1, C and D). The native CA structure is in general agreement with the 9 Å cryo-EM maps of the flattened CA hexagonal lattice (24) (fig. S1A) and tubes (2) (fig. S1B). The native CA fold is also in agreement with crystal and NMR structures of full-length CA (1, 3–5, 24), CA_{NTD}, and CA_{CTD} (12, 14, 16, 25, 27, 28). Key interactions between CA_{NTD} and CA_{CTD} that are likely to stabilize the capsid are intrahexamer (CA_{NTD}-CA_{NTD}, CA_{NTD}-CA_{CTD}) contacts between adjacent CAs around the sixfold axis, interhexamer (CA_{CTD}-CA_{CTD}) contacts at the two- and threefold axes (Fig. 1D) (1, 2), and intrasubunit (CA_{NTD}-CA_{CTD}) contacts.

The interhexamer interactions at the twofold axis are clearly defined in native CA (Fig. 1E). They involve multiple residues and water molecules (Fig. 2A). These contacts are reminiscent of but different than those in CA_{CTD} structures (fig. S2A and table S2) (16, 25). They also differ considerably from the original cross-linked hexameric [CA_{XL}, Protein Data Bank identification number (PDB ID): 3H47] (fig. S2B and table S2) and other

CA structures containing W184→A184 (W184A) and M185A (29) at the dimerization interface (1, 18, 19).

Cryo-EM and modeling studies indicate that interhexamer interactions are hydrophobic (2, 25). The higher-resolution native CA structure reveals details of these contacts (Fig. 2B and table S3) that include hydrophilic, water-mediated interactions at the two- and threefold interfaces (Fig. 2, A and B). These waters engage in H-bond interactions with either side chains of conserved residues (S149, E175, and W184 at the twofold interface) or main-chain carbonyls (Q176 at the twofold interface; I201 and A204 at the threefold interface) (Fig. 2, A and B). Thus, the closest intersubunit distance at the threefold interface of native CA is ~6 Å versus ~15 Å in CA_{XL} (fig. S3, A and C). Interhexamer interface waters contribute to capsid stabilization. Of 450 assigned waters per hexamer (75 per monomer), 30 are at the two- and threefold interfaces (three waters for each of the six twofold interfaces and two waters for each of the six threefold interfaces per hexamer). As there are ~250 hexamers per capsid, thousands of water molecules at the interhexamer interfaces should substantially contribute to capsid stability.

To investigate the role of this hydration layer, we used controlled dehydration (30) that alters the water content of crystals. CA crystal dehydration (dCA) contracted the unit cell dimensions by ~3 to 6% (table S1). Superposition of native CA on dCA reveals conformational rearrangement in the relative orientation of CA_{NTD} and CA_{CTD}, imparted by a hingelike motion with the linker region as a pivot point (fig. S4). These changes correlate with packing differences at the two- and threefold interfaces (Fig. 2, A versus C and B versus D, and fig. S3, A versus B). Hence, dCA hexamers (fig. S3B) arrange tighter (closest distance ~3 Å) than CA_{XL} (fig. S3C) and native CA (fig. S3A) hexamers, creating up to 22 and 15 additional contacts in the dCA two- and threefold interfaces (tables S2 and S3). Solvent-accessible area calculations reveal that dCA has ~150 and ~200% more buried surface than native CA and CA_{XL} at the twofold interface and ~400% more than CA at the threefold interface (table S4). The threefold interface interactions are between K203-T216 and K203-A217 (main-chain atoms) and A204-L205, G206-T216, and P207-E213 (side chains) (table S3 and Fig. 2D). Sequence alignment (fig. S5) shows that almost all two- and threefold residues are entirely conserved (205, 206, 213, 216; 207 is highly conserved). Consistent with our data, A204 substitutions result in noninfectious virions with unstable and/or abnormal cores (2, 7). We also observe variations in the CA_{NTD}-CA_{NTD} and CA_{NTD}-CA_{CTD} intrahexamer contacts (table S5) that alter H-bond networks and water-mediated or hydrophobic interactions (Fig. 3A). Mutagenesis studies also confirm the importance of interface residues for core morphology and stability, DNA synthesis, and infectivity (tables S2 and S3) (1, 6–11, 26, 31).

A notable difference between native CA and dCA involves interactions of R143 in CA_{NTD} with E175 or Q176 in CA_{CTD} helix H8 within the same

[1]Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, USA. [2]Department of Chemistry, University of Missouri, Columbia, MO 65211, USA. [3]Department of Molecular Microbiology and Immunology, University of Missouri School of Medicine, Columbia, MO 65211, USA. [4]Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA. [5]Department of Biochemistry, University of Missouri, Columbia, MO 65211, USA. [6]Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, Charlottesville, VA 22908, USA.
*Corresponding author. E-mail: sarafianoss@missouri.edu

subunit (Fig. 3, A and B). The R143-Q176 interaction in native CA is replaced with R143-E175 in dCA. This change repositions helix H9, leads to the loss of the W184-bound water, and alters the H9-H9′ twofold interhexamer interactions (Fig. 3B). Alternative interactions of R143 with CA_{CTD} were observed in engineered unliganded (PDB ID: 3H4E) or ligand-complexed structures (PDB ID: 4U0E) crystallized in orthorhombic (1) or

hexagonal space groups, respectively (18). This finding, along with virological studies (7, 18), highlights the functional importance of R143.

Capsid plasticity may allow therapeutic intervention by stabilizing nonproductive uncoating and assembly intermediates or disrupting the interhexamer hydration layer. Residues at interhexamer interfaces are conserved among clades (7, 9, 32), providing a high genetic fragility that

