Genes behind embryonic aneuploidy *pp. 180 & 235*

Closing the Central American Seaway early *pp. 186 & 226*

Chemical imaging of membrane lipids *p. 211*

# Science

$10
10 APRIL 2015
sciencemag.org

AAAS

## *CubeSats take flight*

Cheap, miniature satellites democratize space *p. 172*

# CONTENTS

183, 199, & 200
Rules of protein degradation

10 APRIL 2015 • VOLUME 348 • ISSUE 6231

## NEWS

### IN BRIEF



### IN DEPTH






### FEATURES

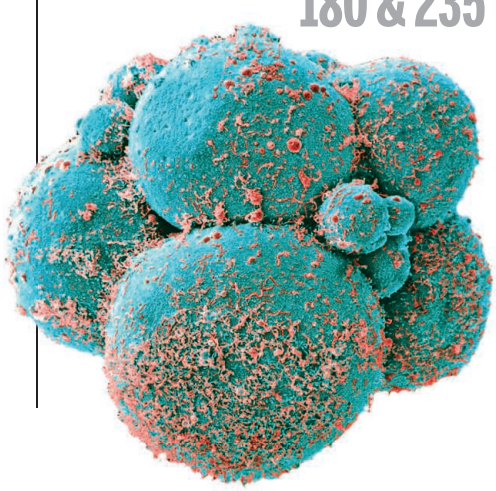


## INSIGHTS

### PERSPECTIVES



180 & 235






### BOOKS *ET AL.*

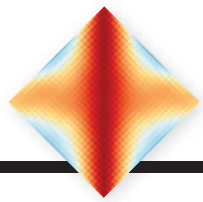




Published by AAAS

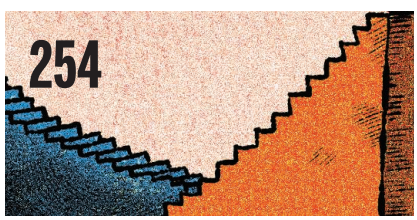
# CONTENTS

10 APRIL 2015 • VOLUME 348 • ISSUE 6231



185 & 207
Where quantum mechanics meets thermodynamics

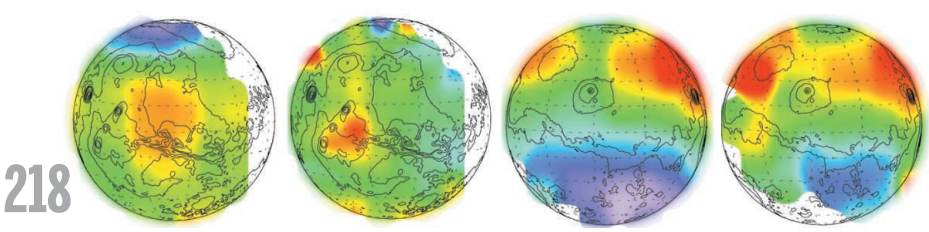







## ON THE COVER


On 27 February 2015, Planet Labs released two CubeSats from the International Space Station, adding to its swarm of Earth-observing satellites, which are no larger than a loaf of bread. Each CubeSat can discern features as small as trees and trucks. But the real power comes from having dozens, even hundreds, in space; the company plans to image the entire Earth every day. See page 172. *Photo: NASA*




**SCIENCE** (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices.  Domestic individual membership and subscription (51 issues): $153 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $1282; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #1254 88122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
**Change of address:** Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request.  *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.



*Published by AAAS*

Science
AAAS
1200 New York Avenue, NW, Washington, DC 20005
Bateman House, 82-88 Hills Road, Cambridge, UK CB2 1LQ



**SUBSCRIPTION SERVICES** For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

**INSTITUTIONAL SITE LICENSES** 202-326-6755 **REPRINTS:** Author Inquiries 800-635-7181 **COMMERCIAL INQUIRIES** 803-359-4578 **PERMISSIONS** 202-326-6765, permissions@aaas.org **AAAS Member Services** 202-326-6417 or http://membercentral.aaas.org/discounts

*Science* serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority of conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and books reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

**INFORMATION FOR AUTHORS** See pages 678 and 679 of the 6 February 2015 issue or access www.sciencemag.org/about/authors





*Published by AAAS*

EDITORIAL

# NIH research: Think globally

The U.S. National Institutes of Health (NIH) has for more than 60 years supported research to improve the health and prolong the lives of people in the United States and around the world. Mean life expectancy worldwide has doubled to more than 70 years, due in large part to medical and public health interventions developed with NIH funding. Now, in the face of serious fiscal constraints, the idea has reemerged from some congressional leaders and disease constituency groups to more closely align NIH funding for disease research with disease burden in the United States. Although the nation must maintain robust research support for diseases that cause illness and death among large numbers of Americans, it would be unwise to deemphasize diseases that exact their largest toll elsewhere in the world. The United States has a vital interest in the health of people around the globe, rooted in an enduring tradition of humanitarian concern as well as in enlightened self-interest. Engagement in global health protects the nation's citizens, enhances the economy, and advances U.S. interests abroad.

***"Engagement in global health protects the nation's citizens, enhances the economy, and advances U.S. interests abroad."***

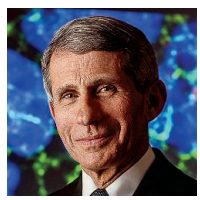
People from all walks of life understand and appreciate the moral imperative to alleviate human suffering, regardless of where it occurs. Polls show that Americans support efforts to improve health in developing countries, both for the sake of those individuals and for the sake of Americans exposed to infectious diseases that transcend national boundaries. The recent outbreak of Ebola virus disease in West Africa, which quickly found its way to the United States, is one more reminder of how global health challenges can become domestic. The concept of medical diplomacy—winning the hearts and minds of people in poor countries by exporting medical interventions, expertise, and personnel to improve their health—also resonates with many Americans, as does reducing instability in places where the United States has substantial economic and political interests.

The U.S. government, the largest funder of global health research and development, has played a central role in developing medical interventions that have saved countless lives in the world's poorest countries. Smallpox has been eradicated, polio nearly eliminated, and important infectious diseases of childhood controlled with vaccines. An extraordinary 7.6 million AIDS deaths were averted in low- and middle-income countries between 2003 and 2013 by the development and distribution of antiretroviral drugs to treat HIV infection. Future products, including improved drugs for tuberculosis, treatments for parasitic diseases, vaccines for malaria, and new strategies to prevent and treat HIV infection could save millions more lives. Also, studying such complex diseases provides new insights that can advance how we diagnose, treat, and prevent other health challenges, including many commonly seen in the United States. For example, the treatment of hepatitis B virus infection has been revolutionized by antiviral drugs originally developed to treat HIV infection.

History shows that the tools of modern biology offer the opportunity to practically elimate major diseases that sap human health and exacerbate instability in areas where the United States has substantial interests. It is imperative that the nation sustain momentum and work with its global partners to deliver the fruits of global research to the people who need them most, both at home and abroad. Without such a commitment, we may miss opportunities to curtail or even eliminate important diseases such as AIDS and also risk the resurgence of major global health threats such as drug-resistant bacteria, tuberculosis, and malaria, for which new interventions are badly needed.

In 1940, President Franklin D. Roosevelt noted that "NIH speaks the universal language of humanitarianism… [it] has recognized no limitations imposed by international boundaries and has recognized no distinctions of race, of creed, or of color." The NIH—and the United States—must continue to live by these words.

– **Anthony S. Fauci and Francis S. Collins**



*Anthony S. Fauci is director of the U.S. National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD. E-mail: afauci@niad.nih.gov*

*Francis S. Collins is director of the U.S. National Institutes of Health, Bethesda, MD. E-mail: collinsf@mail.nih.gov*

10.1126/science.aab2733

# NEWS

> “We just figured, we’ll go with [a name] that has some meaning to us.”
>
> **Biochemist Tom Schmidlin**, founder and namesake of the recently launched Postdoc Brewing Company. **http://scim.ag/postdocbrew**

## IN BRIEF

### Sheep, goat virus targeted for eradication

A viral disease is threatening goat herds like this one in Tanzania.

Animal health specialists meeting in Abidjan, Ivory Coast, last week agreed to try to rid the world of peste des petits ruminants (PPR), a viral disease devastating goat and sheep flocks throughout Africa, the Middle East, and Asia. Control efforts have fallen short, and the time has come for a “bolder next step,” said José Graziano da Silva, director-general of the Food and Agriculture Organization (FAO) of the United Nations, at the meeting FAO organized with the World Organisation for Animal Health (OIE) to kick off a global eradication program. Also called ovine rinderpest, PPR kills up to 90% of the animals it infects within days. The virus has spread rapidly over the past 15 years and is now in 70 countries, putting 80% of the world’s more than 2 billion goat and sheep at risk. FAO and OIE believe they will need $4 billion to $7 billion over the next 15 years to accomplish their goal. **http://scim.ag/PPRerad**

**AROUND THE WORLD**

### Biomed lab inks hospital deal

LONG ISLAND, NEW YORK | Cold Spring Harbor Laboratory (CSHL), a bastion of basic biomedical research, is making a major foray into more applied drug development. Last week the lab and North Shore-LIJ Health System, a local hospital system, announced a new alliance and a more than $120 million investment aimed at moving basic cancer discoveries into the clinic. Scientists at the not-for profit CSHL have used genomics, RNAi screens, and mouse models to find important cancer drug targets. Now, the lab wants to turn those discoveries into drugs by conducting early stage clinical trials with North Shore-LIJ, which has 16,000 new cancer patients each year in the New York City area. CSHL chief Bruce Stillman says the new venture will not compromise the lab’s basic research. **http://scim.ag/coldsprharb**

### Banana disease spreads

SINDH PROVINCE, PAKISTAN | A fungus that devastates banana plants has been confirmed in Pakistan, raising worries that it might spread to India, the world’s largest producer of bananas. The fungus, a strain of *Fusarium oxysporum* called tropical race 4 (TR4), emerged in the 1990s and has ravaged thousands of hectares in China, Indonesia, Malaysia, and the Philippines. Growers in Pakistan first noted suspicious symptoms, such as brown rot in stems (below), on a small plantation in 2012.

PHOTOS: (LEFT TO RIGHT) © CHRIS MATTISON/ALAMY; © DAVID HANCOCK/ALAMY





Published by AAAS

By January 2014, about 121 hectares had succumbed. A team led by Gerrit Kema of Wageningen University and Research Centre in the Netherlands analyzed samples of the bananas from Pakistan and Lebanon—and confirmed in *Plant Disease* last week that TR4 has now reached both countries. TR4 is soil-borne, so it could spread to India by flooding or human transport. "We are afraid it will show up everywhere," Kema says. TR4 hit Mozambique in 2013 and was confirmed last month in Queensland, Australia.

## Open Access 3.0: Pay the authors

DODEWAARD, NETHERLANDS | Most open-access journals have authors pay an article processing fee to publish a paper. Now, a small online malaria journal based in the Netherlands is promising to pay its authors—€150 for every article it publishes. The idea behind the move, possible thanks to a Dutch funding agency, is both to lure authors to the journal and to drive home the message that academic publishing is too expensive, says the journal's editor, Bart Knols. The upstart journal—which has so far published only 57 papers—is part of MalariaWorld, a website and networking tool that has some 8500 registered users in 140 countries. Two experts review papers submitted to the journal, Knols says; if they disagree, the journal's editors decide whether to publish. The plan is to reward every published paper; multiple authors split the €150. http://scim.ag/MalWorld

NEWSMAKERS

## Three Q's

In 2012, a Harvard University freezer containing more than 50 donated brains from people diagnosed with autism failed. Autism researchers were left with less than 10% of the brain tissue experts say they need to untangle the disorder's biological basis. Last month, the National Institutes of Health and private research foundations launched a collaboration to collect new donations. **David Amaral**, research director of the University of California, Davis, MIND Institute, discussed the project, called the Autism Brain Net.

**Q: Why isn't there enough brain tissue to study autism?**
**A:** There are probably 500 deaths a year of people with autism, and yet the number of brains donated is a small fraction of that. It's vital to get samples from all ages. You also have to make sure [the brain] is from someone who really did have autism.

**Q: How many brains do you have?**
**A:** We did an inventory in January of last year, and counted roughly 20 [intact] brains.

**Q: What questions can the brain tissue answer?**
**A:** The brains of young individuals with autism are enlarged, but we don't know why. Is it too many neurons, too many connections? [To answer that] we have to do the neuropathology.

ILLUSTRATION: LUIS REY

## Was *T. Rex* cousin a cannibal?

The group of ferocious meat-eating dinosaurs known as tyrannosaurs—of which the most famous member is *Tyrannosaurus rex*—may have turned their sharp teeth on each other. A distinctive pattern of tooth marks on the skull and jaw of a 75-million-year-old tyrannosaur discovered at Dinosaur Provincial Park in Alberta, Canada, offers the best evidence of that yet, scientists say. The dinosaur, known as *Daspletosaurus* (pictured above in an artist's reconstruction), was not fully adult when it died, but still weighed about 500 kilograms and was almost 6 meters long. Full-grown *Daspletosaurus* were just a tad smaller than *T. rex*, which lived 7 or so million years later. In a study published this week in *PeerJ*, the researchers conclude from the spacing and shape of the puncture marks on its skull that it was bitten by another tyrannosaur—quite possibly another *Daspletosaurus*—while it was still alive, probably as the result of fraternal infighting. Those wounds healed, but tooth marks on its jaw suggest that it was chomped on by yet another tyrannosaur after it died.

FINDINGS

## How Europeans evolved white skin

Pale skin, and other traits such as tallness and the ability to digest milk as adults, arrived in most of Europe relatively recently, scientists reported last week at the 84th annual meeting of the American Association of Physical Anthropologists in St. Louis, Missouri. The international team of researchers searched DNA from 83 ancient individuals from archaeological

Protesters gathered at Mauna Kea last week to prevent construction of the Thirty Meter Telescope.

## Telescope protesters arrested

Ongoing protests against the construction of the Thirty Meter Telescope (TMT) on the summit of Mauna Kea led to the arrest of a dozen people last week, according to the Hawaii Police Department. About 100 people gathered on 2 April to block construction machinery from reaching the summit, which Native Hawaiians say is sacred ground. The protesters have cited concerns that construction of the $1.4 billion telescope will irreparably harm the environment and Native Hawaiian ancestral graves and sacred places. Protesters also disrupted the TMT's groundbreaking ceremony last October (*Science*, 17 October 2014, p. 284), but last week marked the first arrests. The protesters were released after posting bail. TMT, which will be one of the most powerful ground-based telescopes in the world, is scheduled to achieve first light in 2022.

### BY THE NUMBERS

**1000**

The amount of plastic, in tonnes, in an accumulation zone recently identified within the Mediterranean Sea.

**10%**

Fraction of samples, advertised online as pure human breast milk, that had been adulterated with cow's milk, according to a study in *Pediatrics*.

**$1 billion**

Damages sought in a lawsuit filed last week by 774 plaintiffs against Johns Hopkins University and the Rockefeller Foundation, alleging they deliberately infected hundreds in Guatemala with sexually transmitted diseases between 1945 and 1956.

CREDITS: (PHOTO) AP PHOTO/*HAWAII TRIBUNE-HERALD*/HOLLYN JOHNSON; (IMAGE) JORGE GONZALES

sites throughout Europe for genes under strong natural selection—including traits so favorable that they spread rapidly throughout Europe in the past 8000 years. By comparing the ancient European genomes with those of recent ones from the 1000 Genomes Project, they found five genes associated with changes in diet and skin pigmentation that underwent strong natural selection. The finding offers dramatic evidence of recent evolution in Europe and shows that most modern Europeans don't look much like those of 8000 years ago. **http://scim.ag/Euroskin**

## How the Dutch became tallest

The Dutch are the tallest people on Earth, with an average height of 1.84 meters for men and 1.71 meters for women. Better nutrition and health care help explain a growth spurt of 20 centimeters over 150 years, but natural selection may also have played a role, suggests a study in the *Proceedings of the Royal Society B*. From a database tracking 94,516 Dutch people, behavioral biologist Gert Stulp of the London School of Hygiene & Tropical Medicine and colleagues found that taller Dutch men have more children than those of average height and more of them survive, which likely helps spread genes contributing to tallness. Previous studies showed that tall men in the United States don't have more children; why the Dutch are different is unclear, Stulp says. **http://scim.ag/dutchtall**

## Terror of the late Triassic

North Carolina was a tropical swamp 231 million years ago—and at the top of the food chain was a 3-meter-tall crocodilian ancestor that walked on its hind legs and ate the ancestors of early mammals. A newly analyzed fossil, which includes parts of a skull, spine, and upper forelimb found more than a decade ago in central North Carolina, represents one of the earliest examples of crocodylomorphs, paleontologists reported last month in *Scientific Reports*. The team used high-resolution surface scanners and skulls of close relatives to create a three-dimensional model of *Carnufex carolinensis*, known as the "Carolina butcher." Crocodylomorphs vied with theropod dinosaurs at the end of the Triassic for top predator slots and succeeded—for a time. Only their smaller cousins remained following the end-Triassic extinction event 201.3 million years ago, allowing dinosaurs to take over for the next 135 million years. **http://scim.ag/Carnufex**

PHOTO: REUTERS/CAREN FIROUZ

Deal would allow Iran to continue operating thousands of uranium-enrichment centrifuges, seen with an unnamed official in 2009.

NUCLEAR WEAPONS

# *Iran deal would transform its nuclear infrastructure*

## Secretive Fordow facility could become a stable isotope factory and international accelerator center

*By* **Richard Stone**

In about 2007, Western intelligence analysts believe, Iran began secretly converting a tunnel complex in the Fordow Mountains, near the holy city of Qom, into a facility capable of producing the highly enriched uranium needed to make nuclear weapons. After Iran acknowledged the facility in 2009, some foreign policy hawks started debating whether bunker-busting bombs could destroy the site, a former military base.

Instead, foreign researchers may be waltzing through Fordow's front door before long. According to the outline of a deal reached last week with the United States and five other nations, Iran will convert the sensitive facility into an international research center that would produce isotopes for industry—and may even host one or more small linear accelerators for basic science.

Fordow's emergence from the shadows is one highlight of the tentative agreement, which is designed to dismantle large pieces of Iran's nuclear program in exchange for easing crippling economic sanctions. Other technical elements include reengineering a plutonium-producing reactor and removing or diluting enriched uranium that could otherwise be further enriched to make several atomic bombs. The plan seeks a delicate balance: preventing Iran from building an atomic arsenal while allowing it to retain nuclear R&D. "Iran is still going to have a nuclear program and can do a lot with it," says physicist James Acton, a nonproliferation expert at the Carnegie Endowment for International Peace in Washington, D.C.

A final agreement isn't due until the end of June. U.S. officials do not see any technical showstoppers—but there are plenty of potential political roadblocks. Some U.S. lawmakers and Israel's leaders are skeptical of the emerging deal, which lacks key details. "Probably the most significant issue will be to converge on the exact sequence and triggers" for phasing out sanctions, U.S. Energy Secretary Ernest Moniz told *Science*.

In recent months, Moniz, a physicist, became the lead technical negotiator for the United States, working alongside experts from the four other permanent members of the U.N. Security Council—China, France, Russia, and the United Kingdom—as well as Germany (a group known as the P5+1). Their goal: to slow Iran's "breakout time"—the time needed to produce enough weapons-grade fissile material for one bomb—from an estimated 2 to 3 months to at least a year.

Achieving that goal boils down to blocking what Moniz calls Iran's four pathways to nuclear weapons: two uranium paths through its Natanz and Fordow enrichment facilities, where thousands of centrifuges separate uranium isotopes; a plutonium path that involves the Arak heavy water reactor; and a covert path involving hidden facilities. The challenge has been to block these pathways without shuttering a single nuclear facility—something that Iran insisted was out of the question. Building that nuclear infrastructure "has cost Iran dearly, but it has also become a source of national pride," says Muhammad Sahimi, an Iran-born chemical engineer at the University of Southern California in Los Angeles and an analyst of Iran's nuclear program. By preserving it, Sahimi says, Iran's supreme leader, Ali Khamenei, "can claim that the redline he set has not been crossed."

The Arak 40-megawatt fission reactor, under construction in the central province of Markazi, has been a bone of contention. Iranian officials say it is meant to make radioisotopes for medicine. But simply running the reactor on its natural uranium fuel would yield about 7.7 kilograms of plutonium each year, enough for at least one atomic bomb. To slash that output, nonproliferation experts had suggested changing the fuel to low-enriched uranium (LEU), which would reduce the plutonium yield fivefold. They also urged halving the reactor's power output, cutting plutonium pro-

duction to less than 1 kilogram a year.

To the surprise of many, Iranian officials have—in principle—agreed to even more stringent changes. Not only will Arak switch to LEU fuel, but engineers will also downsize the reactor's calandria—the vessel in which the fuel resides—because LEU fuel requires less space. That downsizing also would make it harder to restore natural uranium fuel to the reactor and produce more plutonium. The changes would lengthen the plutonium breakout time to several years, Moniz estimates. And even that could be a moot point: Iran has pledged to ship all of Arak's spent fuel out of the country for the reactor's lifetime, Moniz says.

An Arak redesign will be worked out in the coming weeks. "They want and we want the P5+1 to certify the redesign," Moniz says. Besides minimizing plutonium production, the new design would increase the reactor's neutron flux—which would make it more efficient at generating medical radioisotopes and "would be good for their research program," Moniz says. Siegfried Hecker, a plutonium specialist at Stanford University in Palo Alto, California, and former director of Los Alamos National Laboratory, has called for developing Arak as an international user facility for medicine and materials science. Although that idea is not in the draft agreement, Moniz says "it's still on the table."

*The notion of making the Arak reactor an international facility "is still on the table."*

**Ernest Moniz**, U.S. energy secretary

Fordow, however, would become what Moniz terms an "international physics center." The facility would be largely gutted, with about two-thirds of its centrifuges removed. The remaining 1044 centrifuges, Moniz says, would be modified to produce molybdenum (Mo) and iridium (Ir) isotopes for use in medicine and engineering. In molybdenum's case, the goal would be to increase the proportion of the isotope Mo-98 in samples, then expose them to neutrons at Arak to make Mo-99. Mo-99 is the parent of technetium-99m, a radioisotope used in more than 20 million nuclear medicine procedures a year. The other enrichment target is Ir-191; adding a neutron makes Ir-192, which is widely used in gamma cameras to check for structural flaws in metal. As a nonproliferation dividend, the modified centrifuges would become so contaminated with isotopes that "those machines would be very much unusable" for making weapons-grade uranium, says a U.S. official.

Each of the P5+1 nations is seeking to carve out other research niches at Fordow. French negotiators have even floated the idea of installing up to three low-energy electron linear accelerators. "The concept is an accelerator center of excellence," says the U.S. official. Opening up Fordow "is a great idea to get [Iran] into the international research community," Hecker adds.

Several other facets of the deal are also aimed at blocking the uranium pathways to a bomb. For starters, Iran would reduce the number of installed centrifuges for enriching uranium from about 19,000 to 5060—all of which, for at least 10 years, would be Iran's less capable, first-generation IR-1 centrifuges. Iran would also soon end research projects involving its more advanced IR-2m and IR-4 centrifuges, which "will be removed and put under seal," Moniz says. And the country's LEU stockpile would be slashed from about 10,000 kilograms to 300 kilograms.

One option for achieving that reduction is to blend the LEU with depleted uranium, which, compared with natural uranium, contains less of the fissile isotope uranium-235. But Frank von Hippel, a physicist and arms control expert at Princeton University in New Jersey, says that approach would "throw away all the enrichment work"—a waste of valuable reactor fuel. A more logical solution, he says, would be to send the excess LEU to Russia for conversion into fuel rods for Iran's Bushehr nuclear reactor, in operation since 2011.

The "single biggest challenge," Acton says, will be eliminating the covert pathway. To improve the odds of detecting secret work, the final agreement will detail an intrusive inspection regime and unprecedented oversight of Iran's purchases for its nuclear program.

Other thorny issues remain. One is Iran's R&D on laser enrichment of uranium, a new route to fissile material that could be even easier to hide than ranks of centrifuges. "That's not something we want to see them pursuing," says the U.S. official. And Acton says it's worth bearing in mind that, until negotiators hammer out the final details, "we don't have a deal." ■

PHOTO: AP PHOTO/RONALD ZAK

GEOCHEMISTRY

# Acid oceans cited in Earth's worst die-off

## Signature of acidification found in Permian extinctions 250 million years ago

*By* **Eric Hand**

Things went really sour for life on Earth 250 million years ago. A team of European geoscientists has found the most direct evidence yet that the world's oceans became sharply acidic at the boundary between the Permian and Triassic periods—when scientists estimate 90% of Earth's species died in the worst mass extinction ever.

The "PT" die-off affected all types of living things, but it hit marine species the hardest—killing off, for instance, the once-ubiquitous trilobites. The new study, published on page 229 of this issue, concludes that the ocean acidification played a major role in the cataclysm. Acidification can kill sea creatures by weakening their ability to produce their calcium-bearing shells, and it is driven by excess carbon dioxide ($CO_2$) dissolving in the ocean. As such, the extinction holds a cautionary lesson for today: Because of $CO_2$ released by burning fossil fuels, oceans could now be acidifying even faster than they did 250 million years ago, although the process hasn't yet persisted nearly as long.

"It's not inconsequential that we are disrupting the earth's carbon cycle at a faster rate than the worst extinction of all time," says Lee Kump, a geochemist at Pennsylvania State University, University Park. "Even if it only exists for a few centuries rather than 10,000 years, rates of change do matter."

Unlike the asteroid-induced extinction that killed the dinosaurs 66 million years ago, most scientists think the even bigger catastrophe at the end of the Permian was homegrown: triggered by a massive bout of volcanism in Siberia that released trillions of tons of carbon into the atmosphere and oceans. Researchers previously found signs that living things endured multiple stresses as a result of the eruptions: global warming, ocean acidification, a drop in dissolved oxygen in the oceans, and even a buildup

of toxic sulfur. But sorting out the relative importance and interdependence of these effects has been difficult.

Now, scientists have better evidence that ocean acidification hit living things hard. It comes from carbonate-bearing limestones in the United Arab Emirates. They formed some 250 million years ago in the shallow waters offshore of Pangaea, the supercontinent at the time, locking in the geochemical signals of the ancient Tethys Ocean. Traditionally, geochemists have used variations in certain carbon isotopes as a sign that a pulse of atmospheric carbon dioxide was entering the ocean and triggering acidification. But in the new *Science* study, researchers analyzed the Tethys rocks for isotopes of boron—a fainter signal, but one that corresponds more directly to ocean acidity. The method works because chemical reactions in seawater cause the ratio of the isotope boron-11 to boron-10 to increase along with the pH. Rocks that precipitate on the sea floor reflect those changing ratios—and with them, the changing acidity of the ocean over time.

The researchers found a drop in the isotopic signal that would have corresponded to a drop of 0.6 to 0.7 pH units—a significant change in seawater chemistry. "This is the first really direct evidence of ocean acidification for this mass extinction," says Matthew Clarkson, a geochemist at the University of Otago, Dunedin, in New Zealand, who led the study.

The boron anomaly comes about 50,000 years after a sharp change in carbon isotopes that has long been thought to signal the start of ocean acidification and the beginning of the extinction. The study team says the gap between the carbon and boron signals suggests that separate pulses from carbon-spewing volcanoes drove a two-stage extinction, an idea first put forward in 2012. First, a slow 50,000-year pulse of carbon entered the air and then the oceans. The researchers hypothesize that the oceans were already somewhat alkaline, allowing them to absorb the incoming $CO_2$ with very little change in pH and modest impacts on marine life. But then a second, sharper pulse of carbon arrived in a period of 10,000 years, overwhelming the oceans' feedback mechanisms.

Clarkson says the scenario explains why paleontologists have found that sea animals like gastropods and bivalves—calcifying creatures that are most susceptible to acidification—suffered their greatest losses in the later stage of the Permian extinction. "The fossil record supports what we're seeing with the geochemistry," he says. Andrew Knoll, a paleontologist at Harvard University, agrees that a late, sharp pulse of acidification could help explain the extinction record.

The Permian-Triassic catastrophe holds mixed messages for Earth today. On the one hand, the pace of acidification was slower than it is now. The study team estimates that, in the acidification event, 24,000 gigatons of carbon were injected into the atmosphere over 10,000 years—a rate of 2.4 gigatons per year—and most of it wound up in the oceans. Currently, scientists estimate carbon from all sources is entering the atmosphere at a rate of about 10 gigatons per year.

On the other hand, today's economically viable fossil fuel reserves contain only about 3000 gigatons of carbon—far shy of the Permian total, even if human beings burn it all. "We're injecting the carbon faster, but it's unlikely that we have as much carbon to inject," says study co-author Tim Lenton, an Earth systems scientist at the University of Exeter in the United Kingdom. But knowing that the Permian was much worse doesn't bring Lenton much comfort. "Biology is pretty smart—it can cope with a certain amount of acidification," he says. "But I suspect there are limits to adaptation. There will be some point at which [species] crack."

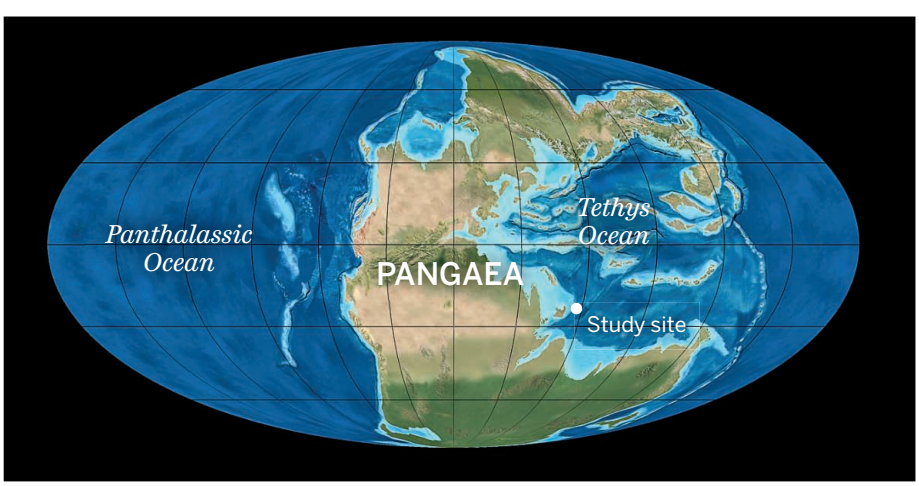
Study co-author Rachel Wood, a geobiologist at the University of Edinburgh in the United Kingdom, wants to make sure that the acidification pulse was more than a regional catastrophe. The team next plans to test rocks that formed on the floor of the Tethys 250 million years ago in present-day Iran and Oman. "We need to establish that this is a global signature," she says. ■

Musandam mountain cliffs in the United Arab Emirates include rocks laid down during a 250-million-year-old catastrophe that killed off 90% of life on Earth.

CREDITS: (PHOTO) D. ASTRATTI; (IMAGE) RON BLAKEY/COLORADO PLATEAU GEOSYSTEMS INC.

## Acid bath

The rocks that recorded the deadly ocean acidification were once the floor of a shallow sea off the coast of the supercontinent Pangaea.



# U.S. agencies fall in line on public access

## Major research funders move to make papers free to all

*By* **Jocelyn Kaiser**

After a decade of often fierce debate over whether the public should have free access to the scientific papers produced by their tax dollars, advocates for so-called open access celebrated a notable victory last month: The National Science Foundation (NSF) unveiled a plan to require its grantees to make their research freely available. NSF's move meant that the federal agencies that provide the bulk of the nation's basic and applied research funding have now complied with a 2013 White House order to make the peer-reviewed papers they fund freely available within 12 months of publication. The order, which applies to federal agencies that spend more than $100 million a year on research and development, will ultimately make hundreds of thousands of scientific papers once hidden behind paywalls available to anyone with an Internet connection (see table).

For the typical U.S. scientist with a federal research grant, the new public access mandate means they will need to follow the example of researchers funded by the National Institutes of Health (NIH). Since 2008, NIH grantees have been required to send their accepted manuscripts or final published paper to the agency's PubMed Central full-text archive. Researchers may deposit the paper themselves through a special Web portal, or the journal may do it for them. At most agencies, those who ignore the mandate will not get credit for their papers during reviews of funding proposals or job performance.

The details of how the public will get access to the full-text papers vary by agency. Some, including NASA, will share papers through a full-text archive similar to PubMed Central, which has grown to hold more than 3 million full-text papers (see graph). Open-access advocates prefer this model because they say it allows for simpler text and data mining across an entire corpus of articles. But two agencies, NSF and the Department of Energy, have heeded concerns from publishers that full-text archives will divert eyeballs from their sites and cut into advertising revenue. They are building portals that provide only a paper's abstract and other metadata; links will take users to the full-text article on the publisher's website. ■

### Public access policies at U.S. science agencies

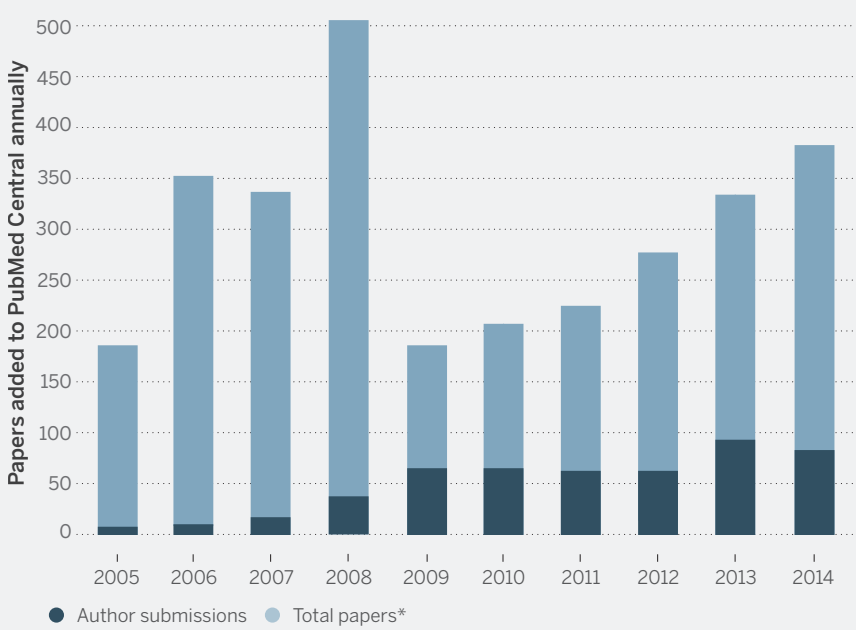
Researchers must make papers freely available within a year, either through abstracts that link to publisher websites or a central full-text archive like PubMed Central.

| AGENCY | BUDGET | MODEL | EXPECTED PAPERS PER YEAR | RESEARCH FUNDED AFTER |
|---|---|---|---|---|
| NIH | $30.3 billion | Central archive | 104,000 | 2008 |
| DOE | $5.07 billion (Office of Science) | Linked abstracts | 20,000–30,000 | 1 October 2014* |
| NSF | $7.3 billion | Linked abstracts | 40,000 | 1 January 2016 |
| DOD | $2.3 billion (basic research) | Central archive | 40,000 | 1 October 2016 |
| USDA | $1.1 billion (agricultural research) | Central archive | 20,000 | 1 January 2016 |
| HHS agencies (besides NIH) | $13.1 billion | Central archive | 6000 | 1 October 2015 |
| NASA | $3 billion (research) | Central archive | 10,000 | 1 January 2015 |
| NIST | $864 million | Central archive | 1300 | 1 October 2015 |

Not yet released: USGS, EPA, NOAA. *Office of Science; other offices 1 October 2015

### A growing trove of full-text biomedical research papers

Since 2008, NIH has required grantees to submit full-text papers to PubMed Central either directly or via the journal. Journals deposit many more papers that are not funded by NIH.




*Includes back issues added from 2006 to 2008 and papers deposited directly by journals

CREDITS: (DATA SOURCE) NIH; DOE; NSF; DOD; USDA; HHS; NASA; NIST; (GRAPH) NIH MANUSCRIPT SUBMISSION SYSTEM; NCBI

PALEONTOLOGY

# *Bully for* Brontosaurus!

## A comprehensive study of one branch of the dinosaur family tree resurrects *Brontosaurus* as a valid genus

*By* **Michael Balter**

What's in a name? Sometimes plenty, as the U.S. Postal Service (USPS) found out in 1989 when it issued colorful dinosaur stamps, including one for *Brontosaurus*. Paleontologists and educators loudly protested that the correct scientific name for the iconic beast was *Apatosaurus*—a fact that even lay dino aficionados and many 8-year-olds took pride in knowing. The Smithsonian Institution accused USPS of favoring "cartoon nomenclature to scientific nomenclature." It didn't help that the stamps were officially launched at Disney World.

Now, a dinosaur-sized study of the family tree of the Diplodocidae, the group that includes such monstrous beasts as *Apatosaurus*, *Diplodocus*, and *Barosaurus*, finds that USPS got it right: The fossils originally called *Brontosaurus* show enough skeletal differences from other specimens of *Apatosaurus* that they rightfully belong to a different genus. The study, published online this week in the journal *PeerJ*, brings the long-banished name back into scientific respectability as a genus coequal with *Apatosaurus*.

"I didn't start out trying to resurrect *Brontosaurus*," says lead author and paleontologist Emanuel Tschopp. He was just trying to better understand the evolutionary relationships among all Diplodocidae as part of his Ph.D. thesis at the New University of Lisbon. Indeed, paleontologists are impressed by the scope of the new study, which included 81 skeletons and 477 skeletal features or characters, far more than any previous analysis. "Emanuel's data set is now the largest published so far" for plant-eating dinosaurs, says Philip Mannion, a paleobiologist at Imperial College London. The name change is likely to stick, adds paleobiologist Paul Upchurch of University College London: "I will be happy to start using *Brontosaurus* again."

How *Brontosaurus* lost its name in the first place is the stuff of dino legend. During the "Bone Wars" of the late 19th century, paleontologists Edward Cope and Othniel Marsh marauded through the American West bitterly competing for dinosaur fossils, which led to some rapid and slapdash descriptions. In 1877, Marsh published a brief note on one of his skeletons, calling it *Apatosaurus* ("deceptive lizard"). Two years later, he published an equally brief report on a supposed new genus, *Brontosaurus* ("thunder lizard"). By 1903, paleontologists had decided that the two beasts were too similar to be divided into separate genera. Because *Apatosaurus* had been named first, it had precedence under the rules of scientific nomenclature. Paleontologists confirmed this conclusion in the 1970s, and know-it-all children have instructed their parents ever since that "*Brontosaurus*" isn't a valid name. (The taxon had its defenders, including evolutionary biologist Stephen Jay Gould in his "Bully for *Brontosaurus*" essay.)


The U.S. Postal Service caught flak for calling these giant plant-eating dinosaurs *Brontosaurus* in 1989, but now scientists say the genus name is valid after all.

In the new analysis, Tschopp looked at every Diplodocidae specimen he could get his hands on. These plant-eating dinos lived in today's North America and Europe between 160 million and 145 million years ago; their long necks and tails stretched up to 35 meters from end to end. Tschopp examined bones in 18 different museums in the United States and Europe and studied photos and drawings from other specimens. The idea was to get a high-resolution family tree by examining many individual specimens, rather than focusing only on fossils that represent an entire species. The study included six specimens of *Apatosaurus excelsus*, as *Brontosaurus* has been called since 1903.

Tschopp and colleagues found that *A. excelsus* differed from the three other recognized species of *Apatosaurus* in at least a dozen key characters across the skeleton. That's as many differences as seen between genera long accepted as distinct, such as *Diplodocus* and *Barosaurus*. For example, a true *Apatosaurus* has a bulkier neck and was "even more robust than *Brontosaurus*," Tschopp says. But *Brontosaurus* also has some features that *Apatosaurus* lacks, such as a rounded expansion of one edge of its shoulder blade and a longer bone in its ankle.

Given these anatomical differences, "It would be unfair to lump *Apatosaurus* and *Brontosaurus* together," Upchurch says. Mannion agrees, and also endorses the team's finding that another Diplodocidae genus, *Eobrontosaurus*, should be included within *Brontosaurus*, giving the genus a few additional species to call its own.

Some paleontologists have reservations. "It's going to force us to ask questions about what we really mean by genus and species in a paleontological context," says paleontologist John Whitlock of Mount Aloysius College in Cresson, Pennsylvania. "Is it more useful to distinguish specimens as *Brontosaurus* and *Apatosaurus* than it is to distinguish *A. excelsus* from other species of *Apatosaurus*? I don't know, but I hope it's the start of a conversation." He and others, including Tschopp himself, note that the characters used aren't cut-and-dried and could be scored differently by others.

Upchurch thinks this kind of detailed taxonomy could help resolve questions such as how diverse dinosaurs were just before they went extinct about 66 million years ago. Others welcome the resurrection of an icon. "*Brontosaurus* has a prominent place in the public imagination," Mannion says. "It can only be a good thing that it is back with us. … It shows that science develops through time and that it's possible to change our minds, even about long-held views." ■

IMAGE: LEFTERIS PAPAULAKIS/SHUTTERSTOCK.COM

Brazil's government has moved to protect the dusky grouper—at least on paper.

CONSERVATION

# *Brazil roils waters with moves to protect aquatic life*

## Fishing moratorium draws fire from industry and scientists

*By* **Herton Escobar**, *in São Paulo, Brazil*

Although Brazil is famous for its biodiverse rainforests, the animal pictured on its most valuable banknote (100 reais) is a fish from the ocean: *Epinephelus marginatus*, the dusky grouper. The high-denomination banknote is a fitting place for the species, which is prized for its juicy white meat. Too prized, it turns out. The dusky grouper is one of about 100 species of commercially valuable fish and other aquatic organisms recently listed as threatened by the Brazilian Ministry of the Environment (MMA). As of 16 June, every one of them, along with more than 300 other species, could be off-limits for fishing, to the distress of industrial and artisanal fishermen alike.

The new "red lists" of endangered species were published on 17 December 2014, representing 5 years of work by more than 1000 Brazilian scientists. But within weeks, the prospect of a fishing ban sparked political and scientific discord. The Ministry of Fisheries and Aquaculture says the list will severely impact the national fishing economy; together with industry it is pressuring MMA to reevaluate the status and rules for dozens of species, including grouper, parrotfish, sharks, and rays. Some scientists and conservationists also worry that a blanket moratorium will backfire. "Fishermen won't stop fishing," says Ronaldo Francini-Filho, a reef and fish biologist at the Federal University of Paraíba. "They will either continue to fish illegally or shift the pressure to other species."

MMA's December ruling will make it illegal to fish for any of the listed species, which include both fresh- and saltwater animals. In principle, species in the "vulnerable" category—as opposed to "endangered" or "critically endangered"—can still be fished, as long as regulations are in place to guarantee that the take is sustainable. The problem is that these regulations don't exist for most species, including the dusky grouper, so they, too, are in line to become off-limits.

Brazil's governance of fisheries is fragile at best. "We don't even know how many fishing vessels are operating," says biologist Monica Brick Peres, director of Oceana Brasil and a leading expert on fisheries management. Many fear that because of weak enforcement, the moratorium could increase the level of illegal fishing rather than take pressure off the endangered species. At the same time, the lack of data from fishing fleets could deprive scientists and policymakers of the ability to monitor what is happening in the ocean. "To make fishing sustainable you need to work with data, and that is the first thing we don't have," Peres says.

Poor fishermen prize the guaiamum crab.

Better than a blanket prohibition, she says, would be to develop monitoring systems and work in partnership with fishermen to develop science-based recovery plans that are specific to fishing stocks, not species. "Publishing the red list is an important first step to diagnose the problem, but it doesn't solve it. We need to take measures to get those species off the list, and just telling people they can't fish is not likely to achieve that," she says.

The fisheries ministry is negotiating with MMA to exempt some commercially valuable species from the ban. MMA has convened a group of experts to help it develop sustainable fishing plans for 34 lucrative species in the "vulnerable" category, and has agreed to extend the 16 June deadline if those plans cannot be finished by then. It might even exempt one "critically endangered" species, the guaiamum crab, which is still abundant in some areas and represents an important source of protein and income for low-income fishermen.

But the Environment Ministry is unlikely to relent on other endangered or critically endangered species. Ugo Vercillo, general coordinator of research and monitoring at the Chico Mendes Institute for Biodiversity Conservation (ICMBio), the MMA organ responsible for drafting the red lists, says industry is exaggerating the economic consequences of the ordinance and that management plans for vulnerable species will alleviate most of the impact. "It would be great to take some names off the list, but if we do that, it will be based on scientific data, not political pressure," he says.

Experts add that a moratorium may be the only hope for survival for some species, such as highly endangered sharks and rays, which have slow reproductive cycles. "There is just no way to exploit these species sustainably anymore," says Otto Gadig, a marine biologist at São Paulo State University.

Protecting these species is not so easy, Vercillo concedes. Targeted fishing of sharks and rays is already illegal in Brazil, but they are snared as by-catch in other fisheries, such as bottom trawling for shrimp and tuna longlines. At ICMBio, environmental authorities are not sure how to control the by-catch without interfering with legal fishing for other species that are not on the endangered list. "We don't have a solution for this yet," Vercillo says. "But we have to make sure these animals don't enter the market." ■

*Herton Escobar writes for* O Estado de São Paulo.

PHOTOS: (TOP TO BOTTOM) IMAGEBROKER/ALAMY; JEAN DELARD DE RIGOULIÈRES/FLICKR

CANCER BIOLOGY

# Infectious cancer found in clams

## A single leukemia afflicts shellfish beds across hundreds of kilometers, puzzling biologists

*By* **Erik Stokstad**

Cancer cells aren't contagious. That's a rule in biology. But this precept has broken down spectacularly in the sand flats and shallows of North America's east coast, where a fatal leukemia-like condition has spread throughout soft-shell clams from Prince Edward Island in Canada all the way to New York state, perhaps via free-floating cancer cells. Columbia University's Stephen Goff, one of the researchers who describe the extensive transmission of this cancer in *Cell* this week, calls it "a super metastasis, spreading to whole new animals."

"This is a remarkable paper," says Charles Walker, a molecular biologist at the University of New Hampshire in Durham, who also studies the clam disease. "It's very interesting and provocative." The discovery adds one more example to the two known cases of transmissible cancer: a venereal disease in dogs and facial tumors that Tasmanian devils spread by biting each other. And it suggests that supercancers spreading from individual to individual may not be as rare as was thought.

The clam disease, called disseminated or hemic neoplasia, was first noticed in soft-shell clams (*Mya arenaria*) in New England in the 1970s. Sick clams suffer from rapidly proliferating immune cells in their circulatory fluid. Many die in outbreaks, but researchers say there is no health risk to humans who eat diseased ones. A similar condition afflicts other bivalves, including oysters, mussels, and cockles.

Carol Reinisch, who studied the disease at the Marine Biological Laboratory in Woods Hole, Massachusetts, thought a virus might be involved, so a few years ago she contacted Goff, an expert on leukemia viruses in mice. He initially checked samples of dead clams for reverse transcriptase, an enzyme associated with cancer-causing retroviruses. The samples were "wildly positive," he says, but the transcriptase hadn't come from a virus. Instead, he and colleagues discovered, the enzyme was originating from the genome of the cancer cells, where it was encoded by a fragment of added DNA called a retrotransposon. Retrotransposons have the unusual ability to copy themselves and insert into new locations within the genome.

The researchers were struck by several oddities about the retrotransposon, which they dubbed Steamer after the clam's common name. Cancer cells in clams from Maine and Prince Edward Island all had roughly the same number of copies of the retrotransposon, and they had a lot of them, approximately 150 copies per cell. The clam's healthy tissue contained fewer than 10 copies, Goff's team reported in the *Proceedings of the National Academy of Sciences* last year. Even more weird: Although retrotransposons tend to jump around the genome, the distribution of copies of Steamer was similar from one clam cancer to the next. The tantalizing implication was that the cancer had not arisen independently in each individual, as almost always happens in other animals, but had spread among the clams. "It was so wild and so exciting; we didn't believe it for a long time," Goff recalls.

But postdoc Michael Metzger, also at Columbia University, subsequently found more examples of similarly located Steamer insertions. And a close comparison of subtle genetic changes called single-nucleotide polymorphisms (SNPs) has now confirmed that the cancer cells' DNA does not match the genomes of the clams they afflict. The seemingly inescapable conclusion: All the cancer cells derive from one primordial immune cell gone rogue, which then spread from clam to clam.

It's possible that the disease may have slowly spread from one population to another along hundreds of kilometers of coastline, says invertebrate pathologist Bruce Barber, a visiting professor at Eckerd College in St. Petersburg, Florida. In addition, the outbreaks could have hopscotched along the coast if people accidentally stocked clam beds with infected clams.

The age of the initial mutations is unknown, but a handful of genetic differences between cancer cells in the U.S. and Canadian clams suggests they happened sufficiently long ago for the cell lineages to start diverging. Researchers also wonder how the disease is transmitted. Perhaps sick clams release cancer cells during spawning, when injured, or after death, Goff and his colleagues speculate. The cancer cells may infiltrate healthy clams as they filter feed, somehow slipping across their digestive tracts. Another puzzle is how the cancer cells—which are foreign tissue—escape the clams' immune systems, says Elizabeth Murchison, a geneticist at the University of Cambridge in the United Kingdom who studies the transmissible cancers in dogs and Tasmanian devils.

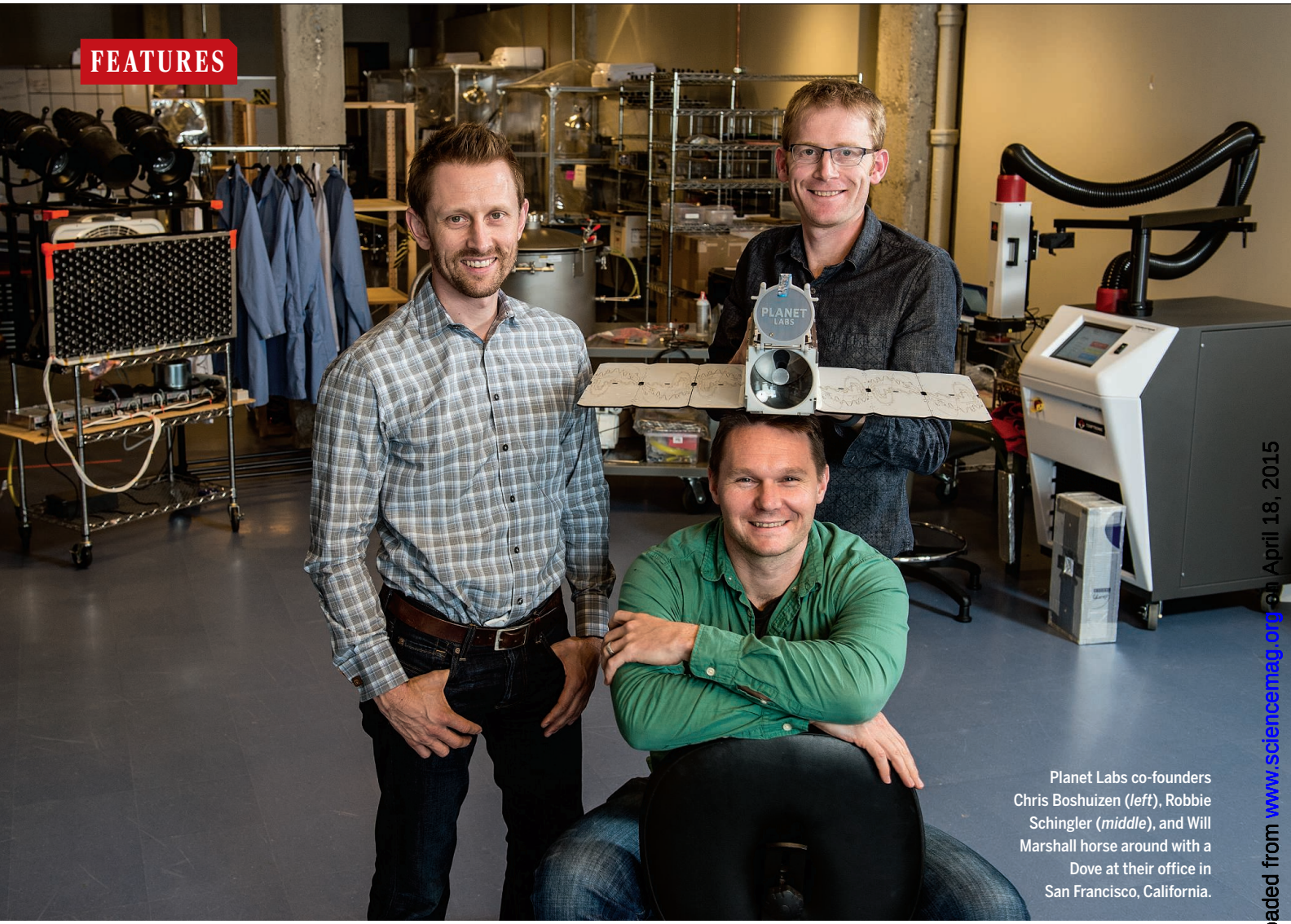
Goff and his colleagues are now checking whether similar diseases in other kinds of shellfish are also caused by transmissible cancer cells. "It could be that this type of disease might not be so rare as we thought," Murchison says. ■

### The 1000-kilometer cancer

Cancer cells from soft-shell clams in four locations likely originated from a single mutant cell, which infected other clams. The leukemia-like disease occurs from northern Newfoundland to the Chesapeake Bay, nearly all of the clam's range.


CREDITS: (MAP) ADAPTED FROM METZGER *ET AL. CELL* **161** (9 APRIL 2015) © 2015 ELSEVIER INC.; (PHOTO) MICHAEL J. METZGER

FEATURES




Planet Labs co-founders Chris Boshuizen (*left*), Robbie Schingler (*middle*), and Will Marshall horse around with a Dove at their office in San Francisco, California.



# STARTUP LIFTOFF

## How flocks of small, cheap satellites, hatched in Silicon Valley, will constantly monitor a changing Earth

*By* **Eric Hand**, *in San Francisco, California*

Vats of homebrewed porter and brown ale ferment under a lunchroom table. In the corner lie a drum kit and guitar, property of Hank and the Doves, the company's pop cover band. Emma the dog roams—and sheds—freely. In some ways, Planet Labs is your typical Silicon Valley startup. But it's not where you'd expect to see the precision assembly of space satellites. "In terms of overall cleanliness, we just don't care," says co-founder Chris Boshuizen, who wears a droopy Santa hat in preparation for an office holiday party on this rainy December day in San Francisco.

Boshuizen pushes aside strips of clear vinyl sheeting and enters what he calls the "clean enough" room. He stomps on a mat of sticky tape that helps eliminate static charges that could zap satellite electronics—a rare precaution. Beyond another line of tape, no alcohol is allowed. There a shelf is stocked with the company's product: space telescopes no bigger than a loaf of bread.

Two dozen of these telescopes, called Doves, already orbit the Earth, imaging the

PHOTO: DEANNE FITZMAURICE

ground with a resolution good enough to pick out treetops, roads, and buildings. Another 14 are set to ride into orbit next week on a SpaceX cargo rocket. Although heftier spacecraft can spy on Earth with higher resolution, few can match the repeat rate at which one craft in Planet Labs' swarm passes over the same patch of ground. If the company can get between 150 and 200 Doves in orbit, it will be able to take a daily snapshot of the entire planet. This time-lapse flipbook will reveal flooding on rivers, logging in forests, and road building in cities, as they happen.

Change is the name of the game—and the main attraction for researchers and commercial clients. Boshuizen says the company has contracts in place that are worth more than the $135 million in venture capital funding it has so far received.

From the shelf, Boshuizen grabs a Dove. Though it is destined for orbit, he handles it as roughly as he would the phone in his pocket. He doesn't worry about damage, because the inexpensive electronics inside it are similar to the phone's. The nonchalance is all part of the plan. Plenty of Doves have fizzled out in orbit or fallen back to Earth and burned up in the atmosphere. But that's OK, because more Doves are always ready to take wing.

For decades, engineers have been building satellites like bespoke Swiss watches, sparing no expense and spending years to perfect them. More than 1000 people in 17 countries have had a hand in building the $8 billion James Webb Space Telescope, more than a decade in the making. Once launched in 2018, the Webb telescope will never be made again. At Planet Labs, Doves are mass produced to the brink of disposability; in 2015, the company is aiming for a sustained production rate of 25 Doves a month. Boshuizen says the cost of building and launching each Dove is well under $1 million.

Mass production lets engineers experiment with the satellites and launch them in aggressive, iterative cycles. In less than 2 years, Planet Labs has gone from Dove 1.0 to sending version 12 into space. Yet the ratio of total Doves launched (99) to employees (105) is still about one. "It changes because we keep hiring," Boshuizen says. "But we also keep building satellites."

Much of Planet Labs' success can be attributed to a decision to squeeze their ambitions into a very small box, a specific form factor called a CubeSat. Named after their characteristic 10-centimeters-a-side size, CubeSats were first launched in 2003 as an educational tool—a way for graduate students to get something in space (see sidebar, p. 176). The little boxes beeped and did little else. But fueled by the massive investment in consumer electronics, the size and cost of most satellite components—radios, computers, solar panels—have plummeted even as their capabilities have exploded. At the same time, CubeSat builders have found new and cheaper ways to get into orbit: packed into spare payload space around larger satellites, or stowed on cargo flights to the International Space Station. In 2014, a record 132 CubeSats were launched—and 93 of them were Doves, according to Jonathan McDowell, an astronomer at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, who tracks satellite launches.

## I spy with my little eye

Each Dove must fit within a "3U" box measuring 10 centimeters by 10 centimeters by 34 centimeters, or three CubeSats' worth of space. The telescope is pointed with the help of a star tracker, reaction wheels, and electromagnetic torquers.

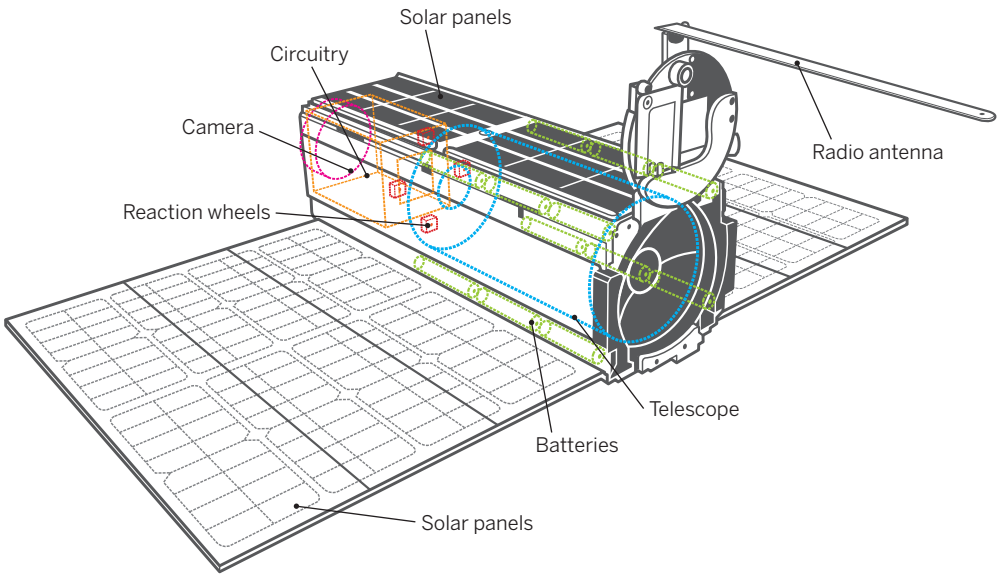



ILLUSTRATION: JAMESON SIMPSON

> ***"What happens when anyone on the planet can understand the state of the world?"***
>
> **Robbie Schingler**, Planet Labs

McDowell says CubeSats are gripping the aerospace industry and changing the way business—and science—is done. "Now you're seeing not just student projects, but CubeSats deployed by the military, by space agencies—doing real jobs," he says. Jordi Puig-Suari, the co-founder of the CubeSat standard, concurs. "These little guys are finally ready to do serious missions," he says. And one upstart company is leading the charge, he adds: "Planet Labs is the darling of the CubeSat community."

**IN THE LATE 2000s**, small satellites were still a curiosity, but a buzz was in the air at NASA's Ames Research Center near Mountain View, California. Boshuizen and another physicist, Will Marshall, recall a senior Ames engineer waving around a government-issued smart phone, declaring that it had more computing power than the average satellite. Why not just launch the smart phone? he asked. "We eventually took him seriously," Boshuizen says. "We got a smart phone, stuck it in a vacuum chamber, and it still worked." It turned out that costly "rad-hard" parts, built to withstand the vacuum of space and its radiation environment, weren't so important.

"Perhaps you've been lied to with this whole notion that things need to be space-qualified," Boshuizen says. After all, he notes, astronauts on the barely shielded space station use iPads. (The more intense radiation of deep space, however, might pose a threat to future CubeSat missions to other planets, which would venture outside Earth's protective magnetic field.)

In 2010, Boshuizen and Marshall assembled a PhoneSat team at Ames. They took an HTC Nexus One smart phone out of its case, reprogrammed its Android operating system, and added extra batteries and a radio that would downlink pictures to Earth. Total cost: about $3500. In 2013, the first three PhoneSats were launched. Two of them lacked solar panels, but they took and

sent back pictures in the week before their batteries ran down. "We bet the farm on this idea that we could launch a phone into space and that it would work," Boshuizen says. "And it did."

Even before the PhoneSats were launched, Boshuizen and Marshall began talking about taking the idea further with a third friend, Robbie Schingler, who was working for the chief technologist at NASA headquarters in Washington, D.C. The fast-talking trio knew each other from long before: They had met at a 2002 Space Generation Summit in Houston, a workshop for young space evangelists. "The day I showed up [there], I decided they were my people," Boshuizen says.

In December 2010, they registered their own company, initially called Cosmogia. They talked about many possible business models, including a constellation of CubeSats that would provide global Internet. They ultimately decided that an Earth-imaging mission carried fewer risks and was a service for which there is growing business demand. "Frankly, we chose to do the remote-sensing mission because we had a higher probability of success," Schingler says.

Even though the team had no funding and no way of paying anyone, Ames engineers like Ben Howard were eager to join. "They had a big idea for what to do with these CubeSats," he says. "I also didn't think they were crazy."

For most of 2011, they worked out of the "Rainbow Mansion," a six-bedroom house in Cupertino, California, that Marshall and Schingler had rented in 2006 and packed with like-minded people. With house policies achieved by consensus and weekly salons given by the likes of physicist Roger Penrose and Internet activist Lawrence Lessig, the house is like a "hippie commune except it's mostly filled with tech geeks," Howard says.

Like so many Silicon Valley startups, the group began working out of the garage. They pushed camping equipment to the side and hung a stop sign from the ceiling to keep random Rainbow residents out. The team worked on couches, hunched over laptops loaded with computer-aided design software. One person might design a part and get it 3D-printed, while another would work on orbital simulations.

In designing the Doves, the group chose a frame the size of three stacked classic CubeSats: 10 centimeters square and 34 centimeters long. They found that Questar telescopes—used by hobbyists for more than half a century—fit the space perfectly, so they ordered a couple of custom ones fashioned out of Invar, a thermally stable nickel-iron alloy. They pointed Questars out the garage door to test settings on megapixel cameras attached to the rear of the scopes. Occasionally, they would cross the bay to Lick Observatory, set up a receiver antenna, and test the strength of their radio. "We were not naive," Howard says. "We had come from NASA and had seen the so-called right way to design a satellite. We were very aware that we were doing something different and risky."

Soon, they had assembled mostly off-the-shelf parts into a working satellite that they were willing to send to space. With their personal savings, the three co-founders booked a $250,000 spot on the maiden flight of Orbital Sciences' Antares rocket, in April 2013.

A fire in a Brazilian field is captured by a Dove (*bottom*) a day after the same spot was imaged by Landsat (*top*).

Michael Safyan, another early employee, says he didn't sleep much during the 6 days that Dove-1 lived. The team was ecstatic when a 4.5-meter radio dish in Chilbolton, U.K., captured the first image: a patch of forest in the Pacific Northwest, in such good focus that they could count the trees. "We had no idea if the thing would even turn on," Safyan says. "It was such a huge validation that we were on the right track."

By that time, the fledgling company had moved out of the garage and into its San Francisco office. More Doves entered orbit, and the images they took began to stitch together the patchwork quilt of Earth. The Planet Labs team started to notice changes. The boundaries of pit mines expanded. In Rio de Janeiro, they saw that favelas had been bulldozed in the weeks leading up to the 2014 World Cup. They even noticed plumes of smoke over a patch of ground in California—an incipient forest fire. "It was just 10 minutes old," Howard says.

Investors began to line up, and customers began to sign contracts. Planet Labs has publicly announced only a few, but Boshuizen talks about several areas of commercial interest. A primary one is agricultural: the ability to monitor the productivity of fields. Environmental compliance is another—for instance, mining companies wishing to show that they have restored an area to the correct standard. A third area is in commercial mapping: By monitoring the growth of roads and homes, Planet Labs can identify areas where Internet mapping companies need to concentrate their data-gathering efforts.

Two Doves are expelled from the International Space Station on 27 February 2015.

Plenty of scientists would also like to get their hands on Planet Labs data, says Curtis Woodcock, a geographer at Boston University and co-leader of the science team for Landsat, the venerable series of Earth-monitoring satellites operated by the U.S. Geological Survey (USGS). Woodcock says officials routinely ask Landsat scientists what the satellites could do better, and the answer is always the same: Make more frequent passes over each patch of Earth. Daily snapshots could help home in on the

PHOTOS: (TOP TO BOTTOM) USGS/NASA LANDSAT; PLANET LABS

PHOTO: NASA

moments when forests green up in the spring. "It's becoming more volatile and has shifted earlier in many places as a result of changing climate," Woodcock says. Another scientific application, he says, is nailing the timing of snowmelt, which is hard to predict and has big implications for water management.

But with an entire earth to survey and just two working satellites, Landsat 7 and 8, an 8-day repeat rate is the best Landsat can manage, Woodcock says. To be sure, the big orbiters have 11 spectral bands to the Doves' three basic color bands and a heritage that ensures that images are precisely calibrated from one mission to the next. "The scientific community functions in the measurement domain," Woodcock says. "Whether Planet Labs is going to make it from the picture world to the measurement world is still up in the air."

But Doves, with 3- to 5-meter resolution, already outperform the 15-meter resolution of Landsat 8—although other Earth-imaging services can do even better. DigitalGlobe, for example, provides satellite imagery that has a resolution of better than a meter, but tasking the company's fleet of six truck-sized satellites to get a new image for a specific area can take a week or two and is expensive. Another company, SkyBox, plans to launch a constellation of two dozen satellites the size of mini-refrigerators—still small by typical aerospace standards, but much bigger than the Doves—that are also capable of reaching submeter resolution. (Google paid $500 million to buy Skybox last June.)

Still other imaging companies are developing drones for jobs such as high-resolution monitoring of oil and gas pipelines. But Boshuizen says the Doves can complement those efforts—for example, by flagging changes that drones or more capable satellites can then examine in more detail. "I view them as our customers, not competitors," he says.

Although Planet Labs is trying to make money, the three co-founders exude humanitarian idealism. For example, they say they hope someday to make Planet Labs data free—perhaps, Boshuizen says, by charging only for the newest images or for access to large numbers of images, as Google Maps does. They say their tiny telescopes could allow watchdog groups to monitor environmental degradation, or let human rights groups keep tabs on the size of refugee camps and the movements of marauding militias. Schingler says the Doves will empower people who have never before had access to daily geospatial data. "What happens when anyone on the planet can understand the state of the world?" he asks. Power dynamics will shift markedly, Marshall says: "Not everyone's going to like it. We're very cognizant of that."

Planet Labs' way of doing business could be equally transformative, McDowell says. The whole notion of doing Earth observation via constellations of small satellites poses a threat to the old order, in which large aerospace companies build expensive, large satellites for agencies like NASA. For example, Landsat 8, launched in 2013, was built by Orbital Sciences at a total cost of $855 million.

USGS is planning a replacement Landsat mission, and budget realities mean it will have to be cheaper than its predecessors. Boshuizen says Planet Labs does not intend to formally compete to replace Landsat but would be happy to sell its data to USGS. The company has also discussed other missions: CubeSats stuffed with sensors rather

than telescopes. "You could do radar, lidar; you could do GPS augmentation or replacement," Boshuizen says. "We could move on to do things like earth science, atmospheric science, science in low-Earth orbit; we could do astrophysics and heliophysics, in part."

Pete Worden, director of NASA Ames, is proud that the people and ideas behind Planet Labs were incubated there—even if their visionary approach comes back to haunt NASA centers. "Applying Silicon Valley to aerospace is the most revolutionary thing that's happened probably since Goddard built his rocket," he says. "It's not surprising that it started here. But it's spreading."

**THE FINAL MONTHS OF 2014** brought major changes to Planet Labs, and one big challenge. It started on 28 October, when 26 Doves—known as Flock 1d—were scheduled to ride into space aboard an Orbital Sciences Antares cargo rocket bound for the space station. The launch, from the Wallops Flight Facility in Virginia, was the company's seventh.

At the Planet Labs office in San Francisco, Schingler made pancakes in the lunchroom—a launch-day tradition at the company—and served them to 70 or so employees and guests who had gathered to watch a NASA TV webcast. Also in keeping with tradition, Marshall delivered a prelaunch briefing. The Doves, he explained, were locked inside special deployers, wrapped in bubble wrap, and strapped down in the capsule, alongside food, water, and experiments destined for the space station. He reminded his listeners that launches were uncertain: Delays could occur at the last second, failure was always possible, and everything was out of their hands.

After this Antares cargo rocket exploded on 28 October 2014, Planet Labs scrambled to get new Doves into orbit.

PHOTO: NASA

At 3:22 p.m. Pacific Time, the rocket lumbered off the launch pad. Fifteen seconds later, one of the engines exploded; then a larger fireball engulfed the entire rocket. "There were gasps of shock," Marshall recalls. In a single moment, the company had lost a huge chunk of its assets.

For an old-school space mission, it would have been a shattering blow. In 2009, for example, another Orbital Sciences rocket fell into the ocean while attempting to deliver into space the Orbiting Carbon Observatory (OCO),

# Thinking inside the box

*By* **Eric Hand**

Why a 10-centimeter cube? The trademark size of the CubeSat emerged somewhat accidently, recalls Jordi Puig-Suari, an aerospace engineer at California Polytechnic State University in San Luis Obispo. Student-built satellites date back to the 1980s, but they were often unwieldy, stuffed with dubious hardware from RadioShack and auto parts stores. Not only were they expensive to launch, but commercial rocketeers were also wary of packing them alongside primary payloads.

In 1999, Puig-Suari met with Bob Twiggs, at the time an aerospace engineer at Stanford University, to discuss ways of getting more student projects into space. "We had to do something to get more opportunities to launch these things," recalls Twiggs, now at Morehead State University in Kentucky. They focused on slimming down the spacecraft, because weight drives up the cost of reaching orbit.

Over lunch at a sandwich shop in San Luis Obispo, Twiggs and Puig-Suari sketched out options on a napkin. They thought hard about the potential capabilities of a 10-centimeter cube with a mass limit of 1 kilogram—the size and weight of a liter of water. Clad in solar cells, the cube would eke out perhaps a watt of power, enough to power a small computer and a radio: "a Sputnik," Puig-Suari says. Back at Stanford, Twiggs found the perfect life-size demonstration model: a plastic box used for storing the insanely popular stuffed animals known as Beanie Babies. A standard was born.

In 2003, the first six student projects rode a Russian Eurockot into orbit, for about $30,000 a pop. Companies quickly sprang up, selling standard components, such as a chassis or radio, allowing university teams to focus on instrumentation. Many groups settled on the "3U" form factor—three CubeSats' worth of space, the size of Planet Labs' Doves—as the ideal balance of compactness and capability.

Early on, the biggest single expense was the rocket ride. Here, too, Twiggs and Puig-Suari offered an important innovation: Poly Picosatellite Orbital Deployers, or P-PODs. These spring-loaded boxes keep CubeSats quarantined from the main payload, then fling them into space. With time, the P-POD and its kin earned the trust of the world's major rocket launchers. Even U.S. military and spy agencies now accommodate CubeSats if a rocket has thrust to spare.

In recent years, launch prices have stayed put around $100,000 for a 1U CubeSat, Twiggs says. There are free ways to space, too: In 2010, NASA started subsidizing a dozen or more launches per year with its CubeSat Launch Initiative, which takes CubeSats to the International Space Station, where they are released into space. Reaching orbit will become even easier when a handful of companies develop dedicated CubeSat launchers, Puig-Suari says. One of the first may be Super Strypi, a rail-launched U.S. military rocket that is scheduled to carry a dozen CubeSats to orbit in October.

The CubeSat trend soon caught the attention of Therese Moretto Jorgensen, a program director for space weather research at the U.S. National Science Foundation (NSF) in Arlington, Virginia. "They were seen very much as toys," she says. "But we decided to give it a go." In 2008, she organized NSF's first solicitation for CubeSats grants. She received about 30 proposals for the grants, which typically award 3-year funding of $900,000—enough to build a CubeSat but a tiny fraction of a conventional spacecraft's cost.

The first NSF winner to fly was the Radio Aurora Explorer (RAX), launched in 2010. It sought to understand turbulent bubbles in the ionosphere by observing how radio waves from ground radar stations scattered off the bubbles into space. The RAX team found that bubbles were smaller and more confined than expected. "It produced this beautiful data," Jorgensen says. "It gave us a lot of confidence to keep doing things."

an important climate-monitoring satellite. It took 5 years to return a copy of OCO to orbit.

Planet Labs, however, still had Doves in space and a production line to make more. Boshuizen took stock: There were enough parts on hand to make 10 Doves immediately. Schingler began calling officials at NASA and at NanoRacks, the company that built the deployers, to see how quickly he could get on a subsequent launch. He worked to transfer regulatory licenses to the replacement satellites. He secured room for two Doves on the next SpaceX resupply mission, scheduled for December. That meant building and boxing up the satellites in record time. The team finished in 9 days. "The team did a major sprint," Schingler says. "It was a herculean effort."

Two Doves, covered in personalized graffiti ("NBD, all sats burn up *some*day"), were locked in padded Pelican cases and shipped to NanoRacks' facility in Webster, Texas. Boshuizen sneaked special features onto each of the spacecraft: One would gather three times as many pixels over a larger field of view, and the other had an experimental infrared filter on top of the three needed for color pictures.

The SpaceX launch was delayed twice. When it finally took place—on 10 January 2015 at Cape Canaveral in Florida—Planet Labs was already in the middle of its next big project: boxing up to move. The new headquarters, a few blocks away, boasted exposed brick walls and wood beams, breathing room for the growing staff, and space for brand-name diagnostic equipment such as a ThermoStream, a machine for pumping hot air into chambers to test the resilience of circuits and other equipment. The dog-friendly policy was being rescinded (primarily because of a new employee's allergy, Boshuizen says).

The launch itself signaled another transformation for the company: In contrast to October's drama of loss and recovery, it was *routine*. Schingler watched from a laptop at another of his communal compounds, in the jungle of Costa Rica. Marshall was in Florida, watching the predawn launch of the SpaceX Falcon 9 rocket in person for the first time. In San Francisco, Boshuizen went out for dinner and returned to the soon-to-be-old office around midnight. With only a handful of employees hanging around so late, there was no need for a prelaunch briefing. Or pancakes. Boshuizen, lead guitarist for Hank and the Doves, picked up his guitar and started to noodle around. "What else do you do on a Friday night?" he asks. "We were just sitting around chatting and drinking." A few more employees trickled in on their way home from San Francisco's bars and clubs.

Finally, the countdown came at 1:47 a.m. Pacific Time. Boshuizen, remembering how hard his team had worked to get the two Doves on board, watched nervously as the Falcon 9 roared off the launch pad, a streak of orange in the humid Florida night. He kept his eyes fixed on the video screen for another 10 minutes, until both the first and second stages had cut out—until the big bird had carried his two little birds to space. "Once it got out of the atmosphere, I could relax," he says. "I knew it would be fine." ■

DATA SOURCE: JONATHAN MCDOWELL, HARVARD-SMITHSONIAN CFA

Many early CubeSats tackled problems in space weather, an ideal field for in situ observations, but other areas of science are opening up as teams shrink and refine instruments. In January, ExoCube launched with a mass spectrometer, the first ever on a CubeSat, designed to measure the composition of the exosphere and ionosphere. In March, the Micro-sized Microwave Atmospheric Satellite was released from the space station. It will carry a microwave radiometer, an instrument typically found on billion-dollar weather satellites, which can map the 3D thermal structure of storms or the presence of sea ice. And in 2013, the NSF funded the Optical Profiling of the Atmospheric Limb CubeSat. When it flies in a few years, it will contain a "hyperspectral" instrument that will measure the temperature of the outermost atmosphere during magnetic storms by imaging it in 100 different spectral bands. Similar instruments could one day be aimed at the ground to map vegetative health, ocean algae blooms, or mineral deposits.

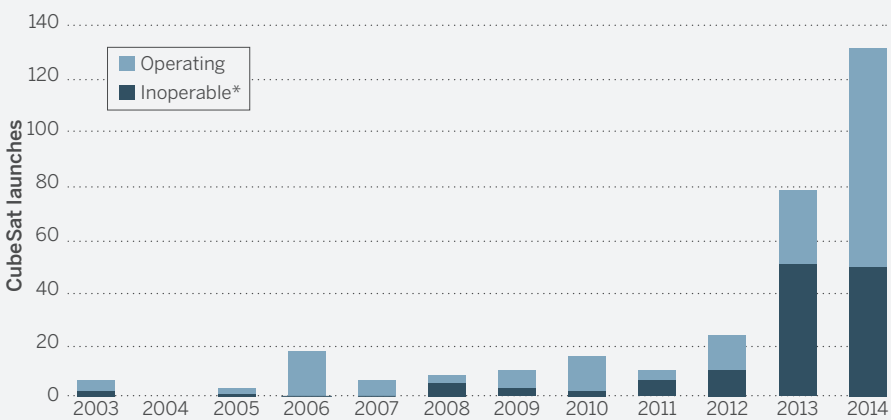
Some scientists think CubeSats can play a role far beyond low-Earth orbit. NASA has begun work on Mars Cube One, twin 6U CubeSats that will hitch a ride on InSight, a Mars lander scheduled for launch next year, and help relay its data back to Earth. Other NASA projects plan to equip CubeSats with solar sails—thin reflective sheets that use sunlight for propulsion—to travel to near-Earth asteroids and the moon. The moon probe, called Lunar Flashlight, would also use its sail to reflect light into the permanently shadowed craters at the moon's poles.

Jonathan McDowell, a researcher at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, who has studied the history of CubeSats from the start, says the small spacecraft are laying siege to the status quo, in which large teams of scientists build big, budget-busting satellites. "You end up with huge satellites that do 10 different things because you have 10 different stakeholders supporting them. That drives up complexity and development time and the cost of failure."

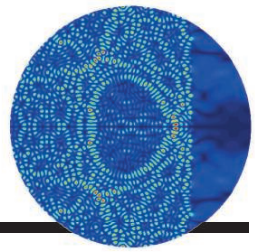
They are also opening space to new participants. According to McDowell, the 29 countries that have launched CubeSats include developing nations such as Vietnam, Peru, and Ecuador. In 2013, a team of high school students in Virginia launched one. Thanks to CubeSats, space is "no longer the domain of the large governments or corporations," says Bruce Yost, deputy manager of the small spacecraft integrated product team at NASA's Ames Research Center in Mountain View, California. He calls it "the democratization of space." ■

## The swarm cometh

In 2014, a record 132 CubeSats were launched—more than in their first decade.




*CubeSats that have reentered the atmosphere, are dead in orbit, or failed to launch. As of 10 March 2015.



PERSPECTIVES

MEDICINE

# Transatlantic lessons in regulation of mitochondrial replacement therapy

The UK has approved MRT for clinical use, but the discussion has just begun in the U.S.

*By* **I. Glenn Cohen,[1]* Julian Savulescu,[2] Eli Y. Adashi[3]**

Mutant mitochondrial DNA (mtDNA) gives rise to a broad range of heritable clinical syndromes (*1*). A cure for those affected remains out of reach (*1*). However, recently developed mitochondrial replacement therapy (MRT) has raised the prospect of disease-free progeny for women carriers (*2*–*4*). Moreover, the feasibility of replacing mutant oocytic or zygotic mtDNA with a donated wild-type counterpart in humans has now been firmly established (*2*–*4*). In the United Kingdom, legislation regulating the clinical application of MRT, now 10 years in the making, has recently been approved by the House of Commons (*5*) and the House of Lords (*6*). The regulatory vetting of MRT in the United States, under way for a year, remains a work in progress (*7*). Here, we compare and contrast the regulatory history of MRT in the United Kingdom and the United States and examine potential lessons learned.

POLICY

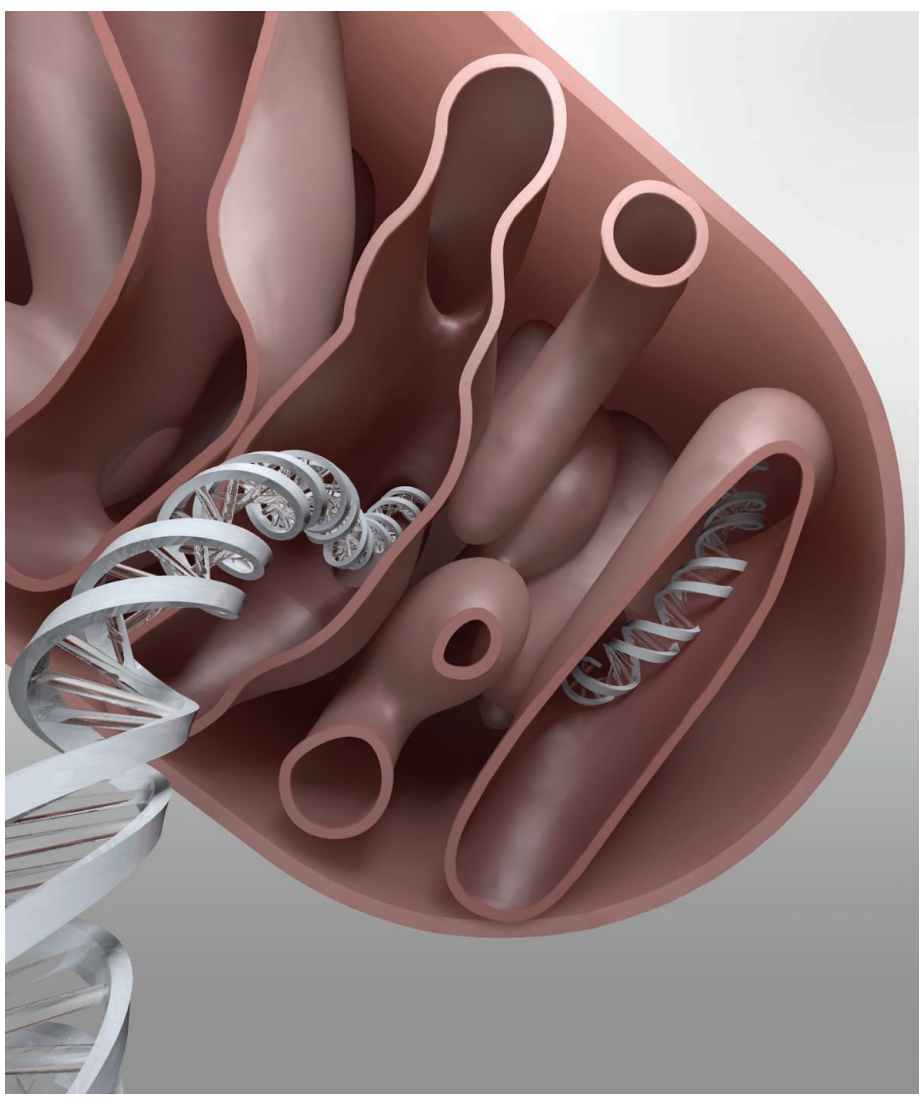


ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

**THE UK REGULATORY EXPERIENCE.** In the United Kingdom, matters relating to reproductive technologies are wholly governed by the Human Fertilisation and Embryology Authority (HFEA), an independent regulatory agency established by Parliament through the Human Fertilisation and Embryology Act of 1990 (HFE Act). Among its responsibilities, the HFEA licenses and monitors human embryo research. In 2005, the HFEA issued a research license to the Newcastle Centre for Mitochondrial Re-

search to "investigate the feasibility of using IVF-based techniques to prevent transmission of mitochondrial disease" (*8*, *9*). At that time, the conduct of clinical trials of MRT was prohibited by the HFE Act. However, in 2008, Parliament amended the law to empower the HFEA to license and monitor clinical trials of MRT subject to subsequent parliamentary approval.

In 2011, the HFEA convened an expert panel to review the "effectiveness and safety of mitochondrial transfer" (*10*). The panel concluded that the "evidence currently available does not suggest that the techniques are unsafe" (*10*). Updated reports issued in 2013 and 2014 reached similar conclusions. In 2012, the HFEA launched a public consultation process on the social and ethical implications of MRT, the outcome of which revealed "general support for permitting mitochondria replacement in the UK, so long as it is safe…and is done so within a regulatory framework" (*11*). The Nuffield Council on Bioethics and the Ethics Committee of the British Medical Association reached similar conclusions. On 3 February 2015, the House of Commons voted to approve the conduct of MRT under the auspices of the HFEA (*5*). The House of Lords followed suit on 24 February. Clinical trials could commence as early as 29 October 2015.

Ethical questions discussed in the United Kingdom revolved around the safety of the procedure for newborns, the matter of parenthood, and the notion that MRT represents germline modification and thus crosses a "red line." Although a full review of the ethical discussions in the United Kingdom is outside the purview of this policy piece, we would point out that several members of Parliament noted that MRT is more about changing a "battery pack" than it is about genetic modification (*5*). Although the term "three-parent" reproduction has been used in the public debate, as per the recent parliamentary vote, mitochondrial donors have not been accorded legal parental status. It follows that the identities of prospective donors are not to be disclosed to "mitochondrial donor-conceived persons" (*5*).

On the matter of germline modification, during the House of Commons and House of Lords debates, some MRT opponents suggested that the legislative action under consideration would violate the European Union Directive on clinical trials that states that "[n]o gene therapy trials may be carried out which result in modifications to the subject's germ line genetic identity." Proponents who won the day argued that HFEA approval of MRT would not violate this directive and/or that the directive did not apply to MRT, but their view is thus far untested in court (*5*).

**THE U.S. REGULATORY EXPERIENCE.** The United States, unlike the United Kingdom, lacks a specific governmental entity whose sole charge it is to regulate reproductive technologies. In the case of MRT, jurisdiction has been asserted by the Food and Drug Administration (FDA) Office of Cellular, Tissue, and Gene Therapies of the Center for Biologics Evaluation and Research, whose task it is to oversee "human cells used in therapy involving the transfer of genetic material by means other than the union of gamete nuclei" (*12*). Under FDA's existing regulations, approval of the therapeutic use of MRT will require the conduct of phased clinical trials pursuant to an Investigational New Drug application (IND) (*12*).

***"The regulatory paths of the United States and the United Kingdom toward MRT have diverged considerably…"***

The FDA has never officially considered or approved early-phase clinical trials of MRT. In fact, its 2 July 2013 draft guidance for industry for the "design of early-phase clinical trials of cellular and gene therapy products" (published for public comment) makes no mention of MRT (*13*). The first tremor in this status quo was felt early in 2014 when the FDA convened the Cellular, Tissue, and Gene Therapies Advisory Committee to discuss "oocyte modification in assisted reproduction for the prevention of transmission of mitochondrial disease" (*7*). Although there was no official conclusion, Reuters quoted the chairman of the committee as summarizing the session in the following way: "Several panelists felt 'there was probably not enough data in animals … to move on to human trials without answering a few additional questions'" (*14*).

The FDA has since commissioned an ad hoc committee of the Institute of Medicine (IOM) to weigh in on the "Ethical and Social Policy Considerations of Novel Techniques for Prevention of Maternal Transmission of Mitochondrial DNA Diseases" (*15*). Two of the planned committee sessions will be open to the public. No further FDA action is expected until the IOM report has been released, which is estimated to be 19 months after the start date of September 2014. In the interim, any relevant INDs submitted would remain on hold.

The recent debate over MRT in the House of Commons and elsewhere went well beyond safety considerations to explore ethical arguments and matters of conscience. In the United States, some of these same issues have been discussed in the popular press, but a full consideration of the ethics and not just the science of MRT in the United States has yet to occur.

**TRANSATLANTIC LESSONS.** The regulatory paths of the United States and the United Kingdom toward MRT have diverged considerably, with the United Kingdom clearly further along in the decision-making process, having approved the clinical use of MRT. Why is there such a transatlantic difference? This is not an academic question, because MRT represents but one of a growing complement of novel reproductive technologies, many of which will require expert regulatory adjudication.

The first transatlantic distinction of note concerns the regulatory agency charged with overseeing extant and leading-edge reproductive technologies. In the United Kingdom, the regulatory adjudication of MRT has been relegated to the HFEA. Absent a comparable U.S. agency, the task of regulating MRT has fallen to the FDA, the mandate of which encompasses all therapeutics. Even the FDA Office of Cellular, Tissue, and Gene Therapies—the designated proximate overseer of MRT—is entrusted with a diverse portfolio of cellular, tissue, and gene therapeutics. It follows that UK regulators, unlike their U.S. counterparts, view MRT as a circumscribed outgrowth of related and highly familiar technologies (e.g., in vitro fertilization) rather than as a therapeutic (*5*). In contrast, U.S. regulators are proceeding on the premise that MRT constitutes a drug and/or a biological product. We posit that the narrow framing of MRT by the HFEA is partially responsible for the expert nature of the UK regulatory paradigm.

The second transatlantic dissimilarity harkens back to the prevailing national values and mores regarding reproduction and especially research involving the derivation and destruction of a human embryo. In the United States, scientific research with human embryos has long been controversial, and federal funding in this area has been constrained. This is in part due to the fact that, in the United States, such research has been tangled in electoral and religious contestation over abortion. No such impediments cloud the regulatory adjudication of MRT in the United Kingdom, where public funding for human embryo–derived stem cell generation and human embryo research

[1]Harvard Law School, Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics, Harvard University, Cambridge, MA, USA. [2]Oxford Uehiro Centre for Practical Ethics, University of Oxford, Oxford, UK. [3]Warren Alpert Medical School, Brown University, Providence, RI, USA.
*Corresponding author: igcohen@law.harvard.edu

through day 14 of development has been legalized.

The third material transatlantic difference concerns the relative weight assigned to public consultation on regulatory issues of substance. The UK public consultation process was an extensive outsourced multimethod (e.g., surveys and workshops) effort on a national scale lasting 6 months. The U.S. regulatory approach has thus far been largely limited to a conversation among experts, with relatively brief sessions open to the public (*7*).

A fourth transatlantic variance revolves around the framing of MRT as a beacon of national scientific prowess. For better or worse, the parliamentary debate has proceeded with an air of national pride. Even those opposed to MRT noted their admiration for the world-class work of the Newcastle group (*2, 5*). We

***"This is not an academic question, because MRT represents but one of a growing complement of novel reproductive technologies, many of which will require expert regulatory adjudication."***

believe that this national sense of pride may have swayed some votes in support of MRT. No such sentiment has been sweeping the United States, even though U.S. scientists have made equally vital contributions to this field of inquiry (*3*, *4*).

**CONCLUSIONS.** This examination of the different approaches taken to the regulation of MRT in the United Kingdom and the United States leads us to reexamine the wisdom of burdening the FDA with the regulatory adjudication of MRT, as opposed to adopting an HFEA-like paradigm (*16*). In the eyes of some, the regulatory oversight of reproductive technologies in the United States leaves much to be desired. Yet others are content with the status quo, in which reproductive technologies are not directly licensed (as in the United Kingdom) but instead are left to what can be characterized as self-regulation by the medical profession and its representative associations. However, with the MRT challenge looming and others not too far behind, it may be time to renew the national conversation as to the rules that should govern this terrain. Understandably, the outcome of such conversation is far from certain. Because some forms of MRT involve embryo destruction, approval in the United States will be embroiled in the prolife/prochoice divide (*17–19*). It remains an open question whether an initiative to reform the regulatory oversight of reproductive technologies in the United States can be realized without capsizing under its own weight and the force of the political winds. ■

## REFERENCES AND NOTES


1. O. Russell, D. Turnbull, *Exp. Cell Res.* **325**, 38 (2014).
2. L. Craven *et al.*, *Nature* **465**, 82 (2010).
3. M. Tachibana *et al.*, *Nature* **493**, 627 (2013).
4. D. Paull *et al.*, *Nature* **493**, 632 (2013).
5. www.parliament.uk, Parliamentary business; Publications & records; Human Fertilisation and Embryology, 3 February 2015; http://www.publications.parliament.uk/pa/cm201415/cmhansrd/cm150203/debtext/150203-0002.htm#15020348000001
6. G. Vogel, Mitochondrial gene therapy passes final U.K. vote, *Science*, ScienceInsider, 24 February 2015, ; http://news.sciencemag.org/biology/2015/02/mitochondrial-gene-therapy-passes-final-u-k-vote10.1126/science.aaa7899
7. U.S. Food and Drug Administration, Advisory Committees, 2014 Meeting Materials, Cellular, Tissue and Gene Therapies Advisory Committee, 25–26 February 2014; www.fda.gov/AdvisoryCommittees/CommitteesMeetingMaterials/BloodVaccinesandOtherBiologics/CellularTissueandGeneTherapiesAdvisoryCommittee/ucm380047.htm
8. Human Fertilisation and Embryology Authority, HFEA grants licence to Newcastle Centre at LIFE for Mitochondrial Research, 8 September 2005; www.hfea.gov.uk/671.html
9. Human Fertilisation and Embryology Authority, Newcastle Fertility Centre at Life, R0152: Pluripotency, reprogramming and mitochondrial biology during early human development, 24 September 2014 ; www.hfea.gov.uk/1564.html
10. Human Fertilisation Embryology Authority, Scientific review of the safety and efficacy of methods to avoid mitochondrial disease through assisted conception (2011); www.hfea.gov.uk/docs/2011-04-18_Mitochondria_review_-_final_report.PDF
11. Human Fertilisation and Embryology Authority, Mitochondria replacement consultation: Advice to Government, March 2013.
12. U.S. Food and Drug Administration, Vaccines, Blood, and Biologics. Letter to Sponsors/Researchers, Human cells used in therapy involving the transfer of genetic material by means other than the union of gamete nuclei, 6 July 2001; www.fda.gov/BiologicsBloodVaccines/SafetyAvailability/ucm105852.htm
13. Department of Health and Human Services, Food and Drug Administration, *Fed. Regist.* **78**, 39736 (2013); www.gpo.gov/fdsys/pkg/FR-2013-07-02/pdf/2013-15797.pdf
14. S. Begley, U.S. FDA weighs evidence on producing 'three-parent' embryos, Reuters, 25 February 2014; http://uk.reuters.com/article/2014/02/25/us-usa-health-ivf-idUKBREA101WL20140225
15. The National Academies, Current Projects System, Ethical and Social Policy Considerations of Novel Techniques for Prevention of Maternal Transmission of Mitochondrial DNA Diseases (2014); www8.nationalacademies.org/cp/projectview.aspx?key=IOM-HSP-14-25.
16. E. Parens, L. P. Knowles, *Hastings Cent. Rep.* **33**, S1 (2003).
17. I. G. Cohen, E. Y. Adashi, *N. Engl. J. Med.* **364**, e48 (2011).
18. C. B. Cohen, *Renewing the Stuff of Life: Stem Cells, Ethics, and Public Policy* (Oxford Univ. Press, Oxford, 2007).
19. P. Toynbee, This isn't about three-parent babies: It's about saving families needless misery, *The Guardian*, 3 February 2015; www.theguardian.com/commentisfree/2015/feb/03/three-parent-babies-families-religious-mps-vote-mitochondrial-replacement


10.1126/science.aaa8153

DEVELOPMENT

# *Aneuploidy and mother's genes*

## A human genetic variant found at high frequency is associated with reduced fertility

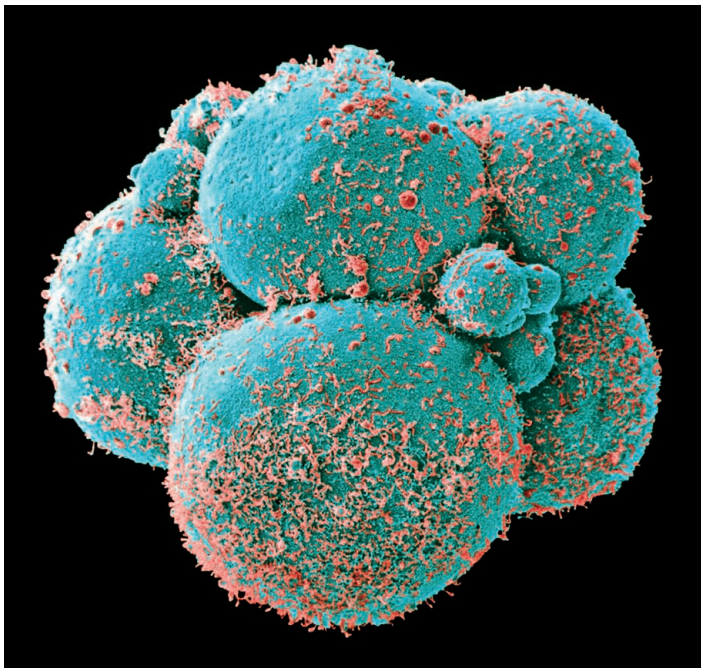
*By* **Samuel H. Vohr** *and* **Richard E. Green**

Biology, ecology, and culture have shaped human genetic variation over thousands of generations. Technology now allows us to know the sequence of our genomes and to act on this knowledge. Which genes did a child inherit from either parent? With the direct-to-consumer genome scan products now available, this question can be answered at the cost of a few hundred U.S. dollars and a few milliliters of spit. Which fertilized embryo is free of genetic and genomic abnormalities? By combining in vitro fertilization, preimplantation genetic screening, and whole-genome scans, this is also now possible to assess (*1*, *2*). But what if genes themselves select potential children? On page 235 of this issue, McCoy *et al.* (*3*) indicate that this may be the case. The authors describe paradoxical results of a genomic study of thousands of preimplantation human embryos and their parents. They turn up a maternal-effect genetic variant that occurs at high frequency in many populations, that was likely under positive selection in our recent past, and that dramatically decreases embryonic viability.

Since its first use in England in 1977, in vitro fertilization has become a mainstream reproductive technique, responsible for 2 to 3% of babies in developed countries (*4*). In vitro fertilization can be coupled with preimplantation genetic screening, wherein one or more genetic assays are performed on cells taken from an early-stage embryo before implantation (see the figure). This combination provides further options to many would-be parents. It allows a sneak peek at the genetic makeup of an embryo and is routinely used to screen for chromosomal abnormalities or disease-causing genetic defects when the parents are known or suspected carriers (*2*).

Recently, preimplantation genetic screen-

ing has been broadened to scan entire genomes of preimplantation embryos by means of whole-genome amplification and genotyping arrays (*5*). This allows for a hypothesis-free and more comprehensive genetic assessment of an early-stage embryo than was previously possible. But it also comes with data quality issues, such as "allelic dropout," in which one or both alleles of a gene fail to be amplified in a whole-genome amplification step that is required when input DNA is limited to a single cell. This technical problem can be mitigated, however, by using genomic data from the parents together with that from the embryo. Using this approach, aneuploidies—the presence of an atypical complement of chromosomes—can be detected at fine resolution. It also becomes possible to determine which of the two parental chromosomes is aneuploid.



**Preimplantation.** A colored scanning electron micrograph of a human embryo at the eight-cell stage, 3 days after fertilization. At this stage, the embryo has not yet implanted in the uterus.

In both in vitro and in vivo fertilizations, aneuploidy is surprisingly common and often leads to early pregnancy loss (*6*). It has long been known that maternal age is positively correlated with aneuploidies (*6*) and that these are typically meiotic aneuploidies (i.e., present from the beginning, in the unfertilized egg of the mother). Against this background of common maternal meiotic aneuploidy, it has been difficult to detect and measure the rates of other types of aneuploidies such as paternal meiotic aneuploidies (present in the sperm that fertilized the egg) or later, mitotic-derived aneuploidies that arose in early cell divisions after fertilization.

To distinguish between the different sources of aneuploidy, McCoy *et al.* analyzed genotype data from sets of embryos and both parents to search for evidence of chromosome gain and loss. Reasoning that losses and gains that exclusively affect maternally inherited chromosomes in the embryo are meiotic in origin, they were able to identify these cases and confirm their known association with maternal age. Because there is a low rate of aneuploidy in sperm, the authors further reasoned that aneuploidies affecting paternally inherited chromosomes can generally be attributed to mitotic errors and focused on those. Tellingly, they found that these paternal aneuploidies often affect multiple chromosomes and that their frequency of occurrence is not affected by the age of the mother. Given that some women produce aneuploid embryos more often than others (*7*) and that the mitotic machinery in early embryogenesis is largely maternally derived, is it possible that mitotic aneuploidies could be affected by the genes of the mother?

Armed with a large cohort to test this hypothesis, McCoy *et al.* searched maternal genomes for variants associated with various types of aneuploidy. They found no association between the mother's genotype and rates of meiotic errors. However, they found a strong genetic association between the mother's genotype and the rate of observed mitotic errors. Surprisingly, the variant most closely associated with high rates of embryonic mitotic errors, the single-nucleotide polymorphism rs2305957, is found at high frequency in populations across the world. Several genes are tightly linked to the high-risk variant, but the authors singled out the *Polo-like Kinase 4* (*PLK4*) as a candidate causal gene because of its known role in the centrosome cycle (*8*), a process whose disruption can cause aneuploidy. (Centrosomes are structures that organize microtubules into the mitotic spindle that orchestrates the separations of duplicated chromosomes during cell division). Underscoring the importance of the maternal genotype in this region, mothers with the high-risk variant contributed fewer later-stage (5-day) embryos, presumably because there is an elevated rate of embryonic mortality associated with these aneuploidies.

How could a genetic variant so strongly associated with reduced fecundity have risen to such high frequency across the world? A possible clue is that this genomic region was previously identified in a scan for positive selection by contrasting genetic variation found in the Neandertal genome with that present in humans today (*9*). The scan has maximum power to identify episodes of positive selection that occurred in the time since human ancestors split from Neandertals but before human population differentiation—that is, between roughly 400,000 years ago and 100,000 years ago. McCoy *et al.* provocatively speculate that a genetic variant that reduces fecundity may provide a selective advantage by obscuring paternity. This interpretation may make sense in a species with low fecundity and heavy parental investment, such as humans.

Around 30% of natural human conceptions do not go to full term (*10*). Armed with reproductive and genetic tools for inquiry, it is now possible to start unraveling this mystery. The results described by McCoy *et al.* are cause for both optimism and uncertainty about the future. If some human genomes carry a legacy of adaptive reduced fecundity, and people now have the means and motivation to select against these variants, will that happen? If the evolutionary driving force was indeed paternity confusion, does it make sense to purposefully select against this variant now that technology also exists to eliminate paternity confusion? What about the prospect of genetic screening or valuation of egg donors based on the genotype-predicted viability of the resulting embryos? Knowledge of the determinants of embryonic development and viability may add a new turn in the complicated trajectory of human evolution. ■

*Department of Biomolecular Engineering, University of California, Santa Cruz, CA 95060, USA.*
*E-mail: ed@soe.ucsc.edu*

IMAGE: YORGOS NIKAS/SCIENCE SOURCE

## REFERENCES

1. M. Bayefsky, *Hastings Cent. Rep.* **45**, 7 (2015).
2. P. R. Brezina, W. H. Kutteh, *BMJ* **350**, g7611 (2015).
3. R. C. McCoy *et al.*, *Science* **348**, 235 (2015).
4. T. Kirby, *Lancet* **376**, 1293 (2010).
5. D. S. Johnson *et al.*, *Hum. Reprod.* **25**, 1066 (2010).
6. T. Hassold, P. Hunt, *Nat. Rev. Genet.* **2**, 280 (2001).
7. J. D. Delhanty, J. C. Harper, A. Ao, A. H. Handyside, R. M. Winston, *Hum. Genet.* **99**, 755 (1997).
8. E. A. Nigg, *Curr. Opin. Cell Biol.* **10**, 776 (1998).
9. R. E. Green *et al.*, *Science* **328**, 710 (2010).
10. A. J. Wilcox *et al.*, *N. Engl. J. Med.* **319**, 189 (1988).