makes CA an attractive antiviral target (9). Among >20 CA-targeting compounds, PF74 is the most potent antiviral (12–17). To address how PF74 alters interhexamer interactions affecting capsid stability and reverse transcription, we solved the crystal structure of PF74 with native CA (CA_{PF74}). PF74 binds across two CA monomers in CA_{PF74}, similar to recent cross-linked structures (18, 19) (Fig. 4A and fig. S6A) but with the indole moiety
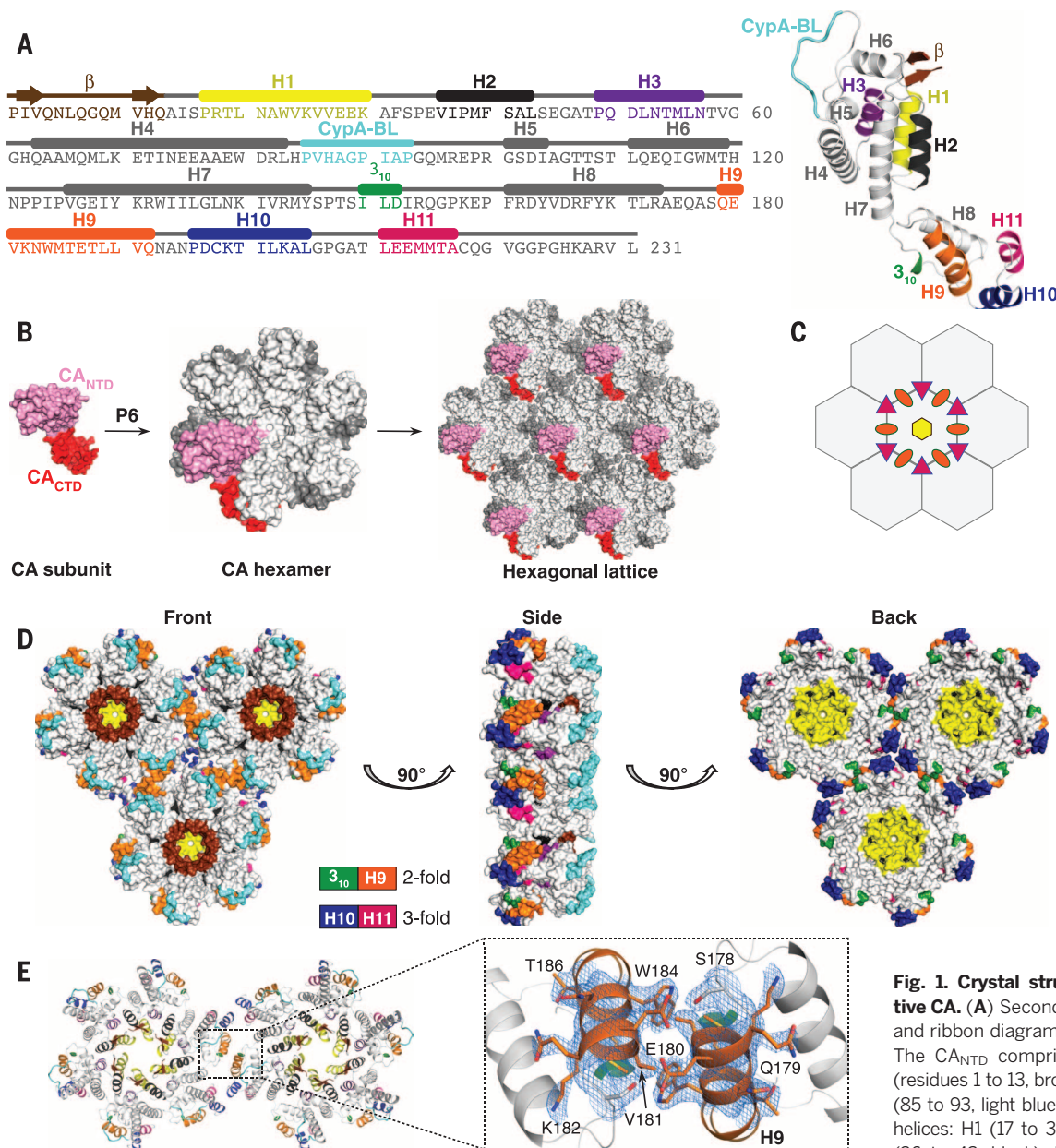
Fig. 1. Crystal structure of native CA. (A) Secondary structure and ribbon diagram of native CA. The CA_{NTD} comprises β hairpin (residues 1 to 13, brown), CypA-BL (85 to 93, light blue), and seven α helices: H1 (17 to 30, yellow); H2 (36 to 43, black); H3 (49 to 57, purple); and H4 (63 to 83), H5 (101 to 104), H6 (111 to 119), and H7 (126 to 145) in gray. The CA_{CTD} comprises the 3_{10} helix (residues 150-152, green) and four α helices: H8 (161 to 173, gray), H9 (179 to 192, orange), H10 (196 to 205, blue), and H11 (211 to 217; pink). (B) Application of sixfold crystallographic symmetry generates the native CA hexagonal lattice. A single native CA molecule is shown in surface view representation (pink, CA_{NTD}; red, CA_{CTD}). (C) The six native CA subunits in a hexamer are related by sixfold crystallographic symmetry (yellow hexagon); CA subunits from neighboring hexamers are related by twofold (orange ovals) and threefold (pink triangles) crystallographic symmetry, shown at the interhexamer interfaces. (D) Orthogonal views of three native CA hexamers colored as in (A). The hexamers are stabilized by interactions at the sixfold (brown, β hairpin; yellow, H1; black, H2; purple, H3), twofold (green, 3_{10}; orange, H9), and threefold (blue, H10; pink, H11) interfaces. (E) Interhexamer interactions at the twofold interface. Interpretable electron density is now observed for all residues at the twofold interface of native CA (2.4 Å; $2F_{obs} − F_{calc}$; σ = 1.2), including residues 176 to 187, which were previously disordered in cross-linked hexamer structures.

arranged differently than in CA$_{NTD-PF74}$ (Fig. 4, B and C). In addition to the interactions with CA$_{NTD}$ helices H3, H4, H5, and H7 observed in CA$_{NTD-PF74}$ (14), PF74 also contacts CA$_{CTD}$ H8 and H9 from a neighboring subunit in CA$_{XL-PF74}$ and CA$_{PF74}$ (18, 19). In CA$_{XL-PF74}$, H9 mutations W184A and M185A leave part of the H9-H9′ twofold interface disordered, and there are no interactions at the neighboring threefold region. These interactions are now observed in CA$_{PF74}$. We also observed

differences in PF74 contacts with CA$_{CTD}$ residues at or near dimerization helix H9 (Y169 and K182 in CA$_{PF74}$ versus Q176, S178, and E179 in CA$_{XL-PF74}$) (Fig. 4, B and C) (18, 19). PF74 binding results in subtle structural changes at the remote three- and twofold regions (Fig. 4D and fig. S6, B and C). These variations are reminiscent of allosteric changes in the CA$_{CTD}$ complex with antiviral peptide CAI binding near the CA$_{CTD}$-CA$_{CTD}$ interface (16). Specifically, we observe changes in hydrophobic and