10.1126/science.aab0877

CELL BIOLOGY

# Finding nascent proteins the right home

## A protein complex prevents promiscuous targeting of nascent polypeptides in the cell

*By* **Günter Kramer, D. Lys Guilbride, Bernd Bukau**

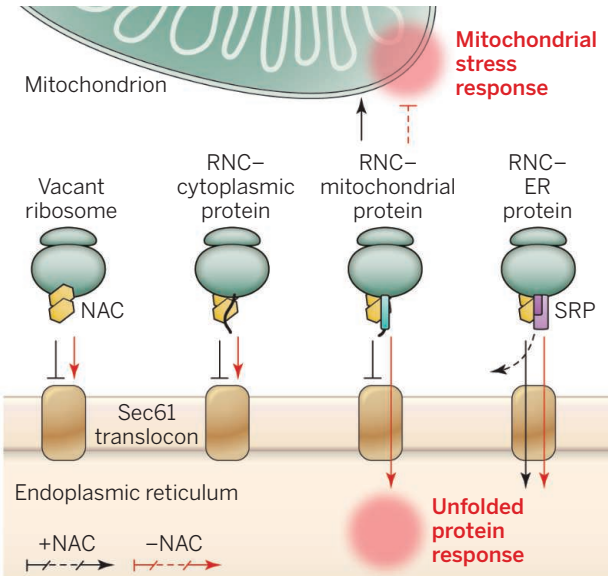
Cells must deliver the thousands of polypeptides they synthesize every minute to various specific subcellular locations. Precisely how this happens has been a topic of intense research, and some controversy, for the past 20 years, as critical components of the translation and translocation machineries—the ribosomes and the signal recognition particle (SRP)—do not confer full target discrimination. On page 201 of this issue, Gamerdinger *et al.* (*1*) elegantly demonstrate how the nascent chain associated complex (NAC) enhances the specificity of the metazoan protein-sorting machinery to provide more discriminatory targeting for newly synthesized proteins in vivo.

For the majority of secretory proteins destined for the endoplasmic reticulum (ER), the targeting process begins cotranslationally, before nascent polypeptide chains are complete. The SRP transiently associates with translating ribosomes, enabling SRP to search for and bind with high affinity to hydrophobic signal sequences on the emerging polypeptide amino termini. Ribosome-bound SRP binds the SRP receptor in the ER membrane for delivery of the selected nascent chains to the Sec61 membrane import translocon, which forms the translocation pore. However, ribosomes can also bind to the ER translocon indiscriminately and with high affinity, independently of SRP. Furthermore, SRP can bind, though less strongly, to nascent chains that do not have ER signal sequences, and these represent the majority of cellular proteins. What, then, ensures target discrimination and fidelity?

In 1994, Wiedmann and co-workers identified NAC (*2*), present in equimolar amounts to ribosomes. NAC transiently associates with the large ribosomal subunit close to the peptide tunnel exit and binds a broad spectrum of ribosome-nascent chain complexes (RNCs) (*3*). In vitro NAC association sterically prevents indiscriminate binding of RNCs to the ER membrane, regardless of whether the nascent chain has a signal sequence or not (*4*, *5*). Addition of SRP, however, specifically rescues ER targeting of signal-containing RNCs (*6*). NAC also prevents SRP association with nonsignal RNCs, which preempts mistargeting of nonsignal proteins to the ER translocon (*7*). NAC binding to the ribosome therefore allows selection of signal-containing proteins by SRP. This led to a model where NAC mediates exclusion of SRP from RNCs of nonsignal proteins, and exclusion of indiscriminate ribosome binding to the ER. Together, these two NAC activities ensure faithful sorting of signal peptide–containing nascent proteins to the ER (*8*).

However, later studies using different experimental settings found no inhibitory effect of NAC on the membrane interaction of isolated RNCs (*9*, *10*). Although these latter results were subsequently proposed to result from the use of nonphysiological concentrations of NAC (*11*), the basis for the discrepancy has remained moot. Embryonic lethality of NAC deletion in flies and worms hampered attempts to verify NAC function in vivo. Studies in yeast, where NAC deletion is nonlethal, showed that NAC coordinates early recruitment of SRP to ribosomes translating ER proteins and modulates SRP substrate selection, but NAC mutation does not result in aberrant or indiscriminate targeting to the ER (*3*, *12*), further clouding understanding of NAC function in vivo.

Gamerdinger *et al.* examine the role of NAC in metazoans using *Caenorhabditis elegans* as a model and present data validating the earlier in vitro findings (*2*). Reducing the expression of NAC in adult worms results in global mistargeting of ribosomes translating non-ER substrates to the ER, without reducing specific SRP-mediated ER targeting (see the figure). NAC depletion also leads to ER-associated degradation (ERAD) of primarily mitochondrial proteins, indicating ER internalization of mistargeted mitochondrial polypeptides. Thus, in the absence of NAC, and independently of SRP, some newly synthesized proteins with mitochondrial targeting sequences override the proofreading activity of the Sec61 translocon, which is thought to fully open only upon specific interactions with ER signal sequences (*13*). Overexpression of NAC, by contrast, hinders SRP binding to translating ribosomes and slows SRP-mediated ER targeting. NAC, therefore, sharpens the specificity of the protein sorting machinery by preventing SRP-independent interactions of RNCs with the Sec61 translocon, and also by eliminating indiscriminate SRP-mediated targeting of RNCs to the ER membrane (*3*). Selectivity of ER targeting, rather than relying purely on SRP and the proofreading function of the Sec61 translocon, instead centers around NAC function at the ribosome. NAC acts to repress RNC interaction with the translocon; SRP overcomes NAC repression by interaction with RNCs containing signal-sequence proteins. Finally, Gamerdinger *et al.* show that NAC depletion in vivo shortens the life span of *C. elegans* and provokes both ER- and mitochondrial-specific stress responses, suggesting that NAC depletion–induced mistargeting profoundly perturbs basic protein




**Sweet Child O´ Mine—where do we go now?** In *C. elegans*, NAC prevents vacant ribosomes and ribosome-nascent chain complexes (RNCs) translating cytoplasmic or mitochondrial proteins from mistargeting to the ER membrane. Upon depletion of NAC activity, the proofreading activity of the Sec61 translocon suffices to prevent cytoplasmic proteins, but not many mitochondrial proteins, from import into the ER.

*Center for Molecular Biology of the University of Heidelberg (ZMBH) and German Cancer Research Center (DKFZ), DKFZ-ZMBH Alliance, Im Neuenheimer Feld 282, D-69120 Heidelberg, Germany. E-mail: bukau@zmbh.uni-heidelberg. All authors contributed equally to this work.*

ILLUSTRATION: P. HUEY/*SCIENCE*

homeostasis and cell physiology.

Precisely how SRP and NAC interact at the ribosome remains unclear. Gamerdinger *et al.* propose that SRP and NAC compete for overlapping binding sites on the ribosomal protein uL23 (*4*), and that NAC binds translating ribosomes, unless an emerging signal sequence provides a selective binding advantage to SRP. Other in vitro data suggest that there is an alternative NAC ribosome binding site near eL31 (*5*, *12*), and that both NAC and SRP concomitantly bind the ribosome (*12*). SRP could quickly scan translating ribosomes irrespective of NAC presence, until an emerging signal sequence triggers strong SRP binding and NAC release. How other ribosome-bound chaperones and enzymes involved in the folding and processing of nascent chains affect the selection of NAC versus SRP also remains unclear. NAC apparently directly influences cotranslational import of proteins into mitochondria in yeast (*14*), possibly explaining induction of the mitochondrial stress response upon NAC depletion seen in the *C. elegans* study.

The in vivo work by Gamerdinger *et al.* establishes and further defines a central process in protein biogenesis for metazoan cells, and corroborates much of the earlier in vitro work done by Wiedmann. Systematic approaches such as proteome-wide interaction profiling of nascent chains are now needed to elucidate the dynamics and interplay of SRP, NAC, and other ribosome-associated factors at the ribosome. Finally, the Deuerling-Wiedmann model (see the figure) of antagonistic "sort and countersort" reflects a recurring principle of check and countercheck common to a number of biological mechanisms. Such systems provide a calibrated equilibrium between two opposing functions that enhances accuracy and efficiency in decision-making processes within living cells. ■

## REFERENCES


1. M. Gamerdinger, M. A. Hanebuth, T. Frickey, E. Deuerling, *Science* **348**, 201 (2015).
2. B. Wiedmann, H. Sakai, T. A. Davis, M. Wiedmann, *Nature* **370**, 434 (1994).
3. M. del Alamo *et al.*, *PLOS Biol.* **9**, e1001100 (2011).
4. R. D. Wegrzyn *et al.*, *J. Boil. Chem.* **281**, 2847 (2006).
5. M. Pech, T. Spreter, R. Beckmann, B. Beatrix, *J. Boil. Chem.* **285**, 19679 (2010).
6. B. Lauring, G. Kreibich, M. Weidmann, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9435 (1995).
7. B. Lauring, H. Sakai, G. Kreibich, M. Wiedmann, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5411 (1995).
8. I. Moller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13425 (1998).
9. A. Neuhof, M. M. Rolls, B. Jungnickel, K. U. Kalies, T. A. Rapoport, *Mol. Biol. Cell* **9**, 103 (1998).
10. D. Raden, R. Gilmore, *Mol. Biol. Cell* **9**, 117 (1998).
11. I. Möller *et al.*, *FEBS Lett.* **441**, 1 (1998).
12. Y. Zhang *et al.*, *Mol. Biol. Cell* **23**, 3027 (2012).
13. B. Jungnickel, T. A. Rapoport, *Cell* **82**, 261 (1995).
14. C. Lesnik, Y. Cohen, A. Atir-Lande, M. Schuldiner, Y. Arava, *Nat. Commun.* **5**, 5711 (2014).


10.1126/science.aab1335

BIOCHEMISTRY

# *Details of destruction, one molecule at a time*

## Protein ubiquitination and destruction by the proteasome is examined at the single-molecule level

*By* **David Komander**

Essential cellular processes, such as cell division, rely on the coordinated destruction of proteins. The predominant means of accomplishing this involves a large cellular machine, the proteasome (*1*). Proteasomal degradation ensues when proteins are modified with ubiquitin, a small protein, that has many different roles (*2*). This tagging involves a carrier protein (an E2 ubiquitin-conjugating enzyme) and a substrate-determining protein (an E3 ligase). For example, during the cell division cycle, a large multiprotein E3 ligase, the anaphase-promoting complex/cyclosome (APC/C), utilizes two E2 enzymes, UBE2C and UBE2S, to target proteins for destruction (*3*). On pages 199 and 200 of this issue, two Research Articles by Lu *et al.* focus on these reactions and illuminate, at the single-molecule level, the process of ubiquitination by APC/C (*4*), as well as the recognition and subsequent destruction of APC/C substrates by proteasomes (*5*). Both studies substantially enrich our knowledge of ubiquitination and degradation, reveal new properties of APC/C and the proteasome, and challenge established concepts about the ubiquitin-proteasome system.

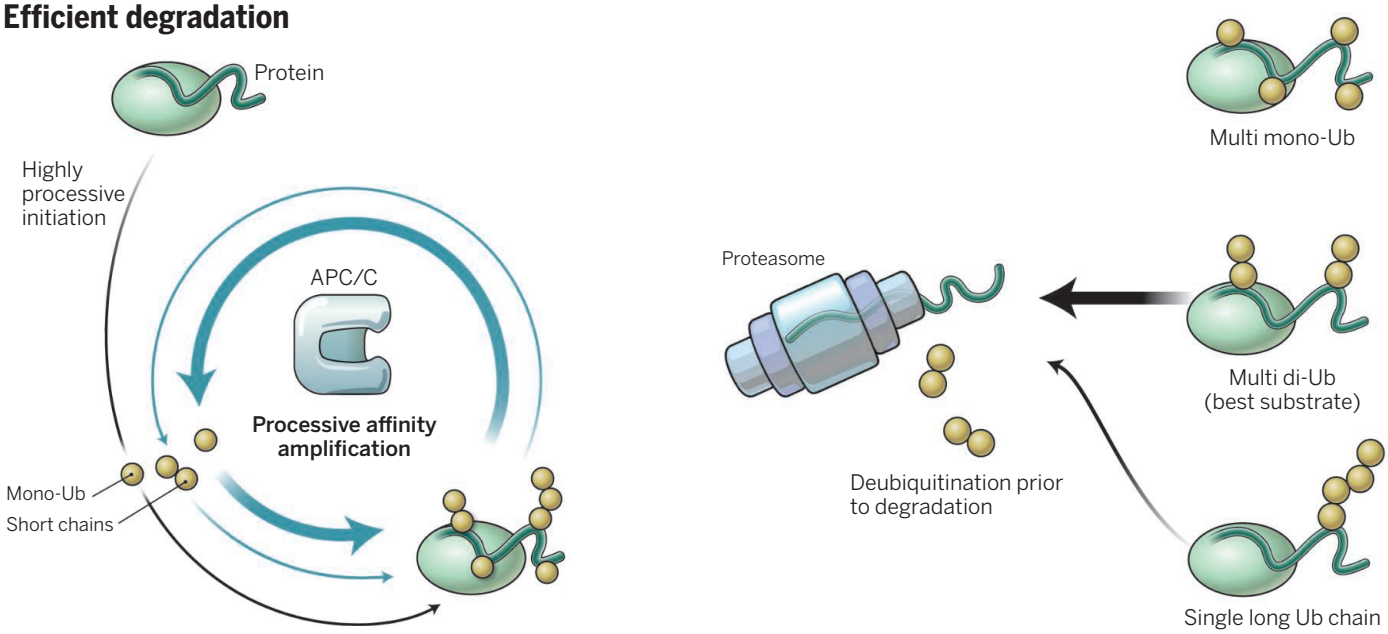
In one study, Lu *et al.* (*4*) immobilized fluorescently labeled APC/C substrates on a glass slide and then exposed the slide to APC/C. Interaction between an APC/C and a substrate, and the subsequent attachment of fluorescent ubiquitin to the substrate, were analyzed using total internal reflection fluorescence (TIRF) microscopy. In the other study, the authors analyzed interactions between immobilized fluorescently labeled proteasomes and a range of substrates containing chains of fluorescent ubiquitin of defined length and composition. In both studies, the fluorescently labeled ubiquitin allowed reporting on the number of ubiquitin moieties attached to the substrates. The approach has enabled a kinetic description of the ubiquitin transfer reaction, revealed determinants for substrate engagement by the proteasome, and delineated the mechanism occurring within the proteasome that couples the initiation of protein degradation with the removal of ubiquitin from the substrate.

The APC/C has a difficult task. It needs to precisely and quickly identify proteins for disposal for the cell cycle to proceed. For this purpose, the APC/C utilizes short, low-complexity recognition sequences in its substrates, which it binds to with the help of coactivators (*3*). Because these sequences are present in roughly one-third of the cell's entire protein repertoire ("proteome"), it is unclear how the APC/C distinguishes potential substrates. Nonetheless, once APC/C selects a substrate, it is ubiquitinated and degraded within minutes.

Lu *et al.* (*4*) find that the first encounter of the APC/C with a substrate leads to efficient mono-, di-, and triubiquitin modification on multiple sites (lysine residues), driven by UBE2C. This was observed both with purified components, but remarkably also with endogenous APC/C activity from cell lysates. Intriguingly, the authors found that after initial ubiquitination, affinity of the substrate for the APC/C is increased. This indicates that there may be unknown ubiquitin receptors on the APC/C, and recent insight into the APC/C from structural biology (*6*) should facilitate their identification. Moreover, based on the observed interaction with ubiquitinated substrates, the authors propose a feedforward-like mechanism called "processive affinity amplification" (see the figure), which ensures that substrates marked for destruction are kept in a ubiquitinated state, while ubiquitinatable decoy un-substrates can be selected against by the APC/C. Despite multiple encounters and higher

> ***"Both studies ... challenge established concepts about the ubiquitin-proteasome system."***

*Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK. E-mail: dk@mrc-lmb.cam.ac.uk*




**Optimizing demise.** APC/C maintains an ubiquitination level suitable for protein destruction (**left**). Those tagged with multiple short chains are superior proteasome substrates (**right**).

affinity for APC/C, preubiquitinated substrates are not efficiently modified beyond the first set of ubiquitins; chain elongation by UBE2S is comparatively inefficient. This suggests that reassociation with the APC/C may serve to simply "top up" ubiquitination to keep substrates primed for proteasomal degradation. However, the latter result is somewhat inconsistent with findings that an APC/C substrate binding event leads to processive ubiquitin amplification by UBE2S (*7*, *8*).

The observed APC/C activity is consistent with previous mass-spectrometry analysis on cyclin B (*9*), a protein involved in pushing the cell through the mitosis phase of the cell cycle, but raises questions about whether a such modified protein is a good proteasome substrate. Importantly, the other study by Lu *et al.* (*5*) compares substrates modified with four monoubiquitins, two diubiquitins, or one tetraubiquitin molecule, showing that the protein modified with two diubiquitins is the superior proteasome substrate. This overturns a paradigm in the ubiquitin-proteasome field stating that a proteasome substrate must harbor a tetraubiquitin chain to be degraded (*10*), but is consistent with substantial structural data (*11*, *12*) that have failed to identify a tetraubiquitin receptor on the proteasome. Correlating proteasome residence times with the number of ubiquitins on a substrate, Lu *et al.* (*5*) reveal cooperative binding for the first three ubiquitin molecules, and linear, stochastic increase in residence time with additional ubiquitin. Although the cooperativity presumably originates from engaging separate receptors for ubiquitin, the subsequent affinity increases are likely due to avidity effects.

Why then are multi-monoubiquitinated proteins not degraded by the proteasome? Indeed, the single-molecule dwell-times of multi-mono- and polyubiquitinated proteins at the proteasome are similar. Lu *et al.* (*5*) find that a ubiquitin chain, irrespective of length, has to be present to activate proteasomal degradation. The first step in degradation is the initiation of translocating a protein into the degradation chamber. This requires adenosine 5′-triphosphate (ATP). Locking the proteasome in an ATP-bound state improved residence times for substrates modified with chains of ubiquitin, suggesting that in this state, the proteasome exposes a chain receptor near the entry channel.

The proteasome does not degrade ubiquitin, but rather recycles it, and for this, it employs several different deubiquitinases (*1*). The deubiquitinase Rpn11 is located near the entry channel (*11, 14*). Closing the circle, Lu *et al.* (*5*) study Rpn11–mediated deubiquitination of substrates at the proteasome, at single-molecule resolution, and show that Rpn11 releases the complete short chains of ubiquitin in a coordinated fashion as the substrate is pulled into the proteasome. This mode of Rpn11 activation has been suggested recently (*13*, *14*).

The findings of Lu *et al.* (*4*, *5*) reveal how protein degradation can be used as a rapid and efficient means to regulate cellular processes. The proteasome has a quantitative requirement, in needing enough ubiquitin to establish sufficient residence times on its ubiquitin receptors, as well as a qualitative requirement in processing only substrates modified with ubiquitin chains, which are essential to initiate degradation. It makes perfect sense that the APC/C focuses on ensuring that the minimal requirements are met for its multiple substrates, as such efficiency is likely important to coordinate the fundamental processes of cell division.

The idea that proteasomal degradation relies on nondiscriminative bulk modification of proteins, rather than single long chains, may rationalize many findings at odds with prior models (*15*), and suggests that single-chain ubiquitination events could be repurposed for alternative and nondegradative processes. ■

## REFERENCES


1. D. Finley, *Annu. Rev. Biochem.* **78**, 477 (2009).
2. D. Komander, M. Rape, *Annu. Rev. Biochem.* **81**, 203 (2012).
3. I. Primorac, A. Musacchio, *J. Cell Biol.* **201**, 177 (2013).
4. Y. Lu, W. Weiping, M. W. Kirschner, *Science* **348**, 1248737 (2015).
5. Y. Lu, B.-H. Lee, R. W. King, D. Finley, M. W. Kirschner, *Science* **348**, 1250834 (2015).
6. L. Chang *et al.*, *Nature* **513**, 388 (2014).
7. K. E. Wickliffe *et al.*, *Cell* **144**, 769 (2011).
8. H.-J. Meyer, M. Rape, *Cell* **157**, 910 (2014).
9. D. S. Kirkpatrick *et al.*, *Nat. Cell Biol.* **8**, 700 (2006).
10. J. S. Thrower *et al.*, *EMBO J.* **19**, 94 (2000).
11. G. C. Lander *et al.*, *Nature* **482**, 186 (2012).
12. E. Sakata *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1479 (2012).
13. E. J. Worden *et al.*, *Nat. Struct. Mol. Biol.* **21**, 220 (2014).
14. G. R. Pathare *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2984 (2014).
15. K. Flick *et al.*, *Nat. Cell Biol.* **6**, 634 (2004).


10.1126/science.aab0931

ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

PHYSICS

# Assembling a complex quantum ensemble

## A cold-atom technique is used to identify hidden constraints in thermodynamic ensembles

*By* **I. B. Spielman**

Statistical mechanics provides a systematic approach for predicting the behavior of systems when only partial information is present. The key assumption underlying all of statistical mechanics is that every allowed configuration of the system, called a microstate, occurs with equal probability. This approach, valid for both classical and quantum systems, requires only minimal additional information for it to be astonishingly predictive. On page 207 of this issue, Langen *et al.* (*1*) describe a system in which this paradigm is violated and conventional statistical mechanics fails; almost all of the usually allowed microstates are inaccessible via the system's dynamics.

In a classical gas, a microstate describes the position and momentum of every constituent particle. In practical experiments, such complete information is unavailable. Instead, we might wish to know aggregate properties, such as the typical number of particles at some position (the spatial distribution function). The problem is how to predict these properties. Each specific distribution can be described equally well by a great many microstates; in statistical mechanics, we select the distribution that is compatible with as many microstates as possible, thereby maximizing the entropy.

Still, slightly more information is required to make useful predictions. For an isolated system, the total internal energy must be conserved. Of all possible microstates, almost none have any given energy. The allowed microstates are now selected from a reduced set called the microcanonical ensemble. The entropy must be maximized at fixed energy, which is achieved by minimizing the free energy. This constrained minimization is performed by introducing a new a parameter called a Lagrange multiplier (in this case, the temperature). The key point is that while the system's total energy is constrained, the energy of any single particle is not.

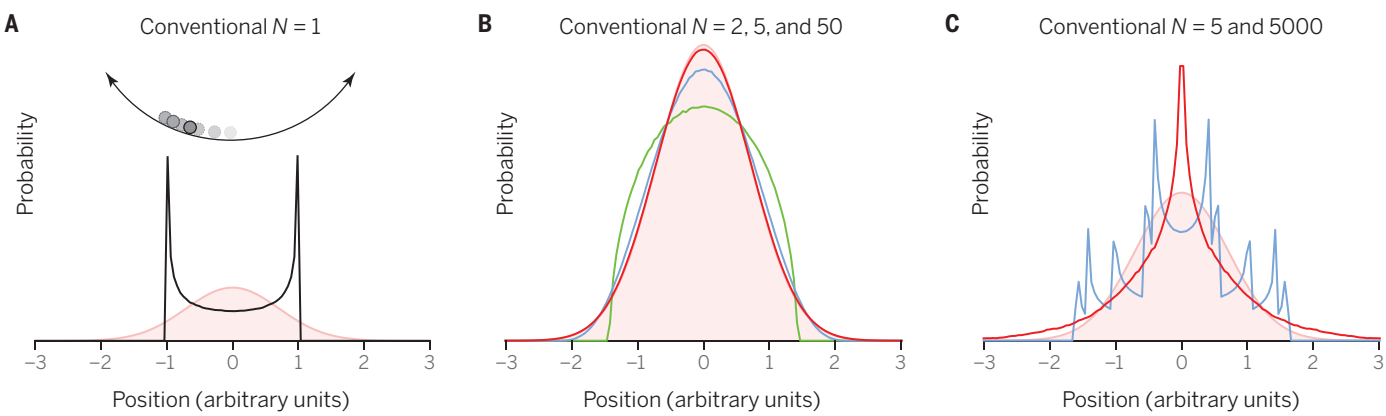
This concept is well illustrated by a single particle moving in a one-dimensional (1D) harmonic potential—for example, describing the oscillatory motion of a pendulum. The spatial distribution of this single particle (see the figure, panel A) is strongly peaked, resulting from the constraint that the dynamics of just one particle is nearly fully determined by its energy. The corresponding distributions for 2, 5, and 50 particles, with the same average per-particle energy but with only the total energy constrained, show that these distributions approach the well-known Maxwell-Boltzmann distribution depicted by the pink curve (see the figure, panel B).

Because of their remarkable environmental isolation, ultracold atoms can have additional conserved quantities. Good examples include the conserved angular momentum of rotating atomic gases in highly symmetric potentials, or the conserved magnetization of bosonic and fermionic systems. These constraints have allowed experimental access to new phases of matter (*2*, *3*). Each of the symmetry-induced conservation laws contributes new terms into the suitable free energy. For example, conserved angular momentum introduces an effective magnetic field, whereas conserved magnetization introduces a spin-dependent chemical potential. In these examples, a particular symmetry of the system adds a single new constraint in total, not per particle.

Langen *et al.* studied a 1D atomic Bose gas in which the standard thermodynamic paradigm fails and the internal dynamics contains hidden conserved quantities with the number of constraints proportional to the system size; such a system is said to be integrable. The so-called generalized Gibbs ensemble (GGE) (*4*) is the thermodynamic

*Joint Quantum Institute (University of Maryland and National Institute of Standards and Technology), Gaithersburg, MD 20899, USA. E-mail: spielman@nist.gov*

## Ordering by number




**Taking up positions.** Distribution functions for particles in a 1D harmonic potential are shown; in each case, the pink curves are the conventional Maxwell-Boltzmann distribution. (**A**) Distribution for a single particle along with sketch of 1D harmonic motion. (**B**) Microcanonical distributions for 2 (green), 5 (blue), and 50 (red) particles. (**C**) Generalized Gibbs ensemble for 5 (blue) and 5000 (red) particles with integrable dynamics.

ensemble that incorporates all of these conserved quantities. Each conserved quantity is associated with its own temperature; a 5000-particle system might require 5000 different temperatures to specify the distribution function.

Our harmonic oscillator example for 5 and 500 particles, where each oscillator separately conserves energy, shows two central points (see the figure, panel C). First, these distributions look nothing like the standard Maxwell-Boltzmann distribution (pink); second, the distributions depend on the system's initial conditions, with peaks set (as in panel A) by the initial energy of each particle. If the per-particle energy had been distributed differently between particles, then the distribution would be peaked at different points. This initial state effect was first observed in a 1D Bose gas (*5*) in which an atomic "Newton's cradle" was created.

***"...cold-atom experiments can be used to study the thermodynamics of systems with internal constraints and can provide the understanding required to predict the outcome of an experiment."***

The physical origin of these conservation laws is simple and depends on the central assumption that each microstate is occupied with equal probability. Usually, physicists argue that collisions between particles are effective in moving the system among all allowed microstates. For 1D systems with binary, local, and elastic collisions (such as between billiard balls), this assumption fails. If we consider just two particles with momentum $p_1$ and $p_2$ colliding at some position $x$, then there are only two possible outcomes that conserve energy and momentum: Either the dynamics continues unchanged ($p_1 \rightarrow p_1$ and $p_2 \rightarrow p_2$), or it continues with the roles of the particles swapped ($p_1 \rightarrow p_2$ and $p_2 \rightarrow p_1$). In either case, the distribution function is unaltered by the collisions.

Langen *et al.* first prepared individual ultracold 1D Bose gases with about 5000 atoms and split these into two parallel decoupled 1D systems, mapping each atom into a coherent superposition of being in the left and the right system. As a result, each of these 1D subsystems was initialized very far from equilibrium. The quantum mechanical phase difference was then observed along the length of the split system as it approached equilibrium. Langen *et al.* found that after some time the split system reached an equilibrium state, but not one predicted by a standard thermodynamic ensemble.

To identify the details of their experiment's underlying ensemble, the authors made a comparison with the higher-order correlations of the measured phase difference (essentially asking how the phase differences at various points are related) up to 10th order. These distributions were found to be qualitatively different from those predicted by standard thermodynamics. Instead, Langen *et al.* considered the resulting quasi-equilibrium in terms of the predictions of the GGE and found that two separate temperatures were required before the ensemble was consistent with their experimental data within the technical noise.

According to the eigenstate thermalization hypothesis (*4*), many-particle quantum systems generically evolve to distribution functions resembling the microcanonical ensemble for local observables, such as the spatial distribution functions discussed here. It requires special effort, such as in Langen *et al.*'s experiment, to find initial states going beyond this.

Langen *et al.* have shown that cold-atom experiments can be used to study the thermodynamics of systems with internal constraints and can provide the understanding required to predict the outcome of an experiment. One might have expected some 5000 constraints to be required to reconcile experiment and theory; experimentally one was insufficient, but two sufficed. In light of the eigenstate thermalization hypothesis, can the impact of all of these constraints ever be felt, or is a small subset always sufficient to describe the outcome of realistic experiments? ■


### REFERENCES AND NOTES

1. T. Langen *et al.*, *Science* **348**, 207 (2015).
2. V. Schweikhard, I. Coddington, P. Engels, V. P. Mogendorff, E. A. Cornell, *Phys. Rev. Lett.* **92**, 040404 (2004).
3. Y. Shin, M. W. Zwierlein, C. H. Schunck, A. Schirotzek, W. Ketterle, *Phys. Rev. Lett.* **97**, 030401 (2006).
4. M. Rigol, V. Dunjko, V. Yurovsky, M. Olshanii, *Phys. Rev. Lett.* **98**, 050405 (2007).
5. T. Kinoshita, T. Wenger, D. S. Weiss, *Nature* **440**, 900 (2006).



### ACKNOWLEDGMENTS

I thank M. Rigol for helpful conceptual discussions. Supported by Army Research Office's atomtronics Multidisciplinary University Research Initiative (MURI), Air Force Office of Scientific Research Quantum Matter MURI, NIST, and the NSF through the Physics Frontier Center at the Joint Quantum Institute.

10.1126/science.aaa8070


GEOLOGY

# An early start for the Panama land bridge

## The land bridge between North and South America formed 10 million years earlier than previously thought

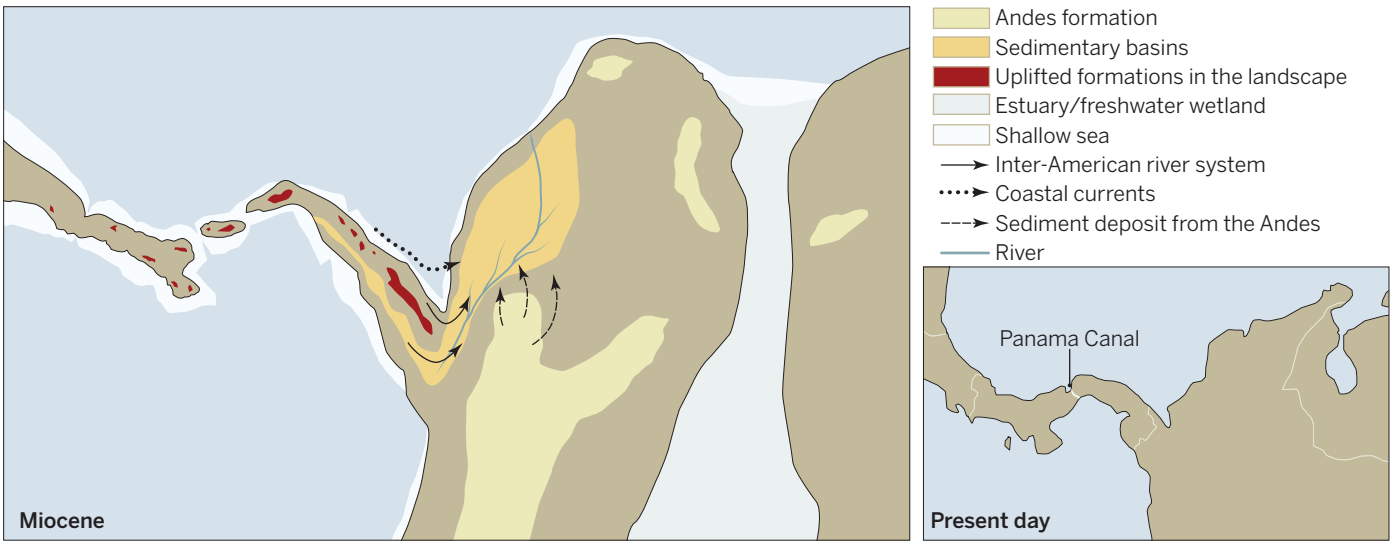
*By* **Carina Hoorn** *and* **Suzette Flantua**

The birth of the Panama land bridge, which connects the Americas, has been associated with one of the biggest biological exchanges in Earth history as numerous species migrated from one continent to the other (*1*). Nevertheless, the timing of formation of the land bridge is still much debated (*2*). On page 226 of this issue, Montes *et al.* (*3*) propose that the Central American Seaway, which separated South and North America, closed about 15 to 13 million years ago, more than 10 million years earlier than previously thought (*4*), with important implications for ocean circulation, climate, and biotic exchange.

The authors studied ancient river deposits found in northern Colombia. Through a combination of geochronological analysis and geological mapping, they determined the age, origin, and pathway of this river. They show that the river system began to flow about 15 to 13 million years ago and could not have originated in any other place than the volcanic arc of Panama. This means that a terrestrial connection must have existed between northern Colombia and the river's source area at this time. A marine connection, the authors state, could only have consisted of transient shallow channels west of today's Panama Canal (see the figure). With this, they contest alternative scenarios (*4*), which suggest that a deep-sea strait separated the Americas until 3 million years ago, preventing massive biotic exchange before this time.

Montes and colleagues first questioned the generally accepted young age of the seaway closure in 2012 (*5*). Before this, however, some geologists and biologists had already

## Connecting the Americas




**Panama then and now.** According to Montes *et al*. (*3*), a Panama land bridge with an "Inter-American" river system formed during the Miocene (15 to 10 million years ago). Support for this land bridge comes from fossils of herald animals that crossed the land bridge in either direction well before 3 million years ago (*6*). (**Inset**) Map of Central America today.

suspected that the land bridge might have emerged earlier than presumed. Paleontologists found that an array of "herald animals" such as camels, peccaries, horses, beardogs, and rhinoceroses were at the bridge by 15 million years ago and that giant sloths, terror birds, and plants (among others) passed the bridge in the northern direction well before 3 million years ago (*6*). Bacon *et al.* have also inferred early exchange of plants between the continents from dated molecular phylogenies (*7*).

Various scenarios were invoked to explain these observations, but none of them were conclusive. A crucial point of controversy has been whether early biotic exchange reflects a terrestrial connection or was a result of random long-distance dispersals (*7*) or perhaps climatic cooling (*8*). With excavations well on their way for renewal of the Panama Canal, a wealth of new paleontological data (*2*) and comparative molecular analyses (*7*) is enforcing the idea of an early land bridge.

Other evidence for an early emergence of the land bridge comes from the magmatic rocks in Panama and northern South America. Changes in the geochronological composition of the Panama volcanic arc at about 24 million years ago mark its collision with the South American island continent (*9*). This collision represented the end of a 2- to 3-thousand-km oceanic voyage of the Caribbean plate on which embryonic Panama was situated (*10*) and formed the prelude to complex deformation at the triple junction of the Caribbean, South American, and Nazca plates. In essence, subduction processes built the fundaments of Panama and ultimately determined the pace of its emergence (*10*).

The formation of the Panama land bridge was not an isolated event. As the Caribbean plate scraped along South America, it reconfigured the northern South American landscape (*9*), resulting in the characteristic trifurcate shape of the northern Andes. Shoaling of the Central American Seaway also coincided with large changes in the South American lowlands corresponding to today's Amazonia, where a mainly fluvial landscape changed into a vast wetland that developed as a consequence of deep subsurface processes, the rising Andes, and related climate change (*11*). Wetland formation, contemporary marine influence in Amazonia, and an extended inland coastline (*12*) are all related to the Caribbean collision and closure of the Central American Seaway.

Pushing back the age of the Panama land bridge by more than 10 million years is noteworthy. However, an early terrestrial connection reconciles an array of seemingly odd results, such as the early dispersal of freshwater fishes between South and Central America (*13*) and shoaling, changes in oceanic currents, and deep-water exchange between the Pacific and Atlantic between 12 and 7 million years ago (*6*, *14*, *15*). These events are difficult to explain if the Central American Seaway did not close until about 3 million years ago.

The geological evidence presented by Montes *et al.* lends support to the idea that the Great American Biological Interchange (GABI) between the Americas started millions of years earlier than commonly assumed. But why did many organisms wait before migrating around 3 million years ago? Molnar (*8*) has suggested that the peak in biological exchange about 3 million years ago was a result of climatic cooling and the formation of savannas suitable for herbivore migration. A full understanding of the dynamics of the GABI will require better knowledge of the early land bridge and its environments. Data are also needed on the existence and duration of any intermittent transoceanic connections elsewhere along the narrow strip of land that separates the Atlantic and the Pacific. ■

*Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1090 GE Amsterdam, Netherlands.*
*E-mail: m.c.hoorn@uva.nl*

### REFERENCES AND NOTES

1. L. G. Marshall *et al.*, *Science* **215**, 1351 (1982).
2. R. Stone, *Science* **341**, 230 (2013).
3. C. Montes *et al.*, *Science* **348**, 226 (2015).
4. A. G. Coates, R. F. Stallard, *Bull. Mar. Sci.* **89**, 801 (2013).
5. C. Montes *et al.*, *GSA Bull.* **124**, 780 (2012).
6. E. G. Leigh *et al.*, *Biol. Rev. Camb. Philos. Soc.* **89**, 148 (2014).
7. C. D. Bacon, A. Mora, W. L. Wagner, C. A. Jaramillo, *Bot. J. Linn. Soc.* **171**, 287 (2013).
8. P. Molnar, *Paleoceanography* **23**, PA2201 (2008).
9. D. W. Farris *et al.*, *Geology* **39**, 1007 (2011).
10. L. M. Boschman *et al.*, *Earth Sci. Rev.* **138**, 102 (2014).
11. C. Hoorn *et al.*, *Science* **330**, 927 (2010).
12. M. Boonstra *et al.*, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **417**, 176 (2015).
13. E. Bermingham, A. P. Martin, *Mol. Ecol.* **7**, 499 (1998).
14. P. Sepulchre *et al.*, *Paleoceanography* **29**, 176 (2014).
15. K. H. Nisancioglu, M. E. Raymo, P. H. Stone, *Paleoceanography* **18**, PA000767 (2003).

### ACKNOWLEDGMENTS

We thank D. van Hinsbergen and A. Antonelli for discussions and valuable comments.

10.1126/science.aab0099

DRUG DEVELOPMENT

# Are trade secrets delaying biosimilars?

## Regulations for approving biologic drugs thwart the market for would-be competitors

By W. Nicholson Price II[1] and Arti K. Rai[2]

On 6 March 2015, the United States Food and Drug Administration (FDA) approved, under the Biologics Price Competition and Innovation Act (BPCIA), a biosimilar of filgrastim (Neupogen), for treating chemotherapy-caused neutropenia (*1*). Although this action represents a step toward cheaper medical treatments, it masks systemic problems. Not only has it taken 5 years since the BPCIA's passage (*2*), but economists estimate that even by 2020, biosimilar competition will reduce consumer prices only modestly (*3*). Why will price competition be so lacking? One key reason is the barrier to competitive entry created by trade secrecy in biologics manufacturing.

Unlike chemically synthesized small-molecule drugs, biologics are isolated from natural sources or produced in cells. They include therapeutic proteins, DNA vaccines, and monoclonal antibodies. As drug companies turn away from the steep challenge of finding breakthrough small-molecule drugs, biologics are becoming increasingly important to modern medicine. Unfortunately, biologics, especially recombinant therapeutic proteins, are extremely expensive, costing consumers and insurance companies almost $100 billion annually in the United States (*4*).

Congress passed the BPCIA as a policy response, creating a pathway for firms to develop biosimilars, the biologic equivalent of generic small-molecule drugs. To be approved, a biosimilar, also called a "follow-on" biologic, must be shown to be highly similar to an already approved biologic (the "innovator" product) in that there are no clinically meaningful differences in safety, purity, and potency.

For purposes of price reduction, the BPCIA is unlikely to be sufficient. According to most estimates, biosimilar prices will be only about 25% less than innovator prices, with the result that total costs to consumers will be reduced by under $25 billion—cumulatively—by 2020. (*3*). Substantial price reduction won't occur even though a dozen best-selling biologics face patent expiry in the United States by 2020.

[1]University of New Hampshire School of Law, Concord, NH 03301, USA. [2]Duke Law School, Durham, NC 27705, USA. E-mail: nicholson.price@law.unh.edu; rai@law.duke.edu

The key hurdle to competitive entry by biosimilar manufacturers, and thus to price reduction, is trade secrecy in the biologics manufacturing process. To understand the barrier this trade secrecy poses, one must understand the interaction of the science of manufacturing biologics with the FDA regulatory process.

Biologics are generated in living cells and their production requires many idiosyncratic choices, including cell-line selection, growth media and conditions, and purification methods (*5*). These factors shape the final biologic, especially with respect to biochemical modifications such as glycosylation (*6*).

Thus, a biologic is sensitive to its production pathway, and changes can alter its effectiveness, stability, and safety (*7*), with potentially serious health implications.

Moreover, at least for complex biologics, the predominant view of the scientific community, shared by the FDA (*8, 9*), is that fundamental knowledge has not advanced to the point where product characteristics can fully be identified and characterized through analysis of the final product. From this point of view, path-dependent definitions—which focus more on how a biologic was made than what it does—are inevitable.

Certainly, views regarding the state of the art in analytical characterization techniques are contested, particularly with respect to simple biologics. Notably, Sandoz—which submitted the filgrastim biosimilar application to the FDA—argued that analytical techniques have advanced to the point where biosimilar manufacturers should frequently be able to rely almost entirely on them (*10*). Nonetheless, even Sandoz's case for the relatively simple biologic filgrastim relied heavily on clinical trials and extensive European patient experience with its product (*11*).

Debate about the FDA's process-based definitions is not new. What has previously gone unnoticed, however, is the manner in which FDA's process-based definitions reinforce the trade secrecy barrier that pervades biologics manufacturing. Innovator manufacturers keep the details of their processes as trade secrets. Thus, would-be competitors, faced with the FDA's process-based definitions, must attempt to reverse-engineer complex and idiosyncratic manufacturing techniques—for example, exactly how the innovator manufacturer generates and purifies its protein. This task ranges from merely expensive to nearly impossible and creates much of the cost barrier for biosimilar entrants. Fundamentally, the FDA definition strengthens trade secrecy and enables innovator firms to potentially create very long-lasting monopolies far longer than the explicit, carefully calibrated monopolies provided by patent law and FDA regulatory exclusivity.

More perniciously, a link between regulation and trade secrecy decreases incentives to generate fundamental knowledge

PHOTO: SIGRID GOMBERT/SCIENCE SOURCE

about innovator biologics themselves. When competitors face major hurdles in fully understanding a biologic's production and characterization, the trade secrecy–protected monopoly on that biologic can continue indefinitely. Would-be biosimilar applicants have incentives to develop stronger analytical tools, but do not have access to the extensive information possessed by innovator firms and the FDA. This lack of fundamental knowledge weighs down the industry, blocking both competition and further innovation.

Regulation could, however, provide a potential solution. The FDA could play an important role in mediating disclosure by originator manufacturers. Even without congressional action, the FDA may be able to offer incentives for disclosure such as accelerated review. Congressional action could provide the FDA the ability to offer other incentives, such as longer exclusivity periods, or to mandate disclosure in limited circumstances.

Additional procedural burdens would be small. Although trade secrecy, particularly in complex areas like biologics manufacturing, often involves tacit knowledge—difficult to codify and thus transfer (*12*)—originator manufacturers must already codify and submit the relevant manufacturing details to the FDA. Disclosure of these regulatory submissions would not only drive competition but would also provide a rich source of knowledge to support additional work, including fundamental research into the science of how to develop and manufacture biologics. ■

## REFERENCES AND NOTES


1. S. Tavernise, A. Pollack, FDA Approves Zarxio, Its First Biosimilar Drug, *New York Times*, 6 March, 2015.
2. B. Falit, S. Singh, T. Brennan, *Health Aff.* **34**, 294 (2015).
3. H. Grabowski, R. Guha, M. Salgado, *Health Aff.* **33**, 1048 (2014).
4. *The Global Use of Medicines: Outlook Through 2016* (IMS Institute for Healthcare Informatics, 2012).
5. S. Ozturk, W.-S. Hu, *Cell Culture Technology for Pharmaceutical and Cell-Based Therapies* (CRC Press, Boca Raton, FL, 2005).
6. H. Li, M. d'Anjou, *Curr. Opin. Biotechnol.* **20**, 678 (2009).
7. B. Sharma, *Biotechnol. Adv.* **25**, 325 (2007).
8. U.S. Food and Drug Administration, Scientific Considerations in Demonstrating Biosimilarity to a Reference Product (2012); www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291128.pdf.
9. U.S. Food and Drug Administration, Quality Considerations in Demonstrating Biosimilarity to a Reference Protein Product (2012); www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291134.pdf.
10. M. McCamish, G. Woollett, *Clin. Pharmacol. Ther.* **90**, 405 (2011).
11. M. McCamish, Global Health Biopharm and Oncology Injectables Development, Sandoz, Presentation to the Oncologic Drugs Advisory Committee, 7 January 2015.
12. D. J. Teece, *Economic Performance and the Theory of the Firm* (Edward Elgar, Cheltenham, UK, 1998).


10.1126/science.aab1684

OPTICS

# A chaotic approach clears up imaging

## A laser that emits bright, incoherent light provides an ideal light source for imaging

*By* **Harald G. L. Schwefel**[1] *and* **Hakan E. Türeci**[2]

Lasers appear to be ideal light sources for a variety of projection and imaging systems because of their spectral brightness and their ability to produce a beam of light that can be tightly collimated to travel long distances. Lasers owe these extraordinary properties to a quality called coherence. Yet, lasers are not widely used in imaging and projection applications, because the coherence of laser light is just too extreme. Spatiotemporal coherence of the imaging source leads to artifacts such as speckle, caused by the uncontrolled scattering of laser light and multipath interference that degrade the image considerably. Redding *et al.* (*1*) now report how a semiconductor laser based on a chaotic cavity can offer a "compact" solution to this problem. The availability of such low-cost, on-chip semiconductor lasers and the possibility to electrically modulate them make such lasers attractive light sources for a variety of applications, ranging from compact projectors to optical coherence tomography.

Most modern-day sources for imaging and projection are low-coherence sources, either thermal sources or light-emitting diodes (LEDs). However, their low power is an obstacle for application in high-speed imaging and wide-area projection. The ideal illumination source would combine high output power of a laser with the lower spatiotemporal coherence of an LED.

Chaotic microcavity lasers can fill this gap. In these lasers, the feedback is provided by the specular reflection of rays at the boundary of a micrometer-sized dielectric body (see the figure, panel A). This principle for trapping light rays is the same as in the better-known whispering-gallery resonators, except that because of the deformation of the boundary from a rotationally symmetric shape, the motion of light rays in chaotic microcavities exhibits chaos, similar to the dynamics of a billiard ball bouncing in a deformed pool table.

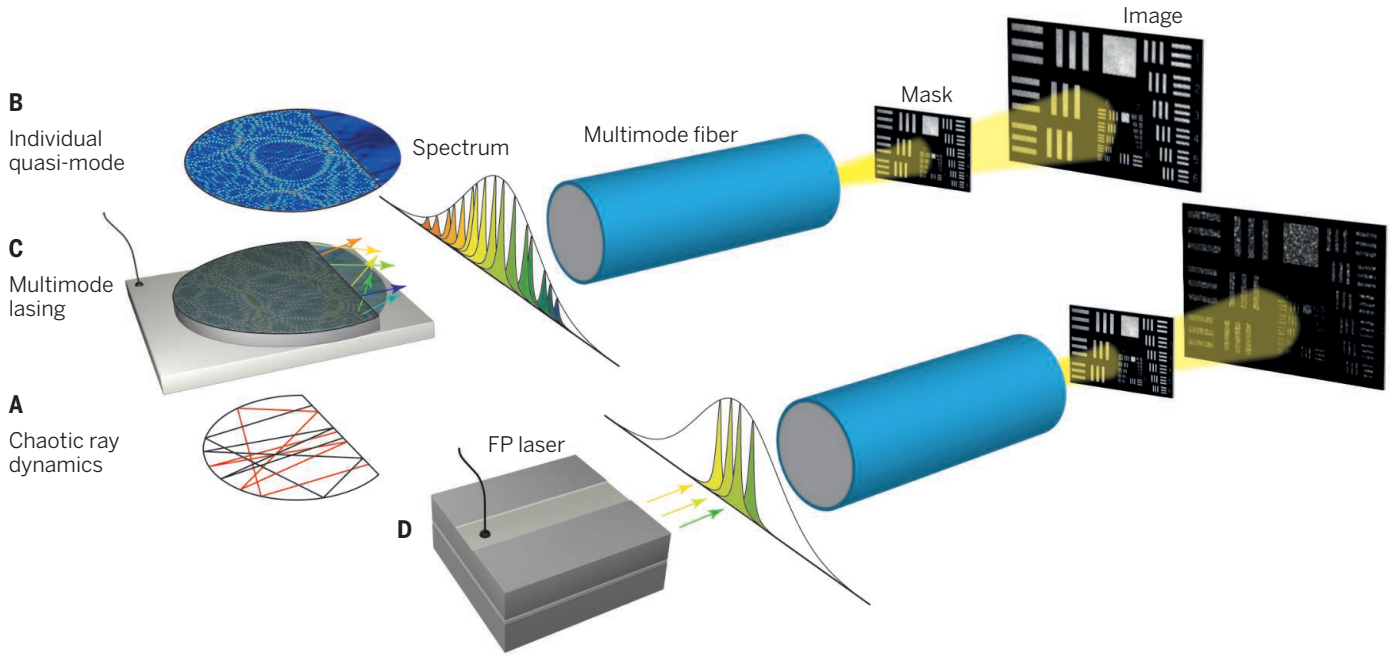
Such chaotic resonators, introduced as model systems to study quantum manisfestations of classical chaos (*2*), have spurred the extension of optical resonator theory to complex resonators (*3*). The findings of Redding *et al.* represent a beautiful example of basic science ultimately providing an elegant solution to a technologically relevant problem. In that sense, the chaotic laser has been a "solution looking for a problem."

The key idea (*4*, *5*) is to use a highly multimode laser source to partially suppress coherence. In the proof-of-principle demonstration, Redding *et al.* pumped the laser cavity to oscillate in about 1000 modes. Usually a vast amount of engineering and fabrication effort is spent in designing robust single-mode sources for various applications such as spectroscopy and optical communications. Thus, getting a laser to oscillate in many modes may appear to be a simple matter. However, for a compact laser source like the one used in (*1*) to lase in not just a few but in hundreds of modes, the laser cavity must create a distribution of laser thresholds that is as narrow as possible, so that the number of lasing modes can be maximized before the useful range of the injection current is reached.

> ***"The key idea is to use a highly multimode laser source to partially suppress coherence."***

There are three critical requirements to accomplish this goal. The first is a large cavity so that the mode density is high, and as a consequence, a large number of modes are within the gain bandwidth of the active medium. Second, these modes must have comparable lifetimes. Both of these requirements can be satisfied with certain designs of chaotic (*1*) and random microcavities (*4*). Redding *et al.* opted for the former because of its simplicity of fabrication and lower in-

[1]*Max Planck Institute for the Science of Light, 91058 Erlangen, Germany and Department of Physics, University of Otago, Dunedin, New Zealand.* [2]*Department of Electrical Engineering, Princeton University. Princeton, NJ 08544 USA. E-mail: harald.schwefel@mpl.mpg.de; tureci@princeton.edu*




**Imaging without the speckle.** The D-shaped resonator of Redding *et al*. emits thousands of modes of slightly different wavelengths that are not phase-coherent, but deliver high power. (**A**) The ray motion in such a resonator is chaotic. The individual linear quasimodes [one example is shown in (**B**)] display quasirandom fluctuations above a uniform background. These patterns may contain additional structure that arise from "scars" (*14*) of periodic ray orbits. (**C**) For above-threshold operation, this device displays multimode emission from hundreds of modes, a few of which are shown in the spectrum of emission. The total emission has substantially reduced phase coherence, resulting in an image for which speckle formation is suppressed, over and above that of an edge-emitting (Fabry-Perot, FP) laser (**D**) that lases in a few modes.

trinsic losses. The third and far more subtle requirement is to suppress spatial hole burning as much as possible. In homogeneously broadened sources, once a mode turns on, it depletes the gain where it has high intensity and suppresses other modes with similar spatial patterns. A lasing mode may never turn on because of such interactions (*6*), thus suppressing highly multimode lasing.

To narrow down the design space, Redding *et al*. resorted to a recently developed theoretical framework, the steady-state ab initio laser theory (SALT) (*7*). This approach was originally developed to address spatial hole burning interactions in a way that provides direct physical insight and is computationally efficient. SALT accurately describes characteristics of lasing in a variety of spatially complex nanostructured photonic media (*6*, *8*–*10*).

Not all chaotic laser cavities are alike, and the "degree of chaoticity" can vary. Only a few well-studied cavity shapes are known to be maximally chaotic. For a generic cavity shape, the phase space for ray dynamics (a map of allowed trajectories) is neither fully regular ("integrable") nor fully chaotic ("ergodic"), but somewhere in between ("mixed"), with certain initial conditions resulting in quasi-integrable ray trajectories and others in quasi-chaotic, phase-space filling motion. It is a general expectation of quantum chaos theory that in the semiclassical limit, where the wavelength of lasing is much smaller than the size of the cavities, the resulting wave patterns of an ergodic cavity are pseudo-random, filling the entire cavity uniformly up to Gaussian fluctuation for the field amplitude and random phases (*11*).

Redding *et al*. hypothesized that a maximally chaotic cavity should satisfy all three requirements set forth for maximally multimode lasing. Carrying out numerical calculations with SALT, they show that a candidate microcavity is D-shaped. Such a shape is known to be chaotic (*12*), and the maximally chaotic phase space is achieved with a segment cut at exactly half the radius. The individual modes of such a cavity (see the figure, panel B) are fully delocalized over the entire cavity area, suppressing spatial hole burning interactions. The resulting highly multimode laser is demonstrated to have substantially reduced speckle noise in imaging a test pattern (see the figure, panel C), as compared to a ridge waveguide Fabry-Perot cavity (see the figure, panel D).

Finding microcavity designs to maximize the number of lasing modes in a given pump power range is a fascinating theoretical problem. The mathematical complexity derives from three aspects. First, the phase space is generically nonintegrable for ray dynamics in the semiclassical limit (panel A). Second, the resulting modal patterns are spatially complex in a driven and dissipative wave system (panel B). Third, there is competition between the modes to monopolize the gain medium in which they reside (panel C). Even when the ray dynamics is maximally chaotic, however, the resulting spatial mode patterns can have structure. Well-known examples include "quantum scars" (*13*), fascinating remnants of quantum interference in the semiclassical limit, that were a central theme in quantum chaos theory in 1990s and first observed in microcavity lasers in the early 2000s (*14*). Another dimension of this problem is that a steady state, explicitly assumed in this work, may not exist. It is to be expected that this optimization problem may offer a rich playground for theoretical physicists and photonic engineers at the crossroads of quantum chaos theory, nonlinear dynamics, and nonlinear optics, with great potential benefits for future novel sources with engineered coherence that may provide the next generation of battery-powered handheld or wearable personal projectors. ■

## REFERENCES AND NOTES


1. B. Redding *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 1304 (2015).
2. J. U. Nöckel, A. D. Stone, *Nature* **385**, 45 (1997).
3. H. Cao, J. Wiersig, *Rev. Mod. Phys.* **87**, 61 (2015).
4. B. Redding *et al*., *Nat. Photonics* **6**, 355 (2012).
5. M. Nixon *et al*., *Opt. Lett.* **38**, 3858 (2013).
6. L. Ge, O. Malik, H. E. Türeci, *Nat. Photonics* **8**, 871 (2014).
7. H. E. Türeci *et al*., *Phys. Rev. A* **74**, 043822 (2006).
8. H. E. Türeci *et al*., *Science* **320**, 643 (2008).
9. S. L. Chua *et al*., *Opt. Express* **19**, 1539 (2011).
10. P. Stano, P. Jacquod, *Nat. Photonics* **7**, 66 (2013).
11. M. V. Berry, *J. Phys. Math. Gen.* **10**, 2083 (1977).
12. V. Doya *et al*., *Opt. Lett.* **26**, 872 (2001).
13. E. J. Heller, *Phys. Rev. Lett.* **53**, 1515 (1984).
14. N. B. Rex *et al*., *Phys. Rev. Lett.* **88**, 094102 (2002).


10.1126/science.aaa7409

BOOKS *et al.*

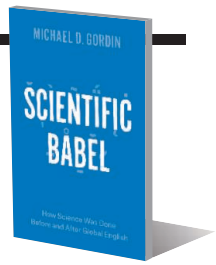
LINGUISTICS

# *Speaking of science*

## How English became the language of science

*By* **Lynn K. Nyhart**

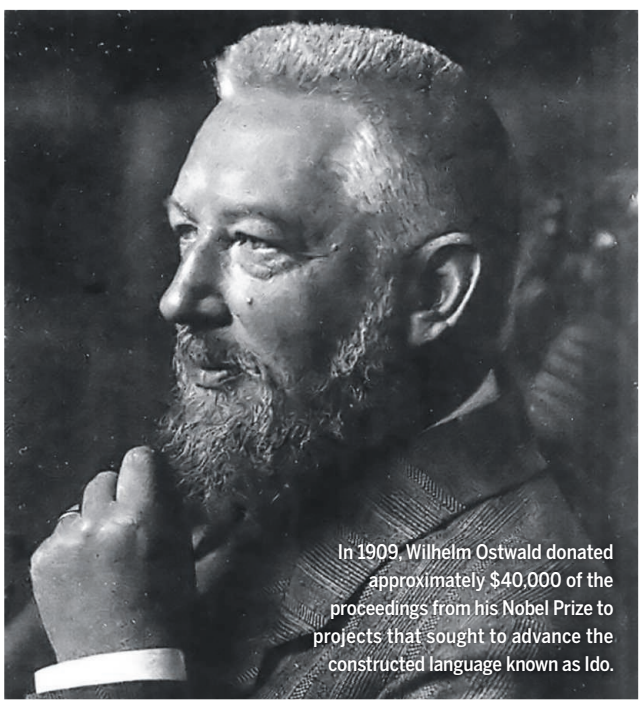
**Scientific Babel**
**How Science Was Done Before and After Global English**
*Michael D. Gordin*
University of Chicago Press, 2015. 421 pp.




"You are able to read this sentence." Surely this assertion, which opens *Scientific Babel*, is self-evident. But as Michael Gordin shows in his sweeping historical account of the languages used by scientists, a profound question lurks just under its surface. How is it that readers interested in science today—those likely to pick up this book—are virtually all able to read English? One might imagine that the United States has simply reaped the benefits of the British empire's earlier aggressive colonization efforts and used its own more recent dominance to bend the scientific world to "global English." But this is not Gordin's take on the story; in fact, he maintains that nothing was inevitable about the outcome.

German, French, and English were considered the "big three" scientific languages in the 19th and early 20th centuries. To be sure, even in the late 19th century, more scientific publications appeared in English than in any other language, but the others were close. German even surpassed English and French briefly in the 1910s and 1920s (*1*). As Gordin shows, however, the geopolitical consequences of the two World Wars caused the use of German in scientific communications to plummet, further facilitating the rise of English.

For Gordin—an expert on the histories of chemistry, Russian science, and the Cold War—the communication problems faced by Russian scientists in the 19th century were especially poignant. As an example, Gordin offers the story of the chemist Dmitri Mendeleev, whose formulation of the periodicity of the chemical elements was cast into doubt when an assistant substituted the word "phased" for "periodic" in a German translation of his research. Moreover, in an age when important new ideas were often published first in books, publishers were reluctant to risk the expense of translating a book without the support of local scientists. But how were the scientists to know whether to offer their support if they couldn't read the original language? (Language barriers were also an issue for Japanese scientists in the early 20th century, although Gordin laments that linguistic incapacity prevented him from including a detailed discussion of their struggles.)



In 1909, Wilhelm Ostwald donated approximately $40,000 of the proceedings from his Nobel Prize to projects that sought to advance the constructed language known as Ido.

In the late 19th century, schemes abounded for a new "vehicular" or "auxiliary" language—one that was native to no one and neutral to all, easy to learn and use. Such a language seemed especially fitting for science, because scientific knowledge was itself presumed to be neutral and universal. Thus arose a series of constructed languages. First came Volapük, a language with roots in English and German. Volapük swirled into vogue for a few years in the late 1880s but was quickly succeeded by Esperanto, a language with Romance, Germanic, and Slavic influences. Although Esperanto would become the most successful constructed language (some children of mixed linguistic parentage even learned it as a native language), its use in science was challenged in 1907, in favor of a modified version called Ido. But with the onset of World War I, Ido's most vocal champion, the physical chemist Wilhelm Ostwald, would drop the idea of a neutral universal language in favor of "Weltdeutsch" (World-German), a language that was not at all neutral—nor, in the end, successful.

During the Cold War, Americans and Russians were eager to learn what scientific advances were buried in the opposition's technical literature. Both governments invested heavily in machine translation, a novel solution enabled by the advent of the computer. But machine translation required much coding and testing at the front end, before any actual translations could result. More immediate action was needed. Whereas the Russians ramped up their abstracting from foreign journals to an unprecedented level, the Americans took another approach: direct, cover-to-cover translation of Soviet journals. The story of how the profit-oriented American approach worked (at the expense of the translators) is one of the many gems of this book.

*Scientific Babel* offers a close-to-the-ground perspective that reveals how short-lived language movements, translation efforts, and publication schemes have interacted with global geopolitics to produce larger historical trends. This emphasis on historical contingencies invites us to reconsider the dominance of English in science today and may ruffle the complacency of those who imagine that it is permanent.

*The reviewer is in the Department of the History of Science at the University of Wisconsin–Madison, Madison, WI 53706, USA. E-mail: lknyhart@wisc.edu*

### REFERENCES AND NOTES

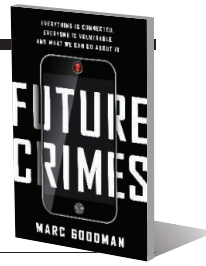
1. Gordin draws heavily on a graph published by Ulrich Ammon (*2*), reproducing it on p. 6 and referring back to it repeatedly.
2. U. Ammon, *Appl. Linguistics Rev.* **3**, 333 (2012).

10.1126/science.aaa6730

PHOTO: WIKICOMMONS

IMAGE: MAGNILION/ISTOCKPHOTO

**Future Crimes**
**Everything Is Connected, Everyone Is Vulnerable, and What We Can Do About It**
*Marc Goodman*
Doubleday, 2015. 462 pp.




TECHNOLOGY

# *Cyber crime 2.0*

## In a world where everything is connected, how will we balance security and privacy?

*By* **Dov Greenbaum**

In *Future Crimes*, author Marc Goodman takes the reader on a well-researched whirlwind tour of Internet-based crime. According to Goodman, the shadowy "Crime Inc." (his catch-all term for all that is malicious on the Internet) is a virtually unstoppable force that, combined with the trend toward a global Internet of Things, leaves us more susceptible to criminal activity than ever.

Although we may legitimately fear Big Brother, and particularly the National Security Agency, which seems to be able to surveil us as it pleases, Goodman reminds us that we must also remain vigilant against criminal enterprises that trade in our identities, credit card numbers, and other personal and private information. According to Goodman, we are very much complicit in the vulnerable position in which we find ourselves. Goaded on by free software services, we wantonly share our most intimate and private information online for others to gather and sell, ignoring the maxim "If you're not paying for it, you're not the customer, you're the product."

Alarmingly, it turns out that all this time online hasn't made us any less gullible. According to Goodman, even sophisticated users can be duped into believing falsified information presented on our trusted devices. In one example, Goodman recounts how, in August 2000, a 23-year-old community college student successfully manipulated the stock market by posting a fake press release about an investigation into a Nasdaq-traded company to an Internet newswire. Investors lost over a hundred million dollars within minutes when the press release went viral.

The title of the book is somewhat of a misnomer. In addition to discussing the future of Internet crime, Goodman provides many engaging examples of past and current online criminal activities. In many of these examples, Goodman describes how bad actors have exploited a number of everyday devices. For example, in one relatively benign case, hackers commandeered webcams, alarms, computer peripherals, and even refrigerators to send out hundreds of thousands of malicious e-mail messages. Goodman instills a fear of imminent disaster by encouraging us to consider what could have happened had the devices been part of more critical infrastructure, such as Internet-enabled artificial limbs, automobiles, avionics, or advanced traffic control systems.

There are numerous hardware and/or software implementations that could be employed by the typical layperson to avoid becoming a victim of the criminal capers described in the book, including, at minimum, encrypting our e-mail, keeping current with security patches and antivirus software, and employing robust aliases. However, Goodman glosses over many of these tools. Perhaps this is because the benefits associated with these technologies come at a relatively high price, requiring users to sacrifice conveniences such as easy-to-remember passwords and user-friendly, but less rigorous, security protocols.

For all the good that security-promoting technologies have to offer, they can't seem to shake negative connotations. The very use of encryption technologies, for example, has been viewed by some as incriminating (*1*). Recently, the United Kingdom's Prime Minister David Cameron threatened to shut down popular messaging applications that employ encryption, including Facebook's WhatsApp and Apple's iMessage, arguing that these sites represent safe havens for terrorist communications (*2*). Such a knee-jerk reaction to established and wildly popular technologies would harm freedom of expression and would likely prove ineffective. Not only could terrorists and criminals turn to a myriad of alternative solutions to encrypt their communications, there is also no indication that being able to intercept encrypted messages is of any value in thwarting terrorism (*3*). Such governmental efforts tend to create additional privacy concerns for the general public, can disincentivize innovation in these and related areas of technology, and can result in additional costs—for example, those associated with developing compliance technologies, which are then passed on to the consumer.

There is no easy solution for finding the balance between security and privacy. The best we can hope for is to encourage discussions early on in the development of new technologies such that reasonable concerns are adequately addressed without impeding innovation. Books like *Future Crimes* will be helpful in starting these conversations.

*The reviewer is at the Zvi Meitar Institute for Legal Implications of Emerging Technologies, Radzyner Law School, Interdisciplinary Center, Herzliya, Israel; and Yale University, Department of Molecular Biophysics and Biochemistry, New Haven, CT, USA. E-mail: dov.greenbaum@yale.edu*

### REFERENCES

1. N. Kayyali, K. Rodriguez, Security is not a crime—Unless you're an anarchist, Electronic Frontier Foundation, 16 January 2015; https://www.eff.org/deeplinks/2015/01/security-not-crime-unless-youre-anarchist.
2. C. Bennet, Britain could ban Snapchat, WhatsApp, *The Hill*, 13 January 2015; http://thehill.com/policy/cybersecurity/229307-britain-could-ban-snapchat-whatsapp.
3. Privacy and Civil Liberties Oversight Board, *Report on the Surveillance Program Operated Pursuant to Section 702 of the Foreign Intelligence Surveillance Act*, 2 July 2014; www.pclob.gov/library/702-Report.pdf.

10.1126/science.aaa7298

LETTERS

*Edited by* **Jennifer Sills**

# A new privacy debate

A STEADY STREAM of large-scale data breaches has focused attention on privacy and led to calls for anonymity, especially for collections of sensitive health data. Meanwhile, recent research has demonstrated—again—that true anonymization of an individual's data is virtually impossible ("Unique in the shopping mall: On the reidentifiability of credit card metadata," Y.-A. de Montjoye *et al.*, Reports, special section on The End of Privacy, 30 January, p. 536). Any policy focused on protecting patient privacy via anonymization will render crucial data useless for clinical and public health research, negate billions of dollars in data infrastructure and analysis investments, and cause real harm by slowing the pace of medical progress. Privacy policies geared to exceptions instead of the norm, or that ignore the breadth and diversity of the many fields using identifiable data practice, will not be efficient.

We can protect individual privacy without sacrificing the potentially transformative insights that large collections of personally identifiable data provide. First, we must acknowledge that relying on anonymization algorithms to scrub our personal information from these data resources is not currently a viable solution. These methods can be circumvented by individuals both internal and external to the organization and anonymization process. Moreover, while failing to protect our identities, they also distort much of the information needed to make the data useful for fields such as medicine and public health.

Second, it is crucial that policies are sensitive to the different intents and practices of the many diverse fields using identifiable data. For example, the analysis of personally identifiable data in clinical and public health research is typically designed to statistically aggregate and compare large groups of people. Consider, for example, a statistical model built to predict a patient's likelihood of responding to cancer treatment. Such models are built by combining identifiable patient information from large numbers of individuals. The output of this research, however, is only valuable if it is generalizable beyond a single individual—de facto anonymization. These life-saving methods, therefore, can be used in practice without risking exposure of personally identifiable information. Sweeping policies that prevent these sorts of ethical and proper use cases risk derailing entire fields, such as public health and medicine.

Third, we must develop a stronger culture of individual participation along with greater transparency in activities that use individually identifiable data. In the health care field, for example, the Patient-Centered Outcomes Research Initiative (PCORI) (*1*, *2*), Agency for Healthcare and Research Quality (AHRQ) (*3*), and Institute of Medicine (IOM) (*4*) have been helping develop new policies regarding patient engagement and research dissemination, which are substantially changing the landscape regarding patient data. These funding initiatives should be expanded, and successful models should be widely shared.

Our debate, therefore, should not be focused on the efforts that protect privacy via anonymization. While research in that direction should continue, we must recognize that anonymization as a precondition to storing or collecting personal information is not a viable policy solution. Instead, we should focus our attention toward requiring safeguards on improper storage, distribution, or exploitation of personal data, and on developing a culture of trust and transparency surrounding the use of such data resources.

***Anne-Marie Meyer[1]* and David Gotz[2]***

[1]Gillings School of Global Public Health and Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. [2]Department of Information Science and Carolina Health Informatics Program, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA.

*Corresponding author. E-mail: ameyer@unc.edu

**REFERENCES**

1. D. Hickam *et al.*, Eds., "The PCORI methodology report" (PCORI, 2013); www.pcori.org/assets/2013/11/PCORI-Methodology-Report.pdf.
2. PCORI, What We Mean By Engagement (www.pcori.org/content/what-we-mean-engagement).
3. Agency for Healthcare Research and Quality, "Findings and lessons from the enabling patient-centered care through health IT grant initiative" (Westat Under Contract No. HHSA 290200900023I, AHRQ Publication No. 13-0011-EF, Rockville, MD, 2013).
4. L. Olsen, R. S. Saunders, J. M. McGinnis, Eds., "Patients charting the course: Citizen engagement in the learning health system" (The Learning Health System Series, Institute of Medicine, 2011); www.iom.edu/Reports/2011/Patients-Charting-the-Course-Citizen-Engagement-in-the-Learning-Health-System-Workshop-Summary.aspx.

# Assessing data intrusion threats

Y.-A. DE MONTEJOYE *et al.*'s Report "Unique in the shopping mall: On the reidentifiability of credit card data" (special section on The End of Privacy, 30 January, p. 536) led to a widespread media sensation proclaiming that reidentification is easy with only a few pieces of credit card data (*1*–*3*). Although we agree with de Montejoye *et al.* that data disclosure practices must be responsibly balanced with data privacy and utility, we are concerned that the study's findings reflect unrealistic data intrusion threats. Making policy decisions based on the conclusions from this work would thus be hasty and could lead to the abandonment of modern data protection standards, with negative consequences to privacy, research, and society.

Some media confusion stems from the paper's use of the term "reidentify"; credit card metadata were not actually linked to any personal identities. Instead, it was assumed that an intruder could obtain data about identity, geography, time, and price to reidentify all targeted consumers. Yet this scenario requires some very strong assumptions about the attacker that are unlikely to be realized in practice. First, the study did not demonstrate the extent to which the necessary identifying information could be obtained reliably for any consumer. Second, the study neglects to acknowledge that when the data come from a fraction of the general population, unique purchase data in the sample will often not be unique in the larger population. Given that the undisclosed country's population was likely much larger than 1.1 million, the paper's data uniqueness measure is likely a substantial overestimate of risk. Third, the study's risk estimates are further inflated because they did not include cash or other banks' credit card purchases.

The research communicated in this paper is critical to moving forward privacy discussions about data sharing. However, we stress that claims about reidentification must be based on models that realistically and correctly account for the probability, as well as the possibility, of attacks.

***Daniel Barth-Jones,[1]*Khaled El Emam,[2]***

Xiangqian Li and Jiaqi Liu demonstrate a physical exam during the opening skit at Healthcare Day.

OUTSIDE THE TOWER

# Acting to build trust

My 5-foot 9-inch, 160-pound frame hunches over in pain. "I am Macho Man," I declare to the crowd, "but because of my back pain, I can't even pick up a strawberry!" Suddenly, a white-coated angel appears beside me. I look up in awe. "Can you relieve my pain?" Thus begins the opening sketch at Healthcare Day in Shanghai, China.

Since 2012, I and others from Zhongshan Hospital at Fudan University have been holding Healthcare Day in Community, a quarterly event in which medical scientists from different departments provide health counseling and services for local residents.

As the skit continues, the white-coated angel explains the physiology behind back pain and shows me, Macho Man, how to prevent muscle strain. He models the correct way to lift a heavy load and demonstrates how a physical exam would diagnose the problem. Then he leaves the stage and walks through the audience, looking for other people in need.

During Healthcare Day, community members have the opportunity to ask scientists about genes, tissue engineering, cancer, and cutting-edge medical research. We advocate a healthy lifestyle and give relevant advice about exercise habits, weight control, and proper diet. This event builds trust between the community and medical scientists. We hope more medical scientists and doctors in China will join us in stepping out of the hospital and sharing health science with the public. ***Jiaqi Liu***

Department of Plastic Surgery, Zhongshan Hospital, Fudan University, 200032, Shanghai, China. E-mail: liujiaqi1213@yahoo.com

PHOTO: XIANGQIAN LI, ZHONGSHAN HOSPITAL, FUDAN UNIVERSITY

***Jane Bambauer,[3] Ann Cavoukian,[4] Bradley Malin[5]***

[1]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032, USA. [2]College of Law, University of Arizona, Tucson, AZ 85721, USA. [3]Privacy and Big Data Institute, Ryerson University, Toronto, ON, M5B 2K3, Canada. [4]Children's Hospital of Eastern Ontario Research Institute and University of Ottawa, Ottowa, ON, K1H 8L1, Canada. [5]Department of Biomedical Informatics, Vanderbilt University, Nashville, TN 37212, USA.

*Corresponding author. E-mail: db2431@Columbia.edu

**REFERENCES**

1. N. Singer, "With a few bits of data, researchers identify 'anonymous' people," *New York Times* (29 January 2015); http://bits.blogs.nytimes.com/2015/01/29/with-a-few-bits-of-data-researchers-identify-anonymous-people/.
2. R. Jacobson, "Your 'anonymous' credit card data is not so anonymous, study finds," *PBS NewsHour* (29 January 2015); www.pbs.org/newshour/rundown/anonymous-credit-card-data-anonymous-study-finds/.
3. D. Coldewey, "'Anonymous' credit card data can still give you away," *NBC News* (29 January 2015); www.nbcnews.com/tech/tech-news/anonymous-credit-card-data-can-still-give-you-away-n296446.

## *Response*

BARTH-JONES *ET AL.* claim that our findings "reflect unrealistic data intrusion threats" We strongly disagree and argue that Barth-Jones *et al.*'s Letter is instead a superb illustration of why deidentification is not "a useful basis for policy" (*1*).

A simple and real example of our attack model is a bank sharing metadata for its 1.1 million customers in anonymized form with a third party for analysis. If the third party is able to obtain additional information—such as loyalty program data if the third party is a retailer—that data could be used to reidentify an individual and all the rest of his or her purchases.

Barth-Jones *et al.*'s Letter exemplifies the intrinsic issue with deidentification. One can always, as Barth-Jones *et al.* have, artificially lower the estimated likelihood of reidentification through the use of arbitrary and debatable assumptions.

First, Barth-Jones *et al.* have consistently considered an intrusion to be a breach of privacy only if "all targeted customers" are reidentified (*2*). This is an unrealistic definition of breach of privacy. Second, Barth-Jones *et al.* assume that it is "very unlikely" for an attacker to be able to collect geolocalized information about an individual. At best, this is a striking underestimation of the current availability of identified data. Possible sources would include manually collected clues about an individual we know (e.g., receipts or branded shopping bags) (*3*); having access or collecting from public profiles people's check-ins at shops or restaurants on Yelp, Foursquare, or Facebook (*4*); or having access to a retailer's database or to a database of geolocalized information such as the one collected by smartphone applications (*5*), WiFi companies, and virtually any carriers in the world. Third, Barth-Jones *et al.* assume that an attacker cannot know whether an individual is a client of a bank and is therefore in the data set. This is again an assumption that artificially lowers the estimated, and thus perceived, risks of reidentification without changing at all the actual risk for people in the release data set. Fourth, the fact that an individual might occasionally pay cash only means that an attacker would need a few more points.

Estimated probabilities of reidentification are not a useful basis for policy, and we stand by our comment that "the open sharing of raw [deidentified metadata] data sets is not the future" (*6*).

***Yves-Alexandre de Montjoye* and Alex "Sandy" Pentland***

Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*Corresponding author. E-mail: yvesalexandre@demontjoye.com

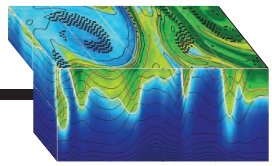
**REFERENCES**

1. President's Council of Advisors on Science and Technology, *Big Data and Privacy: A Technological Perspective* (PCAST, Washington, DC, 2014), pp. 38–39.
2. D. C. Barth-Jones, "Press and Reporting Considerations for Recent Re-Identification Demonstration Attacks: Part 2" (http://blogs.law.harvard.edu/billofhealth/2013/10/01/press-and-reporting-considerations-for-recent-re-identification-demonstration-attacks-part-2-re-identification-symposium/).
3. L. Sweeney, *Int. J. Uncertainty, Fuzziness Knowledge-Based Syst.* **10.05**, 557 (2002).
4. Wallaby, "Is anonymous financial data anonymous?" (www.walla.by/blog/110651700144/is-anonymous-financial-data-anonymous).
5. CNIL, "Mobilitics, season 2: Smartphones and their apps under the microscope" (www.cnil.fr/english/news-and-events/news/article/mobilitics-season-2-smartphones-and-their-apps-under-the-microscope/).
6. J. Bohannon, *Science* **347**, 468 (2015).

# RESEARCH

Downwelling drags particulate and dissolved carbon into the deep

*Omand* et al., *p. 222*



## IN *SCIENCE* JOURNALS

*Edited by* **Caroline Ash**

Mountain gorillas are highly inbred

PRIMATE GENOMICS

### Genomes in the mist

The mountain gorilla is an iconic species that is at high risk of extinction. Xue *et al.* have sequenced 13 gorillas from two different populations to probe their genetic diversity. The genomes show large tracts of homozygosity and the loss of highly deleterious genetic variants, indicating population bottlenecks and inbreeding. This loss of genetic diversity appears to have started over 20,000 years ago and may have been caused by changes in climate and human-associated effects. — LMZ

*Science*, this issue p. 242

PROTEIN TARGETING

### Sorting out cell sorting

The sorting of proteins into appropriate compartments is vital for proper cell function. After its discovery 20 years ago, various potential roles for nascent chain-associated complex (NAC) in the specificity of intracellular protein sorting have been proposed. Until now, no clear phenotypes have been discovered. Gamerdinger *et al.* explored the role of NAC in protein translocation in *Caenorhabditis elegans* (see the Perspective by Kramer *et al.*). They found that NAC prevents the importation of incorrect cargo, such as mitochondrial proteins, into the endoplasmic reticulum. — SMH

*Science*, this issue p. 201; see also p. 182

PROTEOSTASIS

### Giving protein folding a helping hand

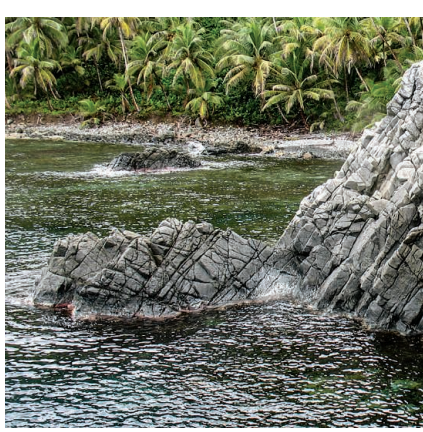
The reversible phosphorylation of proteins controls virtually all aspects of cell and organismal function. Targeting phosphorylation offers a broad range of therapeutic opportunities, and thus kinases have become important therapeutic targets. As targets, phosphatases should be as attractive, but in fact they are more challenging to manipulate. Das *et al.* have found a safe and specific inhibitor, called Sephin1, that targets a regulatory subunit of protein phosphatase 1 in vivo. Sephin1 binds and inhibits PPP1R15A, but not the related regulatory phosphatase PPP1R15B. In mice, Sephin1 prolonged a stress-induced phospho-signaling pathway to prevent the pathological defects of the unrelated protein-misfolding diseases Charcot-Marie-Tooth 1B and amyotrophic lateral sclerosis. — SMH

*Science*, this issue p. 239

PALEOCEANOGRAPHY

### Early closing between oceans

The Central American Seaway, which once separated the Panama Arc from South America, may have closed 10 million years earlier than is believed. Montes *et al.* report that certain minerals of Panamanian provenance began to appear in South America during the Middle Miocene, 15 to 13 million years ago (see the Perspective by Hoorn and Flantua). The presence of the minerals indicates that rivers were flowing from the Panama Arc into the shallow marine basins of northern South America. One interpretation of this finding is that large-scale ocean flow between the Atlantic and Pacific had ended by then. If this is true, then many models of paleo-ocean circulation and biotic



IMAGES: (TOP TO BOTTOM) OMAND *ET AL.*; © JOE MCDONALD/CORBIS; DAVID FARRIS





Published by AAAS

exchange between the Americas need to be reconsidered. — HJS

*Science*, this issue p. 226; see also p. 186

DRUG DISCOVERY

## Allergy drug inhibits viral infection

A drug used to dry up a runny nose and itchy eyes may be repurposed for treating hepatitis C virus (HCV). This viral infection often goes undetected, but it can exacerbate liver diseases, including cancer. The fact that allergy-relieving antihistamines can treat HCV was uncovered by He *et al.* in a screen of a library of approved drugs. Among these, the first-generation antihistamine chlorcyclizine showed highly specific anti-HCV activity in vitro and in mice with "humanized" livers, without evidence of drug resistance, a common problem with antivirals. Moreover, chlorcyclizine synergized with other anti-HCV drugs such as ribavirin, sofosbuvir, and interferon-α. Antihistamines are widely available, safe, and inexpensive: ideal candidates for use in HCV-endemic countries. — MLF

*Sci. Transl. Med.* **7**, 282ra49 (2015).

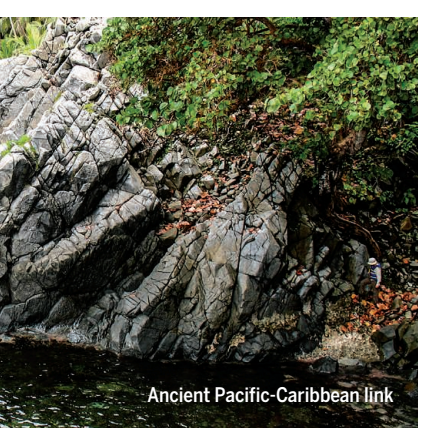
MARS ATMOSPHERE

## Mapping Mars' water history

We know the water cycle on Earth is complex. Neither is it simple on Mars. Infrared maps of water isotopes made by Villanueva *et al.* show the distribution of $H_2O$ and "semiheavy" water (HDO: deuterated water containing a mixture of hydrogen isotopes) across Mars. HDO enrichment varies with time and location; for example, irregular isotopic signals associate with different terrain features. The measurements also allow seasonal sublimation levels of the northern ice cap to be estimated and thus could be used to reveal past climate behavior. — MMM

*Science*, this issue p. 218

PHOTO: © ROSEMARY ROBERTS/ALAMY



Ancient Pacific-Caribbean link

BIOANALYSIS

## Imaging lipid composition

Chemical imaging of cell membranes can be performed with matrix-assisted laser desorption/ionization mass spectrometry (MALDI), but low ionization efficiency often leads to a signal dominated by the main lipid components, such as abundant phosphatidylcholine species. Soltwisch *et al.* used a tunable laser for post-ionization of neutral species to boost the signal for other membrane components, such as cholesterol and phospho- and glycolipids. Imaging of cells and tissues with these methods allows differentiation based on a more extensive chemical signature. — PDS

*Science*, this issue p. 211

HUMAN GENETICS

## Chromosome number varies in humans

Pregnancy loss is often associated with a loss of chromosome number, a condition known as aneuploidy. When examining aneuploid embryos during in vitro fertilization cycles, McCoy *et al.* found a large genomic region associated with defects in maternal chromosome number (see the Perspective by Vohr and Green). This region contains a gene, *Polo-like Kinase 4* (*PLK4*), that is known to affect chromosome segregation and has variants that correlate with an increased rate of maternal aneuploidy. Surprisingly, such variants occur at relatively high levels in human populations and may be under positive selection. — LMZ

*Science*, this issue p. 235; see also p. 180

# IN OTHER JOURNALS

*Edited by* **Sacha Vignieri** *and* **Jesse Smith**

Smoking increases inflammation and worsens infection

INFLAMMATION

## How smoking makes infection worse

Smoking not only increases your risk of cancer but also increases inflammation and slows down recovery from infections, especially in people who have chronic obstructive pulmonary disorder (COPD). People with severe COPD have elevated levels of the cytokine interleukin-33 (IL-33), a secreted protein that promotes airway inflammation. To determine whether IL-33 triggers exaggerated inflammatory responses in COPD, Kearley *et al.* exposed mice lacking IL-33 to cigarette smoke and then infected them with influenza. A deficiency in IL-33 protected mice from excessive inflammation and weight loss. Cigarette smoke led to elevated production of IL-33 and altered expression of its receptor in exposed mice, causing lung inflammation to amplify. — KLM

*Immunity* **42**, 566 (2015).

CELLULAR MECHANICS

## The mechanics of cellular left and right

Cells need to distinguish between left and right to interact during collective movement or embryonic development. To reveal the underlying cellular mechanics, Tee *et al.* studied the organization of actin in human cells using fluorescence, electron microscopy, and computational simulations. Actin fibers forming the cellular skeleton rearranged themselves from a symmetric to an asymmetric pattern through interplay between two types of fibers. The unidirectional tilting of the radial fibers and swirling of the transverse fibers were driven by

COMMUNITY ECOLOGY

## Measuring effective dispersal

Dispersal fundamentally shapes ecological communities and maintains biodiversity, yet it is extremely difficult to quantify. Estimating seed dispersal effectiveness (SDE) requires measurement of both the number of seeds dispersed and the probability of successful germination. González-Castro *et al.* laboriously measured the SDE of frugivorous lizards and birds as dispersers of seeds over 6 years in plant communities on the oceanic island of Tenerife. Their results allowed for comparisons between the SDE of the two main frugivore groups for different plant species and different communities (such as woodland and shrubland), revealing the networks of mutualistic interactions that underpin community persistence. — AMS

*Ecology* **96**, 808 (2015).

***Gallotia galloti* is an important seed disperser on Tenerife**

contractile stress and rotational growth. The actin-crosslinking protein controlled the clockwise or anticlockwise dynamics of the actin network, establishing the left-right asymmetry of the cell. — MSM

*Nat. Cell Biol.* 10.1038/ncb3137 (2015).

CRIME AND PUNISHMENT

## Rethink jail for juvenile justice

The U.S. criminal justice system is strikingly punitive: The incarcerations-to-convictions ratio is 70% larger than that of the next highest country. The U.S. juvenile incarceration rate is five times larger than that of any other country. But we don't randomly assign juveniles to jail, and thus it's difficult to isolate the impacts of incarceration on later-life outcomes, versus the impacts of underlying socioeconomic, cognitive, and other factors that influence juvenile criminal behavior as well as education, employment, and other outcomes. Studying 10 years of data on over 35,000 Chicago juvenile offenders, Aizer and Doyle noticed that judges were randomly assigned to juvenile cases, and judges had different tendencies to sentence incarceration versus probation. Assignment to a high–incarceration-rate judge decreased the likelihood of juveniles completing high school and increased the likelihood of adult incarceration. — BW

*Quart. J. Econ.* 10.1093/qje/qjv003 (2015).

NEURODEGENERATION

## Deconstructing cell death in MS

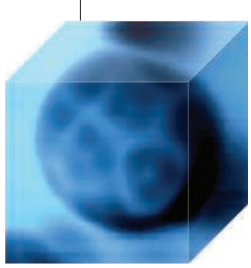
Oligodendrocytes are often referred to as "support" cells for neurons, but in fact they play a critical role in the transmission of nerve impulses. These cells produce the myelin sheath that surrounds and protects axons in the central nervous system. In multiple sclerosis (MS), this myelin sheath erodes because of an inflammatory reaction that triggers the death of oligodendrocytes. Several distinct mechanisms of cell death exist, and understanding which one underlies oligodendrocyte death could lead to new therapies. Studying mouse models of MS, Ofengeim *et al.* find that oligodendrocytes die by a regulated process called necroptosis. A small-molecule inhibitor of a protein kinase that mediates necroptosis prevented oligodendrocyte death in the mouse models. — PAK

*Cell Rep.* **10**, 1836 (2015).

SURFACE IMAGING

## Submolecular resolution in 3D

Surface probe microscopy has recently achieved submolecular resolution with metal tips made atomically sharp by decorating them with adsorbed molecules such as carbon monoxide. This method works well with adsorbed molecules that are small or flat. Moreno *et al.* now show how to achieve subatomic resolution in atomic force microscopy with commercial silicon tips, for three-dimensional structures such as absorbed fullerene molecules or the step edges of oxide surfaces.



The van der Waals interaction is mapped with the tip a few nanometers above the surface in a closed feedback loop. A second higher-resolution scan with an open feedback loop follows this map with a constant offset distance that is adjusted to provide high resolution. — PDS.

*Nano Lett.* 10.1021/nl504182w (2015).

PHYSICS

## Pairing in an off-kilter atomic gas

In a process called Cooper pairing, two electrons of opposite spin and momentum form a pair and join a larger "condensate" to flow effortlessly through a superconducting material. In cold atomic gases, the atomic state plays the role of the spin. But what happens if there is a different number of atoms in the two spin states, so that pairing cannot be perfect? Ong *et al.* tackled that question for $^{6}$Li atoms trapped in a weakly coupled array of pancake-shaped clouds. As they cranked up the interaction between the atoms and increased the relative number of minority spins, the gas in each pancake separated into a core of paired spins surrounded by the majority atoms. — JS

*Phys. Rev. Lett.* **114**, 110403 (2015).





Published by AAAS

## ALSO IN *SCIENCE* JOURNALS

*Edited by* **Caroline Ash**

QUANTUM GASES

### Detecting multiple temperatures

Most people have an intuitive understanding of temperature. In the context of statistical mechanics, the higher the temperature, the more a system is removed from its lowest energy state. Things become more complicated in a nonequilibrium system governed by quantum mechanics and constrained by several conserved quantities. Langen *et al.* showed that as many as 10 temperature-like parameters are necessary to describe the steady state of a one-dimensional gas of Rb atoms that was split into two in a particular way (see the Perspective by Spielman). — JS

*Science*, this issue p. 207; see also p. 185

EARTH HISTORY

### Ocean acidification and mass extinction

The largest mass extinction in Earth's history occurred at the Permian-Triassic boundary 252 million years ago. Several ideas have been proposed for what devastated marine life, but scant direct evidence exists. Clarkson *et al.* measured boron isotopes across this period as a highly sensitive proxy for seawater pH. It appears that, although the oceans buffered the acidifiying effects of carbon release from contemporary pulses of volcanism, buffering failed when volcanism increased during the formation of the Siberian Traps. The result was a widespread drop in ocean pH and the elimination of shell-forming organisms. — NW

*Science*, this issue p. 229

QUANTUM COMPUTING

### Benchmarking quantum simulation

Finding a solution to a problem often amounts to an optimization problem and thus can be recast in terms of the lowest-energy state of a system. To find such ground states, mathematical methods based on annealing were developed. To reach the ground state more quickly than with the earlier classical methods, a quantum-mechanical approach was proposed; however, the evidence for quantum speed-up is contradictory. Heim *et al.* show that the results depend on how the problem is described and how the optimization routine is implemented. This development should be valuable for benchmarking quantum machines. — ISO

*Science*, this issue p. 215

CARBON FLUX

### Down with atmospheric carbon dioxide

How does the ocean move carbon from surface waters to its deep interior? Current understanding is that carbon dioxide is removed from the atmosphere by phytoplankton that are eaten, and in turn their predators die and sink into deep water and seafloor sediments. In addition to this route, Omand *et al.* show that downwelling caused by ocean eddies 1 to 10 km across can deliver much of the carbon produced in spring to the deep sea. The eddies entrain small particles and dissolved organic carbon to augment the flux of large sinking particles. — HJS

*Science*, this issue p. 222

BIOCHEMISTRY

### Single-molecule assay of ubiquitylation

Many biological processes in cells are regulated by ubiquitin peptides that are attached to proteins. Measurement of single fluorescent molecules in cell extracts can be used to trace the kinetics of such reactions. Lu *et al.* refined assay conditions to follow ubiquitination by an E3 ubiquitin ligase (see the Perspective by Komander). They visualized the activity of the anaphase-promoting complex (APC), a ubiquitin ligase critical for control of the cell division cycle. The processive initial reaction catalyzed by the APC was replaced by slower reactions. The results show how small, commonly occurring recognition motifs can guide specific and highly controlled enzymatic events. In a companion paper, Lu *et al.* explored how the number and arrangement of added ubiquitin chains affected the interaction of ubiquitylated proteins with the proteasome (a protein complex that recognizes ubiquitylated proteins and degrades them). The extent of ubiquitylation determined the strength of interaction of a substrate protein with the proteasome, and the arrangement of the ubiquitin chains determined the movement of the protein into the proteasome and thus the rate of degradation. — LBR

*Science*, this issue p. 199, p. 200; see also p. 183

COMETARY FORMATION

### Making comets in the cold

The speciation of nitrogen compounds in comets can tell us about their history. Comets are some of the most ancient bodies in the solar system and should contain the nitrogen compounds that were abundant when they formed. Using the ROSINA mass spectrometer aboard the Rosetta spacecraft orbiting comet 67P/Churyumov-Gerasimenko, Rubin *et al.* found molecular nitrogen at levels that are depleted compared to those in the primordial solar system. Depletion of such a magnitude suggests that the comet formed either from the low-temperature agglomeration of pristine amorphous water ice grains or from clathrates. — NW

*Science*, this issue p. 232

EARLY WARNING SYSTEMS

### Crowdsourcing early warnings of natural disasters

Cellular phones and similar electronic devices contain powerful technologies that can be harnessed for location-based sensing of the environment. Minson *et al.* show how displacement data from consumer-grade Global Navigation Satellite System (i.e., GPS) sensors and Inertial Navigation System (e.g., accelerometer) sensors acquired from widely distributed cell phones can be combined to produce earthquake and tsunami early warnings. Such rapidly acquired crowdsourced data could be a valuable way of improving public safety and facilitating disaster responses in earthquake-prone regions. — KVH

*Sci. Adv.* 10.1126/2ciadv.1500036 (2015).

SYSTEMS BIOLOGY

### Networking death signals to fight cancer

Selectively killing cancer cells without inducing resistance is the holy grail of cancer therapy. Cancer cells are particularly sensitive to cell death triggered by the secreted protein TRAIL. Unfortunately, some cancer cells evade TRAIL-induced death and develop resistance by rewiring their signaling networks. So *et al.* used proteomic data to map a protein interaction network of the kinases that affected TRAIL-induced cell death. The modeling of information flow through the network revealed potential targets that could be exploited to develop combination therapies with TRAIL to kill cancer cells and prevent resistance. — NRG

*Sci. Signal.* **8**, rs3 (2015).

RESEARCH ARTICLE SUMMARY

BIOCHEMISTRY

# Specificity of the anaphase-promoting complex: A single-molecule study

Ying Lu, Weiping Wang, Marc W. Kirschner*

**INTRODUCTION:** Regulation of biological processes is generally thought to be conveyed by structurally complementary interactions between molecules, resulting in a bound state that is the most energetically stable. However, the limited size of most recognition motifs poses a very general question as to how specificity is actually achieved. In the cell cycle, timing of events from exit of mitosis through passage into $G_1$ is regulated by the anaphase-promoting complex (APC), an enzyme that catalyzes ubiquitin conjugation to multiple lysines on substrate molecules, for proteasome-mediated degradation. The difference in how APC ubiquitylates its substrates leads to the ordering of events. APC-substrate interactions are also mediated through very short and redundant sequences, such as the well-studied D and KEN boxes found in 69% of human proteins. It is unclear how the APC can distinguish a relatively small number of substrates from nonsubstrates having the same interaction motifs. At a total concentration of approximately 1 mM, many of these nonsubstrates should compete for APC binding, thus reducing the rate of ubiquitylation and degradation of the real targets.

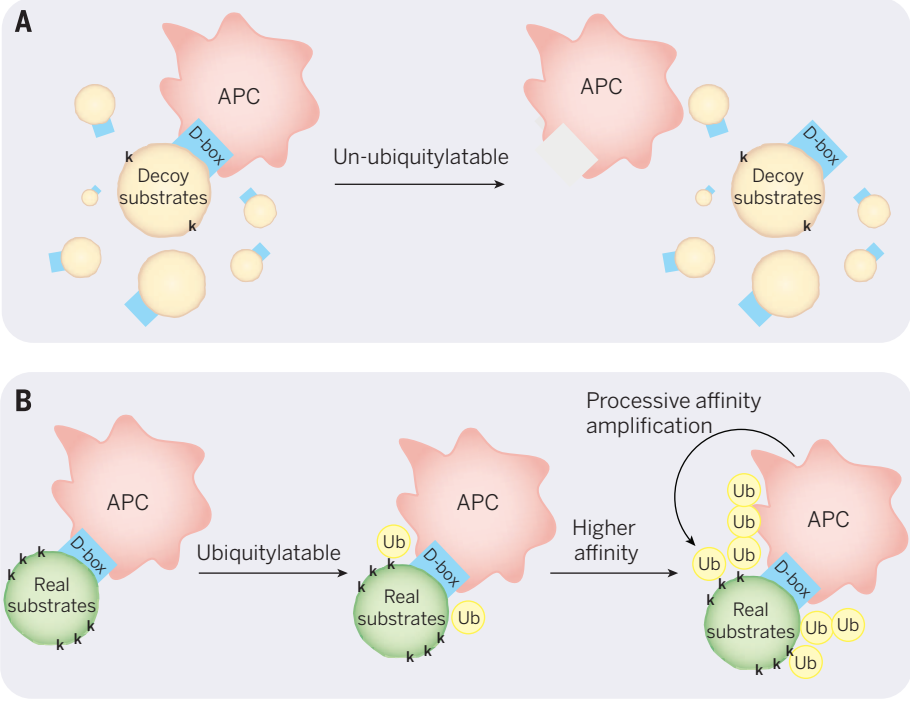
**RATIONALE:** Simple binding based on structural complementarity seems inadequate to explain the specificity in APC-mediated ubiquitylation. We therefore looked beyond a single binding step to the extensive network of reactions that take place at multiple sites on a substrate molecule and in a cellular environment with competing reactions. To describe these, we developed single-molecule fluorescence assays, capable of identifying multiple reaction intermediates, and applied these methods to the study of the kinetics of APC-mediated ubiquitylation in both cytosolic extracts and in purified reaction systems. Our goal was to understand how the APC selects the right substrates among competing molecules and generates effective ubiquitin configurations that can be recognized for degradation.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1248737

**RESULTS:** In this assay, the rates of ubiquitylation and APC-substrate interaction can be measured simultaneously and traced to individual substrate molecules. We find that the APC-mediated ubiquitylation initially involves a highly processive reaction conjugating several ubiquitins on a substrate molecule, followed by multiple encounters and reactions at a slower rate. The initial ubiquitylation greatly enhances substrate-binding affinity with the APC in subsequent reactions, by both increasing the on-rate and decreasing the off-rate. Results of kinetic studies in bulk assays further substantiate this positive feedback mechanism.

**CONCLUSION:** The process of converting the reactivity of ubiquitin on a substrate to binding affinity can complement weak recognition by the D/KEN-box motifs. Cycles of positive feedback achieve high specificity for substrates with short recognition motifs in a cellular environment of competing molecules. This process, called processive affinity amplification (PAA), iteratively increases conjugated ubiquitins on substrates with both the recognition motifs and receptor lysine residues. Reaction schemes similar to PAA may be responsible for maximizing specificity, while maintaining efficiency in other biological processes. ■




**Positive feedback enhances specificity and efficiency in APC-mediated ubiquitylation.** The intrinsic affinity between APC and D-box–containing decoy substrates is low, which prevents APC from inhibition by a high concentration of those substrates in the cytosol (**A**). The initial ubiquitylation on real substrates is highly processive, likely due to their preponderance in the availability and reactivity of receptor lysines. Conjugated ubiquitins enhance substrate's affinity for the APC and promote further ubiquitylation, enabling proteasomal degradation (**B**).

## RELATED ITEMS IN *SCIENCE*

Y. Lu, B. Lee, R. W. King, D. Finley, M. W. Kirschner, Substrate degradation by the proteasome: A single-molecule kinetic analysis. *Science* **348**, 1250834 (2015).
D. Komander, Details of destruction, one molecule at a time. *Science* **348**, 183–184 (2015).

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: marc@hms.harvard.edu
Cite this article as Y. Lu *et al.*, *Science* **348**, 1248737 (2015).
DOI:10.1126/science.1248737

RESEARCH ARTICLE

BIOCHEMISTRY

# Specificity of the anaphase-promoting complex: A single-molecule study

Ying Lu, Weiping Wang, Marc W. Kirschner*

**Biological processes require specific enzymatic reactions, paradoxically involving short recognition sequences. As an example, cell-cycle timing depends on a sequence of ubiquitylation events mediated by the anaphase-promoting complex (APC) based on short redundant motifs. To understand the origin of specificity, we designed single-molecule fluorescence assays that capture transient ubiquitylation reactions. We find that the APC-mediated ubiquitylation involves a highly processive initial reaction on the substrate, followed by multiple encounters and reactions at a slower rate. The initial ubiquitylation greatly enhances the substrate's binding affinity in subsequent reactions, by both increasing the on-rate and decreasing the off-rate. We postulate that these cycles of positive feedback enable high specificity for substrates with short recognition motifs in a complex cellular environment.**

Regulation of biological processes requires that proteins distinguish the proper binding partners or substrates from nonsubstrates. Emil Fischer famously viewed the substrate as the key and the enzyme as the lock. There are many proven examples of specificity arising from such complementary binding interactions. However, specificity cannot be attributed completely to differential binding affinity alone, as has been demonstrated by the discrepancy between the error rate of protein synthesis and the measured binding affinities of the correct and incorrect aminoacyl–tRNAs on mRNA. In a conceptual theory for resolving such discrepancies, called "kinetic proofreading," the energy from adenosine 5′-triphosphate or guanosine 5′-triphosphate hydrolysis is used to substantially reduce error rates (*1*, *2*). The kinetic proofreading mechanism has been invoked to explain the high degree of specificity of various biological processes governed by relatively weak binding interactions (*3*, *4*).

The importance of recognition and specificity in protein degradation is increasingly appreciated. During mitosis, cyclin A, cyclin B, securin, polo kinase, and UbcH10 are ubiquitylated sequentially by the anaphase-promoting complex (APC) and degraded by the proteasome, contributing to an ordered execution of cell-cycle events (*5*, *6*). The precise timing of protein degradation at cell-cycle transitions demands high specificity in APC-substrate interaction. Paradoxically, these interactions are mediated by exceedingly short recognition sequences, such as the D [RXXLXXXX(N)] and KEN (KEN) boxes (where X is any amino acid) (*7*–*9*). These recognition sequences are highly redundant, appearing in 69% proteins in the human genome, whereas only 50 to 100 real substrates for the APC are expected from genome-wide studies (*10*). This preponderance of "decoy substrates" poses an additional challenge. At a total concentration of ~1 mM, they should compete for APC binding and reduce the rate of ubiquitylation and degradation of real substrates. Yet, this does not appear to happen. APC substrates, such as securin and cyclin B, can be efficiently degraded in a few minutes in yeast and mammalian cells. Although these small recognition sequences are required, they are not sufficient to confer susceptibility when transferred to other proteins (*7*). Thus, real APC targets must provide other specificity features, even though no additional sequence features are discernible. Therefore, the simple "key-lock" mechanism appears to be inadequate to explain the specificity in APC-mediated ubiquitylation. To explain the required specificity, we looked beyond single binding interactions to the network of reactions that take place in the cellular environment, where there are competing substrates and competing reactions.