water-mediated interactions at the CA$_{PF74}$ two-fold region that affect the buried surface area (tables S2 and S4). Also, changes of the CA$_{PF74}$ H10 helices (root mean square deviation ~1 Å with native CA) lead to their convergence at the threefold region, with the distance between A204 main-chain oxygens changing from 5.5 to 3.6 Å. In turn, this change leads to a displacement of the water molecules seen in native CA (Fig. 4D and fig. S6C) and a 150% increase in buried surface
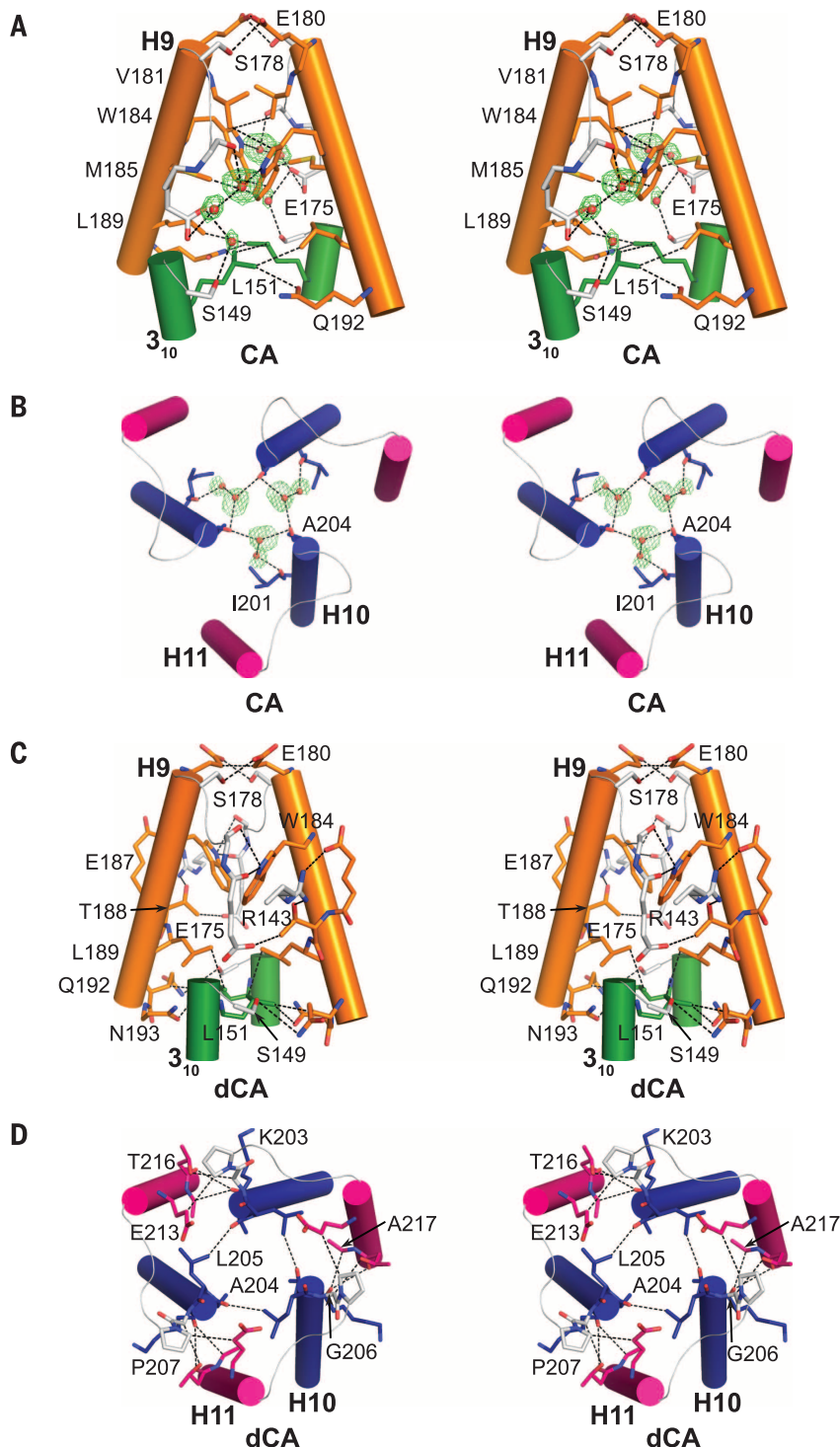
Fig. 2. Interhexamer interactions at the twofold and threefold interfaces. Stereo views of CA$_{CTD}$ regions that are related by twofold [(**A**) and (**C**)] or threefold [(**B**) and (**D**)] symmetry from native CA [(A) and (B)] and dCA [(C) and (D)]. The twofold interface comprises helices H9 and 3$_{10}$ in (A) and (C). The threefold interface comprises helices H10 and H11 in (B) and (D). Ordered water molecules (spheres) at the twofold (A) and threefold (B) interfaces of native CA are modeled in 2.4 Å simulated annealing omit $F_{obs} − F_{calc}$ electron density maps at σ = 2.5 (green mesh). B-factors of refined waters were between 30 to 50 Å^2 and matched well with those of interacting atoms from S149, E175, Q176, W184, I201, and A204. Helices are shown in cartoon representation; black dashed lines between residues (in sticks) indicate that they are within ~4 Å of each other. No water molecules are present at the twofold and threefold interfaces of dCA [(C) and (D)].
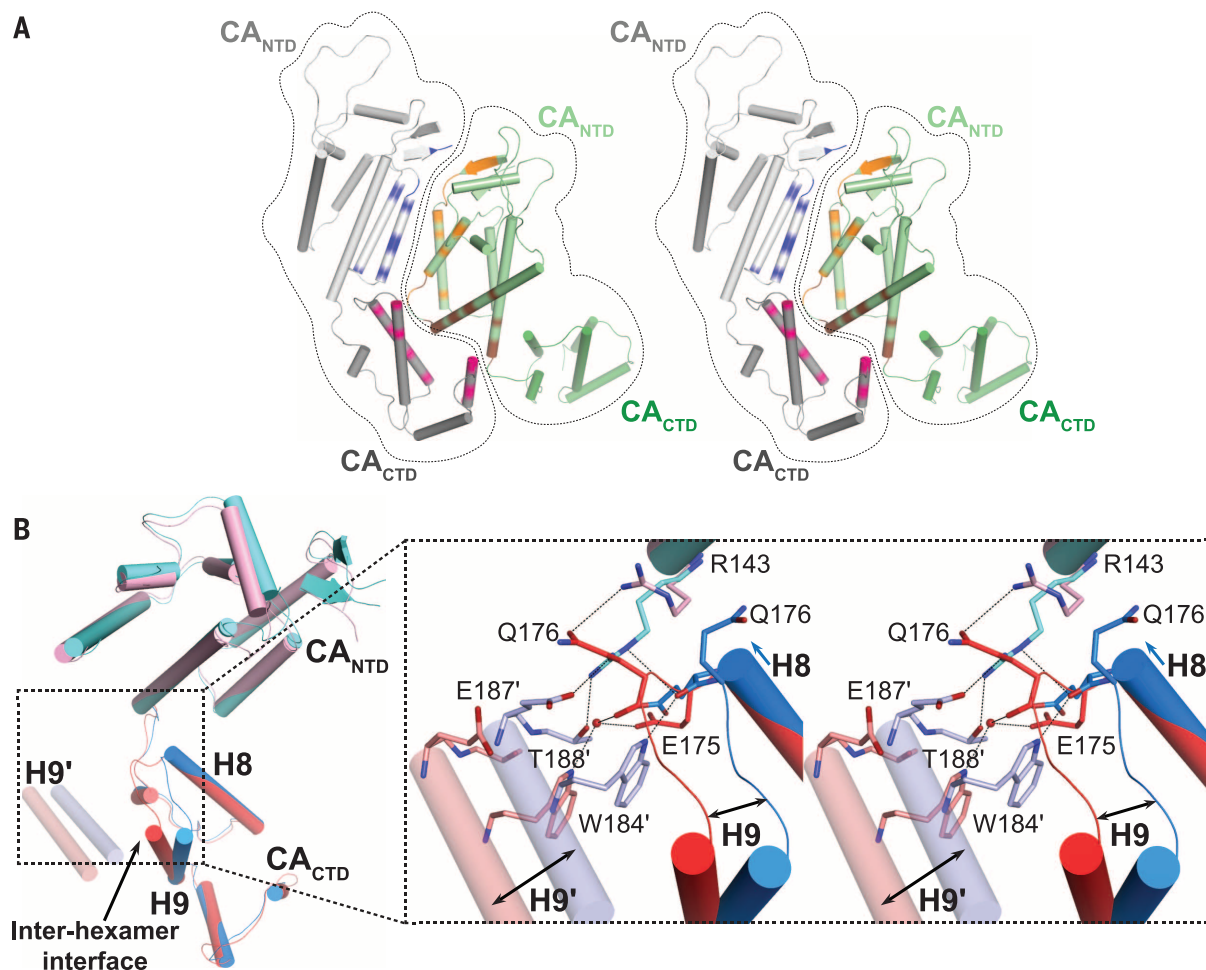
Fig. 3. Changes at the intrahexamer interfaces. (**A**) Stereo view of intrahexamer CA$_{NTD}$-CA$_{NTD}$ and CA$_{NTD}$-CA$_{CTD}$ intersubunit interfaces. Two neighboring CA subunits are shown as cartoons outlined for clarity (CA$_{NTD}$s in lighter colors than the corresponding CA$_{CTD}$s). Sites of varying interactions among native CA, dCA, CA$_{XL}$, and CA$_{PF74}$ are marked in blue and orange for CA$_{NTD}$-CA$_{NTD}$ interfaces or brown and pink for CA$_{NTD}$-CA$_{CTD}$ interfaces. (**B**) Intrasubunit rearrangement linked to changes at the twofold interface. Enlarged stereo view of the boxed region shows changes in the position of H9 helices in neighboring subunits (marked with prime symbols). Least-squares superposition (residues 143 to 174 and 192 to 219) of dCA (cyan, CA$_{NTD}$; darker blue, CA$_{CTD}$) on native CA (pink, CA$_{NTD}$; red, CA$_{CTD}$) is shown. Crystal dehydration results in a slight extension of helix H8 (small blue arrow), interaction of R143 with the main-chain E175 carbonyl instead of Q176, and repositioning of the helix H9 (black arrow). In dCA, W184′ from the H9′ helix (light blue for dCA and light pink for CA) forms a hydrogen bond with main-chain E175 carbonyl from a neighboring subunit, whereas in native CA, W184′ interacts with Q176 and the side chain of E175 through water-mediated contacts. Moreover, in dCA, R143 also interacts with E187′ and T188′ from the neighboring subunit, thus becoming a part of the twofold interface. Black dashed lines connect residues or waters that interact through hydrogen bonds.