The APC interacts with substrates transiently and recurrently, conjugating complex and heterogeneous ubiquitin configurations on substrates molecules (*11*). These features make it difficult to understand the reaction kinetics and mechanisms with conventional biochemical methods. In this study, we have developed single-molecule (SM) fluorescence assays capable of identifying multiple reaction intermediates. Using these assays, we could examine reactions at individual steps to study APC-mediated ubiquitylation in both purified systems and in crude cell extracts. We distinguished different modes of substrate binding, reflected in different kinetics for initial ubiquitin conjugation and subsequent elongation steps. We identified a reaction pathway based on the evolving ubiquitylation status of the substrate that can explain how substrate specificity and reaction efficiency can be achieved in a network of low-specificity interactions.

Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA.
*Corresponding author. E-mail: marc@hms.harvard.edu

## Results

### *A single-molecule ubiquitylation assay*

To observe reactions of single molecules, we used ubiquitin and substrates chemically labeled with different chemical fluorescent dyes (Fig. 1, A and B). The labeled proteins were incubated in concentrated extracts of human HeLa cells in the $G_1$ phase of the cell cycle. Endogenous APC was labeled with a fluorescently tagged antibody; the substrates were C-terminally labeled with biotin and tethered with streptavidin to polyethylene glycol and albumin passivated glass slides. Because each ubiquitin molecule was conjugated with a single fluorophore at its N terminus, the number of conjugated ubiquitins on a substrate molecule could be measured as the total fluorescence intensity. A major problem in developing the SM methods for ubiquitylation in Fig. 1B was the background fluctuation from fluorescently labeled ubiquitin in solution and the nonspecific binding of fluorescent species to the slide surface. We used combinatorial strategies to increase the signal-to-noise ratios and specificity of the detection (fig. S1); these allowed us to distinguish the addition and removal of single ubiquitins at a 95% confidence level in the presence of 3 μM fluorescent ubiquitin (materials and methods), near its physiological range of 5 to 10 μM. The chemical fluorophores with their high flux of photons, and a relatively long exposure time of 3 s afforded by the natural slow reaction rates, along with strong laser excitation, allowed us to reduce Poisson noise. Adding 10 mM imidazole greatly reduced nonspecific binding of all fluorescent species and lowered the background fluctuation without affecting kinetics of ubiquitylation (fig. S2). Intensity values of single molecules were calculated by a method of local background subtraction (figs. S3 and S4 and materials and methods).

To distinguish substrate-specific ubiquitylation signals from nonspecific binding of ubiquitylated proteins from cell extracts, we limited our analysis to ubiquitylation events that were coincident with binding of the fluorescently labeled APC to a fluorescently labeled substrate. Because of the sparseness of nonspecific binding events, the chance of background binding of ubiquitin or ubiquitylated proteins coinciding with the labeled APC was less than $10^{-5}$ (materials and methods). The accuracy and linearity of the ubiquitylation measurement were assessed and confirmed by photobleaching of preformed ubiquitin chains, a process that randomly inactivates single fluorophores (Fig. 1C and figs. S5 and S29 to S30). The APC antibody labeling was sufficiently stable to allow observation of relatively transient substrate-APC interactions (Fig. 1D). So that each event of ubiquitin conjugation was recorded, we completely depleted endogenous free ubiquitin from extracts by adsorption to beads containing the E2-UbcH10, without affecting ubiquitylated

proteins (fig. S6); the extracts were then supplemented with 3 μM Alexa Fluor 467 (alexa647)–labeled ubiquitin (fig. S7).

### Ubiquitylation mediated by the APC at a single-molecule level

Geminin is an APC substrate that functions in control of DNA replication. When geminin was immobilized on a slide as the substrate, we observed a greatly enhanced ubiquitin fluorescence signal, compared to background, that was coincident with APC binding events (fig. S8 and movies S1 and S2). Overall ubiquitylation level was strongly reduced in the presence of the APC inhibitor Emi1, suggesting that APC is the primary E3 for geminin ubiquitylation (fig. S9). We recorded time traces of APC binding and ubiquitylation on single geminin molecules. In cell extracts, the ubiquitylation signal was dynamic, showing features that can be attributed to enzymatic deubiquitylation in the extract, proteasomal degradation of the substrate, and a slow rate of photobleaching (half-life = 8.6 min; fig. S10). The SM method fully reported on the heterogeneity in the reactions: After the binding of an APC molecule, ubiquitylation on individual geminin molecules occurred with variable rates and reached different numbers of conjugated ubiquitins at the moment of APC disassociation. We present a representative raw trace of these features in Fig. 2, A and B, and 20 more raw traces for each substrate [geminin, the cyclin B–N-terminal domain (cyclin B–NT), and cyclin A in cell extracts; cyclin B–NT, securin, and K64–cyclinB–NT in a purified system] in the supplementary materials (figs. S31 to S36). The SM kinetics of ubiquitylation on geminin and cyclin B–NT were very similar. Each is known to be a processive substrate, where processivity is defined as the average number of ubiquitins conjugated to a substrate molecule, regardless of its degree or configuration of ubiquitylation, in each round of APC binding (*4*).

Several observations indicate that the ubiquitin signal on immobilized substrates represents bona fide ubiquitylation by the APC, rather than the binding of ubiquitylated APC or ubiquitylation by other E3s. We did not observe a sharp decrease of ubiquitin signal when the APC dissociated from the substrate, suggesting that the APC was not ubiquitylated. Either the APC inhibitor Emi1 or methylation of lysines on substrates, which prevents ubiquitylation, strongly reduced the ubiquitin signal (figs. S9 and S11). In about 22% cases, we observed increased ubiquitin fluorescence in the absence of detectable APC binding. This may have been caused by either APC conjugated to an unmodified (and hence nonfluorescent antibody) or free APC unconjugated to antibody. However, we cannot exclude a low level of APC-independent ubiquitylation. Ubiquitylated substrates on the slide were competent to interact with the proteasome (labeled with a fluorescent antibody) in the cell extract, suggesting that ubiquitin configurations that are competent for proteasomal degradation had been formed during the reaction (fig. S12).

In the SM traces, there were always two phases of ubiquitylation. The first was highly processive, soon after the initial binding of the APC, resulting in multiple ubiquityl residues transferred to a substrate molecule within 5 s (the time resolution of the experiment) (Fig. 2A and figs. S31 and S32). The second phase was much less processive, as the rate of ubiquitin transfer was lower and was interrupted by APC disassociations. Rebinding of APC was also frequently unproductive ("u" in Fig. 2, A and B), generating no detectable ubiquitin conjugations. Processive ubiquitylation could reoccur after the substrate had been partially or fully deubiquitylated. To illustrate the biphasic ubiquitylation statistically, we aligned 315 traces of cyclin B molecules ubiquitylated in cell extracts by their initiation points and plotted the median number of conjugated ubiquitins as a function of time. The processive initial reaction transferred an average of five ubiquityl moieties, followed by a slower transfer rate of ~1 ubiquitin per minute (Fig. 2C and fig. S13). The actual transfer rate may be slightly higher due to signal loss by photobleaching (fig. S10).

To distinguish the kinetics of monoubiquitylation from that of chain elongation, we replaced the endogenous free ubiquitin in cell extract with zero-lysine ubiquitin (Ub0K) to suppress chain formation. In this case, three rather than five ubiquityl moieties were transferred in the initial reaction, suggesting that some short ubiquitin chains were formed during the initial ubiquitylation when wild-type ubiquitin (wtUb) was used (Fig. 2C). In the second (or slower) phase, the rate of ubiquitin conjugation using wtUb was three times higher than that with Ub0K, suggesting that chain elongation occurred mainly after the initial processive reaction (fig. S13). It is unlikely that Ub chains were transferred en bloc (*12*, *13*), because the vast majority of the E2 UbcH10 was charged with a single ubiquitin (fig. S14).

To generalize this result, we compared histograms of the maximum number of ubiquitins conjugated to individual cyclin B molecules during the experiment using wtUb or Ub0K. For cyclin B, wtUb provided twice the level of ubiquitylation as Ub0K (Fig. 2E), consistent with earlier results by mass spectrometry (*11*). Kinetics of binding of the APC was substrate-dependent and strongly affected by substrate ubiquitylation status (see below). On average, 4.1 APC-binding cycles were recorded on a cyclin B molecule in a 15-min experiment, and 3.7 for geminin (Fig. 3A, fig. S15, and movie S3). To understand the contribution of each successive APC-binding cycle to substrate ubiquitylation, we analyzed the median number of ubiquitins, before, within, and after, the binding of APC on cyclin B–NT molecules in the cell extract. The first APC-binding cycle conjugated on average seven ubiquitins, including both the initial processive and the secondary distributive phases (fig. S16). Subsequent APC-binding cycles conjugated fewer ubiquitins but were important for chain elongation and for maintaining cyclin B in various ubiquitylated states in the presence of active deubiquitylation. To reduce the complexity due to deubiquitylating and proteasomal activities in the cell extract, we performed the

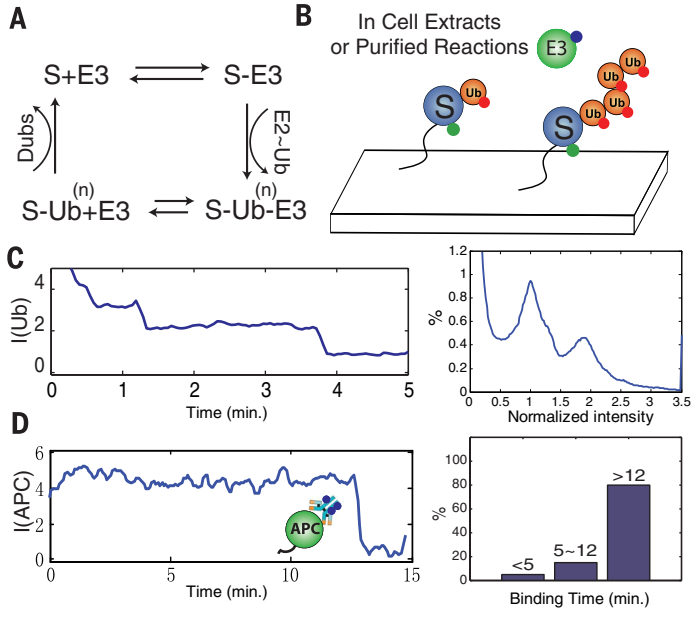



**Fig. 1. The single-molecule ubiquitylation assay.** (**A**) Major reaction pathways in the experiments. S: substrates; Dubs: deubiquitylating enzymes. (**B**) Schematics of the experimental setup. Substrates (S), ubiquitin, and the APC were fluorescently labeled. Substrates were also C-terminally biotinylated and immobilized on slides. Substrate-coated slides were incubated with functional cell extracts or purified APC reaction components. The kinetics of ubiquitylation and APC-substrate interaction were monitored on a TIRF microscope. (**C**) A representative trace of photobleaching a fluorescent polyubiquitin chain synthesized in an E2-25K reaction (more traces are in figs. S29 and S30). Right: Pairwise intensity difference distribution of 168 photobleaching traces. The position of the first peak has been normalized to "1." The photobleaching experiment was performed in the absence of free ubiquitin. (**D**) Biotinylated, purified APC molecules on the slide were incubated with fluorescent antibody against APC4 to analyze the stability of antibody-APC interaction. Left: a representative trace; right: histogram of apparent interaction durations.

same experiments in a purified system, composed of purified APC and other recombinant components. As in the cell extract, we found that for the processive substrates, securin and cyclin B–NT, ubiquitylation was biphasic and required multiple encounters with the APC (Fig. 2B, figs. S34 to S36, and movie S4).

Differences in the processivity of ubiquitylation have been used to explain the order of degradation for APC substrates during the cell cycle (*4*). Multiple steps in the reaction pathway could contribute to processivity, and these can be identified by SM methods. We compared a well-characterized processive substrate cyclin B and a distributive substrate cyclin A for their kinetics of ubiquitylation and APC binding using SM assays. Cyclin A–APC interactions were transient, lasting for on average 7 s compared to 30 s for cyclin B–NT in cell extracts (Fig. 3A). Accordingly, less ubiquitin was transferred to cyclin A in the initial phase or accumulated during the experiment (Fig. 2, C and E). Nonetheless, ubiquitin transfer was still highly efficient on cyclin A molecules once APC bound (Fig. 2A and fig. S33). Therefore, the processivity of cyclin A seems to be limited by relatively short APC interaction and also perhaps by cyclin A having fewer ubiquitylatable lysine residues.

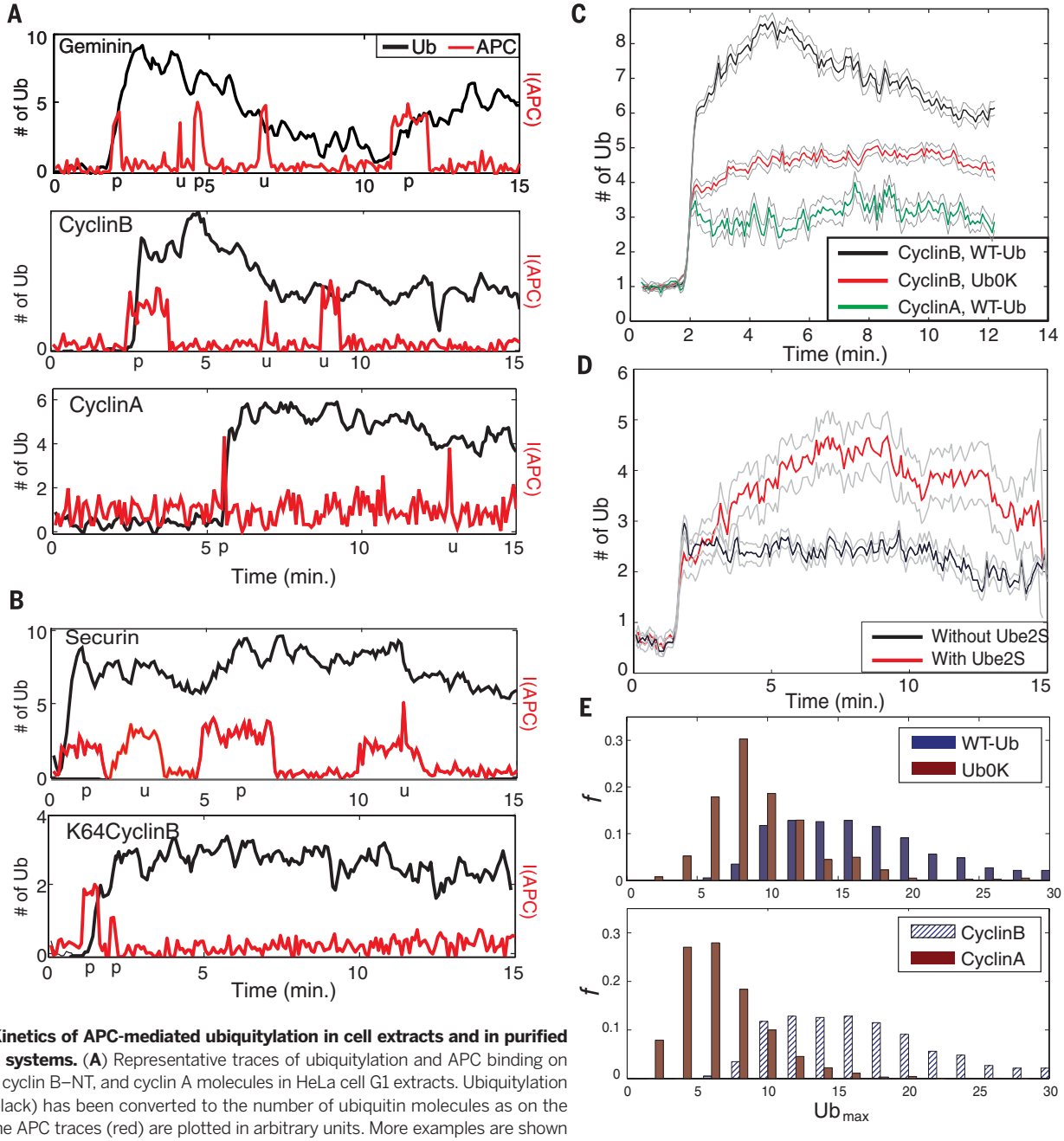



**Fig. 2. Kinetics of APC-mediated ubiquitylation in cell extracts and in purified reaction systems.** (**A**) Representative traces of ubiquitylation and APC binding on geminin, cyclin B–NT, and cyclin A molecules in HeLa cell G1 extracts. Ubiquitylation signal (black) has been converted to the number of ubiquitin molecules as on the *y* axis. The APC traces (red) are plotted in arbitrary units. More examples are shown in the supplementary materials. p: productive binding; u: unproductive binding. (**B**), as in (A), but using wt-securin and single-lysine K64–cyclin B–NT in purified APC ubiquitylation reactions. (**C**) 315 traces of ubiquitylation on individual cyclin B–NT molecules in cell extracts with either wt-Ub (black) or zero-lysine-Ub (Ub0K; red), or cyclin A with wt-Ub (green) were aligned by the ubiquitylation initiation point (arbitrarily set at $t = 2$) of each substrate molecule. The median number of ubiquitins on a substrate molecule is plotted versus time. Gray traces show the boundary of the standard deviation of the mean. (**D**) As in (C), but using single-lysine K64–cyclin B–NT with or without 100 nM Ube2S in purified APC reactions. (**E**) Histogram of the maximum number of ubiquitins on each substrate molecule during a 15-min experiment in the cell extract. Upper, on cyclin B with either wt-Ub or zero-lysine-Ub; Lower, on either cyclin B or cyclin A with wt-Ub.

The elongation of ubiquitin chains on APC substrates is facilitated by another E2, Ube2S. It has been debated whether UbcH10-mediated ubiquitin nucleation or Ube2S-mediated chain elongation is rate-limiting for ubiquitylation (*14*–*16*). Using single-lysine substrate K64–cyclin B–NT that can form only a single ubiquitin chain, we analyzed the two steps individually. In the absence of Ube2S, K64cyclinB-NT essentially received no more than three ubiquityl moieties in an APC reaction (fig. S17). Ube2S promoted chain elongation in a concentration-dependent manner (fig. S17). In the SM assay with Ube2S at its physiological concentration of 100 nM, we found that chain elongation beyond three ubiquitins was usually slow and stepwise, compared to the highly processive initial ubiquitylation mediated by UbcH10 (Fig. 2D). We conclude that UbcH10 initiates monoubiquitylation and short chains with high processivity, but fails to form long ubiquitin chains. Further chain elongation by Ube2S is less processive than the initial ubiquitylation by UbcH10, when these E2s are compared at their physiological concentrations.

## Positive feedback in APC-mediated ubiquitylation

APC-substrate interactions are highly dynamic. The kinetics of enzyme-substrate complex formation is commonly used in conventional assays to obtain the reaction rate constants, such as the off-rate ($k_{off}$), by fitting a simple binding model between molecule A and B, described as $[AB](t) \propto 1 - \exp(-k_{off} * t)$ given $K_d >> [A],[B]$, where [A], [B], and [AB] refer to the concentrations of A, B, and AB, respectively; $K_d$ refers to the dissociation constant. Under this condition, the time scale of the binding curve is dominated by the off-rate, as detailed in materials and methods. Rate parameters obtained in this way should be consistent with the values calculated from analyzing SM traces; a discrepancy may indicate a time-dependent modulation of the interaction in the population.

In cell extracts, the number of APC molecules interacting with a fixed number of substrate molecules increased gradually during the experiment and eventually reached a steady state. This was a common feature for all the processive substrates tested in this study (movies S2, S4, S5, and S7). The parameter needed to fit the APC–cyclin B binding curve (Fig. 3B, blue, and movie S5), as measured by the number of APC molecules on the slide as a function of time, is 0.12 $min^{-1}$. However, this is only 6% of the value for the off-rate obtained from analyzing SM traces of APC–cyclin B interaction in the presence of ubiquitylation (Fig. 3A). Therefore, the rate parameter of the APC–cyclin B binding curve does not represent the actual off-rate, but rather represents the time scale of an evolving factor that modulates the APC–cyclin B interaction, such as the varying state of ubiquitylation of the substrate, which was suggested by the overall parallel of ubiquitylation with the APC binding kinetics (fig. S18). Furthermore, low concentrations of free ubiquitin (the cyan curve in Fig. 3B), where no additional ubiquitin was added, led to reduced amounts of APC–cyclin B complex formation, suggesting that substrate ubiquitylation enhances its affinity with the APC (Fig. 3B and movie S6). This behavior is not an artifact due to the use of fluorescent ubiquitin, because unlabeled ubiquitin added to a Ub-depleted cell extract restored the APC-binding kinetics measured with fluorescent ubiquitin (Fig. 3B and movie S7). This affinity enhancement was also observed in experiments with purified APC, and with other substrates (Fig. 3C), excluding a major role for unidentified cofactors in the cell extract. It is unlikely that E2s charged with ubiquitin facilitate APC binding to substrates, because we observed a reduced amount of substrate-bound APC when we used a cell extract substituted with Ub0K in which the E2s were still charged with ubiquitin (Figs. 3B and 2C and movie S8). Because Ub0K cannot form chains, the result shows that ubiquitin chains on the substrate may facilitate, but are not absolutely required for, enhancement of APC binding. When the elongation factor E2 Ube2S, was added to an in vitro ubiquitylation reaction, it increased the number of conjugated ubiquitins on single-lysine K64–cyclin B–NT and resulted in an enhanced binding of the APC (Fig. 3D). These results suggest that the number of conjugated ubiquitins, rather than the presence of specific ubiquitin configurations, determines the binding enhancement of the APC. In conclusion, the apparent discrepancy between SM and ensemble measurements is resolved if we consider the effect of ubiquitylation on increasing the substrate's affinity with the APC; such a binding enhancement occurs with various ubiquitin configurations, including those without chains.

Because of constitutive deubiquitylation and degradation, ubiquitylated substrates usually represent only a small fraction of the total substrate concentration in the cell extract (fig. S19). Because $K_d$ = [APC] × [$Substrate_{total}$]/[$Substrate_{APC\ bound}$], a small increase in bound APC molecules may reflect a large decrease in $K_d$ for ubiquitylated substrates with the APC. We estimated $K_d$ values for the binding of ubiquitylated and unmodified substrate by the SM assay (materials and methods). Ubiquitylation reduced the $K_d$ of the binding of cyclin B–NT by a factor of 60 and

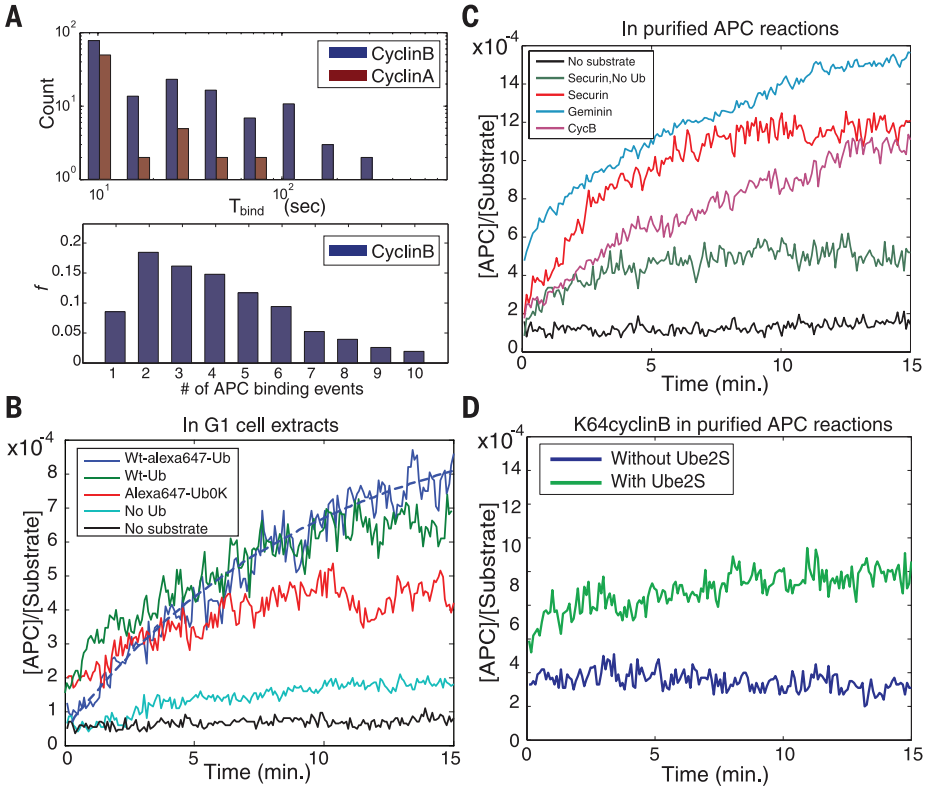



**Fig. 3. Kinetics of APC-substrate interaction.** (**A**) Upper: Histogram of the duration of APC-substrate interaction in the cell extract, plotted on a log-log scale. Only productive binding events have been considered. Lower: Histogram of the number of APC binding events on cyclin B–NT molecules in the cell extract during a 15-min experiment. (**B**) Time traces of the fraction of cyclin B–NT molecules bound by the APC in HeLa cell G1 extracts under the indicated conditions. About 2 × $10^5$ cyclin B–NT molecules were immobilized as the substrate. The dashed line represents a curve fitted using a simple exponential mode described in the text. (**C**) As in (B), but performed in a purified APC ubiquitylation system with wt-alexa647-Ub. Results of cyclin B–NT, geminin, and securin are shown. The same analysis but using K64–cyclin B–NT as the substrate in purified APC reactions, with or without 100 nM E2 Ube2S, is shown in (**D**).

reduced the $K_d$ for geminin by a factor of 40 (Fig. 4A). Consistently shown by SM traces, the interaction of APC with unmodified cyclin B was transient compared to that with ubiquitylated cyclin B and usually characterized by long intervals between successive binding cycles (fig. S20). To separate the contribution of on-rate and off-rate to the $K_d$ reduction, we measured these kinetic constants by SM methods on cyclin B and related them to the ubiquitylation status of the same substrate. We used a purified system to prevent interference from the proteasome and deubiquitylating enzymes. Ubiquitylation simultaneously decreased $k_{off}$ and increased $k_{on}$; together they accounted for the reduction in $K_d$ (Fig. 4B). A similar result was obtained with securin as the substrate (fig. S22).

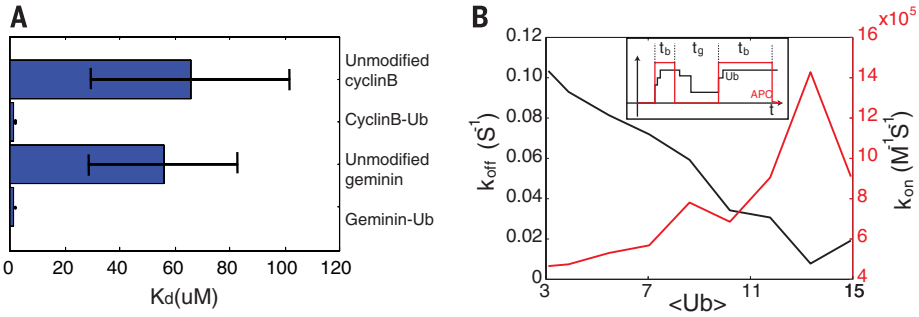



**Fig. 4. Ubiquitylation on a substrate enhances its affinity with the APC.** (**A**) The average $K_d$ values for cyclin B–NT, geminin, and their ubiquitylated forms with the APC in cell extracts. Error bars represent the SD of results from four experiments. (**B**) APC binding rate constants as a function of ubiquitylation on cyclin B–NT in a SM experiment using purified APC components. About 2300 cyclin B–NT molecules were analyzed. Inset: A schematic diagram explaining the calculation of binding rate constants. $k_{off}$ was calculated from APC binding time $t_b$ distribution, and $k_{on}$ was calculated from binding gap time $t_g$. The distribution of $t_b$ and $t_g$ is shown in fig. S21.

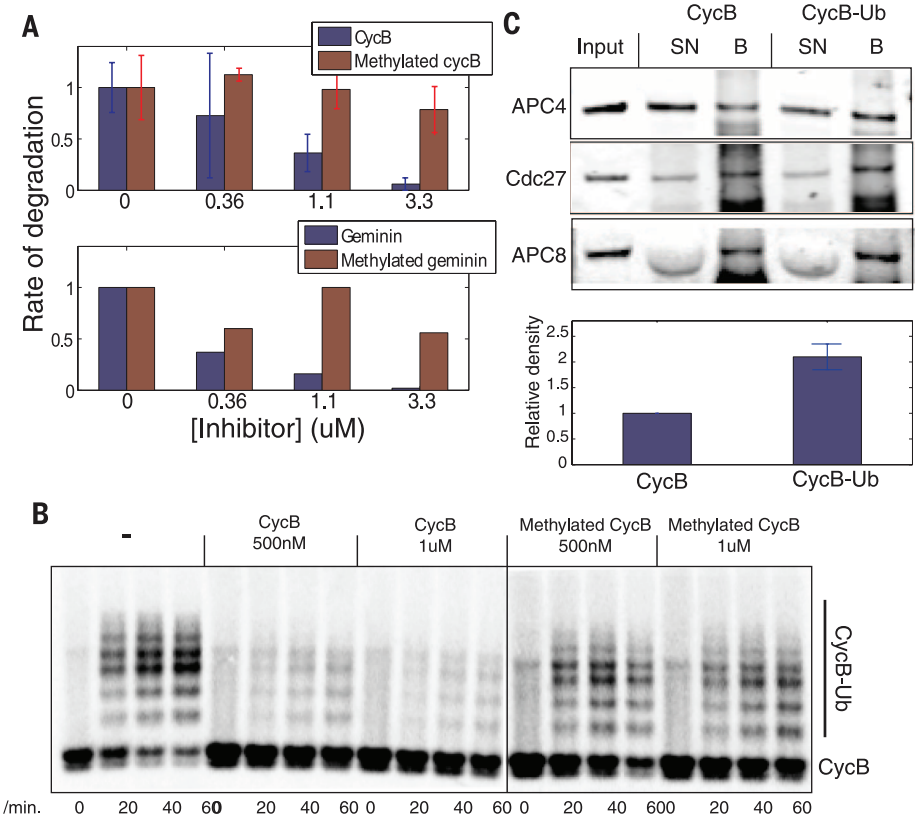



**Fig. 5. Substrate ubiquitylation affects APC binding kinetics in bulk assays.** (**A**) The rate of radioactive geminin degradation in HeLa cell G1 extracts, with indicated concentrations of unlabeled cyclin B–NT, geminin, or their methylated forms as a competitive inhibitor. The first set of data is an experimental replicate with no inhibitor. (**B**) Radioactive cyclin B–NT was ubiquitylated by purified APC in the presence of unlabeled cyclin B–NT or methylated cyclin B–NT as a competitive inhibitor. (**C**) Assay for the effect of substrate ubiquitylation on binding with the APC. Cyclin B–NT (2 μM) was incubated with purified APC reaction components (CycB-Ub) or without E2 (CycB) and was allowed to bind to beads. The amount of APC on beads was quantified by Western blot, shown below. Error bars represent the SD of results from three blots. SN: supernatant, B: beads-bound.

To address potential concerns about the SM experimental design affecting enzyme behaviors at solid-liquid interface, we compared the SM results to conventional assays in bulk solution. Excess APC substrates compete for ubiquitylation by the APC and consequently inhibit proteasomal degradation. Our SM measurements indicate that a protein substrate that is nonubiquitylatable should have a lower affinity for the APC than do ubiquitylated substrates. Lysine methylation completely blocks ubiquitin conjugation on substrates, while preserving structural stability and the D-boxes on the protein. We used cyclin B–NT, geminin, and their methylated forms as competitive inhibitors in degradation assays in cell extracts or as inhibitors of ubiquitylation reactions by purified APC. Indeed, methylated cyclin B and geminin exhibited weaker inhibitory effects in both assays with a low concentration of radioactively labeled cyclin B–NT or geminin as the reporter (Fig. 5, A and B). A consistent but smaller effect was observed with securin as the substrate (fig. S23). We also tested the inhibition of APC by preubiquitylated cyclin B–NT using radioactive cyclin B–NT as the reporter. Preubiquitylated cyclin B–NT was more effective in reducing APC activity in in vitro ubiquitylation assays than unmodified cyclin B–NT, with the difference diminishing at lower concentrations (fig. S24).

To assess directly the effect of ubiquitylation in modulating the interaction with APC, we performed a coprecipitation assay. Ubiquitylated cyclin B–NT interacted more strongly with APC than unmodified cyclin B–NT (Fig. 5C and S25). APC interacted very weakly with substrates, even after substrate ubiquitylation (Fig. 4A). This makes direct binding measurements difficult by conventional methods, which are prone to interference from nonspecific binding (fig. S25). Therefore, we would not expect the binding and competition experiments to quantitatively reproduce the results of SM measurements; nonetheless, both lines of evidence support a stronger interaction of ubiquitylated substrates with the APC.

If ubiquitins conjugated to substrates increase affinity for the APC, increased affinity should facilitate further ubiquitylation on the same substrate molecule during subsequent rounds of binding with the APC. To test this expectation, we fused a single ubiquitin to the N terminus of K48-securin to obtain a homogenous population of a monoubiquitylated substrate. Ub–K48-securin, but not K48-securin alone, was ubiquitylated efficiently in the presence of a high concentration of competitive inhibitors (Fig. 6A). We also tested the effect of conjugated ubiquitins on subsequent ubiquitylation in a reaction without competitive inhibitors by performing a chase experiment on partially ubiquitylated cyclin B–NT. In the middle of an in vitro ubiquitylation reaction on cyclin

B–NT, we added excess hemagglutin (HA)–tagged ubiquitin and measured the distribution of HA-ubiquitin incorporated onto the substrate by Western blot. HA-ubiquitin was first incorporated into high–molecular weight conjugates of the total cyclin B–Ub population. In contrast, it was distributed more uniformly if HA-ubiquitin was added before the reaction (Fig. 6B). It is unlikely that incorporation into the high–molecular weight conjugates is due to partially ubiquitylated cyclin B–APC complex at the point of HA-ubiquitin addition, because APC turnover on cyclin B is much faster than the sampling rate (Fig. 3A). Rather, these results suggest that partially ubiquitylated substrates are the preferred targets for subsequent ubiquitylation by the APC. This conclusion, in conjunction with the effect of conjugated ubiquitins to promote binding with the APC, provides evidence of a positive feedback in the APC reaction.

## Discussion

Specificity in biology is often thought to be conveyed by complementary interactions between chemical species, with the degree of specificity simply reflecting the free-energy difference of binding. However, in certain biological systems, recognition may require more than complementarity. This may be true of substrate specificity in the ubiquitin system, which is largely determined by E3 enzymes. For one E3 enzyme, APC, substrates are characterized by the requirement for short recognition sequences, which are found throughout the genome on many nonsubstrates. Studies of the reactions in this system have been stymied by heterogeneities of the products and challenges in measuring the dynamics. To confront these difficulties, we designed single-molecule fluorescence assays for detecting ubiquitin conjugation and E3-substrate interaction. We used multiple strategies, including minimizing fluorescent nonspecific binding and long exposure times to stabilize the local background fluctuation in total internal reflection fluorescence (TIRF) microscopy due to our use of near-physiological concentrations of free ubiquitin. We could obtain adequate sensitivity and signal-to-noise ratio to observe directly how the APC promoted polyubiquitylation through multiple encounters with a substrate molecule. Using this approach, we examined the mechanism of APC-mediated ubiquitylation reactions in cell extracts and in purified systems and revealed unusual kinetic features; these features provide a basis for understanding how the APC can function efficiently in a complex cytosolic environment with numerous D/KEN-box–containing "decoy" substrates and deubiquitylating enzymes. As discussed below, these experiments suggest a mechanism for how APC can choose substrates among a very large number of proteins with identical recognition sequences.

Substrate molecules are likely to encounter the APC multiple times before degradation. Although we have not accurately measured the rate of proteasomal degradation in the SM assay, the half-life of cyclin B and geminin is about 30 min in bulk assays under our experimental conditions, a time sufficient for multiple encounters to occur. The duration of substrate-APC interactions is very heterogeneous even for molecules with the same number of conjugated ubiquitins, as expected for stochastic processes. Different ubiquitin configurations on the substrate may also contribute to the heterogeneity (*11*). In both extracts and purified systems, we found an initial highly processive ubiquitylation process, involving both multi-monoubiquitylation and chain elongation. We did not resolve the processive phase into single ubiquitin steps. This does not represent a limitation of the SM detection method per se because we could observe stepwise addition of single ubiquitins in certain cases, presumably because of heterogeneity in the reaction rates (fig. S26); rather, it reflected insufficient time resolution due to limitations in the speed of detection.

Our study on a mutant of cyclin B with a single lysine showed a slower ubiquitin transfer rate by the elongation factor Ube2S than the initial ubiquitylation by UbcH10. We expect that Ube2S contributes less to the overall stoichiometry of ubiquitylation in a physiological context where the length of ubiquitin chains is restricted by the action of deubiquitylating enzymes (*11*). However, this small incremental contribution by Ube2S may be critical in selecting the substrate for degradation, because it solely promotes chain elongation.

Single-molecule kinetic studies demonstrate a highly processive initial ubiquitylation for APC-substrates. Furthermore, we find that conjugated ubiquitins bias further ubiquitylation by increasing

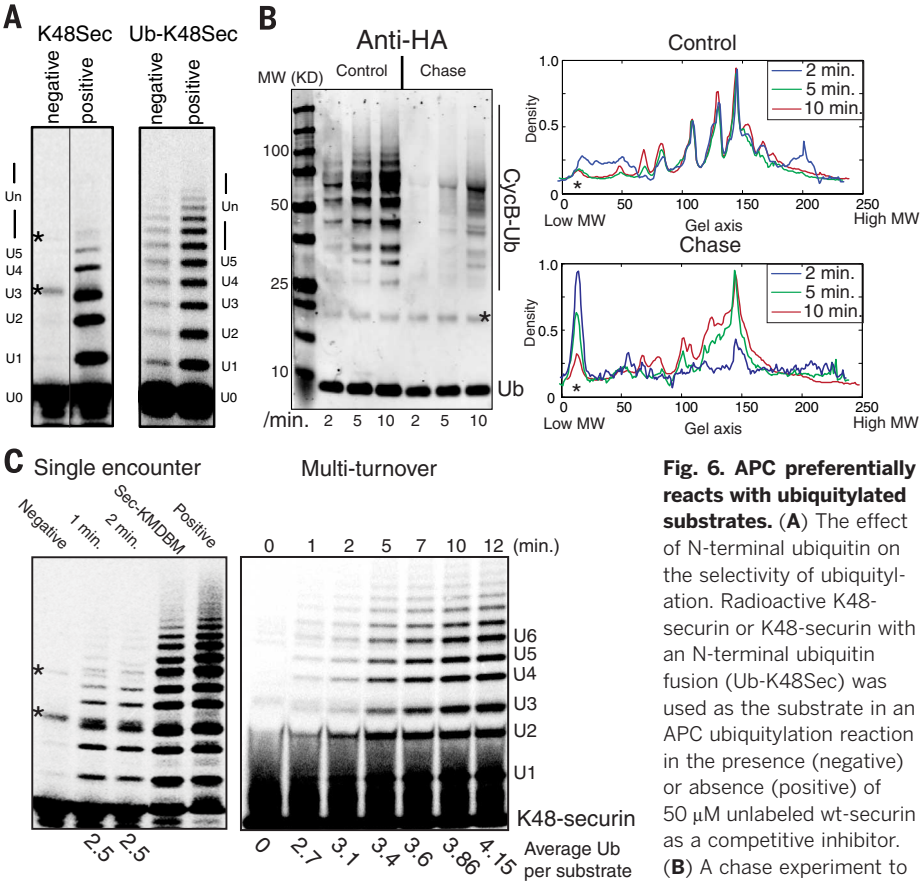



**Fig. 6. APC preferentially reacts with ubiquitylated substrates.** (**A**) The effect of N-terminal ubiquitin on the selectivity of ubiquitylation. Radioactive K48-securin or K48-securin with an N-terminal ubiquitin fusion (Ub-K48Sec) was used as the substrate in an APC ubiquitylation reaction in the presence (negative) or absence (positive) of 50 μM unlabeled wt-securin as a competitive inhibitor. (**B**) A chase experiment to determine the preference of addition of ubiquitin to substrates. Cyclin B–NT (200 nM) was subjected to ubiquitylation by purified APC. After 5 min when ~40% cyclin B–NT had been ubiquitylated, excess HA-ubiquitin was added to the reaction (Chase). In Control, HA-ubiquitin was added at the beginning. Samples were taken at indicated time points and analyzed by Western blot with anti-HA. Line profiles of corresponding lanes at each time point are compared in the right panels. The asterisk (*) indicates a nonspecific band. (**C**) The degree of ubiquitin conjugation in single- and multiple-encounter assays. Radioactive single-lysine K48-securin was used as the substrate in an APC ubiquitylation reaction including 100 nM E2 Ube2S. In the single-encounter assay (left), 50 μM unlabeled wt-securin was added to prevent rebinding of K48-securin to the APC, and sampled after 1 and 2 min. Negative: The competitive inhibitor was added together with labeled substrate. Sec-KMDBM: Using 50 μM wt-securin whose D box and KEN box had been mutated as a competitive inhibitor. Positive: No competitive inhibitor was added. The multi-turnover assay (right) is a standard APC ubiquitylation reaction with no competitive inhibitor. The average number of conjugated ubiquitins on each substrate molecule was quantified using a PhosphorImager.

$k_{on}$ and decreasing $k_{off}$ with the APC. In this way, the chemical reactivity of lysine residues (which themselves do not seem to contribute appreciably to substrate binding) is converted into enhanced affinity with the APC. This would recursively promote the binding and reaction of the substrate with the APC. The decrease in $k_{off}$ may result from ubiquitin interacting directly with the APC. The APC likely contains weak ubiquitin-interacting regions whose identities have not been fully examined; however, in recent studies, ubiquitin has been found to interact with both RING domain and HECT domain of E3 ligases in the E2-ubiquitin-E3 crystal structures (*17*, *18*). Specifically, the RING domain of APC has been shown to contain separate regions interacting with both donor and receptor ubiquitins (*19*). Kinetic studies also show that APC-Ube2S complex can interact and track emerging ubiquitin chains on substrates during elongation (*20*). Modifications of $k_{on}$ have not been typically encountered in biochemistry. In ubiquitylation reactions, an enhanced $k_{on}$ may be a consequence of an increased interaction cross-section between ubiquitylated substrates and the APC, specifically if the latter contains ubiquitin receptors.

The results of our experiments and mathematical modeling suggest a mechanism that would explain the specificity in APC-mediated ubiquitylation. At minimum, specificity should reflect sufficient interaction between enzyme and substrate. The SM measurements indicate that the intrinsic affinity of APC for unmodified substrates is very low, which may reflect the limited size of APC-binding motifs. A low substrate-binding affinity with corresponding high off-rates may be required to ensure the functions of APC. Otherwise, the APC would be completely inhibited by the many D- or KEN-box–containing decoy substrate proteins, the concentration of which reaches about 1 mM in the cell. However, low affinity should result in a low reaction rate and specificity for real substrates, which would result in very low reaction specificity. The SM experiments show that when the APC encounters a real target that contains sufficient ubiquitylatable lysine residues (see below), a wave of processive ubiquitylation occurs, and this leads to higher on-rates and lower off-rates for ubiquitylated proteins with the APC, which makes the ubiquitylated protein a "preferred" target in subsequent binding and reaction cycles (Fig. 7A). Such a positive feedback would amplify the initial weak interaction between the APC and substrate. We termed such process "processive affinity amplification" or "PAA" (Fig. 7A). The pronounced effect of PAA on increasing substrate specificity can be demonstrated with a simple ordinary-differential-equation model simulating substrate-E3 binding, ubiquitylation, deubiquitylation, and degradation steps using physiologically appropriate parameters (Fig. 7B).

PAA should also increase the efficiency of the APC-ubiquitylation system, when it is challenged by deubiquitylating enzymes in the cell. Without PAA, these enzymes would substantially reduce the ubiquitylation and degradation efficiency, especially at high concentrations of substrates (including decoy substrates) (*21*). APC substrate molecules require multiple encounters with the E3 before they accumulate enough ubiquitins for degradation (Fig. 3A and figs. S15 and S16). In the presence of a high concentration of substrates and decoy substrates, there will be increased delays between successive binding cycles, long enough for deubiquitylating enzymes to partially or completely remove conjugated ubiquitin; this would drastically increase the time it takes to accumulate sufficient ubiquitins for degradation. Such an inhibitory effect can be demonstrated in the mathematical model as above, varying the concentration of total substrates (Fig. 7B). At the concentration of total D- or KEN box–containing proteins (about 1 mM) in the cell, APC efficiency for any real substrate drops to less than 1% of its maximum. The PAA mechanism, where partially ubiquitylated substrates bind more quickly and with higher affinity than nonubiquitylated substrates, would increase APC efficiency by focusing the ubiquitylation reactions on a subset of molecules, increasing their complement of conjugated ubiquitin until they are recognized by the proteasome. When we incorporate PAA into the

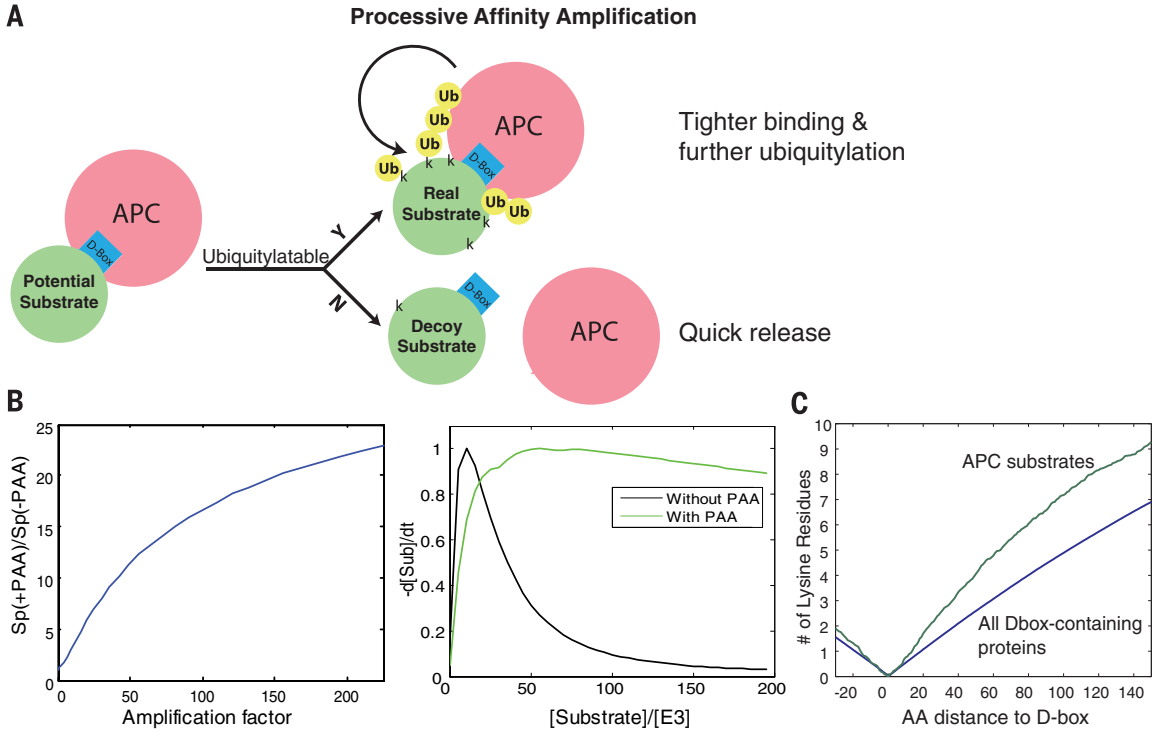



**Fig. 7. Processive affinity amplification enhances the specificity of APC.** (**A**) A schematic diagram showing how the PAA mechanism could enhance the specificity of APC in cytosolic environments. See the text for details. (**B**) Left: The ratio of substrate specificity of the APC in the presence and absence of PAA as a function of the amplification factor, calculated from an ODE model (Methods). Sp: specificity, defined as the ratio of concentration of APC interacting with the real substrates and APC interacting with decoy substrates. The amplification factor represents the maximum fold of $K_d$ reduction due to PAA. Right: The maximum rate of substrate degradation as a function of total substrate concentration, calculated from the ODE model with or without the PAA mechanism. (**C**) The average number of lysine residues within a certain distance ($x$ axis) next to the D box is plotted for 74 published APC substrates from multiple species whose D box has been identified, and for all D-box–containing proteins in human proteome as a comparison.

model, the rate of substrate degradation is much less affected by high substrate concentrations (Fig. 7B). Such "focused" reaction kinetics can also be demonstrated experimentally. In an APC ubiquitylation reaction, the average number of conjugated ubiquitins was more than that obtained in a single encounter assay (using inhibitory concentrations of competitor to prevent substrate-APC reinteraction), suggesting that ubiquitylated species have multiple encounters with the APC. But this was also the case when only a small fraction (5 to 10%) of substrates had reacted (Fig. 6C and fig. S27). In this case, it is unlikely that the small fraction has multiple encounters and reactions with the APC, while the vast majority of substrates does not, unless the APC "focuses" on the small subset by the process of affinity amplification. A similar ordinary differential equation (ODE) model that incorporates the PAA mechanism to simulate this experiment faithfully reproduces the time evolution of the average conjugated ubiquitins, as well as that of each ubiquitylated species (fig. S28).

Information for recognition by the APC must ultimately be encoded somewhere in the sequence or structure of the substrate molecules. The D and KEN boxes provide the initial interaction with the APC. These motifs are not equally accessible by the APC on all proteins, but because their sequences contain argininyl or lysyl side chains, they are likely to be exposed on a protein surface or in disordered regions (*22*). Besides the D and KEN boxes, sufficient lysine residues must be present on substrates to initiate the PAA. Indeed, we found 64% more lysine residues within 60 residues of the D box of published APC substrates, compared to random D-box–containing proteins in the human proteome (*23*) (Fig. 7C). Lysine residues are not equally ubiquitylatable. Ubiquitylatability must certainly be affected by the distance and orientation of these residues to the ubiquitin thioester bond on E2 molecules, and also affected by the p$K_a$ of lysyl side chains, which would be affected by neighboring amino acids (*24*). There are no known sequence motifs for optimized lysine locations, and we might not expect to find them. Therefore, it appears that what determines substrate specificity in APC reactions, after the complementarity achieved by the simple recognition sequences, is the reactivity of lysine residues, likely near the D and KEN boxes. Ubiquitylatable lysine residues alone may not be sufficient to promote high specificity and efficiency for APC substrates in the absence of PAA. The ubiquitin transfer rate on APC substrates dropped to ~1 ubiquitin/minute on average after the initial phase of reaction (Fig. 2C). Without PAA supporting a stronger interaction with the APC, the very low affinity provided by D and KEN boxes with the E3 should not be able to sustain further ubiquitylation for proteasomal degradation in the presence of active deubiquitylation in the cell.

Finding an optimal trade-off of specificity and efficiency is likely to be a central requirement of all biological systems and, therefore, to be under strong selection. It has long been known that mutations in ribosome and DNA polymerase that increase specificity in protein synthesis and in DNA replication decrease the efficiency (or rate) of these processes, whereas mutations that decrease specificity actually increase efficiency (*25*, *26*). Finding the right substrate is basically a search procedure where high on-rates and weak affinity translate into the speed of the search. If the affinity were strong, the APC would spend too much time sampling related sequences on a large number of decoy substrates in the cell. Results from the SM measurements and mathematical modeling explain how the APC may achieve high specificity without sacrificing too much efficiency. The PAA mechanism only increases the affinity of the APC with real substrates, but not with decoy substrates. The resultant stronger binding of the APC with ubiquitylated substrates will not slow down degradation by interfering with proteasome recognition, because even at submicromolar $K_d$, the APC still turns over quickly enough on substrates, compared to the rate of degradation. Consistent with PAA, non-ubiquitylatable cyclin B with lysine residues mutated to arginine when overexpressed does not cause a cell-cycle block in *Schizosaccharomyces pombe* (*27*). This is expected because its interaction with the APC should remain weak, whereas overexpression of a wt–cyclin B, which can be ubiquitylated, inhibits cell-cycle progression. Interaction between APC and K11-linked ubiquitin chains can also be detected in vitro (*28*); however, most substrate-E3 interactions are not stable enough to be detected in the cell. Besides the positive feedback on reaction rates, deubiquitylation and multiple encounters with the APC are essential components in PAA, similar to the requirements in the kinetic proofreading model (*1*).

In the PAA model, APC achieves high substrate specificity with short binding motifs, while maintaining reaction efficiency. Similarly, recognition motifs of many important enzymes, such as the cyclin-dependent kinases (CDKs) and the Polo kinase, are short. For these and many other kinases, phosphorylation often occurs on multiple sites. Phosphorylation, like ubiquitylation, is opposed by hydrolases that reduce the efficiency of the process. In the case of CDK, when associated with the Cks protein, the complex has a higher affinity for substrates that have been phosphorylated by CDK. Thus, the initial phosphorylation could facilitate additional phosphorylation, essentially creating a positive feedback like PAA (*29*). Iterative cycles of binding and reversible modification, found in many systems, may be a common contrivance in biology to increase the specificity of a process, without too much sacrifice in speed.

## Materials and methods

Details of the experimental procedures, statistical analysis, and data sets are provided in the supplementary materials.

### Protein purification, labeling, and modification

Human full-length cyclin B, securin, securin (KMDBM), single-lysine k48-securin, Ub–k48-securin, UbcH10(K119R), N-terminal cyclin B(1–104) from *Xenopus*, single-lysine k64–cyclin B, and geminin from *Xenopus* were expressed in *Escherichia coli* and purified using intein-mediated purification with an affinity chitin-binding tag. For substrates requiring biotinylation, a biotin-containing peptide was ligated to the C terminus of these proteins through protein ligation, with a ~90% ligation efficiency. Biotinylated proteins were reacted with alexa488-maleimide and then purified by gel filtration. Human ubiquitin with a cysteine residue and a His-tag at the N terminus was purified from *E. coli* using cation exchange chromatography and was labeled on cysteine with alexa647-maleimide. HeLa cell G1 extract preparation and APC purification were performed as previously described (*4*). To radioactively label substrates for in vitro ubiquitylation reaction, a PKA site was added to the N terminus and labeled using p33ATP.

### Single-molecule experimental procedures

Glass slides were passivated using 5-kD polyethylene glycol (PEG) + 2.5% 5-kD biotin–PEG in a "clouding-point" solution on amino-silanized slides for 4 hours. We passivated the slides further with 1-kD–PEG–NHS for 1 hour and 10% (w/v) bovine serum albumin (BSA) for 30 min just before the experiment. Streptavidin (0.2 mg/ml) was incubated with the passivated slide for 5 min to immobilize the biotinylated substrates. For experiments with cell extracts, endogenous free ubiquitin was first depleted from the cell extract just before the experiment by incubating with His6-UbcH10 for 20 min at room temperature, followed by Ni-agarose depletion of the conjugates. During this time, 500 nM dylight550–anti-APC4 antibody was added to cell extract to label the APC. Ub-depleted extracts were supplemented with 3 μM alexa647-Ub. A desired amount of biotinylated substrate was mixed with cell extract or in vitro reaction buffer and incubated with a functionalized slide to allow binding before data acquisition at room temperature. Time-lapse data were analyzed with custom-built software in Matlab.

### Degradation and ubiquitylation assay

For degradation assay, 1/10 volume of radioactive substrate from in vitro translation system was mixed with functional cell extract supplemented with energy mix and 10 μM ubiquitin. Reactions were performed at room temperature. For in vitro ubiquitylation assay, PKA-labeled substrates were mixed with 4 to 20 nM APC, 100 nM E1, 2 μM UbcH10, 10 to 30 μM ubiquitin, BSA (2 mg/ml), and the energy-regenerating system in UBAB buffer [25 mM Tris-HCL (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$], incubated at 30°C. Results were quantified using a PhosphorImager.

## REFERENCES AND NOTES

1. J. J. Hopfield, Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 4135–4139 (1974). doi: 10.1073/pnas.71.10.4135; pmid: 4530290
2. J. Ninio, Kinetic amplification of enzyme discrimination. *Biochimie* **57**, 587–595 (1975). doi: 10.1016/S0300-9084(75)80139-8; pmid: 1182215

3. T. W. McKeithan, Kinetic proofreading in T-cell receptor signal transduction. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5042–5046 (1995). doi: 10.1073/pnas.92.11.5042; pmid: 7761445
4. M. Rape, S. K. Reddy, M. W. Kirschner, The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell* **124**, 89–103 (2006). doi: 10.1016/j.cell.2005.10.032; pmid: 16413484
5. J. M. Peters, The anaphase-promoting complex: Proteolysis in mitosis and beyond. *Mol. Cell* **9**, 931–943 (2002). doi: 10.1016/S1097-2765(02)00540-3; pmid: 12049731
6. C. Lindon, J. Pines, Ordered proteolysis in anaphase inactivates Plk1 to contribute to proper mitotic exit in human cells. *J. Cell Biol.* **164**, 233–241 (2004). doi: 10.1083/jcb.200309035; pmid: 14734534
7. M. Glotzer, A. W. Murray, M. W. Kirschner, Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132–138 (1991). doi: 10.1038/349132a0; pmid: 1846030
8. C. M. Pfleger, M. W. Kirschner, The KEN box: An APC recognition signal distinct from the D box targeted by Cdh1. *Genes Dev.* **14**, 655–665 (2000). pmid: 10733526
9. D. Barford, Structural insights into anaphase-promoting complex function and mechanism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 3605–3624 (2011). doi: 10.1098/rstb.2011.0069; pmid: 22084387
10. Y. Merbl, M. W. Kirschner, Large-scale detection of ubiquitination substrates using cell extracts and protein microarrays. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2543–2548 (2009). doi: 10.1073/pnas.0812892106; pmid: 19181856
11. D. S. Kirkpatrick *et al*., Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology. *Nat. Cell Biol.* **8**, 700–710 (2006). doi: 10.1038/ncb1436; pmid: 16799550
12. T. Ravid, M. Hochstrasser, Autoregulation of an E2 enzyme by ubiquitin-chain assembly on its catalytic residue. *Nat. Cell Biol.* **9**, 422–427 (2007). doi: 10.1038/ncb1558; pmid: 17310239
13. W. Li, D. Tu, A. T. Brunger, Y. Ye, A ubiquitin ligase transfers preformed polyubiquitin chains from a conjugating enzyme to a substrate. *Nature* **446**, 333–337 (2007). doi: 10.1038/nature05542; pmid: 17310145
14. N. W. Pierce, G. Kleiger, S. O. Shan, R. J. Deshaies, Detection of sequential polyubiquitylation on a millisecond timescale. *Nature* **462**, 615–619 (2009). doi: 10.1038/nature08595; pmid: 19956254
15. C. W. Carroll, D. O. Morgan, The Doc1 subunit is a processivity factor for the anaphase-promoting complex. *Nat. Cell Biol.* **4**, 880–887 (2002). doi: 10.1038/ncb871; pmid: 12402045
16. A. Williamson *et al*., Regulation of ubiquitin chain initiation to control the timing of substrate degradation. *Mol. Cell* **42**, 744–757 (2011). doi: 10.1016/j.molcel.2011.04.022; pmid: 21700221
17. A. Plechanovová, E. G. Jaffray, M. H. Tatham, J. H. Naismith, R. T. Hay, Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115–120 (2012). doi: 10.1038/nature11376; pmid: 22842904
18. H. B. Kamadurai *et al*., Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT (NEDD4L) complex. *Mol. Cell* **36**, 1095–1102 (2009). doi: 10.1016/j.molcel.2009.11.010; pmid: 20064473
19. N. G. Brown *et al*., Mechanism of polyubiquitination by human anaphase-promoting complex: RING repurposing for ubiquitin chain assembly. *Mol. Cell* **56**, 246–260 (2014). doi: 10.1016/j.molcel.2014.09.009; pmid: 25306923
20. A. Kelly, K. E. Wickliffe, L. Song, I. Fedrigo, M. Rape, Ubiquitin chain elongation requires E3-dependent tracking of the emerging conjugate. *Mol. Cell* **56**, 232–245 (2014). doi: 10.1016/j.molcel.2014.09.010; pmid: 25306918
21. H. J. Meyer, M. Rape, Processive ubiquitin chain formation by the anaphase-promoting complex. *Semin. Cell Dev. Biol.* **22**, 544–550 (2011). doi: 10.1016/j.semcdb.2011.03.009; pmid: 21477659
22. M. Fuxreiter, P. Tompa, I. Simon, Local structural disorder imparts plasticity on linear motifs. *Bioinformatics* **23**, 950–956 (2007). doi: 10.1093/bioinformatics/btm035; pmid: 17387114
23. Z. Liu *et al*., GPS-ARM: Computational analysis of the APC/C recognition motif by predicting D-boxes and KEN-boxes. *PLOS ONE* **7**, e34370 (2012). doi: 10.1371/journal.pone.0034370; pmid: 22479614
24. L. A. Highbarger, J. A. Gerlt, G. L. Kenyon, Mechanism of the reaction catalyzed by acetoacetate decarboxylase. Importance of lysine 116 in determining the pKa of active-site lysine 115. *Biochemistry* **35**, 41–46 (1996). doi: 10.1021/bi9518306; pmid: 8555196
25. M. J. Bessman, N. Muzyczka, M. F. Goodman, R. L. Schnaar, Studies on the biochemical basis of spontaneous mutation. II. The incorporation of a base and its analogue into DNA by wild-type, mutator and antimutator DNA polymerases. *J. Mol. Biol.* **88**, 409–421 (1974). doi: 10.1016/0022-2836(74)90491-4; pmid: 4616089
26. K. Bohman, T. Ruusala, P. C. Jelenc, C. G. Kurland, Kinetic impairment of restrictive streptomycin-resistant ribosomes. *Mol. Gen. Genet.* **198**, 90–99 (1984). doi: 10.1007/BF00328706; pmid: 6394968
27. H. Yamano, C. Tsurumi, J. Gannon, T. Hunt, The role of the destruction box and its neighbouring lysine residues in cyclin B for anaphase ubiquitin-dependent proteolysis in fission yeast: Defining the D-box receptor. *EMBO J.* **17**, 5670–5678 (1998). doi: 10.1093/emboj/17.19.5670; pmid: 9755167
28. M. L. Matsumoto *et al*., K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. *Mol. Cell* **39**, 477–484 (2010). doi: 10.1016/j.molcel.2010.07.001; pmid: 20655260
29. M. Kõivomägi *et al*., Cascades of multisite phosphorylation control Sic1 destruction at the onset of S phase. *Nature* **480**, 128–131 (2011). doi: 10.1038/nature10560; pmid: 21993622



## ACKNOWLEDGMENTS

We thank R. King for providing plasmids and W. Ma, L. Bai, M. Rape, and R. King for commenting on the manuscript. We are grateful to the Nikon imaging center at Harvard Medical School and in particular to J. Waters for technical support; Y.L. is a Damon Runyon Cancer Research Fellow and a Lallage Feazel Wall Fellow. We acknowledge the continuing support of the National Institute of General Medical Sciences (grant no. 5R01GM039023-26).


## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/1248737/suppl/DC1
Materials and Methods
Figs. S1 to S38
References (*30*–*34*)
Movies S1 to S8

20 November 2013; accepted 21 February 2015
10.1126/science.1248737

RESEARCH ARTICLE SUMMARY

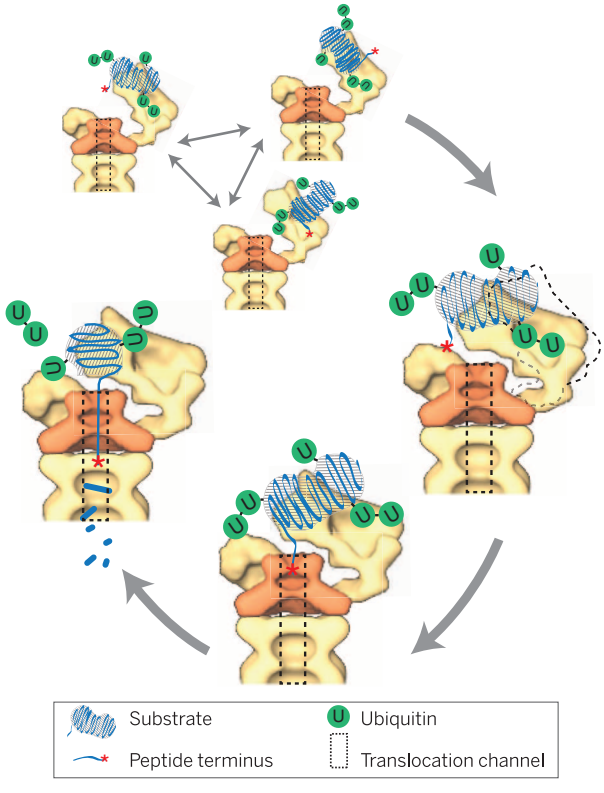
BIOCHEMISTRY

# Substrate degradation by the proteasome: A single-molecule kinetic analysis

Ying Lu, Byung-hoon Lee, Randall W. King, Daniel Finley, Marc W. Kirschner*

**INTRODUCTION:** Protein degradation, mediated by the ubiquitin-proteasome system (UPS), plays a critical and complementary role to transcription, splicing, and translation in the control of gene expression. Effective regulation of the UPS relies on the specificity of substrate recognition, which is conferred largely by the upstream ubiquitylation process. As a specific example, the anaphase-promoting complex (APC) acting on a series of substrates promotes waves of ubiquitylation and degradation, leading to key transitions in the cell cycle. However, in this cascade the proteasome itself also plays a role in the specificity of degradation. There have been various efforts to characterize this role, leading to a commonly held view that a substrate protein must be conjugated with a chain of at least four ubiquitins to be recognized by the proteasome. Yet mass spectrometry studies show that ubiquitin chains on APC substrates (such as cyclin B) contain, on average, only two ubiquitins. But these can have complex ubiquitin configurations created by the multiplicity of ubiquitylated lysine residues. Therefore, the tetraubiquitin chain selection rule may not be generally applicable, and the mechanism by which the proteasome recognizes substrates is still shrouded in mystery.

**RATIONALE:** The multiple lysines on a substrate and on ubiquitin itself generate a large number of possible ubiquitin configurations. To understand the "ubiquitin code" that must be read by the proteasome and converted into a rate of substrate degradation, we examined the kinetics of degradation with defined ubiquitin configurations by conjugating preformed ubiquitin chains on substrate molecules. Further, to reveal the molecular basis through which some ubiquitin configurations promote more efficient degradation than others, we investigated the degradation process using single-molecule (SM) methods that are capable of identifying transient intermediates and measuring their kinetic parameters and sensitivity to ubiquitin configurations.




**Key transitions in the degradation of ubiquitylated substrates by the proteasome.** A polyubiquitylated substrate molecule explores multiple configurations on the proteasome through stochastic binding. Rearrangements of proteasomal subunits, driven by the binding and hydrolysis of adenosine triphosphate, allow a deeper engagement of the substrate with the proteasome, when there are appropriate ubiquitin chain structures on the substrate. As a result of this engagement, the substrate or its terminus is moved closer to the substrate entry port, promoting expeditious initiation of translocation and ensuing degradation.

**RESULTS:** Contrary to the tetraubiquitin chain selection rule, we find that for APC substrates with multiple ubiquitylated lysine residues, diubiquitin chains are more efficient than tetraubiquitin chains in promoting degradation, given the same number of conjugated ubiquitins. Ubiquitin chains are essential for degradation of most substrates. Nevertheless, a multiple monoubiquitylated form of securin, a regulator of chromatid separation, interacts with the proteasome as strongly as securin containing the same number of ubiquitins grouped in chains. By dissecting the degradation process using SM methods, we find that ubiquitin chain structures on substrates promote the passage of a bound substrate into the translocation channel on the proteasome.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1250834

**CONCLUSION:** This systematic study of synthetically constructed ubiquitylated substrates with defined configurations revealed no simple length threshold for ubiquitin chains for degradation by the proteasome. A distributed array of short ubiquitin chains, as appears naturally on APC substrates, is a superior and perhaps optimal signal for degradation; this conclusion will most likely extend to substrates of other E3 ligases. The rate of degradation is an aggregate of two sequential processes: substrate binding and kinetic postbinding events. In the past, it was widely assumed that ubiquitin chains mostly promoted binding to the proteasome. Our SM studies demonstrate that the strength of interaction with the proteasome, for substrates with multiple ubiquitylated lysines, is largely determined by the total number of ubiquitins and is less sensitive to ubiquitin chain configurations. For most substrates, binding alone is not sufficient for degradation. Rather, degradation depends strongly on a process that initiates passage into the substrate translocation channel; this transition, in contrast to binding, is determined by the particular configuration of ubiquitin chains. ■

## RELATED ITEMS IN *SCIENCE*

Y. Lu, W. Wang, M. W. Kirschner, Specificity of the anaphase-promoting complex: A single-molecule study. *Science* **348**, 1248737 (2015).

D. Komander, Details of destruction, one molecule at a time. *Science* **348**, 183–184 (2015).

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: marc@hms.harvard.edu
Cite this article as Y. Lu *et al.*, *Science* **348**, 1250834 (2015).
DOI: 10.1126/science.1250834

RESEARCH ARTICLE

BIOCHEMISTRY

# Substrate degradation by the proteasome: A single-molecule kinetic analysis

Ying Lu,[1] Byung-hoon Lee,[2] Randall W. King,[2] Daniel Finley,[2] Marc W. Kirschner[1]*

**To address how the configuration of conjugated ubiquitins determines the recognition of substrates by the proteasome, we analyzed the degradation kinetics of substrates with chemically defined ubiquitin configurations. Contrary to the view that a tetraubiquitin chain is the minimal signal for efficient degradation, we find that distributing the ubiquitins as diubiquitin chains provides a more efficient signal. To understand how the proteasome actually discriminates among ubiquitin configurations, we developed single-molecule assays that distinguished intermediate steps of degradation kinetically. The level of ubiquitin on a substrate drives proteasome-substrate interaction, whereas the chain structure of ubiquitin affects translocation into the axial channel on the proteasome. Together these two features largely determine the susceptibility of substrates for proteasomal degradation.**

The propensity of a protein for degradation is largely encoded in its state of ubiquitylation (*1*, *2*). The ubiquitylation process results in highly diverse configurations of ubiquitin chains on target proteins (*3*). The 26*S* proteasome recognizes ubiquitylated substrates and degrades them into short peptides (*4*). Tests of defined configurations of ubiquitins on substrates for their influence on recognition and degradation showed that a substrate protein must generally be conjugated with a chain of at least four ubiquitins to interact tightly enough with the proteasome for degradation (*5*). However, no compelling evidence supports the existence of a tetraubiquitin chain receptor on the proteasome. Alternative explanations for the tetraubiquitin chain selection rule mostly rely on the geometric distance between a pair of proteasomal subunits to gauge the length of a ubiquitin chain (*6*, *7*). The requirement for a tetraubiquitin chain for degradation also differs from substrate to substrate, without a predictable relationship to a substrate's function or structure (*8*–*11*).

Exit from mitosis and passage into the $G_1$ phase of the cell cycle requires ubiquitylation mediated by the anaphase-promoting complex (APC) (*12*). Substrates (such as cyclin B, securin, and geminin) typically contain multiple lysine residues, to which ubiquitin moieties are conjugated, providing a very large number of possible combinations of ubiquitin chain configurations (*3*, *8*, *13*, *14*). Mass spectrometry studies indicate that ubiquitin chains on cyclin B molecules generated in reconstituted reactions contain, on average, only two ubiquitins (*3*). Moreover, multi-monoubiquitylation (referring to ubiquitylation on multiple lysine residues without chain formation) on cyclin B leads to efficient degradation. Multiple ubiquitylation sites are commonly found on substrates of other E3 enzymes (*15*, *16*); at least 56% of ubiquitylated human proteins contain more than one ubiquitylation site, though the relevance of these ubiquitylations to protein degradation has not been fully demonstrated (*17*).

Ubiquitins conjugated to the substrate promote interaction with the proteasome; however, binding by itself is not sufficient to initiate degradation. Several known proteasome-interacting proteins (Usp14, Rad23, etc.) are stable (*18*), and autoubiquitylated Cdc34 is not degraded, despite its strong interaction with the proteasome (*19*). To facilitate initiation of substrate translocation, it has been proposed that substrates must have an unstructured terminal region of at least 30 amino acids (*20*); multiple steps occur in the proteasomal degradation process, as suggested by cryo–electron microscopy (cryo-EM) structures (*6*, *7*, *21*). Before or coincident with peptide translocation, conjugated ubiquitins are removed by deubiquitylating enzymes (DUBs) Rpn11, Usp14, and Uch37 on the proteasome (*22*). Whereas Usp14 and Uch37 may have an editing role to tune the rate of proteasomal degradation, they are dispensable for the degradation process. The DUB Rpn11, by contrast, is required for efficient proteasomal activity. Rpn11 is located close to the substrate entry port, where it removes ubiquitin chains en bloc from the translocating peptide (*6*, *23*). It is unclear how the proteasome might use these features to establish selectivity in substrate recognition.

To understand the requirements for efficient protein degradation, we examined the kinetics of degradation of substrates with defined ubiquitin configurations. We found that, for APC substrates with multiple ubiquitylated lysine residues, tetraubiquitin chains were not required for efficient degradation. Rather, given the same number of total ubiquitins on a substrate molecule, diubiquitin chains were more efficient than tetraubiquitin chains in promoting degradation. To elucidate the molecular basis through which some ubiquitin configurations promote more efficient degradation than others, we investigated the intermediate steps in the degradation pathway using single-molecule (SM) methods. For substrates containing multiple lysines, the strength of interaction with the proteasome was determined largely by the total number of ubiquitins and was less sensitive to the ubiquitin configuration. However, substrate binding alone was not sufficient for rapid degradation for most substrates. Rather, degradation depended strongly on the initiation rate of passage into the substrate translocation channel, and this transition was promoted by the presence of ubiquitin chain structures on substrates.

## Results

### *Degradation of defined multiple ubiquitylated substrates*

Most, if not all, APC substrates contain multiple ubiquitylatable lysine residues. To reduce the complexity of these mixed configurations, we sought to control the number, length, and linkage of conjugated ubiquitin chains on wild-type (WT) substrates. Purified APC was used as the ubiquitin ligase to conjugate preformed di- or tetraubiquitin chains of $Lys^{48}$ (K48) linkage, which support most proteasome-mediated degradation in cells (*24*) (Fig. 1A). The chains were methylated on lysines before conjugation to prevent secondary elongation. We then separated the reaction products having different numbers of conjugated ubiquitins by electrophoresis. This approach limits the heterogeneity of ubiquitin configurations on the substrate to combinations of sites accepting a known number of defined ubiquitin chains. We measured the degradation rate of the ubiquitylated substrate for each electrophoretically resolved species by exposure to purified human 26*S* proteasomes that were free of reversibly associating ubiquitin receptors, such as Rad23 (Fig. 1A). Success of this strategy required the absence of interconversion of different ubiquitylated species generated by partial deubiquitylation by the proteasome; this was ensured by removing the DUB Usp14 on the proteasome by salt wash (*25*). Usp14-free proteasomes are known to efficiently degrade substrates, such as cyclin B and Sic1, without generating partially deubiquitylated products (fig. S1) (*8*, *25*, *26*). Another DUB on the proteasome, Uch37, does not appreciably deubiquitylate the substrates used here (*25*). Controls using a general DUB inhibitor (not active against Rpn11) further confirmed that DUB-driven interconversion of ubiquitylated species was unlikely (fig. S2).

We analyzed the rates of degradation of ubiquitylated securin, geminin, and cyclin B–NT (N-terminal fragment from cyclin B), each with

[1]Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA.
[2]Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA.
*Corresponding author. E-mail: marc@hms.harvard.edu

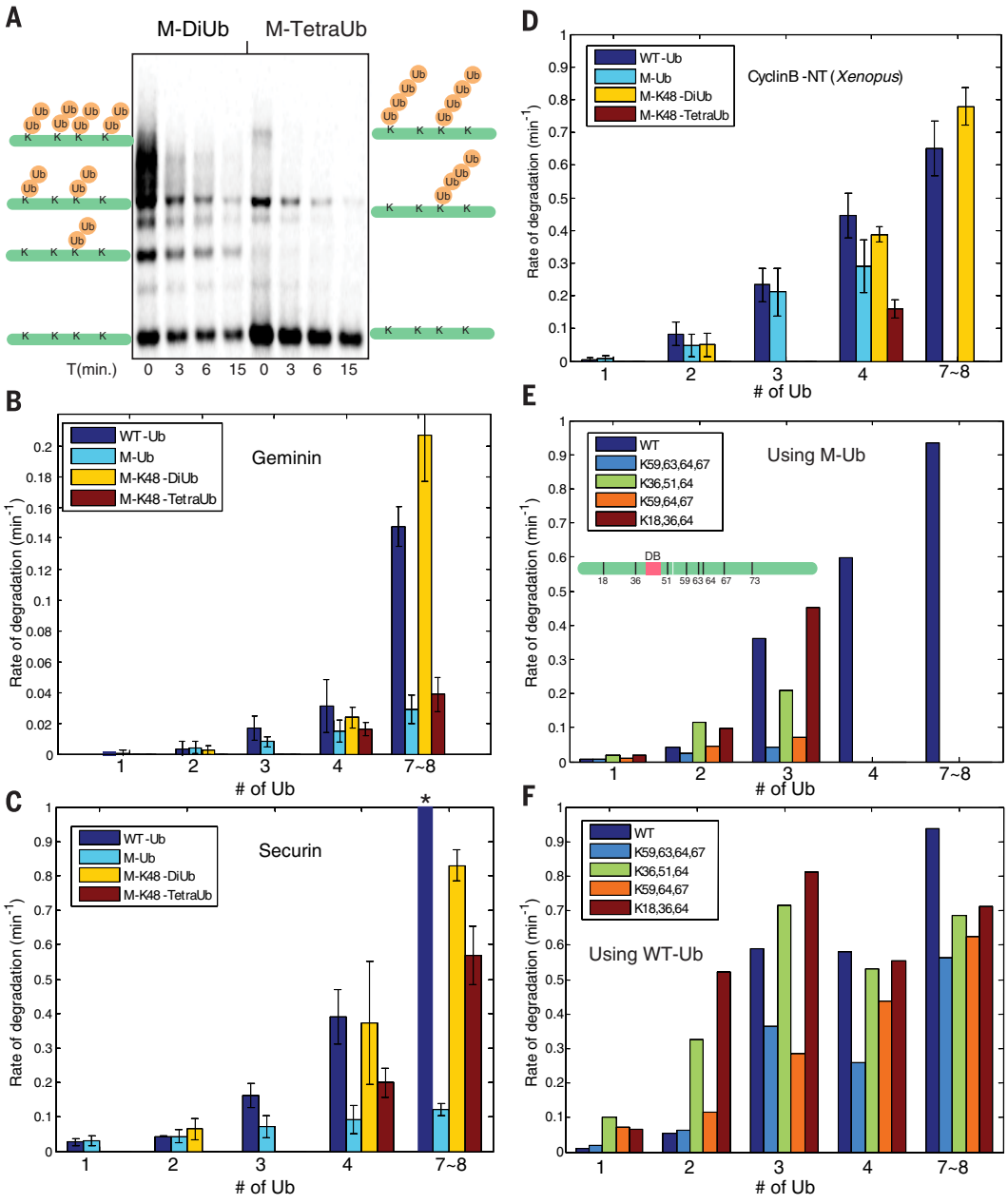



**Fig. 1. Quantitative degradation assay.** (**A**) The assay strategy. Preformed, methylated ubiquitin chains were conjugated to PKA-labeled securin using purified APC and E2 UbcH10. After reaction, the product was subject to degradation by purified 26*S* human proteasome. The decay constant for each ubiquitylated species, separated on a gel, was measured from a time (T) series. Signal from unmodified substrates (Ub0) was used as a control for loading and nonspecific dephosphorylation, which is the main reason for the decrease of Ub0 signals (Materials and methods section). Ub, ubiquitin; K, lysine residue; M-DiUb, methylated diubiquitin chain; M-TetraUb, methylated tetraubiquitin chain. (**B** to **D**) 160 nM geminin, securin, and cyclin B–NT (*Xenopus*) were ubiquitylated using indicated constructs of ubiquitin. Their rates of degradation by 3 nM of purified human 26S proteasome were measured and shown as a function of total ubiquitins per substrate molecule. Error bars represent the SD of three experimental replicates. The asterisk in (C) indicates that the rate for this species is 1.4. The lack of data for certain species is due either to these species not having been tested or to their signals being too weak to quantify. (**E** and **F**) Human cyclin B–NT mutants carrying lysines only at indicated positions were ubiquitylated with either methylated ubiquitin (M-Ub) or WT Ub and tested in a quantitative degradation assay. Original autoradiography for retrieving the rate information is compiled in fig. S8. In (E), the inset shows the location of the D-box (DB) on cyclin B–NT and relevant ubiquitylatable lysine residues identified by mass spectrometry.

a known number of ubiquitin chains of defined length. We incubated these substrates with the 26*S* proteasome at a nonlimiting concentration (*25*) (fig. S3). The measured degradation rates were similar to rates observed in cells (*26*, *27*). The degradation rates were substrate-dependent, even for substrates with the same number of conjugated ubiquitins. For example, highly ubiquitylated securin was degraded fastest, followed by cyclin B and geminin (Fig. 1, B to D). Both methylated K48-diubiquitin chains and WT ubiquitins promoted efficient proteasomal degradation. Conjugation with K48-diubiquitin chains supported a higher rate of degradation than K48-tetraubiquitin chains when the same number of total ubiquitins were conjugated to a substrate molecule (Fig. 1, B to D). The slower degradation associated with tetraubiquitin chains was not due to methylation (fig. S4) nor to potential competition from free tetraubiquitin chains at the experimental concentration, as multidiubiquitylated securin was degraded at the same rate, even after free tetraubiquitin chains were added to the degradation reaction (fig. S4). Faster degradation was observed with diubiquitin chains, even when the two chains were closely apposed (fig. S5); this also held for K11-linked chains (fig. S6). In contrast to cyclin B, multi-monoubiquitylated securin and geminin, generated with either methylated ubiquitin or lysine-free ubiquitin (Ub0K) precluding chain formation, were degraded very slowly compared with constructs with the same ubiquitin stoichiometry but containing chains (Fig. 1, B and C, and fig. S7).

The distribution of ubiquitylated lysines on the substrate may set the degradation rate. Because our method to control ubiquitin configurations does not constrain which lysine residues receive the ubiquitin chains, we explicitly tested the contribution of the sites of ubiquitylation to the degradation rate. Multi-monoubiquitylated cyclin B–NT was degraded as efficiently as WT Ub–conjugated cyclin B–NT (*8*) (Fig. 1D), perhaps because of the specific configurations of lysine targets on cyclin B. We made cyclin B–NT mutants that contain only three or four lysine residues at their original locations, with other lysines mutated to arginines, and studied rates

of degradation of the multi-monoubiquitylated products of those mutants. A clear pattern emerged: Given the same number of total conjugated ubiquitins, the degradation rate strongly depended on the position of the most N-terminal, ubiquitylated lysine residue; the closer to the N terminus, the higher the rate of degradation (Fig. 1E). This pattern was not explained by some mutants being inherently nondegradable, because all mutants were degraded similarly when conjugated to WT Ub at high ubiquitin multiplicity (Fig. 1F) (Ub = 7 to 8). Therefore, moving lysine residues away from the N terminus appeared to progressively convert cyclin B into a substrate whose degradation was more similar to that of securin and geminin. SM studies presented below provide mechanistic insights into these properties.

### *Single-molecule studies of the kinetics of proteasomal degradation*

#### *Proteasome-substrate interactions*

To understand the molecular steps that distinguish different ubiquitin configurations, we monitored the interactions of single ubiquitylated substrate molecules with the proteasome. Purified 26*S* proteasomes from human 293 cells were immobilized onto slides with an antibody to the core 20*S* proteasome (Fig. 2A); the surface was passivated with polyethylene glycol (PEG) and albumin to reduce nonspecific binding. To correlate the SM behavior with the extent of ubiquitin conjugation, each ubiquitin molecule was chemically labeled with a DyLight 550 fluorophore at the N terminus. Fluorescent labeling of ubiquitin had no measurable effect on the kinetics of ubiquitylation and degradation in bulk reactions (fig. S9). After ubiquitylation, the total fluorescence intensity of a substrate molecule was measured by total internal reflection fluorescence (TIRF) microscopy, from which the number of conjugated ubiquitins was calculated. Accuracy and linearity of this method were assessed and confirmed by photobleaching experiments, a process that randomly inactivates single fluorophores (*28*). Background fluctuation was less than 0.2 ubiquitin level; more than 90% of substrate-binding events can be identified with less than 30% uncertainty in measuring the number of conjugated ubiquitins (fig. S10). This uncertainty was principally due to residual uneven illumination.

Ubiquitylated substrates transiently interacted with the 26*S* proteasome. Only background levels

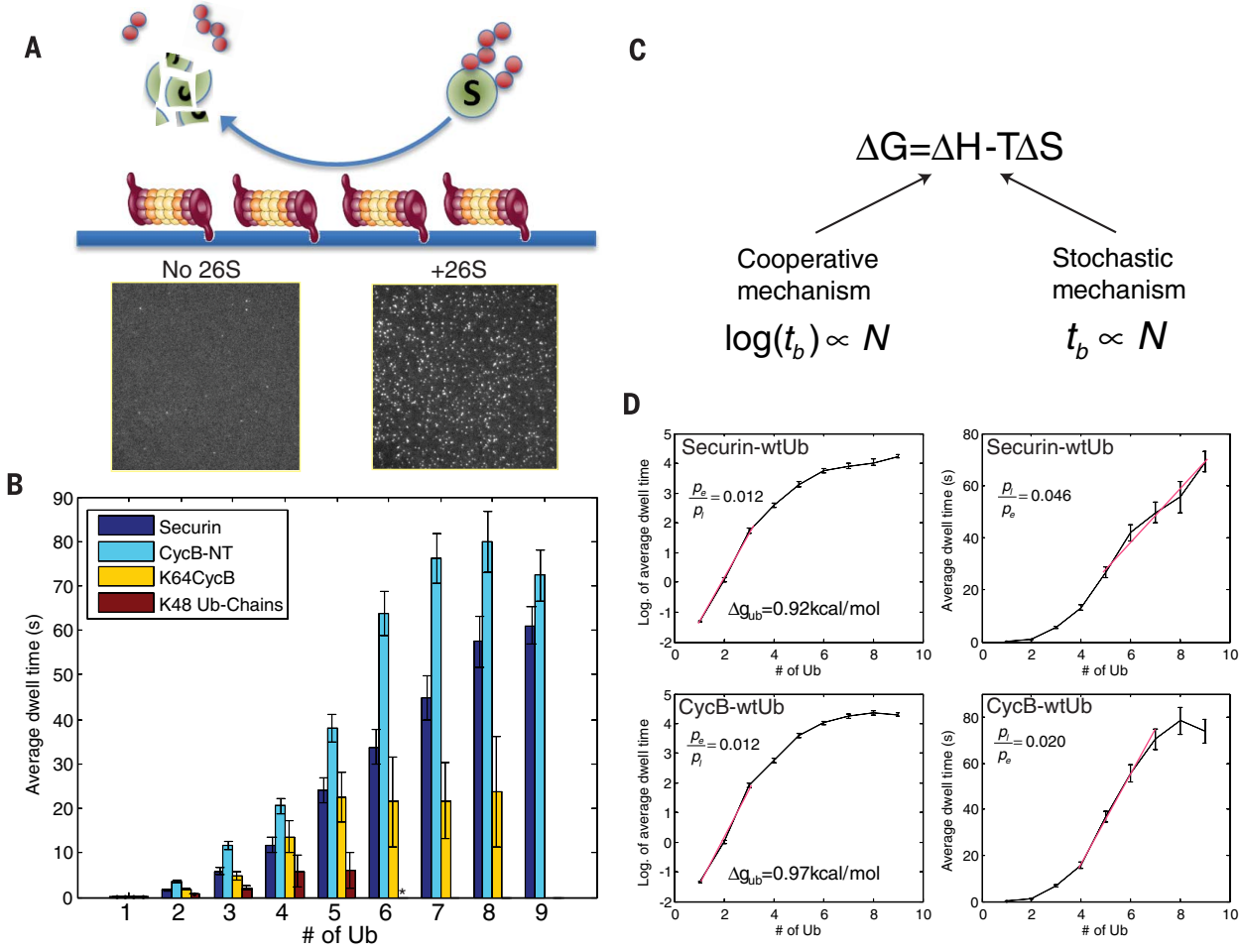



**Fig. 2. Proteasome-substrate interaction kinetics by the SM method.** (**A**) Schematics showing the experimental design, where purified 26*S* proteasome was immobilized on passivated coverslip using anti-20*S* antibody. Ubiquitin was fluorescently labeled and conjugated to substrates in solution. Sample pictures capture ubiquitin signals on the surface with either 26*S* proteasome (+26*S*) or antibody only (no 26*S*). (**B**) Average dwell time on the proteasome for different substrates (or free Ub chain) with varying numbers of conjugated ubiquitins, measured by the SM method. The distributions of individual dwell times are shown in fig. S15. Error bars represent SD of the mean. The asterisk indicates that data for chains of 6 to 9 ubiquitins are not shown due to insufficient events. (**C**) Cooperative and stochastic mechanisms affect binding enthalpy and entropy, respectively. By each mechanism, the expected relationship between dwell time $t_b$ and the number of ubiquitins $N$ is shown below. ΔG, change in Gibbs free energy; ΔH, change in enthalpy; T, temperature; ΔS, change in entropy. (**D**) Dwell time on the proteasome (right), or its logarithm (left), for securin and cyclin B versus the number of conjugated ubiquitins. Red lines show linear fitting. The ratio of $P$ values by fitting the red segment (in linear scale) with either a linear model ($p_l$) or an exponential model ($p_e$) is shown for each plot. $\Delta g_{ub}$ is the binding free energy per ubiquitin on the proteasome, calculated from the slope of the binding curve.

of binding were observed, either if the 20*S* proteasome was substituted for the 26*S* proteasome or if the substrate was omitted (fig. S11). Interaction with the proteasome requires a hydrophobic patch spanning $Leu^{8}$-$Ile^{44}$-$Val^{70}$ on ubiquitin and is compromised by substitutions for $Ile^{44}$ (*29*). In the SM assay, this mutation abrogated the binding of fluorescent ubiquitin to the proteasome (figs. S11 and S12). These results indicate that the interaction occurred predominantly at ubiquitin receptors on the 26*S* proteasome; the kinetics depended on the amount of ubiquitylation of the substrate and was insensitive to excitation laser intensities, excluding low-level fluorophore photobleaching as a complicating factor (figs. S13 and S14). To minimize systematic errors, each experiment and its controls were performed on the same slide with the same batch of proteasome and substrates.

To further validate the SM assay, we compared the affinity of ubiquitylated substrates with the proteasome using SM methods to published results measured in bulk assays. We determined the average dwell time of ubiquitylated cyclin B, securin, cyclin B with a single lysine residue (K64–cyclin B), or free ubiquitin chains on the proteasome, as a function of the total number of conjugated ubiquitins (Fig. 2B). At the same total number of ubiquitins, substrate-anchored ubiquitin chains interacted more strongly with the proteasome than did free chains, consistent with published results (*5*).

Higher ubiquitin stoichiometry consistently led to longer dwell times, but the quantitative relationship between dwell time and the amount of ubiquitin unexpectedly differed from substrate to substrate. Thus, although the dwell time for free ubiquitin chains or K64–cyclin B carrying a single ubiquitin chain plateaued at four to five ubiquitins, the dwell time for multi-lysine substrates, such as cyclin B and securin, increased continually as more ubiquitins were added. In a conventional competition assay, the $K_i$ of a tetraubiquitin chain is 170 nM (*5*) [$K_i \approx K_d = k_{off}/k_{on}$ ($K_i$, inhibition constant; $K_d$, dissociation constant; $k_{off}$, off rate; $k_{on}$, on rate)]. The $K_d$ value measured by the SM method is 210 ± 60 nM, given an estimated $k_{on} = 8.5 \times 10^5\ M^{-1}\ s^{-1}$ (*30*). A $Lys^{6}$→$Ala^{6}$ (K6A) mutation on ubiquitin (UbK6A) weakens the interaction of ubiquitylated substrates with the proteasome (*29*, *31*). Consistent with this, the dwell time of UbK6A-securin on the proteasome was shorter than that of WT Ub–securin in the SM assay (fig. S16). The interaction between substrate backbone and the proteasome appears to be minimal, because the dwell time for substrates conjugated with a single ubiquitin was very short (dwell time < 300 ms) (Fig. 2B).

The dwell time obtained from SM measurements includes both a period for the initial interaction with the proteasome and, for productive interactions leading to degradation, a period for translocation and degradation. To assess the contribution of each step to the observed dwell time, we blocked enzymatically active sites of the proteasome with inhibitors that act downstream of the initial binding events (fig. S17). Neither ubiquitin-aldehyde, which inhibits deubiquitylation by Uch37, nor epoxomicin, which inhibits substrate proteolysis, changed the dwell time. Using 1,10-phenanthroline to inhibit Rpn11 activity and proteasomal degradation also caused little change in the dwell time (figs. S18 and S19). Therefore, dwell time primarily reports on the initial binding event between the ubiquitylated substrate and the proteasome.

### *Cooperative and stochastic features of proteasome-substrate interactions*

Single-molecule studies can provide insights into mechanisms of the initial proteasome-substrate interaction, an important step for deconstructing the specificity of degradation. Only two proteasomal subunits, Rpn10 and Rpn13, contain ubiquitin-binding domains: Rpn10 has two UIM domains, and Rpn13 has one Pru domain (*4*). Together, they could maximally engage three ubiquitins at once. If a substrate molecule simultaneously bound to the three domains, this would constitute a type of cooperative or avidity binding process (Fig. 2C). Binding could also be enhanced by a "stochastic" mechanism, in which larger numbers of ubiquitins enhance the affinity of binding by increasing local ubiquitin concentrations, without requiring simultaneous interactions with different receptor proteins. These two mechanisms can be distinguished by dwell-time analysis, because a cooperative mechanism primarily affects the enthalpic component of the free energy of binding, whereas a stochastic mechanism should change only the entropic contribution. Kinetically, the cooperative mechanism should result in an exponential increase of dwell time with the number of conjugated ubiquitins, but a stochastic mechanism would be characterized by a linear increase, reflecting mass action (Fig. 2C) (*32*).

For both cyclin B–NT and securin, the dwell time initially increased exponentially as a function of the number of conjugated ubiquitins, up to a total of three (Fig. 2D), consistent with a cooperative mechanism (Fig. 2C). By contrast, further ubiquitylation (from four to nine ubiquitins) led to a linear increase in dwell time, consistent with a stochastic mechanism (Fig. 2, C and D). From the slope of the binding curve on a semilog plot, we calculated the mean free energy of binding per bound ubiquitin to the proteasome: 0.92 kcal/mol for securin and 0.97 kcal/mol for cyclin B, suggesting weak interactions. The cooperative behavior at low-ubiquitin stoichiometry and the linear behavior at higher levels provide keys to understanding how ubiquitin configurations are discriminated.

### *The role of the ubiquitin chain structure in proteasome-substrate interactions*

Ubiquitin chains interact more strongly with the proteasome than do ubiquitin monomers (*5*). Most proteasomal substrates require conjugation of ubiquitin chains for degradation. As we showed, multi-monoubiquitylated securin (or geminin) was degraded much more slowly than securin with ubiquitin chains (Fig. 1, C and B). The simplest explanation would be that this difference is due to weaker interaction of the multi-monoubiquitylated forms with the proteasome. To test this explanation, we measured the interaction between Ub0K-conjugated securin and the proteasome and compared it to securin linked to WT ubiquitin. Surprisingly, Ub0K-conjugated securin and WT Ub–conjugated securin showed almost identical dwell times on the proteasome when we compared substrates with the same total number of ubiquitins (Fig. 3A). This result was not caused by surface immobilization of the proteasome because it also occurred in binding assays performed in bulk, measuring interactions with the ubiquitin receptor Rpn10 on the proteasome (fig. S20). Likewise, cyclin B, which can be degraded without forming ubiquitin chains, had a similar binding affinity whether conjugated with WT Ub or Ub0K (Fig. 3B).

Although degradation of most substrates appears to be much more efficient if chains are formed, our results show that the presence of chains does not alter the affinity of substrate binding to the proteasome. Instead, affinity is primarily determined by the total number of ubiquitins on a substrate molecule, irrespective of their configurations except in the special case of single-chain substrates. Nevertheless, the increase in degradation rate when chains are present implies that ubiquitin configuration has an important role in the fate of the substrate after its binding to the proteasome. Therefore, we used SM methods to examine how the presence of ubiquitin chains affects various steps in the degradation process, with the goal of understanding how ubiquitin chains stimulate substrate degradation.

During the process of degradation, the proteasomal subunits in the 19*S* regulatory particle alternate among different adenosine triphosphatase (ATPase)–driven conformational states (*33*–*35*). To test whether the recognition of ubiquitin chains may require a particular proteasome conformation, we locked the proteasome in the adenosine triphosphate (ATP)–bound state with the nonhydrolyzable analog adenosine 5′-*O*-(3-thiotriphosphate) (ATP-γ-S) (fig. S18) and studied the proteasome's interaction with ubiquitylated substrates. There was a consistent increase of ~25% in dwell time for binding of WT Ub–conjugated cyclin B to the ATP-γ-S proteasome, as compared with proteasomes in ATP buffer. This increase in affinity was not observed when cyclin B was conjugated with Ub0K (Fig. 3, C and D). As a more sensitive metric for this effect, 29% of the binding events between ATP-γ-S proteasome and WT Ub–conjugated cyclin B lasted longer than 10 s, whereas only 10% of binding interactions had such persistence under the three control conditions (fig. S21). Another substrate, securin, gave similar results (fig. S22). Proteasomes in buffer containing adenosine diphosphate (ADP) showed weaker interactions with ubiquitylated cyclin B than those in ATP-containing buffer, and the binding of WT Ub–conjugated cyclin B to ADP proteasomes was not different from that of Ub0K-conjugated cyclin B

(fig. S23). We thus conclude that the discrimination of different configurations of ubiquitin on a substrate is enhanced in the ATP-bound state of the proteasome and therefore may result from rearrangement of the dynamic conformations of the proteasomal subunits. This discrimination was observed as a small binding enhancement, which by itself is unlikely to explain the generally large degradation rate enhancement for Ub chain–containing substrates. Perhaps the rearrangement of proteasomal subunits may expose or activate a hidden ubiquitin chain receptor that contributes to substrate discrimination. Two alternative mechanisms for the enhancement of protein degradation in the presence of ubiquitin chains—allosteric opening of the gate of 20*S* complex and stimulation of proteasomal ATPase activity (*36*–*38*)—were found in control studies to have little effect when tested with ubiquitylated Ube2S as the substrate (figs. S24 and S25). Therefore, these mechanisms are unlikely to explain the much larger effect of ubiquitin chains in promoting degradation (Fig. 1, B and C).

### *The role of ubiquitin chains in initiating substrate translocation*

If binding strength was the only determinant for degradation, the degradation rate should be proportional to the dwell time on the proteasome, as predicted by a probabilistic model of chemical reactions. For Ub0K-conjugated securin, this appeared to be the case over most of the range of ubiquitylation (Fig. 3E), where the rate versus dwell time was linear. By contrast, the degradation rate of WT Ub–conjugated securin increased superlinearly with dwell time (Fig. 3E), suggesting that ubiquitin chains, in addition to providing interaction with the proteasome, may also increase the efficiency of commitment of a substrate to degradation once bound to the proteasome [i.e., a turnover number ($k_{cat}$) effect] (Fig. 3F). Notably, at high modification levels, WT Ub was up to 40 times more efficient in promoting securin degradation than Ub0K, per unit dwell time.

In addition to binding kinetics, SM traces can provide critical mechanistic information about events that follow binding, such as substrate translocation and degradation by the proteasome, in the form of the kinetics of ubiquitin removal from the substrate. We observed two different types of signals after binding of substrate to the proteasome: (i) structureless, in which there was a complete loss of ubiquitin within 200 ms (the sampling rate of the camera), and (ii) stepped, in which the ubiquitin signal decreased in several discrete steps with very short intervals (Fig. 4A). These effects were not attributable to photobleaching or the limited speed of camera detection (figs. S26 and S27).

Although the likely explanation for structureless events is dissociation of substrate from the proteasome before ubiquitin removal, the stepped decrease of the ubiquitin signal could represent a series of deubiquitylation events catalyzed by Rpn11, the only effective DUB in our samples (Fig. 4B). In support of this interpretation, inhibitors of Rpn11 activity, such as 1,10-phenanthroline and ATP-γ-S, almost completely eliminated the stepped decreases (fig. S28). When cyclin B–NT was conjugated with multiple diubiquitin chains, we observed with each step the removal of two ubiquitins rather than one (fig. S29). Thus, the proposal that Rpn11 cleaves ubiquitin chains en bloc (*23*) is strongly supported at the SM level. Accordingly, we did not observe the stepped loss of ubiquitin signals for ubiquitylated K64–cyclin B, a single-lysine substrate that carries only one ubiquitin chain (fig. S27). Presumably, if all ubiquitins on a multi-lysine substrate happened to form a single chain, only structureless events would be observed. However, such cases should be rare for APC substrates, because the E2 UbcH10 forms long ubiquitin chains with low efficiency.

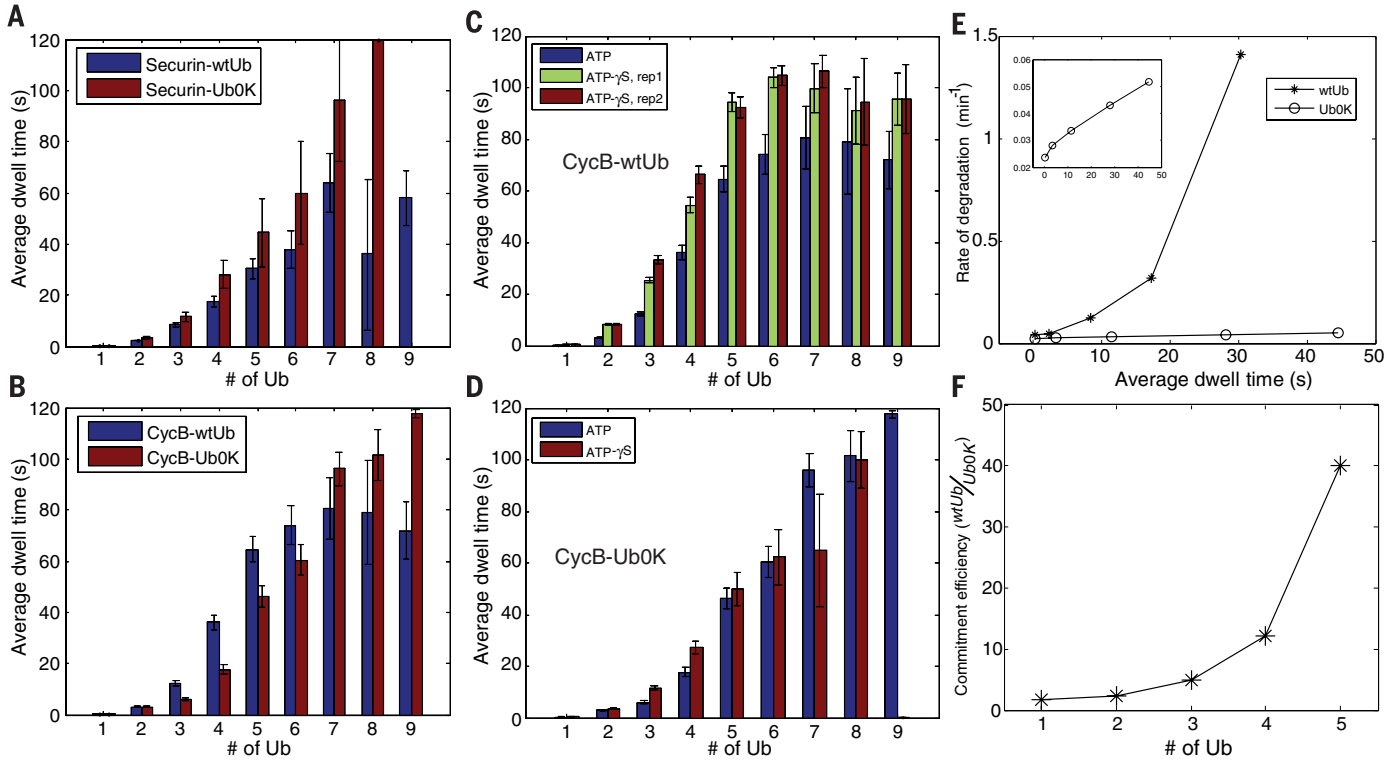



**Fig. 3. Interaction with the proteasome is mainly determined by the total number of ubiquitins on a substrate, insensitive to Ub chain structures.** (**A** and **B**) Securin or cyclin B–NT was ubiquitylated by APC with either WT Ub or Ub0K and was tested for interaction with the proteasome, as in Fig. 2B. (**C** and **D**) Interaction of WT Ub– or Ub0K-conjugated cyclin B with the proteasome in the presence of ATP or ATP-γ-S. rep1 and rep2 are two experimental replicates. The data for the ATP-proteasome experiment are plotted for comparison; these data are identical to those shown in (B). Cyclin B–ATP data in (C) are identical to those in (B). In (A) to (D), error bars denote SD of the mean. (**E**) Degradation rate and proteasomal dwell-time relationship for WT Ub– and Ub0K-conjugated securin. Degradation rates were measured in the quantitative degradation assay (Fig. 1C), and dwell time on the proteasome was measured using the SM method. The inset shows the Ub0K result on a smaller *y* axis. (**F**) Ratio of "commitment efficiency" for securin–WT Ub over securin-Ub0K, as a function of total conjugated ubiquitins. Commitment efficiency is defined as the degradation rate divided by the dwell time.