area at the threefold interface of CA$_{PF74}$ (tables S3 and S4). Hence, the CA$_{PF74}$ structure provides insights into the mechanism by which high concentrations of PF74 change intra- and interhexamer interactions and affect core stability (tables S2, S3, and S5).

Obal *et al.* report extensive structural heterogeneity of CA monomers in the bovine leukemia virus CA hexamers (*33*). Similarly, comparison of the present native CA, dCA, and CA$_{PF74}$ HIV-1 structures also suggests a CA conformational plasticity, indicating that capsid structural variability is a common feature among multiple retroviruses. Variability in inter- and intrahexamer contacts is also proposed in HIV-1 capsid molecular models (*2, 24*).

The nature of CA structural variability goes beyond mere side-chain rearrangement. A likely key structural determinant important for the stabilization of variable interfaces is the presence of structured waters observed at strategic regions (*1*), including the two- and threefold interfaces (Fig. 2, A and B). This hydration layer could function as an extension of the CA structure, contributing to surface complementarity among flexible CAs. It is likely adaptable and facilitates nearly isoenergetic structural rearrangements contributing to quasi-equivalent structural variability.

Similar to previous reports (*1*), we observed changes in the relative orientation of CA$_{NTD}$ and CA$_{CTD}$ (~10° rotation in fig. S4) that may cause a tilt among neighboring hexamers and contribute to the surface curvature of the capsid. Different interhexamer tilt angles would be anticipated at the narrow and broad ends of the capsid or between laterally or longitudinally positioned hexamers. Given the asymmetric nature of capsid, it is likely that no two CAs are identical in an HIV-1

capsid (*2, 24*). Hence, CA plasticity allows a wide range of conformations that contribute to the structural variability of the HIV-1 asymmetric core.

CA pliability allows interactions of capsids with multiple host factors. For HIV-1, these include TRIM5α, CPSF6, MxB, cyclophilin, NUP153, and NUP358 [reviewed in (*34, 35*)]. Such diverse interactions afford a functional versatility reminiscent of a Swiss Army knife. Moreover, the variability in subunit distances at the threefold interface (~3 to 15 Å) (fig. S3) may affect permeability to deoxynucleoside triphosphates (footprint diameter ≤10 Å). This may be relevant to reverse transcription, although nucleotide trafficking may also occur through imperfections in the malleable core structure. Extensive variability may become an Achilles heel for HIV by providing opportunities for pocket targeting, core destabilization, or stabilization of nonproductive structural intermediates.
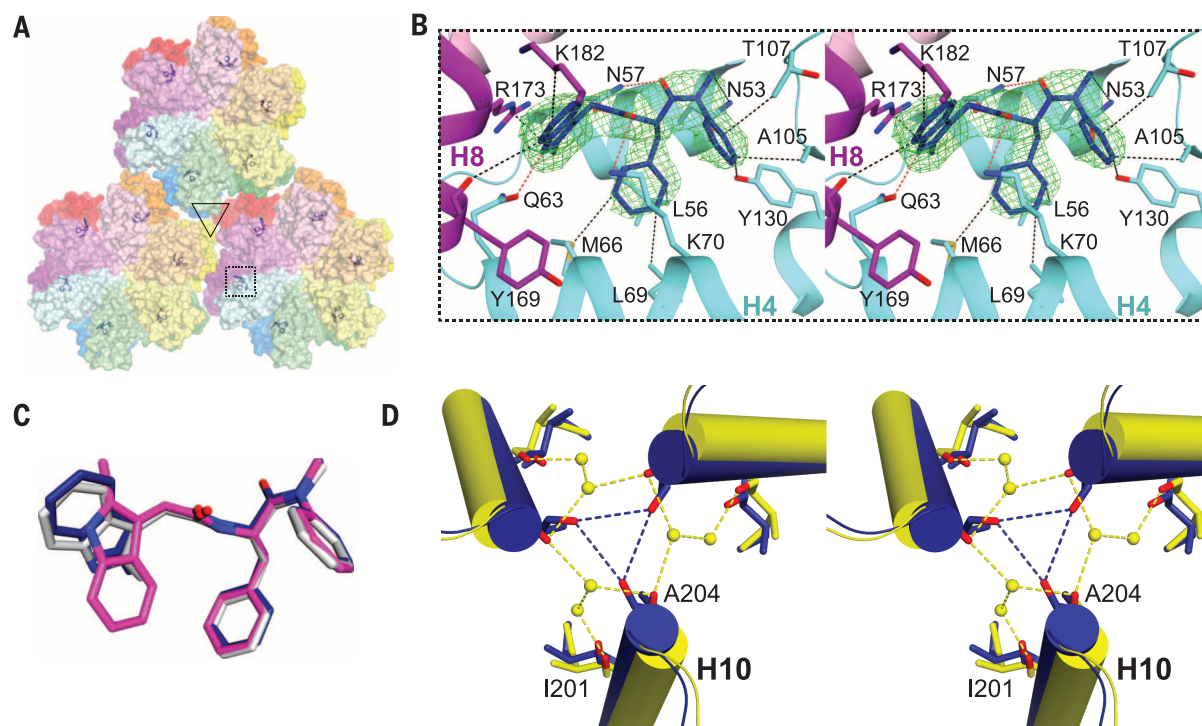
Fig. 4. Effects of PF74 on HIV-1 CA structure. (**A**) PF74 binding at the CA$_{NTD}$-CA$_{CTD}$ interfaces of neighboring subunits within a CA hexamer. Top view of CA$_{PF74}$ hexamers is shown (side view is shown in fig. S6A). There is one PF74 molecule bound to every CA subunit at sites that are distant to the threefold interface (black triangle). CA$_{NTD}$s and the corresponding CA$_{CTD}$s are colored by the same colors (light and dark, respectively). (**B**) Close-up stereo view of the PF74 binding site [small box in (A)]. CA$_{NTD}$ and a CA$_{CTD}$ of a neighboring subunit bind PF74 (light blue and purple ribbons). H9 is omitted for clarity. PF74 is modeled in a 2.7 Å simulated annealing omit map (σ = 2.5). Black dashed lines indicate interactions within ~4 Å, and red dashed lines indicate H-bond interactions. (**C**) PF74 conformations in CA$_{PF74}$ (blue), CA$_{XL-PF74}$ (PDB ID: 4U0E; PF74 in gray), and CA$_{NTD-PF74}$ (PDB ID: 2XDE; PF74 in magenta). PF74 in PDB ID 4QNB is almost identical to 4U0E and is omitted for clarity. (**D**) Stereo view of the threefold interface of CA$_{PF74}$ superposed onto native CA (aligned on residues 1 to 219). Helices H10 of CA$_{PF74}$ and CA are in blue and yellow, respectively. CA waters are shown as yellow spheres; no waters were present in CA$_{PF74}$.

REFERENCES AND NOTES

1. O. Pornillos et al., Cell **137**, 1282–1292 (2009).
2. G. Zhao et al., Nature **497**, 643–646 (2013).
3. O. Pornillos, B. K. Ganser-Pornillos, M. Yeager, Nature **469**, 424–427 (2011).
4. L. Deshmukh et al., J. Am. Chem. Soc. **135**, 16133–16147 (2013).
5. S. Du et al., J. Mol. Biol. **406**, 371–386 (2011).
6. J. Jiang et al., Virology **421**, 253–265 (2011).
7. U. K. von Schwedler, K. M. Stray, J. E. Garrus, W. I. Sundquist, J. Virol. **77**, 5439–5450 (2003).
8. B. K. Ganser-Pornillos, U. K. von Schwedler, K. M. Stray, C. Aiken, W. I. Sundquist, J. Virol. **78**, 2545–2552 (2004).
9. S. J. Rihn et al., PLOS Pathog. **9**, e1003461 (2013).
10. X. Meng et al., PLOS Pathog. **8**, e1002886 (2012).
11. T. R. Gamble et al., Science **278**, 849–853 (1997).
12. B. N. Kelly et al., J. Mol. Biol. **373**, 355–366 (2007).
13. C. Tang et al., J. Mol. Biol. **327**, 1013–1020 (2003).
14. W. S. Blair et al., PLOS Pathog. **6**, e1001220 (2010).
15. F. Curreli et al., Bioorg. Med. Chem. **19**, 77–90 (2011).
16. F. Ternois, J. Sticht, S. Duquerroy, H. G. Kräusslich, F. A. Rey, Nat. Struct. Mol. Biol. **12**, 678–682 (2005).
17. L. Lamorte et al., Antimicrob. Agents Chemother. **57**, 4622–4631 (2013).
18. A. J. Price et al., PLOS Pathog. **10**, e1004459 (2014).
19. A. Bhattacharya et al., Proc. Natl. Acad. Sci. U.S.A. **111**, 18625–18630 (2014).
20. T. Fricke, C. Buffone, S. Opp, J. Valle-Casuso, F. Diaz-Griffero, Retrovirology **11**, 120 (2014).
21. K. Peng et al., eLife **3**, e04114 (2014).
22. J. Shi, J. Zhou, V. B. Shah, C. Aiken, K. Whitby, J. Virol. **85**, 542–549 (2011).
23. K. A. Matreyek, S. S. Yücel, X. Li, A. Engelman, PLOS Pathog. **9**, e1003693 (2013).
24. B. K. Ganser-Pornillos, A. Cheng, M. Yeager, Cell **131**, 70–79 (2007).
25. D. K. Worthylake, H. Wang, S. Yoo, W. I. Sundquist, C. P. Hill, Acta Crystallogr. D **55**, 85–92 (1999).
26. V. Bartonova et al., J. Biol. Chem. **283**, 32024–32033 (2008).
27. B. N. Kelly et al., Biochemistry **45**, 11257–11266 (2006).
28. R. K. Gitti et al., Science **273**, 231–235 (1996).
29. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
30. B. Heras, J. L. Martin, Acta Crystallogr. D **61**, 1173–1180 (2005).
31. A. Joshi, K. Nagashima, E. O. Freed, J. Virol. **80**, 7939–7951 (2006).
32. S. Dimonte, M. Babakir-Mina, S. Aquaro, Virus Genes **48**, 429–437 (2014).
33. G. Obal et al., Science **349**, 95–98 (2015).
34. L. Hilditch, G. J. Towers, Curr. Opin. Virol. **4**, 32–36 (2014).
35. Z. Ambrose, C. Aiken, Virology **454-455**, 371–379 (2014).

Will you be meeting a Nobel Prize winner this December?

(If you have a recent PhD you could be.)