E2 UbcH10 acts broadly to increase the number of sites of mono- or diubiquitylation on multiple lysines (*3*).

Kinetic studies also reveal the sequence of events in degradation. Before the very processive phase of deubiquitylation, there was always a delay. The distribution of delay time was not exponential, as would be expected for a single-step reaction. Instead, the distribution corresponded closely to a Γ distribution with a modal delay of 2 s (fig. S30), suggesting intermediate rate-limiting steps between initial binding and the onset of deubiquitylation.

Rpn11 sits at the substrate entry port, where it has a critical role in removing ubiquitin chains on translocating peptides (*6*, *39*–*42*). Translocation may signal irreversible engagement of the substrate into the degradation process. This tight coupling between deubiquitylation and translocation was supported by the absence of deubiquitylated but nondegraded substrates in our bulk assays (fig. S1). We therefore used the processive deubiquitylation mediated by Rpn11 as a kinetic indicator for substrate translocation. To test whether the substrate was cleaved by the proteasome contemporaneously with deubiquitylation, we labeled both ends of cyclin B–NT with DyLight 550. We observed a stepped decrease of fluorescence on cyclin B that was sensitive to proteasome inhibition by MG132 and was coincident with the deubiquitylation, suggesting co-translocational activity of Rpn11 (figs. S31 and S3). We compared the fraction of stepped events among all binding events for WT Ub–conjugated securin and Ub0K-conjugated securin to assess their relative probability of undergoing deubiquitylation and translocation steps. At increased amounts of conjugated ubiquitin, this probability increased dramatically for securin conjugated with WT ubiquitin capable of forming chains. By contrast, it remained low if Ub0K, which precludes chain formation, was used (Fig. 4C). This difference was large enough to account for the 12-fold degradation rate disparity between WT Ub–conjugated and Ub0K-conjugated securin. Multi-monoubiquitylated securin was not intrinsically refractory to deubiquitylation, which still occurred processively in the few cases of stepped events on multi-monoubiquitylated securin (fig. S32). Thus, ubiquitin chains on substrates appear to specifically promote initiation of translocation and degradation.

## Discussion

The multiple lysines on a substrate and on ubiquitin itself generate a large number of possible ubiquitin configurations. These configurations represent a "ubiquitin code" of unknown degeneracy that must be read by the proteasome and converted into a rate of substrate degradation. We have approached this problem with two methods: construction of substrates with defined ubiquitin configurations and SM techniques. We found that tetraubiquitin chains are not essential for rapid proteasomal degradation of APC substrates, which would explain why a tetraubiquitin receptor on a proteasome has not been found. In fact, ubiquitin chains on cyclin B, and possibly other APC substrates, are typically short (*3*), and multiple ubiquitylatable lysine residues are a common feature of these substrates. A distributed array of short ubiquitin chains appears to be a superior and perhaps an optimal signal for proteasomal degradation; this conclusion could probably extend to substrates of other E3 ligases. Although a single ubiquitin chain may be sufficient for degrading certain substrates, such as Sic1 mutants and IκB (*16*, *43*), increasing the number of ubiquitylated lysine residues of the canonical single-chain substrate β-galactosidase greatly accelerates its degradation (*44*). Similarly, WT cyclin B was degraded faster than mutants with fewer lysines at the same total amount of ubiquitylation (Fig. 1F) (Ub = 7 to 8). Besides K48 chains, the APC also establishes K11 and K63 linkages on substrates (*3*). We found that K48 chains promoted more efficient degradation than K11, K27, and K63 chains (fig. S6).

By studying cyclin B mutants, we found that proximity of the first ubiquitylated lysine to the N terminus was associated with faster degradation (Fig. 1E), suggesting that the degradation rate is sensitive to the position of ubiquitylated lysine residues. There was a correspondence between long dwell times and elevated rates of degradation (fig. S33A), suggesting that ubiquitin chain position could control the rate of degradation, at least partially, through controlling affinity with the proteasome. Single-lysine mutants of cyclin B with a ubiquitin chain at different positions had indistinguishable binding kinetics to the proteasome, suggesting that monochain and multichain substrates may interact with the proteasome by different mechanisms (fig. S33B). For WT substrates, our current method of constructing defined ubiquitin configurations does not specify the chain positions. The results are understood as a populational average of all actual combinations of positions, of which the vast majority can promote efficient degradation.

Comparison of the $K_d$ values for tetraubiquitin chains measured by our SM methods with those from bulk assays suggests that surface immobilization of the proteasome is unlikely to distort kinetic rate constants. In addition, ubiquitylated securin took, on average, 10 s to complete translocation and possibly the degradation process on the surface-bound proteasome (fig. S34). This result is consistent with time for degrading similar-size proteins, such as dihydrofolate reductase and Sic1, measured in bulk assays (*45*), suggesting that the surface-bound proteasome is unimpaired for unbiased kinetic studies.

The rate of degradation is determined by both binding (i.e., $K_d$ effect) and postbinding (i.e., $k_{cat}$ effect) events on the proteasome. For APC substrates, multiple diubiquitin chains were more efficient degradation signals than tetraubiquitin chains, given the same total number of conjugated ubiquitins (Fig. 1, B to D). The explanation for this distinction may be their different binding strength with the proteasome. Using SM methods, we observed weaker binding if ubiquitins were assembled into a single chain on K64–cyclin B when compared to that for WT cyclin B, which

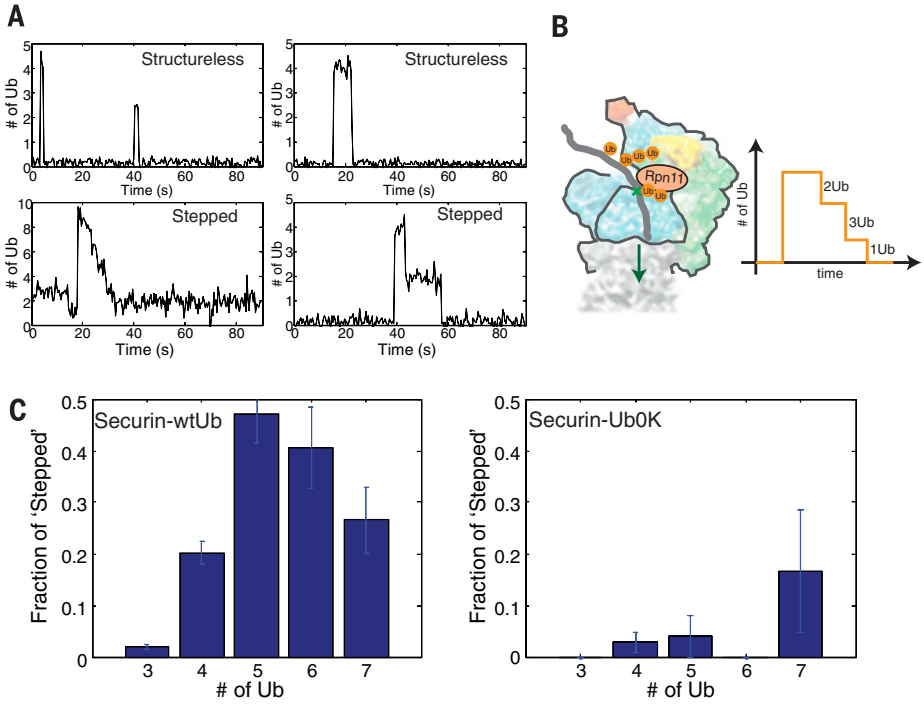



**Fig. 4. Ubiquitin chains on substrates promote translocation initiation.** (**A**) Examples of structureless and stepped traces from a SM experiment on WT Ub–conjugated securin. Stepped traces are due to progressive deubiquitylation by Rpn11, as illustrated in (**B**). (**C**) The fraction of stepped traces, as a function of the number of ubiquitins, for WT Ub ($N$ = 663) or Ub0K-conjugated ($N$ = 133) securin. Error bars represent SD of the mean.

had short and distributive chain configurations (Fig. 2B). High-resolution cryo-EM structures of the proteasome are consistent with the potential effectiveness of multiple short ubiquitin chains. Because proteasomal ubiquitin receptors Rpn10 and Rpn13 are distant from each other (*6*, *7*), distributive configurations of ubiquitins on a substrate molecule might promote the use of more ubiquitin receptors on the proteasome; a single ubiquitin chain might also be less effective due to steric constraint. Furthermore, Rpn10 and Rpn13 may not be the only ubiquitin receptors on proteasome because budding yeast can tolerate mutations of both (*46*). Additional receptors or shuttle factors for the proteasome are also thought to contribute to the binding of ubiquitylated proteins (*47*, *48*) (see below).

The SM binding measurements suggest a model wherein a substrate molecule samples multiple modes of binding during its interaction with the proteasome (Fig. 5). Evidence for such a mechanism comes from measurement of the dwell time as a function of ubiquitylation levels on the substrate. Beyond the ubiquitin-binding capacity of the proteasome, most likely limited to three or four ubiquitins by available ubiquitin receptors, a further increase of binding affinity relies on an increase in the local ubiquitin concentration on the substrate: the stochastic interaction. This stochastic mechanism stabilizes the bound state by increasing its entropy, or the number of microscopic states, because entropy is proportional to the logarithm of the number of these states. In this system, an increase in relevant entropy may occur if the substrate molecule can explore multiple conformations on the proteasome through intramolecular diffusion while remaining associated with the proteasome (Fig. 5). Such dynamic sampling should also increase the likelihood that the peptide terminus would be captured by the substrate entry port on the ATPases, thereby facilitating initiation of translocation. A cooperative process is implied by the exponential increase of dwell time as a function of the number of conjugated ubiquitins, whereas a stochastic process is implied by a linear increase (Fig. 2, C and D). A similar, biphasic binding relationship ($1/K_i$ ~ Ub number) was suggested in an early publication using competition assays, though the interpretation was different (*5*) (fig. S35). An exponential increase of dwell time involving greater than three simultaneous interactions would further increase the discrimination of ubiquitin levels over a linear increase. Why, then, is the process no longer exponential after four ubiquitins? Cooperative mechanisms tend to promote tight binding, which has potential risks for the cell. If a highly ubiquitylated substrate could not be degraded by the proteasome, it would stably block the proteasome. Such an inhibition by stably binding complexes has been proposed to underlie the accumulation of ubiquitylated intermediates in various neurodegenerative diseases (*49*). A linear increase in affinity at high-ubiquitin stoichiometry, though less discriminating, is also less prone to form unproductive, inhibitory substrate-proteasome complexes.

Although binding to the proteasome is a prerequisite for degradation, it does not in itself determine the rate of degradation. For example, WT Ub–conjugated securin and Ub0K-conjugated securin bind equally tightly, but the former is degraded much faster than the latter. The specificity of degradation must also reflect postbinding events. The SM analysis of Rpn11-dependent deubiquitylation indicates that the chain structure of ubiquitin promotes the initiation of translocation, a requirement for degradation. This effect of Ub chains also applies to cyclin B, a special substrate that can be degraded even without ubiquitin chains. We observed a shorter delay between binding and the initial deubiquitylation event for WT Ub–conjugated cyclin B than for Ub0K conjugates, consistent with the translocation-promoting activity of ubiquitin chains (fig. S36).

Most substrate-proteasome encounters do not lead to degradation, especially for substrates with a low number of conjugated ubiquitins (Fig. 4C). Even for highly ubiquitylated substrates, binding events sometimes lasted for tens of seconds without leading to degradation. Thus, there may be a latent state of the proteasome, in which heterogeneous ubiquitin chain conformations might affect deubiquitylation (*50*) or, perhaps more likely, might affect the orientation of a bound substrate, placing the translocation-initiating element far away from the substrate entry port. In this context, the presence of a short flexible domain at a substrate's terminus should substantially accelerate its rate of degradation (*20*, *51*). Therefore, engagement of the translocation-initiating element by force-generating pore loops of the proteasomal ATPases, which are reached via the substrate entry port, may generally be a rate-limiting step in degradation. Translocation initiation has been proposed to underlie "commitment," a hypothetical point at which substrates are irreversibly destined to degradation (*39*). We argue that translocation initiation, sensitive to the configuration of ubiquitin groups on the substrate, is either the commitment step or is closely coupled to it.

To understand how ubiquitin chains promote translocation initiation, we propose a model based on our experimental observations (Fig. 5). Conformational changes of the proteasome, driven by the ATPases quickly transiting through different nucleotide-bound states (*33*, *40*), may activate a ubiquitin chain receptor(s) that participates in substrate recognition (*52*). Candidates for such a receptor include ATPase Rpt5, which can be

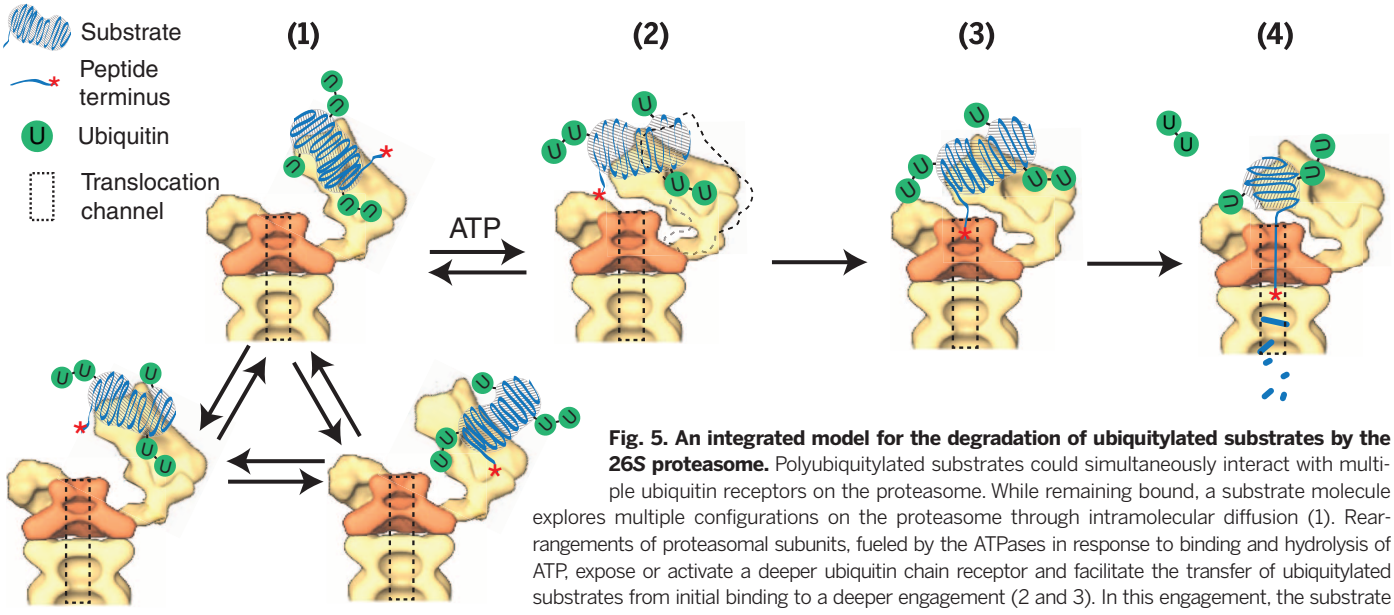



**Fig. 5. An integrated model for the degradation of ubiquitylated substrates by the 26S proteasome.** Polyubiquitylated substrates could simultaneously interact with multiple ubiquitin receptors on the proteasome. While remaining bound, a substrate molecule explores multiple configurations on the proteasome through intramolecular diffusion (1). Rearrangements of proteasomal subunits, fueled by the ATPases in response to binding and hydrolysis of ATP, expose or activate a deeper ubiquitin chain receptor and facilitate the transfer of ubiquitylated substrates from initial binding to a deeper engagement (2 and 3). In this engagement, the substrate or its terminus is closer to the substrate entry port, which expedites translocation initiation (3) and ensuing degradation (4).

cross-linked to bound ubiquitin chains (*53*). The same result can also be explained by rearrangement of ubiquitin receptors (Rpn10 and Rpn13) into a higher-affinity state for ubiquitin chains. It would make sense if the additional ubiquitin chain receptor were closer to the substrate entry port than Rpn10 or Rpn13 to facilitate translocation initiation by engaging the substrate into a deeper conformation after the initial interaction involving mainly Rpn10 and Rpn13 (Fig. 5). Such an intermediate step is indicated by the delay before deubiquitylation or translocation (fig. S30). Consistently, Rpt5 is very close to the substrate entry port, and conformational changes induced by ATP-γ-S dramatically reduce the distance between Rpn10 and Rpt4/5, indicating a possible direct transfer of substrates from initial binding to deeper engagement (*33*, *54*) (fig. S37).

We have shown that there is no simple length threshold for ubiquitin chains for degradation by the proteasome. Rather, there are at least two requirements: a minimal number of ubiquitins to result in tight binding and a certain number or length of chains to promote translocation into the axial channel. The ultimate rate of degradation is probably set by ubiquitin stoichiometry, chain configuration, and properties of the substrate that affect not only the capacity to be ubiquitylated and the configuration of chains but also the orientation of the chains and translocation-initiating elements once bound to the proteasome.

## Materials and methods

### Protein purification and labeling

Recombinant, full-length human securin, Ube2S, GST-Emi1 (297-447), *Xenopus* geminin, *Xenopus* N-terminal cyclin B (amino acids 1 to 104), human N-terminal cyclin B (amino acids 1 to 88), and mutants were purified from *Escherichia coli* cells using an IMPACT kit (NEB, E6901S). A PKA (protein kinase A) site (RRASV) was also placed at the N terminus of the substrates used in degradation assays. Human ubiquitin, UbK48, Ub0K, UBK6A, and UbI44A mutants, each with a cysteine residue and a His tag at the N terminus, were purified from *E. coli* cells using cation-exchange chromatography (GE, 17-1152-01) labeled with DyLight 550 maleimide (Pierce, 62290). After removing unreacted dyes, labeled ubiquitin was subjected to anion-exchange chromatography (GE, 17-1153-01) to separate labeled and unlabeled species. Finally, the N-terminal His tag was cleaved off using thrombin.

Fluorescently labeled ubiquitin chains were synthesized using E2-25K in reactions containing 30 μM DyLight 550–labeled UbK48 and 10 μM DyLight 550–labeled Ub0K; carried out in UBAB buffer [25 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$], 3 mM ATP, 1 μM E1, and 10 μM E2-25K; and incubated at 37°C for 16 hours. The product was purified and separated using anion-exchange chromatography. For binding assays with the proteasome, different fractions of ubiquitin chains were mixed as the sample.

Anti-20*S* antibody (MCP21) was biotinylated using biotin-NHS (Pierce, 20217) and was purified using desalting columns.

Radioactive $^{33}$P-ATP was used to label substrates with a PKA site at the N terminus for in vitro ubiquitylation and degradation assays.

E1, E2s (UbcH10, E2-25K), WT ubiquitin, biotin ubiquitin, methylated ubiquitin, K48-diubiquitin chains, and K48-tetraubiquitin chains were from Boston Biochem. Purified streptavidin was from Invitrogen.

For biotinylating substrates, a biotin-containing peptide (NEB, Bio-P1) was ligated to the substrate C terminus, which had been activated by intein cleavage during purification, resulting in ~90% ligation efficiency.

### Ubiquitin chain methylation

Ubiquitin chain samples were buffer-exchanged into a mixture of 50 mM Hepes-Na (pH 7.5), 150 mM NaCl, and 2% glycerol. Dimethylamine borane and formaldehyde were added to the sample to final concentrations of 20 and 40 mM, respectively. After 2-hour incubation on ice, another 20 mM dimethylamine borane and 40 mM formaldehyde were again added, bringing final concentrations to 40 and 80 mM, respectively, and incubation was continued for 16 hours on ice. Methylated proteins were buffer-exchanged to tris-buffered saline [20 mM Tris-HCL (pH 7.5), 150 mM NaCl].

### APC purification from HeLa cell $G_1$ extract

HeLa cell $G_1$ extract preparation and APC purification were performed as described previously (*13*). Briefly, 2L HeLa S3 cell spinner culture was synchronized at prometaphase using thymidine/nocodazole double block and was then released for 3 hours into $G_1$. Cells were harvested and subjected to nitrogen cavitation in 75% volume swelling buffer [20 mM Tris-HCl (pH 7.5), 5 mM KCl, 1.5 mM $MgCl_2$, 1 mM dithiothreitol (DTT), protease-inhibitor tablet (Roche, 05892953001)]. APC was purified from cell extracts using anti-Cdc27 (AF3.1, custom-made) agarose and eluted using a competitive peptide.

### Affinity purification of the 26S human proteasome

Human proteasomes were affinity-purified on a large scale from a stable human embryonic kidney 293 cell line harboring HTBH-tagged hRPN11 (a gift from L. Huang). The cells were Dounce-homogenized in lysis buffer [50 mM $NaH_2PO_4$ (pH 7.5), 100 mM NaCl, 10% glycerol, 5 mM $MgCl_2$, 0.5% NP-40, 5 mM ATP, and 1 mM DTT] containing protease inhibitors. Lysates were cleared and then incubated with NeutrAvidin agarose resin (Thermo Scientific) overnight at 4°C. The beads were then washed with excess lysis buffer, followed by the wash buffer [50 mM Tris-HCl (pH 7.5), 1 mM $MgCl_2$, and 1 mM ATP]. Usp14 was removed from proteasomes using the wash buffer plus 100 mM NaCl for 30 min. 26*S* proteasomes were eluted from the beads by cleavage, using TEV protease (Invitrogen).

### In vitro ubiquitylation reaction

The APC ubiquitylation reactions were carried out in UBAB buffer [25 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$] containing 5 nM APC, 100 nM E1, 2 μM UbcH10, 2 mg/ml bovine serum albumin (BSA), an energy regenerating system, and either ubiquitin or ubiquitin chains: 10 μM ubiquitin, methylated ubiquitin, or DyLight 550–Ub0K; 5 μM methylated K48-diubiquitin; and 1 μM methylated K48-tetraubiquitin. Reactions were incubated at 30°C. Results were quantified using a phosphorimager. For PKA-labeled substrates, calyculin A (EMD, 19-139), a broad-spectrum phosphatase inhibitor, was added at 10 μg/ml.

Ube2S autoubiquitylation reactions were carried out in UBAB buffer containing 6 μM Ube2S, 0.5 μM E1, and 5 mM ATP. 100 μM ubiquitin, methylated ubiquitin, or 40 μM methylated K48-diubiquitin chain were also added. Reactions were incubated at 37°C for 4 hours.

### Quantitative degradation assay

PKA-labeled substrates (800 nM) were ubiquitylated by the APC using various forms of ubiquitin or ubiquitin chains for 45 min at 30°C. Reaction products were then mixed with proteasome-containing solution (UBAB buffer, 3 nM 26*S* proteasome, 2 mg/ml BSA, 1 μM GST-Emi1, 3 mM ATP, 10 μg/ml calyculin A) at a ratio of 1:4. The degradation reaction was sampled at 0, 3, 6, and 15 min at 37°C and was quantified using a phosphorimager.

To calculate the degradation rate, the intensity of each ubiquitylated species was quantified using ImageJ. These values were normalized to the intensity of the unmodified substrate (Ub0) to correct for loading error and nonspecific dephosphorylation. Finally, traces of the time course of degradation were fitted with an exponential decay function to obtain the rate constant.

### Peptidase assay

Products of Ube2S autoubiquitylation reaction (6 μM) were mixed with proteasome-containing solution (UBAB buffer, 6 nM 26*S* proteasome, 2 mg/ml BSA, 3 mM ATP, 30 μM Z-GGL-AMC, 0.4 mM DTT) at a 1:1 ratio. Hydrolysis of GGL-AMC was monitored at 37°C using a fluorescence spectrometer.

### Binding assay for ubiquitylated substrate

Wild-type securin was ubiquitylated in a standard APC reaction with either WT Ub or Ub0K in a total volume of 20 μl for 1.5 hours at 30°C. The product was incubated with GST-magnetic beads conjugated with GST-Rpn10 for 40 min at 4°C (*8*). After 2× wash with buffer [25 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.5 mM EDTA, 0.05% Tween-20], the supernatant and bead-bound fraction were assayed for securin using anti-securin antibody by Western blot.

### Slide passivation

We followed a basic slide passivation protocol using 5-kD PEG plus 2.5% 5-kD biotin-PEG (LaysanBio, MPEG-SVA-5000; Biotin-PEG-SVA-5000) in a "clouding-point" solution on amino-silanized slides for 4 hours, as described previously (*55*, *56*). In addition, we passivated them again using 1-kD PEG-NHS (Nacocs, PG1-SC-1k) for 1 hour, followed

by 10% (w/v) BSA for 30 min before use. After passivation, slides were assembled into reaction chambers (*55*). Streptavidin (0.2 mg/ml) was incubated with passivated slides for 5 min for immobilizing biotinylated substrates. The quality of slide passivation was tested by incubating with HeLa cell extracts supplemented with DyLight 550–Ub and measuring the level of nonspecific binding.

### Single-molecule proteasome assay

Purified substrates (200 nM) were ubiquitylated by the APC with 5 μM fluorescent ubiquitin (WT Ub, Ub0K, UbK6A, or UbI44A) for 40 min (60 min for K64–cyclin B) at 30°C. The reaction product was diluted 100× into imaging buffer (UBAB, 3 mM ATP, 40 mM imidazole, 5 mg/ml BSA). Imidazole in the imaging buffer serves to reduce the nonspecific binding of fluorescently labeled proteins.

26*S* proteasome (15 nM) and biotinylated MCP21 antibody (15 nM) were mixed and incubated at room temperature for 15 min. The proteasome-antibody mix was loaded onto passivated slides coated with streptavidin. After a brief incubation, unbound protein was washed off and replaced with imaging buffer containing diluted ubiquitylation product. Image acquisition was started immediately with <15 s delay.

For experiments involving ATP-γ-S, ATP in imaging buffer was replaced with 2.5 mM ADP plus 0.5 mM ATP-γ-S. For experiments involving ADP, proteasome was incubated in 3 mM ADP buffer. ADP buffer and ubiquitylation product were also incubated with 30 mM glucose and 2 mg/ml hexokinase for 60 min at 30°C to remove residual ATP in the solution. For experiments involving 1,10-phenanthroline, 3 mM phenanthroline was added to the imaging buffer and proteasomal mixture; the sample was incubated for 20 min at room temperature before performing the experiment.

For the SM degradation assay, human cyclin B–NT was engineered with two cysteine residues close to its N and C termini and was labeled with DyLight 550–maleimide. Doubly labeled cyclin B was enriched via anion-exchange chromatography (HiTrap Q HP, pH 8.0). After ubiquitylation with unlabeled WT ubiquitin, 2 nM cyclin B was studied in the SM setup, as described above. Data analysis was identical to that described for the study of deubiquitylation.

We used a Nikon Ti TIRF microscope equipped with three laser lines of 491 nM (27 mW, full output from objective), 561 nM (30 mW), and 640 nm (59 mW), as well as an Andor DU-897 EMCCD camera. Time series were acquired at 200 ms per frame for 2 min, unless indicated.

### Single-fluorophore intensity calibration

The fluorescence intensity of a single DyLight 550 molecule on ubiquitin was calibrated through photobleaching preformed ubiquitin chains (*27*). Ubiquitin chains were synthesized in a standard E2-25K reaction (see above) containing 30 μM DyLight 550–labeled ubiquitin and 5 μM biotin-ubiquitin in UBAB buffer supplemented with 3 mM ATP, 1 μM E1, and 5 μM E2-25K. The reaction was incubated at 37°C for 16 hours.

The E2-25K reaction product was diluted 10,000× in imaging buffer (above) and loaded on a passivated slide via streptavidin. Photobleaching was observed with the use of a TIRF microscope, illuminated with a laser-intensity level 2.5× higher than that in proteasome experiments (to achieve faster photobleaching).

The uncertainty of measuring ubiquitin stoichiometry is ~30%, as suggested by the standard deviation of the intensity distribution of single fluorescent ubiquitin. This uncertainty is primarily due to uncorrected uneven illumination.

### Single-molecule data reproducibility

To test the reproducibility of dwell time data as in fig. S13, the same experiment with slightly varying concentrations (±50%) of ubiquitylated substrates was repeated three times on the same coverslip but at different positions.

Long-term (<1 year) reproducibility tests of the results of binding measurements carry 10 to 20% variation, probably due to batch-to-batch variability of proteasome samples or laser intensity drifts. Therefore, to minimize systematic variation in the measurements, each experiment with its controls was performed on the same slide using the same batch of proteasome and substrates.

### Data analysis

The workflow of data analysis is illustrated in fig. S38.

### Image processing

Raw images were first corrected for stage drifting after subtracting a uniform-intensity background. Stage drifting was corrected by subtracting the stage motion, which was derived using successive image registration to calculate the shift (step 1 in fig. S38). A custom-built algorithm was applied to identify binding spots of ubiquitylated molecules to the proteasome on the basis of their absolute intensity and local signal-to-noise ratio (step 2 in fig. S38). The spot-identification algorithm is based on finding the local maxima of ubiquitin fluorescence intensity in a field of view by applying a filter requiring the local signal-to-noise ratio to be larger than 6 and no other spots within four pixels around it. In this way, >90% of substrate-binding events can be identified compared with a blind manual identification (fig. S10). The false-positive rate is less than 5%, as shown by the nonsubstrate control, considering that one ubiquitin generates an average of ~20 photons on the camera per frame. After each substrate-binding spot has been recognized in a time lapse, spots are registered along the time axis according to their coordinate to designate substrate-binding events (step 3 in fig. S38). Specifically, if the coordinates of two substrate spots in two consecutive frames are less than or equal to one pixel apart, they are considered to be the same binding event. The absolute intensity of each spot was obtained by fitting the SM diffraction pattern with a two-dimensional Gaussian function, which generates both the signal intensity $I$(s) and the local background level $I$(bg) (step 4 in fig. S38). The signal intensity $I$(s) is then corrected for inhomogeneous illumination (step 5 in fig. S38). Inhomogeneous illumination was corrected by a self-adaptive algorithm that separates the entire field of view into 15-by-15 identical squares and uses the average fluorescent spot intensity in each square as a surrogate for the relative illumination intensity in the center of that square. Because the illumination intensity varies slowly across the field of view, we interpolated the illumination intensity at a given position from the center values in each square. The corrected signal intensity was then converted to absolute ubiquitin numbers by normalizing with the intensity value of a single ubiquitin obtained in the calibration step. The resulting traces were directly used for subsequent analysis or represented in figures, with no further processing.

### Dwell-time analysis

A custom-built algorithm was used to measure the duration of substrate-binding events (dwell time), based on the background noise intensity plus a cutoff of 0.7 Ub level. The binding measurement is insensitive to the choice of the cutoff (0.5 to 0.9 Ub) or laser intensity (fig. S14). The maximal fluorescence intensity (converted to the number of ubiquitins using a calibrated single-fluorophore intensity value) during each binding event was measured and plotted versus the dwell time.

### Stepped-event identification

A custom-built step-detection algorithm was used to identify stepped events (fig. S39). The formal definition of a stepped event is as follows: If two consecutive drops of fluorescent signal (>0.7 Ub) occur within 5 s, this is designated as a stepped event. Also, it is required that each intermediate state must last longer than one frame. Traces (~80%) can be identified correctly as compared with manual identification. Obvious classification errors were later corrected manually.

For counting the fraction of stepped events, only those events whose ubiquitin number was ≥3 were included in the analysis. We also excluded binding events that lasted longer than 90 s (5 to 10% of total binding events).

The same step-detection algorithm was also used in the analysis of photobleaching traces, to extract single-fluorophore intensity values.

### Curve fitting in Fig. 2D

Segments of the curve (indicated in the figure) showing dwell time versus number of ubiquitin relationships were fitted with a linear function or an exponential function. A $\chi^2$ test $\left(\chi^2 = \sum \frac{(y_i - y_i^0)^2}{e_i^2}\right)$ ($y_i$, data point $i$; $y_i^0$, data point on the fitted line; $e_i$, standard error) was performed to obtain the $P$ value for each fitting model.

### Photobleaching experiment on ubiquitylated substrates

Biotinylated securin was ubiquitylated by APC using DyLight 550–Ub and was immobilized on

passivated coverslips via streptavidin. Time-lapse images were acquired under conditions identical to the SM proteasomal assay, but in the absence of proteasome. The total signal intensity in a central region of the field was analyzed to extract photobleaching information (fig. S26).

## REFERENCES AND NOTES


1. C. M. Pickart, M. J. Eddins, Ubiquitin: Structures, functions, mechanisms. *Biochim. Biophys. Acta* **1695**, 55–72 (2004). doi: 10.1016/j.bbamcr.2004.09.019; pmid: 15571809
2. K. D. Wilkinson, Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. *FASEB J.* **11**, 1245–1256 (1997). pmid: 9409543
3. D. S. Kirkpatrick *et al*., Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology. *Nat. Cell Biol.* **8**, 700–710 (2006). doi: 10.1038/ncb1436; pmid: 16799550
4. D. Finley, Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* **78**, 477–513 (2009). doi: 10.1146/annurev.biochem.78.081507.101607; pmid: 19489727
5. J. S. Thrower, L. Hoffman, M. Rechsteiner, C. M. Pickart, Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **19**, 94–102 (2000). doi: 10.1093/emboj/19.1.94; pmid: 10619848
6. G. C. Lander *et al*., Complete subunit architecture of the proteasome regulatory particle. *Nature* **482**, 186–191 (2012). pmid: 22237024
7. E. Sakata *et al*., Localization of the proteasomal ubiquitin receptors Rpn10 and Rpn13 by electron cryomicroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1479–1484 (2012). doi: 10.1073/pnas.1119394109; pmid: 22215586
8. N. V. Dimova *et al*., APC/C-mediated multiple monoubiquitylation provides an alternative degradation signal for cyclin B1. *Nat. Cell Biol.* **14**, 168–176 (2012). doi: 10.1038/ncb2425; pmid: 22286100
9. A. Ciechanover, A. Stanhill, The complexity of recognition of ubiquitinated substrates by the 26S proteasome. *Biochim. Biophys. Acta* **1843**, 86–96 (2014). pmid: 23872423
10. G. Asher, N. Reuven, Y. Shaul, 20S proteasomes and protein degradation "by default". *BioEssays* **28**, 844–849 (2006). doi: 10.1002/bies.20447; pmid: 16927316
11. N. Shabek *et al*., The size of the proteasomal substrate determines whether its degradation will be mediated by mono- or polyubiquitylation. *Mol. Cell* **48**, 87–97 (2012). doi: 10.1016/j.molcel.2012.07.011; pmid: 22902562
12. D. Barford, Structure, function and mechanism of the anaphase promoting complex (APC/C). *Q. Rev. Biophys.* **44**, 153–190 (2011). doi: 10.1017/S0033583510000259; pmid: 21092369
13. M. Rape, S. K. Reddy, M. W. Kirschner, The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell* **124**, 89–103 (2006). doi: 10.1016/j.cell.2005.10.032; pmid: 16413484
14. O. Cohen-Fix, J. M. Peters, M. W. Kirschner, D. Koshland, Anaphase initiation in *Saccharomyces cerevisiae* is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. *Genes Dev.* **10**, 3081–3093 (1996). doi: 10.1101/gad.10.24.3081; pmid: 8985178
15. J. T. Lee, W. Gu, The multiple levels of regulation by p53 ubiquitination. *Cell Death Differ.* **17**, 86–92 (2010). doi: 10.1038/cdd.2009.77; pmid: 19543236
16. M. D. Petroski, R. J. Deshaies, Context of multiubiquitin chain attachment influences the rate of Sic1 degradation. *Mol. Cell* **11**, 1435–1444 (2003). doi: 10.1016/S1097-2765(03)00221-1; pmid: 12820958
17. W. Kim *et al*., Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* **44**, 325–340 (2011). doi: 10.1016/j.molcel.2011.08.025; pmid: 21906983
18. B. Schwanhäusser *et al*., Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011). doi: 10.1038/nature10098; pmid: 21593866
19. S. Elsasser *et al*., Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat. Cell Biol.* **4**, 725–730 (2002). doi: 10.1038/ncb845; pmid: 12198498
20. S. Prakash, L. Tian, K. S. Ratliff, R. E. Lehotzky, A. Matouschek, An unstructured initiation site is required for efficient proteasome-mediated degradation. *Nat. Struct. Mol. Biol.* **11**, 830–837 (2004). doi: 10.1038/nsmb814; pmid: 15311270
21. K. Lasker *et al*., Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1380–1387 (2012). doi: 10.1073/pnas.1120559109; pmid: 22307589
22. M. J. Lee, B. H. Lee, J. Hanna, R. W. King, D. Finley, Trimming of ubiquitin chains by proteasome-associated deubiquitinating enzymes. *Mol. Cell. Proteomics* **10**, R110.003871 (2011). doi: 10.1074/mcp.R110.003871; pmid: 20823120
23. T. Yao, R. E. Cohen, A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* **419**, 403–407 (2002). doi: 10.1038/nature01071; pmid: 12353037
24. D. Finley *et al*., Inhibition of proteolysis and cell cycle progression in a multiubiquitination-deficient yeast mutant. *Mol. Cell. Biol.* **14**, 5501–5509 (1994). pmid: 8035826
25. B. H. Lee *et al*., Enhancement of proteasome activity by a small-molecule inhibitor of USP14. *Nature* **467**, 179–184 (2010). doi: 10.1038/nature09299; pmid: 20829789
26. J. Hanna *et al*., Deubiquitinating enzyme Ubp6 functions noncatalytically to delay proteasomal degradation. *Cell* **127**, 99–111 (2006). doi: 10.1016/j.cell.2006.07.038; pmid: 17018280
27. D. Izawa, J. Pines, How APC/C-Cdc20 changes its substrate specificity in mitosis. *Nat. Cell Biol.* **13**, 223–233 (2011). doi: 10.1038/ncb2165; pmid: 21336306
28. Y. Lu, W. Wang, M. W. Kirschner, Specificity of the anaphase promoting complex: A single-molecule study. *Science* **348**, 1248737 (2015).
29. K. E. Sloper-Mould, J. C. Jemc, C. M. Pickart, L. Hicke, Distinct functional surface regions on ubiquitin. *J. Biol. Chem.* **276**, 30483–30489 (2001). doi: 10.1074/jbc.M103248200; pmid: 11399765
30. See supplementary text ("Estimating the on-rate of tetraubiquitin chains with immobilized 26S proteasome in SM experiments").
31. F. Shang *et al*., Lys6-modified ubiquitin inhibits ubiquitin-dependent protein degradation. *J. Biol. Chem.* **280**, 20365–20374 (2005). doi: 10.1074/jbc.M414356200; pmid: 15790562
32. See supplementary text ("Background information on the stochastic binding mechanism").
33. P. Śledź *et al*., Structure of the 26S proteasome with ATP-γS bound provides insights into the mechanism of nucleotide-dependent substrate translocation. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 7264–7269 (2013). doi: 10.1073/pnas.1305782110; pmid: 23589842
34. R. Beckwith, E. Estrin, E. J. Worden, A. Martin, Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA+ unfoldase. *Nat. Struct. Mol. Biol.* **20**, 1164–1172 (2013). doi: 10.1038/nsmb.2659; pmid: 24013205
35. D. M. Smith, H. Fraga, C. Reis, G. Kafri, A. L. Goldberg, ATP binds to proteasomal ATPases in pairs with distinct functional effects, implying an ordered reaction cycle. *Cell* **144**, 526–538 (2011). doi: 10.1016/j.cell.2011.02.005; pmid: 21335235
36. A. Peth, H. C. Besche, A. L. Goldberg, Ubiquitinated proteins activate the proteasome by binding to Usp14/Ubp6, which causes 20S gate opening. *Mol. Cell* **36**, 794–804 (2009). doi: 10.1016/j.molcel.2009.11.015; pmid: 20005843
37. A. Peth, N. Kukushkin, M. Bossé, A. L. Goldberg, Ubiquitinated proteins activate the proteasomal ATPases by binding to Usp14 or Uch37 homologs. *J. Biol. Chem.* **288**, 7781–7790 (2013). doi: 10.1074/jbc.M112.441907; pmid: 23341450
38. D. Bech-Otschir *et al*., Polyubiquitin substrates allosterically activate their own degradation by the 26S proteasome. *Nat. Struct. Mol. Biol.* **16**, 219–225 (2009). doi: 10.1038/nsmb.1547; pmid: 19169257
39. R. Verma *et al*., Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* **298**, 611–615 (2002). doi: 10.1126/science.1075898; pmid: 12183636
40. M. E. Matyskiela, G. C. Lander, A. Martin, Conformational switching of the 26S proteasome enables substrate degradation. *Nat. Struct. Mol. Biol.* **20**, 781–788 (2013). doi: 10.1038/nsmb.2616; pmid: 23770819
41. E. J. Worden, C. Padovani, A. Martin, Structure of the Rpn11-Rpn8 dimer reveals mechanisms of substrate deubiquitination during proteasomal degradation. *Nat. Struct. Mol. Biol.* **21**, 220–227 (2014). doi: 10.1038/nsmb.2771; pmid: 24463465
42. G. R. Pathare *et al*., Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2984–2989 (2014). doi: 10.1073/pnas.1400546111; pmid: 24516147
43. M. Karin, Y. Ben-Neriah, Phosphorylation meets ubiquitination: The control of NF-[kappa]B activity. *Annu. Rev. Immunol.* **18**, 621–663 (2000). doi: 10.1146/annurev.immunol.18.1.621; pmid: 10837071
44. T. Suzuki, A. Varshavsky, Degradation signals in the lysine-asparagine sequence space. *EMBO J.* **18**, 6017–6026 (1999). doi: 10.1093/emboj/18.21.6017; pmid: 10545113
45. A. Peth, J. A. Nathan, A. L. Goldberg, The ATP costs and time required to degrade ubiquitinated proteins by the 26 S proteasome. *J. Biol. Chem.* **288**, 29215–29222 (2013). doi: 10.1074/jbc.M113.482570; pmid: 23965995
46. K. Husnjak *et al*., Proteasome subunit Rpn13 is a novel ubiquitin receptor. *Nature* **453**, 481–488 (2008). doi: 10.1038/nature06926; pmid: 18497817
47. S. Elsasser, D. Chandler-Militello, B. Müller, J. Hanna, D. Finley, Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome. *J. Biol. Chem.* **279**, 26817–26822 (2004). doi: 10.1074/jbc.M404020200; pmid: 15117949
48. R. Verma, R. Oania, J. Graumann, R. J. Deshaies, Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system. *Cell* **118**, 99–110 (2004). doi: 10.1016/j.cell.2004.06.014; pmid: 15242647
49. S. Imamura, T. Yabu, M. Yamashita, Protective role of cell division cycle 48 (CDC48) protein against neurodegeneration via ubiquitin-proteasome system dysfunction during zebrafish development. *J. Biol. Chem.* **287**, 23047–23056 (2012). doi: 10.1074/jbc.M111.332882; pmid: 22549779
50. Y. Ye *et al*., Ubiquitin chain conformation regulates recognition and activity of interacting proteins. *Nature* **492**, 266–270 (2012). doi: 10.1038/nature11722; pmid: 23201676
51. S. Fishbain, S. Prakash, A. Herrig, S. Elsasser, A. Matouschek, Rad23 escapes degradation because it lacks a proteasome initiation region. *Nat. Commun.* **2**, 192 (2011). doi: 10.1038/ncomms1194; pmid: 21304521
52. A. Peth, T. Uchiki, A. L. Goldberg, ATP-dependent steps in the binding of ubiquitin conjugates to the 26S proteasome that commit to degradation. *Mol. Cell* **40**, 671–681 (2010). doi: 10.1016/j.molcel.2010.11.002; pmid: 21095592
53. Y. A. Lam, T. G. Lawson, M. Velayutham, J. L. Zweier, C. M. Pickart, A proteasomal ATPase subunit recognizes the polyubiquitin degradation signal. *Nature* **416**, 763–767 (2002). doi: 10.1038/416763a; pmid: 11961560
54. F. Förster, P. Unverdorben, P. Sledź, W. Baumeister, Unveiling the long-held secrets of the 26S proteasome. *Structure* **21**, 1551–1562 (2013). doi: 10.1016/j.str.2013.08.010; pmid: 24010714
55. R. Roy, S. Hohng, T. Ha, A practical guide to single-molecule FRET. *Nat. Methods* **5**, 507–516 (2008). doi: 10.1038/nmeth.1208; pmid: 18511918
56. A. Pertsinidis, Y. Zhang, S. Chu, Subnanometre single-molecule localization, registration and distance measurements. *Nature* **466**, 647–651 (2010). doi: 10.1038/nature09163; pmid: 20613725


## ACKNOWLEDGMENTS

We thank L. Bai and W. Ma for commenting on the manuscript. We thank the Nikon Imaging Center at Harvard Medical School for providing instruments. Y.L. is supported by a Damon Runyon Research Fellowship and is a Lallage Feazel Wall Fellow. This work was supported in part by NIH grants GM43601 (to D.F.) and GM66492 (to R.W.K.)

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/1250834/suppl/DC1
Supplementary Text
Figs. S1 to S39
References

14 January 2014; accepted 20 February 2015
10.1126/science.1250834

RESEARCH ARTICLE

PROTEIN TARGETING

# The principle of antagonism ensures protein targeting specificity at the endoplasmic reticulum

Martin Gamerdinger,[1] Marie Anne Hanebuth,[1] Tancred Frickey,[2] Elke Deuerling[1]*

**The sorting of proteins to the appropriate compartment is one of the most fundamental cellular processes. We found that in the model organism *Caenorhabditis elegans*, correct cotranslational endoplasmic reticulum (ER) transport required the suppressor activity of the nascent polypeptide-associated complex (NAC). NAC did not affect the accurate targeting of ribosomes to ER translocons mediated by the signal recognition particle (SRP) pathway but inhibited additional unspecific contacts between ribosomes and translocons by blocking their autonomous binding affinity. NAC depletion shortened the life span of *Caenorhabditis elegans*, caused global mistargeting of translating ribosomes to the ER, and provoked incorrect import of mitochondrial proteins into the ER lumen, resulting in a strong impairment of protein homeostasis in both compartments. Thus, the antagonistic targeting activity of NAC is important in vivo to preserve the robustness and specificity of cellular protein-sorting routes.**

In eukaryotes, a substantial fraction of membrane and secretory proteins are cotranslationally delivered to and translocated across the endoplasmic reticulum (ER) membrane. This transport is mediated by the signal recognition particle (SRP), which binds to hydrophobic signal sequence–containing ribosome-nascent chain complexes (RNCs) and targets them to the Sec61 translocon via the ER membrane-located SRP receptor (SR) (*1*–*6*). However, despite decades of research on SRP-mediated ER targeting, key questions regarding the requirement of additional sorting factors to guarantee accuracy and efficiency of protein transport remained unanswered. One early study discovered that an abundant protein complex termed nascent polypeptide-associated complex (NAC) may be required to prevent SRP from interaction with signal-less RNCs (*7*). This idea has been highly controversial, and the function of NAC in protein targeting, if any, remains unclear. Furthermore, ribosomes show a SRP-independent, high intrinsic affinity for Sec61 in vitro. This raises the question about the identity and the functional importance of a potential negative regulator that prevents unspecific ribosome-translocon interactions to ensure correct RNC sorting in vivo (*8*–*12*).

We used the metazoan *Caenorhabditis elegans* to dissect the factors underlying ER targeting specificity in vivo.

[1]Department of Biology, Institute of Molecular Microbiology, University of Konstanz, 78457 Konstanz, Germany.
[2]Department of Biology, Applied Bioinformatics Laboratory, University of Konstanz, 78457 Konstanz, Germany.
*Corresponding author. E-mail: elke.deuerling@uni-konstanz.de

## NAC depletion shortens life span and induces ER and mitochondrial stress in *C. elegans*

To understand the potential role of the ribosome-associated α,β-heterodimeric NAC in protein transport, we used conditional RNA interference (RNAi)–mediated knockdown in adult animals because deletion of NAC in *C. elegans* is embryonically lethal (*13*). The knockdown of α-NAC and β-NAC, individually and in combination, substantially shortened the life span (Fig. 1A), suggesting an important function in adult *C. elegans*. Next, we used various cell compartment–specific stress reporter strains that express green fluorescent protein (GFP) under a specific stress-inducible promoter. To specifically measure cytosolic stress, we used transgenic animals carrying the *hsp-16.2p::GFP* reporter (*14*). In these worms, GFP fluorescence strongly increased upon heat shock (fig. S1A) (*15*), demonstrating the functionality of this reporter strain. Knockdown of NAC did not increase GFP expression (Fig. 1B and fig. S1B). Moreover, quantitative polymerase chain reaction (PCR) analysis of different mRNAs encoding cytosolic heat shock proteins revealed an inconsistent expression pattern. One gene was up-regulated (*hsp-70*), but the others were unaltered (*hsp-16.41* and *hsp-17*) and even down-regulated (*f44e5.4*, a member of the Hsp70 family), suggesting that NAC depletion did not cause a robust cytosolic stress response (Fig. 1F). However, depletion of NAC in animals carrying the *hsp-4p::GFP* reporter for analysis of ER stress (*16*) strongly increased GFP fluorescence (Fig. 1C and fig. S1C). Furthermore, β-NAC depletion in wild-type worms (Bristol N2) strongly induced the expression of the ER-resident chaperones Hsp-4 and Pdi-3 (Fig. 1E) and revealed enhanced *hsp-4* transcript levels as well as elevated levels of total and spliced *xbp-1* mRNA that encodes the main transcription factor of the ER stress response (Fig. 1F). Depletion of the β-NAC subunit by means of RNAi resulted in the destabilization of the α-NAC subunit (and vice versa) and thus resulted in depletion of the entire complex (Fig. 1E). Thus, NAC depletion provokes an ER stress response in adult *C. elegans*. In addition, GFP fluorescence was slightly increased in animals carrying the *hsp-6p::GFP* reporter (*17*) that specifically monitors mitochondrial stress (Fig. 1D and fig. S1D). Consistent with mitochondrial stress, we detected enhanced expression of the mitochondrial chaperone Hsp-60 in wild-type worms upon knockdown of β-NAC (Fig. 1E). Transcript levels of *hsp-60* as well as of another mitochondrial chaperone, *hsp-6*, were also significantly elevated (Fig. 1F).

Might NAC thus regulate the transport of proteins from the cytosol to the ER and mitochondria? To address this possibility, we impaired the transport pathways to these compartments by depleting the crucial targeting factors SRβ and tomm-20, respectively, and analyzed the resulting stress responses in dependence of NAC. Knockdown of β-NAC together with SRβ increased the ER stress response (Fig. 1, G and H, and fig. S2A), whereas the combined knockdown of β-NAC and tomm-20 did not provoke an enhanced mitochondrial stress response (fig. S2C). Thus, NAC has a critical function in regulating the cotranslational protein transport to the ER but seemingly does not affect the sorting of proteins to mitochondria directly. Consistent with a potential role of NAC in ER targeting, the ER stress response was also increased when β-NAC was depleted together with SRα or together with SRP54, which is an essential subunit of SRP (fig. S2, A and B).