Stockholm in the second week of December is a special place. The city is alive with excitement as it welcomes and celebrates the new Nobel Laureates at the annual Nobel Prize ceremony.

If you are a PhD student, you could be here too – meeting a Nobel Laureate and receiving a rather special prize yourself.

The journal *Science* & SciLifeLab have established The *Science* & SciLifeLab Prize for Young Scientists, to recognize and reward excellence in PhD research and support young scientists at the start of their careers. It's about bright minds, bright ideas and bright futures.

Four winners will be selected for this international award. They will have their essays published in the journal *Science* and share a new total of 60,000 USD in prize money. The winners will be awarded in Stockholm, in December, and take part in a unique week of events including meeting leading scientists in their fields.

"The last couple of days have been exhilarating. It has been an experience of a lifetime. Stockholm is a wonderful city and the Award winning ceremony exceeds my wildest dreams." –Dr. Dan Dominissini, 2014 Prize Winner

Who knows, The *Science* & SciLifeLab Prize for Young Scientists could be a major stepping stone in your career and hopefully one day, during Nobel week, you could be visiting Stockholm in December once again.

This prize is made possible with the kind support of the Knut and Alice Wallenberg Foundation. This Foundation grants funding in two main areas; research projects of high scientific potential and individual support of excellent scientists.

The 2015 Prize is now open. The deadline for submissions is August 1, 2015.
Enter today: www.sciencemag.org/scilifelabprize

The 2015 Prize categories are:

• Cell and Molecular Biology

• Ecology and Environment

• Genomics and Proteomics

• Translational Medicine

Remote Incubator Monitor

Researchers can now be free to monitor and control their automated incubators from virtually any location where Wi-Fi Internet access is available. The Cymon remote monitoring application essentially replicates the Next Generation Display of the Cytomat 10 Automated Incubator on a Samsung 10.1 inch touchscreen tablet, which can be operated remotely. The Cymon application is available as an option for the Cytomat 10 incubator, which includes the application, a tablet, and a wireless router. The Cytomat 10 is a fully automated incubator and storage module for high-capacity cell growth and assay incubation. It is designed to access plates in less than 10 seconds and is equipped with a fully automated decontamination routine. It is the first incubator to feature the company's Next Generation Display, which uses icons to help users monitor and control process set points, temperature, humidity, CO_2, O_2, and logging of up to 1,000 data sets with selectable increments.

Thermo Fisher Scientific
For info: 800-556-2323
www.thermofisher.com/cytomat10

Liquid Handling Robot

The Microlab NIMBUS independent channel liquid handling robot is a small-scale liquid handler designed for space and budget-conscious labs. The NIMBUS system now adds a high-density deck to its repertoire of deck configurations. The new deck design allows up to 20 SBS deck locations all within the same compact footprint of the NIMBUS platform. The high-density deck increases throughput, enables longer walkaway times and utilizes flexible labware carriers for on-the-fly deck changes. The new high-density deck offers the most labware positions available compared to similar sized platforms in the industry of laboratory automation. The NIMBUS system also includes a 9 mm center-to-center independent channel system. The 9 mm spacing allows for random access across the entire deck surface. This spacing makes for easier programming, increased pipetting speed and simplified sample tracking. The 9 mm spacing is available with our four-channel pipetting arm.

Hamilton Robotics
For info: 508-648-5950
www.hamiltoncompany.com

Digital Dispenser

The popular HP D300 Digital Dispenser now allows customers to titrate low volumes of biomolecules in aqueous solutions, offering greater flexibility for drug discovery workflows. Users can choose between dispensing DMSO, for small molecule studies, and aqueous solutions in combination with a surfactant, for the investigation of proteins, antibodies, enzymes, and nucleic acids. Newly launched T8+ and D4+ Dispenseheads, combined with updated software, allow both new and existing customers to benefit from this breakthrough, offering straightforward setup of dose-response curves and synergy studies. New "fluid class" options allow users to optimize liquid dispensing performance to their specific buffers and surfactants, and an Enzyme Profile function enables fast and easy setup of titrations using differing combinations of enzymes, substrates, and other compounds. These innovative features, combined with the system's proven low volume dispensing performance, make it easy to perform complex experiments, such as enzyme characterization, Km determination, and exploring the effects of enzyme inhibitors.

Tecan
For info: +41-(0)-44-922-81-11
www.tecan.com/digitaltitration

Isothermal Titration Calorimetry System

The Affinity ITC Auto is a fully automated, isothermal titration calorimeter for characterizing a wide variety of molecular interactions, including protein-protein binding, drug-ligand binding, small molecule-receptor interaction, and enzyme kinetics. This automated ITC is available with fixed in place cells of either the standard volume (1.0 mL) or low volume (190 μL). For laboratories requiring higher throughput of ITC measurements, this new instrument improves sample throughput with an industry-proven 96-well plate liquid handling system and ensures the utmost reliability in an automated system through all-new innovations to provide the most accurate, reproducible, high-quality ITC data. Intelligent Hardware Positioning is an integral part of the Affinity ITC Auto that adds accuracy and dependability to the automation hardware. The Affinity ITC Auto features the new FlexSpin stirring and AccuShot injection systems and provides the next step in ITC innovations while maintaining instrument flexibility and the highest sensitivity available.

TA Instruments
For info: 302-427-4000
www.tainstruments.com
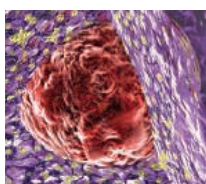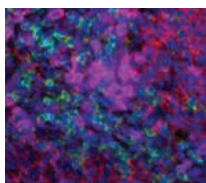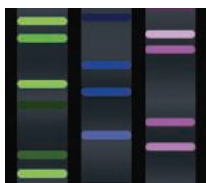
Pipette Tip Racks

The new Rainin TerraRack line of pipette tip racks is an innovative solution to minimizing the problem of plastic tip rack waste. TerraRacks are 50% lighter than conventional tip racks, highly compressible when empty and completely recyclable. TerraRacks are composed largely of PETE (polyethylene terephthalate), exceptionally strong polyester that's easy to recycle and in high demand as a post-waste commodity. The TerraRack shell, with its hinged lid and front latch, is molded as a single PETE unit, making it exceptionally lightweight. A polypropylene tip deck snaps into the shell, adding structural support with minimal weight. TerraRacks have less than half the weight and plastic of conventional racks and, after use, nest inside one another, significantly reducing their volume and disposal costs. Because every rack is presterilized, individually wrapped, and sealed for maximum purity, TerraRacks do not need to be autoclaved.

Mettler Toledo
For info: 800-638-8537
www.mt.com/rainin

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/products/newproducts.dtl for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

West Virginia University

The Department of Microbiology, Immunology and Cell Biology
West Virginia University School of Medicine
Two Open-Rank Faculty Positions in Infectious Disease Research

The WVU Health Sciences Center invites applications for two tenure-track faculty positions (rank open) in the Department of Microbiology, Immunology and Cell Biology: http://medicine.hsc.wvu.edu/micro/. We seek individuals with a PhD, MD or MD/PhD degree who study the molecular mechanisms of infectious diseases toward their prevention and treatment. Dedicated laboratory space and competitive start-up funds will be made available. Applicants should have a strong record of scholarly publications, productive post-doctoral training, and excellent written and oral communication skills. Candidates for an Assistant Professor will be expected to obtain independent extramural funding, while appointments for Associate Professor and Professor will require the transfer of existing funding. The successful faculty will contribute to the undergraduate, graduate and professional teaching of the department.

The WVU Health Sciences Center has a strong commitment to improving the health of people in West Virginia and will benefit all people worldwide. Infections in pediatric patients are of significant interest and successful qualified appointee may hold a secondary appointment in the Department of Pediatrics. This position provides an exceptional opportunity to participate in a robust, interdisciplinary basic and translational research and training programs in a highly collaborative atmosphere. WVU Children's Hospital serves as the premiere site of pediatric care within West Virginia and the surrounding region (http://www.wvukids.com/wvuh/). The WVU HSC is also opening a newly-constructed A/BSL3 facility in the spring of 2016 for research on agents requiring a high-containment laboratory. For more information about WVU HSC and the city of Morgantown, WV: http://medicine.hsc.wvu.edu/micro/working-for-wvu/.

Qualified applicants should submit a single PDF file including: (1) cover letter; (2) curriculum vitae; (3) statement of current and future research goals (4); teaching and mentoring philosophy; (5) a list of three representative publications; and (6) contact information for three or more references to: **Slawomir Lukomski, Ph.D., Chair, Search Committee, c/o Barbara Pritt (bpritt@hsc.wvu.edu), Department of Microbiology, Immunology and Cell Biology, West Virginia University School of Medicine, Morgantown, WV 26506-9177.**

Review of applications will commence in July 2015

West Virginia University is an Affirmative Action/Equal Opportunity Employer. The Health Sciences Center is a tobacco-free campus. West Virginia University is the recipient of an NSF ADVANCE award for gender equity. The university welcomes applications from all qualified individuals, including minorities, females, individuals with disabilities, and veterans.

TENURE-TRACK POSITION in Animal Physiology
W.M. Keck Science Department of Claremont McKenna,
Pitzer, and Scripps Colleges

The W.M. Keck Science Department of Claremont McKenna College, Pitzer College, and Scripps College invites applications for a tenure-track appointment in Animal Physiology at the **Assistant Professor** level to begin July 2016. The department, which houses the biology, chemistry, physics, and environmental analysis faculty for three of the five undergraduate Claremont Colleges, offers innovative and interdisciplinary programs in the natural sciences.

We seek a broadly trained animal physiologist who is committed to excellence in teaching and to working within a cross-disciplinary and integrative department, and who will develop a vibrant research program that fully engages undergraduate students. We are especially interested in applicants whose work involves integrative, comparative/evolutionary, or computational approaches. The position offers opportunities to teach an animal or human physiology course with laboratory, introductory biology, biology courses for non-majors, and to design new courses in the candidate's field. Candidates should hold a Ph.D. degree with focus in physiology or a related field. Postdoctoral experience and a record of scholarly publication are preferred.

Please apply online at website: **https://webapps.cmc.edu/jobs/faculty/faculty_opening_detail.php?PostingID=14027**

Uploaded materials should include a cover letter, curriculum vitae, a statement of teaching philosophy, a description of proposed research with equipment needs, and a diversity statement (of no more than 1 page) outlining the applicant's philosophy for fostering an educational environment that is inclusive of all students. Please also ensure that three reference letters be uploaded. Inquiries about the position should be addressed to **Dr. Donald McFarlane** at dmcfarlane@kecksci.claremont.edu. Additional information about the department can be found at **http://www.kecksci.claremont.edu**. Review of applications will begin **September 21st 2015**, and the position will remain open until filled.

The Keck Science Department of Claremont McKenna, Pitzer, and Scripps Colleges is an Equal Opportunity Employer. In a continuing effort to enrich its academic environment and provide equal educational and employment opportunities, the department actively encourages applications from women and members of other historically under-represented social groups in higher education.

Appointment of Director/Chief Executive

TGAC now seeks to appoint a new Director / Chief Executive who will lead the research Institute in its further development as an internationally-recognised centre of excellence. Reporting to the Board of Trustees, the Director / Chief Executive will combine strategic vision with managerial flair, while also delivering technical expertise and driving innovation. TGAC's new Director will be supported by a well-qualified executive and senior management team with access to a Scientific Advisory Board.

Applications are invited from internationally-renowned scientists with a track record of research success in fields relevant to genomics and computational bioscience. Alongside proven research acumen, applicants must demonstrate significant management skills and inspirational leadership with a strong commitment to knowledge transfer and innovation.

The post-holder will be well connected internationally with demonstrable credibility with senior researchers, research sponsors, government and industrialists.

Applications should consist of a full CV detailing professional qualifications, full employment history, latest remuneration and relevant achievements and should be accompanied by a covering letter describing briefly how candidates meet the criteria in the Person Specification, why the appointment is of interest and what they believe they can bring to the role. These can be uploaded at www. perrettlaver.com/candidates quoting reference 2059. The closing date for applications is midday (BST) on Thursday 30th July 2015.

The full field of candidates will be reviewed in early August and longlisted candidates will be invited for interview with Perrett Laver in mid-August; following this, TGAC will agree a shortlist. Shortlisted candidates will be invited for informal visits to TGAC on the w/c 14th September followed by more formal interviews in early October. Selection of and negotiations with the preferred candidate(s) will take place thereafter.

Perrett Laver

AMERICAS · ASIA PACIFIC · EMEA

Ecologist

The Department of Biology at the University of Pennsylvania invites applications for a tenure-track faculty position in Ecology. The appointment will be at the level of Assistant Professor. We are searching broadly for a community or ecosystem ecologist working at any spatial scale, from local to global, and using theoretical and/or empirical approaches. Areas of interest include, but are not limited to, mechanisms of community assembly, coexistence, diversity, adaptation and the interplay between biotic and abiotic processes in the face of climate and land-use change. We welcome applicants working on any domain of life, from micro- to macroscopic.

We are especially interested in candidates who utilize novel approaches that span levels of biological organization or transcend traditional boundaries.

Penn's Department of Biology has a long-standing tradition of maintaining an integrated research and educational program across all basic biological sciences, from Ecology and Evolution, Plant Sciences, Molecular and Cellular Biology, Genomics, to Neuroscience. The Department values interdisciplinary research, collaboration, and collegiality, and has a vision emphasizing *Life in its Natural Context*.

Candidates are expected to have demonstrated excellence and productivity in research and will be expected to excel in undergraduate and graduate teaching. Interested candidates should submit materials online at **http://facultysearches. provost.upenn.edu/postings/606** and include a curriculum vitae, concise statements of research and teaching interests, a short annotated description of up to five publications, and the name and contact information of at least three referees. Recommenders will be contacted by the University with instructions on how to submit a letter to the website. Review of applicants will begin **September 11, 2015** and continue until the position is filled.

The Department of Biology is strongly committed to Penn's Action Plan for Faculty Diversity and Excellence and to creating a more diverse faculty (for more information see: http://www.upenn.edu/almanac/volumes/v58/n02/ diversityplan.html). The University of Pennsylvania is an Equal Opportunity Employer. Minorities, women, individuals with disabilities, and protected veterans are encouraged to apply.
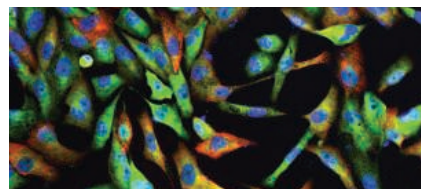
Director of the Life Sciences Institute
UNIVERSITY OF MICHIGAN

The University of Michigan (U-M) seeks a distinguished biomedical scientist and academic leader as its next Mary Sue Coleman Director of the Life Sciences Institute (LSI). The LSI Director provides leadership and strategic direction to the Institute and is a critical leader for biosciences on campus.

The LSI (founded in 1999) is a collaborative, free-standing biomedical research institute located on U-M's Ann Arbor campus, and a cornerstone of the U-M's ~$0.7 billion research program in the life sciences. The LSI is closely linked with U-M's life and biomedical science departments and professional schools, and each LSI laboratory head also holds a tenure-track position in an academic department.

Please visit **http://myumi.ch/lsi-director** for further details including nomination and application instructions. Review will begin immediately and serious applicants are encouraged to respond by August 15, 2015.

Additional information about the University of Michigan and Life Sciences Institute is at **http://www.umich.edu** and **http:// www.lsi.umich.edu**.

The University of Michigan is an equal opportunity/affirmative action employer.

AAAS SCIENCE & TECHNOLOGY POLICY FELLOWSHIPS

"The awareness I gained to the political landscape has been life changing."
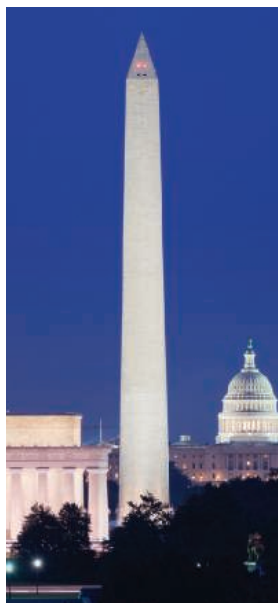
I HAD A MONUMENTAL EXPERIENCE.
YOU CAN TOO.

— Stephanie Adams, PhD
Interdisciplinary Engineering,
AAAS Fellow; Professor and
Department Head, Virginia
Tech Department of
Engineering Education

MAKE A DIFFERENCE.
TRANSFORM YOUR CAREER.

Apply your scientific analysis and technical knowledge to inform policy through assignments in the Legislative, Executive and Judicial Branches.

Stipends from $75,000 to $100,000. Applications due **November 1.**

Details at: **aaas.org/stpf/science2**

Enhancing Policy, Transforming Careers

SCIENCE & TECHNOLOGY **POLICY FELLOWSHIPS**
▲▲AAAS

Find your most interesting question

I t started with a casual remark over coffee. It was the late 1990s, and I was a postdoc in Karin Rabe's ferroelectrics research group, then at Yale University. Ferroelectrics are materials that exhibit a spontaneous electric polarization that can be reversed by applying an electric field. Although they were named after an analogous magnetic phenomenon—ferromagnetism—ferroelectrics were generally thought to not have magnetic properties, which made me, a budding magnetism researcher, an apparently odd fit for the laboratory. That odd juxtaposition, though, turned out to be crucial to my career. "It's a pity there are no magnetic ferroelectrics," my labmate said to me as we guzzled espresso, "because then we could collaborate."

"Woah! Is that really true?" I thought. "Are there really no magnetic ferroelectrics? And if so, why?" It was an intriguing fundamental question, and such materials could be technologically useful if they could be made. It soon became, for me, the most interesting question in the world. So when I took my assistant professorship at the University of California (UC), Santa Barbara, a few months later, I abandoned my plan to work in the well-established field of semiconductor nanostructures and embarked on a single-minded hunt for magnetic ferroelectrics.

My quest for these multiferroics was quixotic, but my timing was fortuitous. First-principles electronic structure theory had matured to the point where materials I thought might be good candidates could be studied. My postdoctoral work had involved extending electronic structure methods to study magnetic systems—exactly what I needed to predict the properties of new multiferroics.

Armed with these tools, I set out to create computer models of magnetic ferroelectrics. I calculated, thought, and calculated some more until, in 2000, I was ready to publish at least a partial answer to the question, "Why are there so few magnetic ferroelectrics?"

Around that time, though, I discovered a downside to having so much scientific fun: Solving difficult problems takes time, and my publication record wasn't strong enough, according to my pretenure review. Also, when you work in a field no one else works in, you don't garner a lot of citations. Kind mentors advised me to start doing something more mainstream—quickly.

Instead I kept going, hoping to see my predictions tested in the lab. I embarked on a pretenure lecture tour, proselytizing about the emerging field and seeking collabora-

Before I knew it, multiferroics were no longer "mine."

tors. My enthusiasm rubbed off, and some excellent experimentalists began making structures I had designed. I gathered the courage to organize the first-ever multiferroics session at the 2001 March meeting of the American Physical Society, which gave the field a further boost. A year later, I was tenured. Phew!

In 2003, in collaboration with the group of Ramamoorthy Ramesh (now at UC Berkeley), we succeeded in producing and understanding thin films of what is now one of the most-studied multiferroic materials, bismuth ferrite. This led to a *Science* paper and an important lesson: A high-profile publication can really help stimulate a field. Soon after, many research groups and industrial labs began working on multiferroics. Before I knew it, multiferroics were no longer "mine." They had a life of their own, with thousands of active researchers around the world.

Frighteningly, I have reached the stage in my career when young people often ask me for advice. My safe and sensible side tells me to pass along the same advice I received: Make a solid contribution to an established field and publish a lot to become known and respected by your community. Save the high-risk stuff until after tenure. But, deep down, I hope young scientists—you—will choose not to follow that advice. I hope instead you will find the question that for you is the most interesting in the world, go after its answer with all your youthful passion, and pioneer your own science revolution. ■

Nicola Spaldin is a professor at ETH-Zurich in Switzerland and winner of the 2015 Körber European Science Prize for "la[ying] the theoretical foundation for the new family of multiferroic materials." For more on life and careers, visit sciencecareers.org. Send your story to SciCareerEditor@aaas.org.