## NAC prevents SRP-independent binding of ribosomes to ER translocons

Next, we separated whole animal extracts into cytosolic and membrane fractions and investigated ribosomal particles through sucrose density gradient sedimentation analysis. As noted previously (*18*), we observed a decline of cytosolic polysome levels in β-NAC–depleted animals (Fig. 2A). Conversely, however, the levels of membrane-bound ribosomes were strongly increased, suggesting that ribosomes were redistributed from the cytosol toward the ER membrane in the absence of NAC (Fig. 2B). Consistently, immunoblot analysis also showed a clear shift of the ribosomal protein RPL-10 from the cytosol toward the membrane fraction in β-NAC–depleted animals (Fig. 2C). 12% of total ribosomes were membrane-attached in control animals, whereas in β-NAC–depleted animals, this ribosome pool increased to 31% (Fig. 2D). RNAi-mediated depletion of Sec61α—the major constituent of the ER translocon, individually and in combination with β-NAC—strongly reduced the levels of ribosomes in the membrane fraction and, concomitantly, increased them in the cytosolic fraction (Fig. 2, C and D, and fig. S3A). Because the Sec61α levels were

unaltered in β-NAC–depleted animals (Fig. 2C and fig. S3C), these data suggest that ribosomes interact to a much greater extent with ER translocons in the absence of NAC.

Next, we knocked down SRP54 to investigate whether the SRP pathway contributed to the increased binding of ribosomes to the Sec61 complex in NAC-depleted animals. As expected, SRP54 depletion shifted ribosomes from the ER membrane toward the cytosol (Fig. 2, C and D, and fig. S3B). Such ribosome redistribution did not occur in animals depleted of both β-NAC and SRP54 (Fig. 2, C and D, and fig. S3B). Because knockdown efficiencies of SRP54 and β-NAC were comparable in the single and double RNAi setups (fig. S3D), these data demonstrate that ribosomes get recruited to the ER membrane SRP-independently.

The increased, SRP-independent binding of RNCs to ER translocons suggested that ribosomes directly interact with Sec61. To test this possibility, we incubated puromycin and high salt-stripped ribosomes as well as microsomes in the presence and absence of NAC and then repurified microsomes by means of centrifugation. In the absence of NAC, ribosomes readily bound to the microsomal fraction, whereas this binding was efficiently blocked when NAC was added (Fig. 2E). The interaction of NAC with ribosomes was necessary for this blocking activity because a ribosome-binding deficient mutant of NAC, $^{RRK/AAA}$NAC (fig. S4), did not prevent binding of ribosomes to microsomes (Fig. 2E). Furthermore, we tested whether NAC also played a role in releasing ribosomes from the ER membrane after translation termination. We extracted native rough microsomes and tested the puromycin-induced release of bound ribosomes in dependence of NAC. As previously shown (*19*), treatment with puromycin only marginally released ribosomes from membranes; however, the addition of NAC strongly enhanced the puromycin-induced ribosome release, whereas $^{RRK/AAA}$NAC had no effect (Fig. 2F). Thus, NAC is a strong negative regulator of ER-transport that prevents direct, SRP-independent interactions of ribosomes with ER translocons.

### NAC overexpression impairs SRP-dependent ER targeting

We reasoned that if NAC generally acts as an ER targeting inhibitor, NAC overexpression could

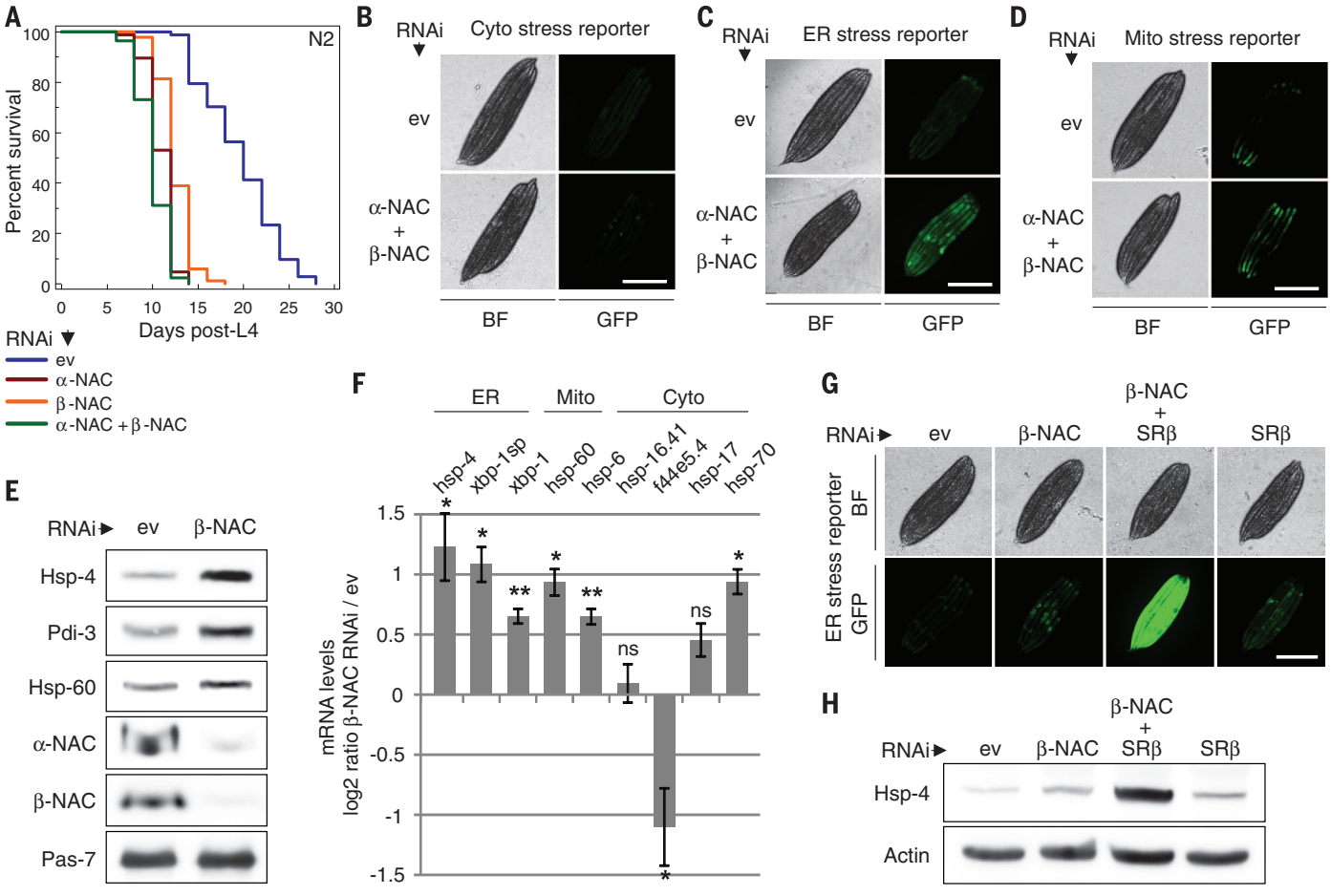



**Fig. 1. NAC depletion shortens adult life span and induces stress in the ER and mitochondria.** (**A**) Life span survival curves of N2 worms grown at 20°C on empty vector control (ev, blue; median life span 20 days), α-NAC RNAi (red; median life span 12 days), β-NAC RNAi (orange; median life span 12 days), or α-NAC + β-NAC RNAi (green; median life span 10 days). (**B**) *hsp-16.2p::GFP* cytosolic stress reporter worms were grown on empty vector control (ev) or α-NAC + β-NAC RNAi. GFP fluorescence was assessed on day 3 of adulthood (fig. S1, A and B). BF, bright-field. Scale bar, 0.5 mm. (**C**) *hsp-4p::GFP* ER stress reporter worms were grown on empty vector control (ev) or α-NAC + β-NAC RNAi, and GFP fluorescence was assessed on day 3 of adulthood (fig. S1C). BF, bright-field. Scale bar, 0.5 mm. (**D**) *hsp-6p::GFP* mitochondrial stress reporter worms were grown on empty vector control (ev) or α-NAC + β-NAC RNAi, and GFP fluorescence was assessed on day 2 of adulthood (fig. S1D). BF, bright-field. Scale bar, 0.5 mm. (**E**) N2 worms were grown on empty vector control (ev) or β-NAC RNAi. On day 3 of adulthood, levels of indicated proteins were assessed with immunoblotting. Immunoblot against Pas-7 served as loading control. (**F**) N2 worms were grown on empty vector control (ev) or β-NAC RNAi. On day 3 of adulthood, mRNA levels of indicated genes were assessed by means of quantitative reverse transcriptase–PCR (RT-PCR). Data are represented as mean ± SD. A Student's *t* test was used to assess significance: $*P < 0.05$, $**P < 0.01$; ns, not significant. *xbp-1sp* = *xbp-1* spliced. (**G**) *hsp-4p::GFP* ER stress reporter worms were grown on empty vector control (ev), β-NAC RNAi, β-NAC + SRβ RNAi, or SRβ RNAi, and GFP fluorescence was assessed on day 2 of adulthood (fig. S2A). BF, bright-field. Scale bar, 0.5 mm. (**H**) N2 worms were grown on empty vector control (ev) or indicated RNAi. On day 2 of adulthood, Hsp-4 levels were analyzed with immunoblotting. Actin served as loading control.

interfere with the binding of ribosomes to ER translocons. We thus established transgenic animals overexpressing both NAC subunits from a strong and ubiquitous promoter (*eft-3p*). The transgenic strains showed a moderate two- to threefold overexpression of NAC (Fig. 3, C and D). All attempts to get strains with a much stronger overexpression by injecting higher amounts of DNA produced no transgenic progeny, indicating that NAC expression over a critical threshold was lethal in *C. elegans*. Consistent with this, the moderately NAC-overexpressing strains were already slightly delayed in development. The levels of cytosolic polysomes were slightly increased in NAC-overexpressing animals (Fig. 3A; quantification is provided in fig. S5A). In contrast, the levels of membrane-attached ribosomes were markedly reduced (Fig. 3B and fig. S5A). Immunoblot analysis of RPL-10 confirmed the membrane-to-cytosol shift of ribosomes, although the levels of Sec61α were unchanged (Fig. 3C). Moreover, Hsp-4 levels were significantly increased in NAC-overexpressing worms, indicating ER stress (Fig. 3D). Thus, NAC overexpression interferes with the targeting of ribosomes to ER translocons, likely by impairing the SRP pathway. To investigate this possibility, we analyzed whether NAC overexpression altered the binding of SRP to ribosomes that would be expected upon targeting inhibition. The total levels of SRP were unchanged in NAC-overexpressing animals (Fig. 3E). However, the binding of SRP to ribosomes was significantly shifted to late polysomes (Fig. 3F), indicating that new rounds of translation initiation on mRNAs encoding SRP substrates occurred in the cytosol because overexpressed NAC delayed the timely targeting of RNCs to the ER. To further explore the effects of NAC overexpression on ER targeting, we depleted critical SRP pathway components by means of RNAi in the first larval stage (L1) so as to compromise protein transport to the ER. The knockdown of SRP54, SRα, and SRβ induced developmental defects in wild-type worms that resulted in a partial larval arrest (fig. S5B). NAC-overexpressing animals were hypersensitive to these RNAi treatments (fig. S5B), whereas NAC-depleted animals were less sensitive (fig. S5C). This suggests that NAC expression above a critical threshold counteracts the activity of the SRP pathway. To corroborate this, we assessed ER targeting directly by analyzing the distribution of four ribosome-associated mRNAs encoding specific SRP substrates (*ile-2*, *sec-61*, *R186.3*, and *hsp-4*) between the cytosol and the ER membrane. As expected, knockdown of SRP54 in control animals strongly shifted these RNCs toward the cytosol (Fig. 3G). NAC overexpression strongly aggravated the targeting defects induced by SRP54 depletion (Fig. 3G). Thus, NAC may act as a general ER targeting inhibitor whose expression levels need to be accurately balanced to allow a specific derepression of ER targeting by the SRP pathway.

### Global mistargeting of ribosomes to the ER in the absence of NAC

Next, we used a genome-wide microarray approach to measure the global distribution of ribosome-associated mRNAs between the cytosol and the ER membrane. Because NAC depletion in *C. elegans* results in a strongly reduced brood size (*18*), we used a temperature-sensitive sterile mutant (SS104) to minimize the risk of artifacts owing to unequal embryonic gene expression. The sterile mutants showed a NAC RNAi phenotype similar to wild-type worms (fig. S6, A to E). To test ER targeting specificity, we isolated cytosolic and membrane-attached ribosomes and quantified the translated mRNAs using DNA microarrays. The majority of mRNAs (93%) that were significantly enriched in the cytosolic fraction of control animals encoded proteins without a predicted N-terminal signal

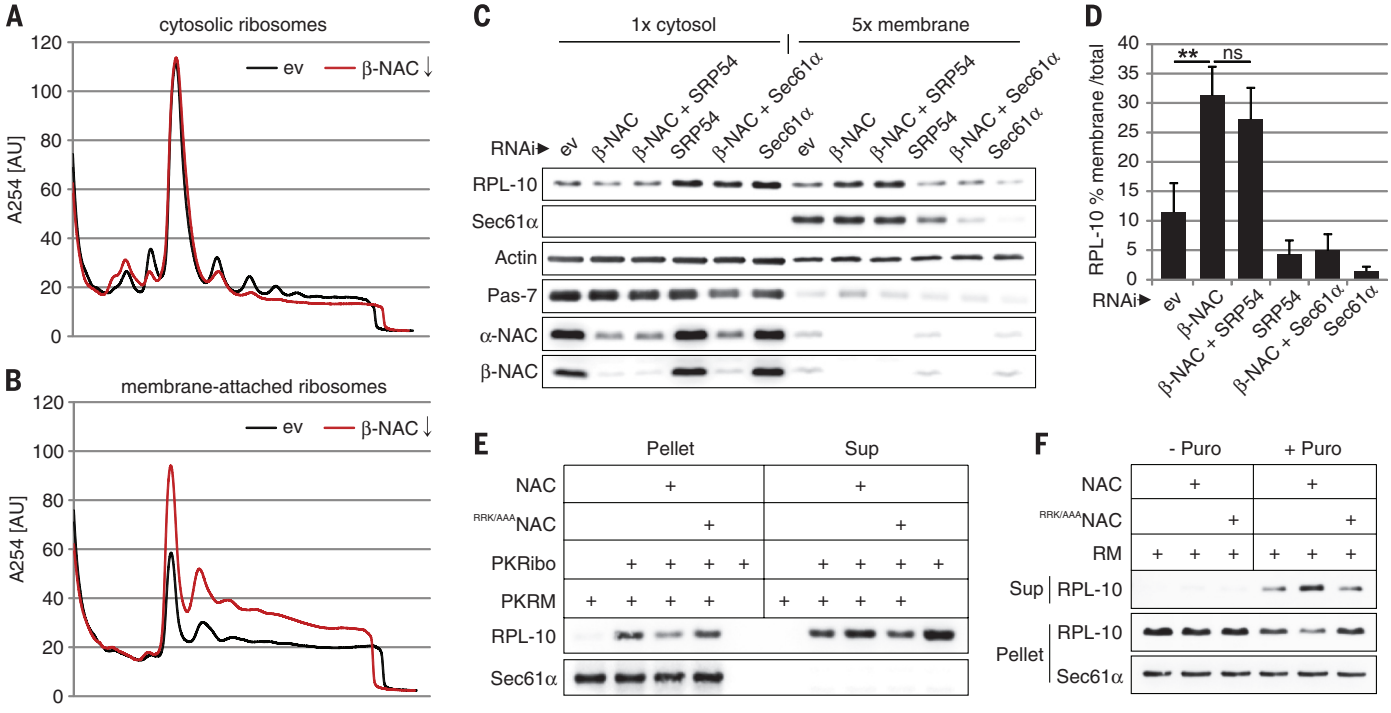



**Fig. 2. NAC prevents SRP-independent ribosome-translocon binding.** (**A**) Cytosolic polysome profiles of day-3 adult N2 worms grown on empty vector control (ev, black) or β-NAC RNAi (red). (**B**) Polysome profiles of membrane-attached ribosomes from animals as in (A). (**C**) N2 worms were grown on empty vector control (ev) or indicated RNAi. On day 3 of adulthood, indicated protein levels in cytosolic and membrane fractions were assessed with immunoblotting. Membrane fraction was loaded in fivefold excess over cytosolic fraction (5× membrane). Actin and Pas-7 served as loading controls (fig. S3, A and B). (**D**) Diagram shows percentage of membrane-bound ribosomes by means of RPL-10 levels in the cytosolic and membrane fractions shown as in (C). Data are represented as mean ± SD. A Student's *t* test was used to assess significance: $^{**}P < 0.01$; ns, not significant. (**E**) Puromycin and high salt-stripped ribosomes (PKRibo) and microsomes (PKRM) were incubated in the presence and absence of recombinant NAC or the ribosome-binding mutant $^{RRK/AAA}$NAC (fig. S4). Microsomes were pelleted, and bound ribosomes were analyzed by means of RPL-10 levels. Sup, supernatant. (**F**) Native rough microsomes (RM) were treated with puromycin (Puro) in the presence and absence of recombinant NAC or $^{RRK/AAA}$NAC. Microsomes were pelleted. Bound and released ribosomes were analyzed in the pellet and supernatant (Sup), respectively, with immunoblot analysis of RPL-10.

sequence or a transmembrane domain (Fig. 4A). Conversely, most mRNAs (88%) that were significantly enriched in the membrane fraction encode proteins containing a predicted N-terminal signal sequence, a transmembrane domain, or both (Fig. 4A). Thus, ER targeting in control animals was accurate, and our fractionation protocol appeared valid. We next compared the distribution of these mRNAs between control and NAC-depleted animals. Scatter plot analysis of $\log_2$-transformed mRNA ratios between membrane and cytosol revealed that the distribution of the membrane-enriched mRNAs was indistinguishable between control and NAC-depleted animals [correlation coefficient ($r$) = 1.0356] (Fig. 4B). However, the cytosol-enriched mRNAs were strongly shifted toward the membrane fraction in NAC-depleted worms, as evident by the lower linear regression slope ($r$ = 0.5533) (Fig. 4B). Thus, NAC depletion does not affect the specific, SRP-mediated targeting of ribosomes to the ER membrane but induces additional unspecific binding of incorrect ribosomes. We verified these data for several mRNAs with quantitative PCR in wild-type worms. In agreement with the microarray data, mRNAs encoding nuclear, cytosolic, and mitochondrial proteins were strongly mistargeted to the ER membrane in β-NAC–depleted animals (Fig. 4C). The additional knockdown of SRP54 did not reduce the levels of these mRNAs at the ER membrane, confirming that mistargeting occurs in a SRP-independent manner. Moreover, β-NAC depletion did not alter the distribution of mRNAs encoding ER proteins, whereas knockdown of SRP54 strongly shifted these mRNAs toward the cytosol (Fig. 4C). Thus, NAC depletion does not impair the activity and specificity of the SRP pathway. The additional knockdown of β-NAC partially restored the targeting defect in SRP54-depleted worms (Fig. 4C), suggesting that in the absence of NAC ribosomes translating SRP substrates bound ER translocons in a SRP-independent manner. This explains why the larval arrest induced by impairment of the SRP pathway was significantly lower in NAC-depleted animals (fig. S5C). Furthermore, although targeting of SRP substrates was partially restored in NAC/SRP54–depleted worms, we consistently observed a further moderate increase of ribosome mistargeting (Fig. 4C). Enhanced mistargeting and at the same time less specific targeting explains why the combined depletion of NAC and essential SRP pathway components provoked such an increased ER stress response (Fig. 1, G and H, and fig. S2, A and B).

### Mitochondrial proteins get mislocalized and degraded by ERAD upon NAC depletion

Knowing that a posttargeting signal sequence recognition and proofreading step takes place at the Sec61 complex (*9*), we analyzed whether or not the mistargeted ribosomes translocate their nascent chains across the ER membrane. We reasoned that potential mislocalized proteins might be degraded by the ER-associated degradation (ERAD) pathway (*20*). To enrich these substrates we depleted Sel-1, an essential component of the ERAD ubiquitin ligase complex (*21*). The knockdown of Sel-1 provoked a moderate ER stress response (fig. S7). The combined knockdown of Sel-1 and β-NAC strongly enhanced the ER stress response, which suggests that the ERAD pathway is indeed activated in NAC-depleted animals to maintain ER protein homeostasis (fig. S7). Many proteins that enter the ER lumen are *N*-glycosylated on asparagine residues in N–X–S/T sequence motifs

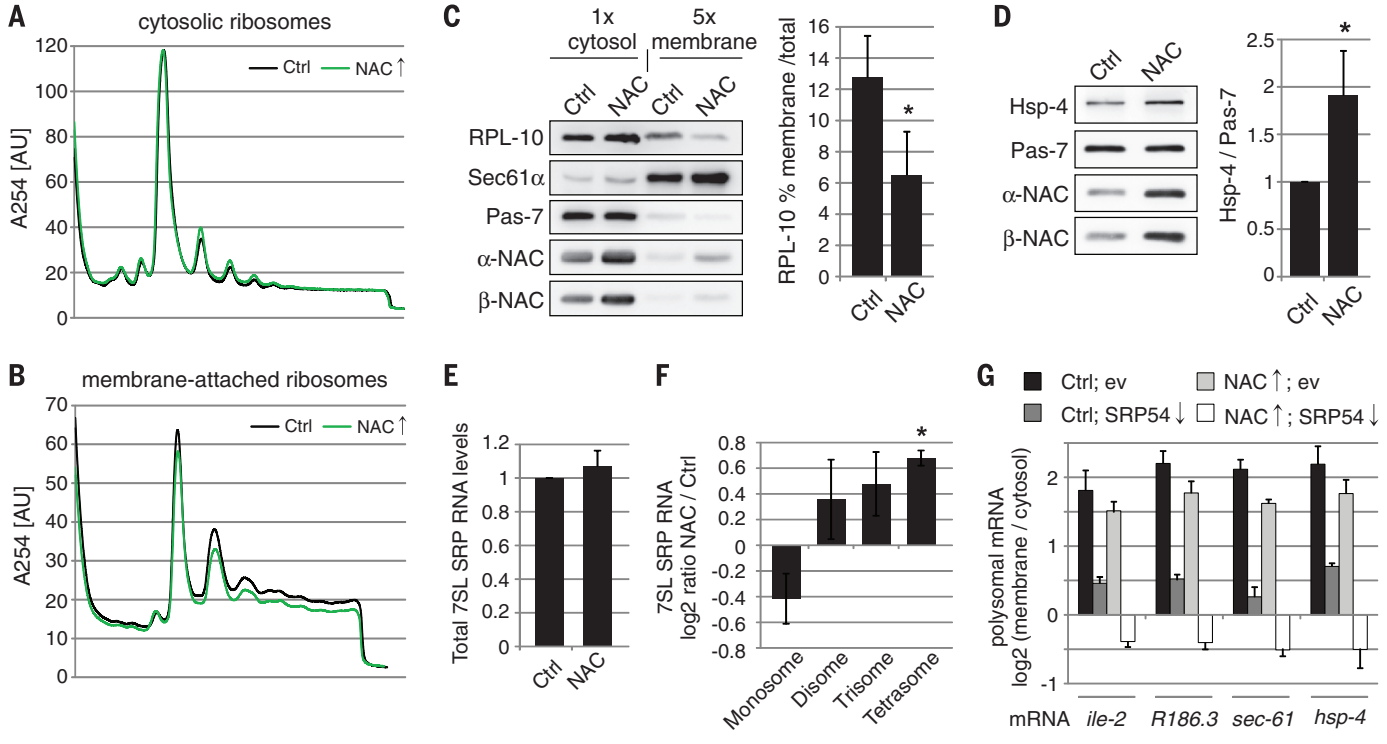



**Fig. 3. NAC overexpression interferes with SRP-dependent ER targeting.** (**A**) Cytosolic polysome profiles of day-2 adult control (Ctrl, black) and NAC-overexpressing worms (green). (**B**) Polysome profiles of membrane-attached ribosomes from animals as in (A). (**C**) Control (Ctrl) and NAC-overexpressing worms were harvested on day 2 of adulthood, and indicated protein levels in cytosolic and membrane fractions were assessed with immunoblotting. Membrane fraction was loaded in fivefold excess over cytosolic fraction (5× membrane). Pas-7 served as loading control. (**D**) Total extracts of animals as in (C) were prepared and indicated proteins were analyzed with immunoblotting. Pas-7 served as loading control. (**E**) Control (Ctrl) and NAC-overexpressing worms were harvested on day 2 of adulthood, and total 7SL SRP RNA levels were assessed by means of quantitative RT-PCR. Data are represented as mean ± SD. (**F**) Monosome, disome, trisome, and tetrasome fractions of ribosome sedimentation analysis shown as in (A) were collected, and 7SL SRP RNA levels were assessed by means of quantitative RT-PCR. Data are represented as mean ± SD. A Student's $t$ test was used to assess significance: $^*P < 0.05$. (**G**) Control (Ctrl) and NAC-overexpressing worms were grown on empty vector control (ev) or SRP54 RNAi. On day 3 of adulthood, the distribution of indicated ribosome-associated mRNAs between cytosol and ER membrane was assessed by means of quantitative RT-PCR. Depicted is the $\log_2$-transformed membrane-to-cytosol ratio for indicated mRNAs. Data are represented as mean ± SD.

(*22*). We speculated that misdirected substrates containing such motifs would be *N*-glycosylated, which could be used as a footprint for mistranslocated proteins. We thus isolated glycosylated proteins using the lectin Concanavalin A (ConA) and analyzed with mass spectrometry (MS) proteins that only appeared in β-NAC/Sel-1–depleted animals. One prominent hit in the MS analysis was the mitochondrial protein Hsp-60, which contains the potential N-glycosylation site N-T-S (Fig. 5A). A substantial amount of Hsp-60 accumulated in the ConA-bound fraction of animals depleted of both β-NAC and Sel-1 (Fig. 5B). Thus, the mitochondrial protein Hsp-60 is mislocalized to the ER and degraded by ERAD in NAC-depleted animals. We then tested a second mitochondrial protein, Atp-2, that contains the predicted N-glycosylation motif N-A-S (Fig. 5A). Similar to Hsp-60, Atp-2 was also strongly enriched in the ConA-bound fraction (Fig. 5B). Treatment of the lysate with the peptide *N*-glycosidase F (PNGaseF) to remove N-glycans before the lectin affinity purification strongly diminished ConA-binding of Hsp-60 and Atp-2, corroborating that these mitochondrial proteins were indeed *N*-glycosylated (Fig. 5C). The highly abundant cytosolic protein actin (Act-1) that harbors the predicted *N*-glycosylation site N-G-S (Fig. 5A) did not accumulate in the ConA-bound fraction (Fig. 5B), although Act-1 RNCs were mistargeted to the ER translocon (Fig. 4C). This could indicate that mitochondrial targeting sequences override the proofreading function of the ER translocon pore. To analyze this more closely, we made use of a split GFP (spGFP) complementation approach (*23*). We engineered different cell compartment–specific spGFP fragments for the cytoplasm, as well as for the ER and mitochondria by fusing specific targeting sequences to spGFP (Fig. 5D). The spGFP fragments, when expressed in the same compartment, complemented only in the expected cell organelle (Fig. 5D). Animals expressing spGFP1-10 in the ER and spGFP11 in the cytosol only showed a weak background signal, and β-NAC depletion did not increase GFP fluorescence (Fig. 5E). However, knockdown of β-NAC strongly increased the GFP fluorescence in animals expressing ER-targeted spGFP1-10 and mitochondria-targeted spGFP11 (Fig. 5F). The GFP fluorescence in these worms showed a typical reticular ER pattern, revealing that the mitochondria-targeted spGFP fragments were misdirected to the ER in the absence of NAC (Fig. 5G). Thus, mitochondrial proteins in particular are incorrectly transported to the ER and subsequently degraded by ERAD. This could explain why NAC depletion induces both ER and mitochondrial stress. NAC is both an essential negative regulator for ribosome-translocon interaction to sustain ER targeting accuracy and is also important to preserve the specificity of cellular protein sorting to mitochondria.

## Discussion

These findings have revealed the essential antagonistic role of NAC in regulating cotranslational protein transport to the ER. Only the combined opposing activities of SRP and NAC guarantee a robust sorting system that ensures the fidelity and specificity of protein translocation in vivo. The importance of NAC is underlined by the mistargeting of noncognate RNCs and by the fact that the in-built proofreading mechanism of the Sec61 translocon (*9*) can be overruled by mistargeted ribosomes translating mitochondrial proteins. The prevailing view has been that only nascent polypeptide chains with sufficiently hydrophobic signal sequences or transmembrane domains can open the translocation channel, whereas nonauthentic substrates would be rejected. In this work, mitochondrial targeting sequences consisting of a stretch of alternating positively charged and hydrophobic amino acid residues forming an amphipathic helix (*24*) could, at least in part, open the ER translocation channel. This is consistent with several studies showing the existence of ambiguous signals directing proteins to both mitochondria as well as the ER (*25*).

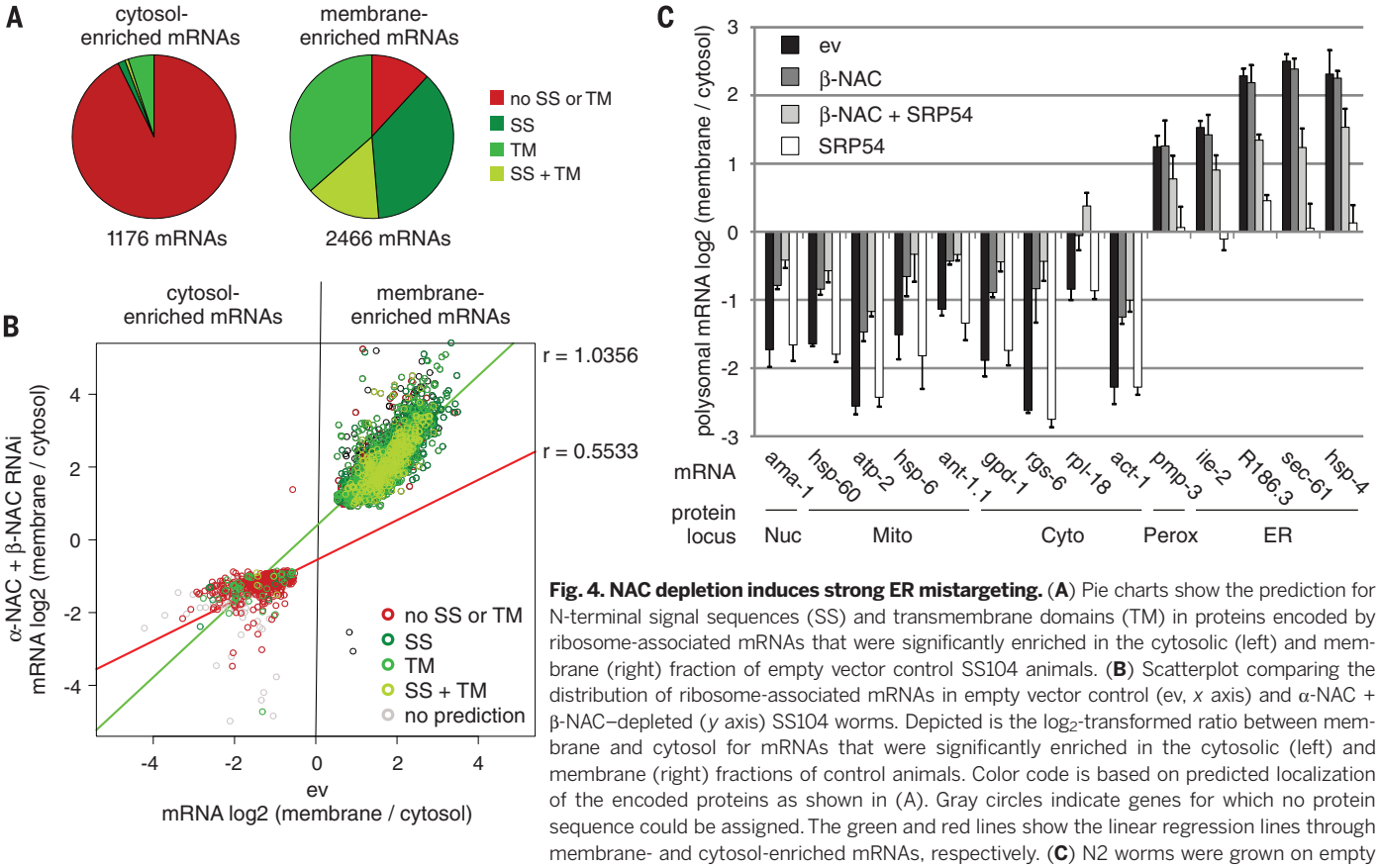



**Fig. 4. NAC depletion induces strong ER mistargeting.** (**A**) Pie charts show the prediction for N-terminal signal sequences (SS) and transmembrane domains (TM) in proteins encoded by ribosome-associated mRNAs that were significantly enriched in the cytosolic (left) and membrane (right) fraction of empty vector control SS104 animals. (**B**) Scatterplot comparing the distribution of ribosome-associated mRNAs in empty vector control (ev, *x* axis) and α-NAC + β-NAC–depleted (*y* axis) SS104 worms. Depicted is the $\log_2$-transformed ratio between membrane and cytosol for mRNAs that were significantly enriched in the cytosolic (left) and membrane (right) fractions of control animals. Color code is based on predicted localization of the encoded proteins as shown in (A). Gray circles indicate genes for which no protein sequence could be assigned. The green and red lines show the linear regression lines through membrane- and cytosol-enriched mRNAs, respectively. (**C**) N2 worms were grown on empty vector control (ev) or indicated RNAi. On day 3 of adulthood, the distribution of indicated ribosome-associated mRNAs between cytosol and ER membrane was assessed by means of quantitative RT-PCR. Depicted is the $\log_2$-transformed membrane-to-cytosol ratio for select mRNAs coding for proteins with destination in the nucleus (Nuc), mitochondria (Mito), cytosol (Cyto), peroxisomes (Perox), and ER. Data are represented as mean ± SD.

Our *C. elegans* study reveals that NAC counteracts the autonomous binding of ribosomes to Sec61 to ensure ER targeting specificity. A role for NAC in ER protein targeting has been indicated by several in vitro analyses (*7*, *11*, *26*). However, this hypothesis was rapidly dismissed by other in vitro studies providing contrary results (*27*, *28*). Moreover, yeast in vivo studies showed no aberrant translocation phenotype upon NAC deletion (*29*, *30*). Why the importance of NAC in ER protein targeting was missed in yeast studies is unclear. The discrepancy could be explained by the fact that yeast cells use a distinct posttranslational ER targeting system, which is seemingly not used in such a pronounced way by higher eukaryotes (*31*). Consistent with this, yeast is the only known organism in which cotranslational ER targeting by the SRP pathway is not essential (*32*). Another possible explanation for the different effects of NAC depletion in yeast and higher eukaryotes could be the potential disparate engagement of available Sec61 channels in the translocation of correct substrates. Our study in *C. elegans* reveals that NAC depletion does not impair the correct targeting of ribosomes to the ER membrane. Rather, vacant translocons that exist in adult worms become occupied by incorrect ribosomes in a SRP-independent manner. It is very likely that in yeast under optimal growth conditions most if not all translocons are saturated, thus counteracting the potential mistargeting of ribosomes occurring in the absence of NAC.

The Sec61 translocon interacts with the ribosome primarily via two cytoplasmic loops of Sec61α as well as with the N-terminal helix of Sec61γ. Both contact the ribosomal proteins uL23 and eL29 as well as the backbone of the 28*S* ribosomal RNA (rRNA) adjacent to the nascent peptide exit site (*33*). The extensive binding surfaces account for the very high inherent affinity ribosomes exhibit for the Sec61 complex (*8*, *9*, *11*). Cross-linking data suggest that NAC contacts the ribosome via the ribosomal protein uL23 (*34*). Because this ribosomal protein is also a major contact point between the ribosome and Sec61, it is very likely that NAC sterically inhibits this high-affinity translocon interaction site. However, recent studies challenge the interaction of NAC with uL23 and suggest that NAC binds to eL31 (*35*, *36*). This ribosomal protein is placed like uL23 near the tunnel exit; thus, NAC could block the functional coupling of ribosomes and translocons also via eL31. In agreement with a function as a general ER targeting inhibitor, NAC is present in equimolar concentration relative to ribosomes, it can interact with virtually all RNCs, and binds reversibly to ribosomes (*29*, *34*, *37*). The reversible binding mode allows a specific derepression of ER targeting. We propose that displacement of NAC from RNCs could be mediated by SRP that, similar to NAC, also contacts the ribosomal protein uL23 at the tunnel exit (*1*). The high affinity of SRP to hydrophobic signal sequences or transmembrane domains in nascent chains (*6*, *27*) probably gives SRP a selective advantage over NAC for ribosome binding. A competition between NAC and SRP for ribosome binding has been previously indicated by several

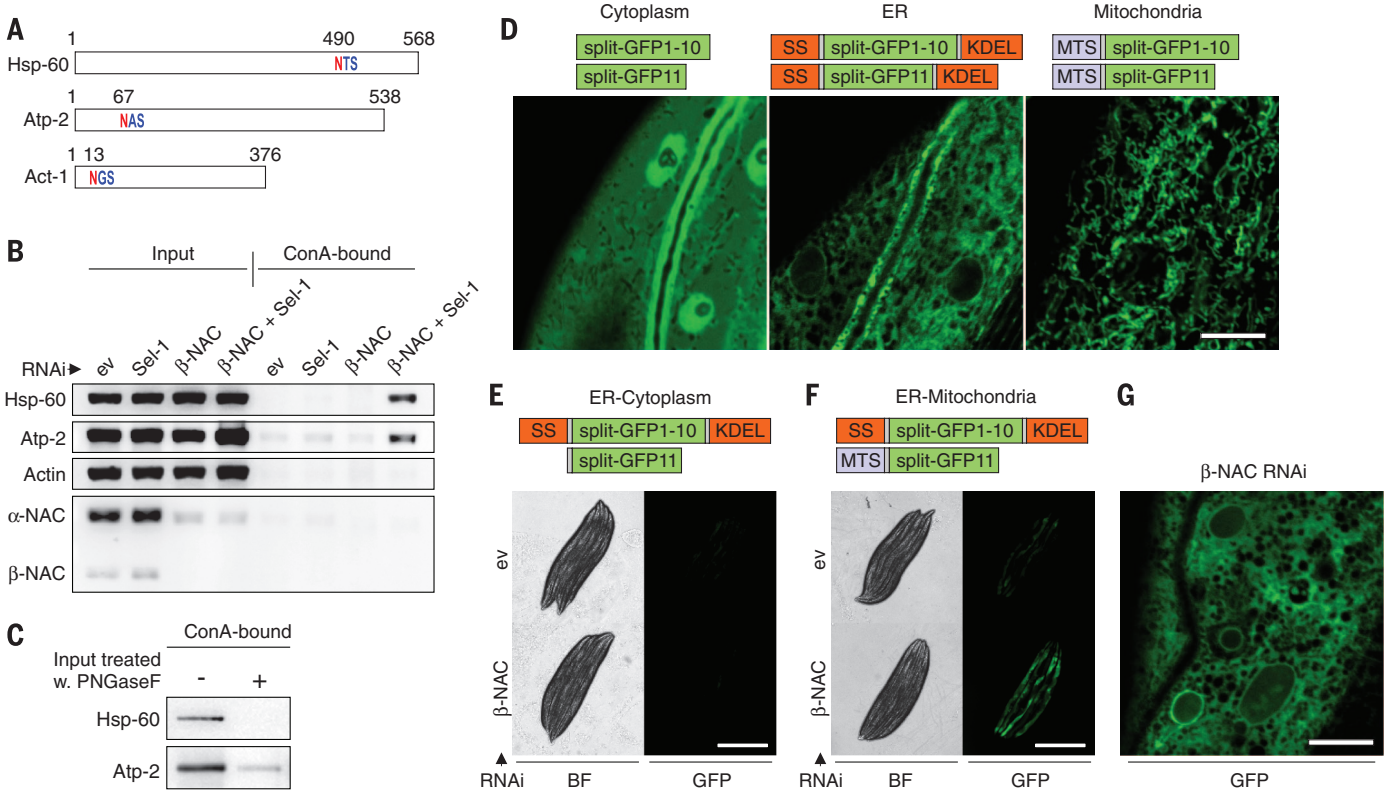



**Fig. 5. Mitochondrial proteins are mislocalized to the ER and degraded by ERAD upon NAC depletion.** (**A**) Schematic illustration showing predicted *N*-glycosylation sites in the mitochondrial proteins Hsp-60 and Atp-2 and the cytosolic protein Act-1. Predictions were performed with NetNGlyc 1.0. (**B**) N2 worms were grown on empty vector control (ev) or indicated RNAi. On day 3 of adulthood, glycosylated proteins were isolated by using the lectin Concanavalin A (ConA), and the levels of indicated proteins in the total extract (Input) and the lectin-bound fraction were analyzed with immunoblotting. (**C**) Total extracts (Input) of N2 worms grown on β-NAC + Sel-1 RNAi were treated or not with the *N*-glycosidase PNGaseF followed by Concanavalin A (ConA) affinity purification. The levels of ConA-bound Hsp-60 and Atp-2 were then assessed with immunoblotting. (**D**) Fluorescent micrographs of intestinal cells in worms expressing complementing split-GFP fragments in the cytoplasm (left), in the ER (middle) and in mitochondria (right). Scale bar, 10 μm. SS, ER-specific signal sequence; KDEL, ER retention motif; MTS, mitochondrial targeting sequence. (**E**) Transgenic worms expressing split-GFP1-10 in the ER and split-GFP11 in the cytoplasm were grown on empty vector control (ev) or β-NAC RNAi, and GFP fluorescence was assessed on day 2 of adulthood. BF, bright-field. Scale bar, 0.5 mm. (**F**) Transgenic worms expressing split-GFP1-10 in the ER and split-GFP11 in mitochondria were grown on empty vector control (ev) or β-NAC RNAi, and GFP fluorescence was assessed on day 2 of adulthood. BF, bright-field. Scale bar, 0.5 mm. (**G**) Fluorescent micrograph of intestinal cells in β-NAC–depleted worms as shown in (F). Scale bar, 10 μm.

in vitro studies (*26*, *36*, *38*). However, at which specific step the inhibitory action of NAC on ER targeting is counteracted by the SRP pathway is unclear.

NAC is evolutionarily conserved from yeast to man and is essential in eukaryotes, except for yeast (*13*, *30*, *39*, *40*). Knockdown of α-NAC in human HeLa S3 cells also activates ER stress responses and causes mitochondrial dysfunction (*41*). Thus, the function of NAC as an ER targeting inhibitor appears to be well conserved during evolution, at least from *C. elegans* to mammalians. Recent studies showed that NAC is sequestered by cytosolic aggregates under protein folding stress conditions (*18*, *42*). This raises the question whether proteotoxic stress in the cytoplasm causes dysfunction of NAC, leading to incorrect sorting of proteins to the ER and mitochondria. A link between cytosolic protein aggregation and ER stress is well established (*43*), and it will be interesting to investigate the role of NAC in this context.

## REFERENCES AND NOTES


1. M. Halic *et al.*, *Nature* **427**, 808–814 (2004).
2. M. Halic *et al.*, *Science* **312**, 745–747 (2006).
3. Y. Nyathi, B. M. Wilkinson, M. R. Pool, *Biochim. Biophys. Acta* **1833**, 2392–2402 (2013).
4. T. A. Rapoport, *Nature* **450**, 663–669 (2007).
5. T. Schwartz, G. Blobel, *Cell* **112**, 793–803 (2003).
6. P. Walter, I. Ibrahimi, G. Blobel, *J. Cell Biol.* **91**, 545–550 (1981).
7. B. Wiedmann, H. Sakai, T. A. Davis, M. Wiedmann, *Nature* **370**, 434–440 (1994).
8. N. Borgese, W. Mok, G. Kreibich, D. D. Sabatini, *J. Mol. Biol.* **88**, 559–580 (1974).
9. B. Jungnickel, T. A. Rapoport, *Cell* **82**, 261–270 (1995).
10. K. U. Kalies, D. Görlich, T. A. Rapoport, *J. Cell Biol.* **126**, 925–934 (1994).
11. B. Lauring, H. Sakai, G. Kreibich, M. Wiedmann, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5411–5415 (1995).
12. A. Prinz, E. Hartmann, K. U. Kalies, *Biol. Chem.* **381**, 1025–1029 (2000).
13. T. A. Bloss, E. S. Witze, J. H. Rothman, *Nature* **424**, 1066–1071 (2003).
14. S. L. Rea, D. Wu, J. R. Cypser, J. W. Vaupel, T. E. Johnson, *Nat. Genet.* **37**, 894–898 (2005).
15. Materials and methods are available as supplementary materials on *Science* Online.
16. M. Calfon *et al.*, *Nature* **415**, 92–96 (2002).
17. T. Yoneda *et al.*, *J. Cell Sci.* **117**, 4055–4066 (2004).
18. J. Kirstein-Miles, A. Scior, E. Deuerling, R. I. Morimoto, *EMBO J.* **32**, 1451–1468 (2013).
19. M. R. Adelman, D. D. Sabatini, G. Blobel, *J. Cell Biol.* **56**, 206–229 (1973).
20. S. S. Vembar, J. L. Brodsky, *Nat. Rev. Mol. Cell Biol.* **9**, 944–957 (2008).
21. F. Urano *et al.*, *J. Cell Biol.* **158**, 639–646 (2002).
22. F. Schwarz, M. Aebi, *Curr. Opin. Struct. Biol.* **21**, 576–582 (2011).
23. S. Cabantous, T. C. Terwilliger, G. S. Waldo, *Nat. Biotechnol.* **23**, 102–107 (2005).
24. W. Neupert, J. M. Herrmann, *Annu. Rev. Biochem.* **76**, 723–749 (2007).
25. O. Yogev, O. Pines, *Biochim. Biophys. Acta* **1808**, 1012–1020 (2011).
26. I. Möller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13425–13430 (1998).
27. A. Neuhof, M. M. Rolls, B. Jungnickel, K. U. Kalies, T. A. Rapoport, *Mol. Biol. Cell* **9**, 103–115 (1998).
28. D. Raden, R. Gilmore, *Mol. Biol. Cell* **9**, 117–130 (1998).
29. M. del Alamo *et al.*, *PLOS Biol.* **9**, e1001100 (2011).
30. B. Reimann *et al.*, *Yeast* **15**, 397–407 (1999).
31. T. Ast, G. Cohen, M. Schuldiner, *Cell* **152**, 1134–1145 (2013).
32. S. C. Mutka, P. Walter, *Mol. Biol. Cell* **12**, 577–588 (2001).
33. R. M. Voorhees, I. S. Fernández, S. H. Scheres, R. S. Hegde, *Cell* **157**, 1632–1643 (2014).
34. R. D. Wegrzyn *et al.*, *J. Biol. Chem.* **281**, 2847–2857 (2006).
35. M. Pech, T. Spreter, R. Beckmann, B. Beatrix, *J. Biol. Chem.* **285**, 19679–19687 (2010).
36. Y. Zhang *et al.*, *Mol. Biol. Cell* **23**, 3027–3040 (2012).
37. U. Raue, S. Oellerer, S. Rospert, *J. Biol. Chem.* **282**, 7809–7816 (2007).
38. T. Powers, P. Walter, *Curr. Biol.* **6**, 331–338 (1996).
39. J. M. Deng, R. R. Behringer, *Transgenic Res.* **4**, 264–269 (1995).
40. D. C. Markesich, K. M. Gajewski, M. E. Nazimiec, K. Beckingham, *Development* **127**, 559–572 (2000).
41. Y. Hotokezaka *et al.*, *Cell Death Differ.* **16**, 1505–1514 (2009).
42. H. Olzscha *et al.*, *Cell* **144**, 67–78 (2011).
43. R. V. Rao, D. E. Bredesen, *Curr. Opin. Cell Biol.* **16**, 653–662 (2004).


## ACKNOWLEDGMENTS

We gratefully acknowledge the help of A. Dillin and K. Steffen to get started with this *C. elegans* project. We thank the *Caenorhabditis* Genetics Center for strains and E. Schulze for the microinjection training. We also thank A. Page, W. Neupert, and C. Bargmann for providing reagents. We thank R. Schloemer, K. Turgay, C. Schlatterer, S. Zboron, and E. Oberer-Bley for their critical discussions and valuable help with the manuscript. This work was supported by research grants from the German Science Foundation (DFG; SFB969/A01) and from Human Frontier in Science Program (HFSP; RGP025) to E.D. Data are deposited in Gene Expression Omnibus under accession no. GSE63928.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/201/suppl/DC1
Materials and Methods
Figs. S1 to S7
References (*44*–*48*)

19 December 2014; accepted 13 February 2015
10.1126/science.aaa5335

REPORTS

QUANTUM GASES

# Experimental observation of a generalized Gibbs ensemble

**Tim Langen,[1] Sebastian Erne,[1,2,3] Remi Geiger,[1] Bernhard Rauer,[1] Thomas Schweigler,[1] Maximilian Kuhnert,[1] Wolfgang Rohringer,[1] Igor E. Mazets,[1,4,5] Thomas Gasenzer,[2,3] Jörg Schmiedmayer[1]***

**The description of the non-equilibrium dynamics of isolated quantum many-body systems within the framework of statistical mechanics is a fundamental open question. Conventional thermodynamical ensembles fail to describe the large class of systems that exhibit nontrivial conserved quantities, and generalized ensembles have been predicted to maximize entropy in these systems. We show experimentally that a degenerate one-dimensional Bose gas relaxes to a state that can be described by such a generalized ensemble. This is verified through a detailed study of correlation functions up to 10th order. The applicability of the generalized ensemble description for isolated quantum many-body systems points to a natural emergence of classical statistical properties from the microscopic unitary quantum evolution.**

Information theory provides powerful concepts for statistical mechanics and quantum many-body physics. In particular, the principle of entropy maximization (*1*–*3*) leads to the well-known thermodynamical ensembles, which are fundamentally constrained by conserved quantities such as energy or particle number (*4*). However, physical systems can contain many additional conserved quantities, which raises the question of whether there exists a more general statistical description for the steady states of quantum many-body systems (*5*).

Specifically, the presence of nontrivial conserved quantities puts constraints on the available phase space of a system, which strongly affects the dynamics (*6*–*9*) and inhibits thermalization (*10*–*12*). Instead of relaxing to steady states described by the usual thermodynamical ensembles, a generalized Gibbs ensemble (GGE) was proposed to describe the corresponding steady states via the many-body density matrix

$$\hat{\rho} = \frac{1}{Z}\exp\left(-\sum_m \lambda_m \hat{\mathcal{I}}_m\right) \qquad (1)$$

(*3*, *11*, *13*, *14*), where $\hat{\mathcal{I}}_m$ denotes a set of conserved quantities and $Z = \mathrm{Tr}[\exp(-\sum_m \lambda_m \hat{\mathcal{I}}_m)]$ is the partition function. The Lagrange multipliers $\lambda_m$ associated with the conserved quantities are obtained by maximization of the entropy under the condition that the expectation values of the conserved quantities are fixed to their initial values. The emergence of such a maximum-entropy state does not contradict a unitary evolution according to quantum mechanics. Rather, it reflects that the true quantum state is indistinguishable from the maximum-entropy ensemble with respect to a set of sufficiently local observables (*5*).

[1]Vienna Center for Quantum Science and Technology, Atominstitut, TU Wien, 1020 Vienna, Austria. [2]Institut für Theoretische Physik, Ruprecht-Karls-Universität Heidelberg, 69120 Heidelberg, Germany. [3]ExtreMe Matter Institute (EMMI), GSI Helmholtzzentrum für Schwerionenforschung GmbH, 64291 Darmstadt, Germany. [4]Wolfgang Pauli Institute, 1090 Vienna, Austria. [5]Ioffe Physico-Technical Institute, 194021 St. Petersburg, Russia.
***Corresponding author. E-mail: schmiedmayer@atomchip.org**

The GGE is a direct generalization of the usual thermodynamical ensembles and is formally capable of describing a wide range of dynamically emerging steady states (*15*). For example, in the case where only the energy is conserved, the GGE reduces to the standard canonical or Gibbs ensemble, with temperature being the only Lagrange multiplier (*4*). Moreover, the GGE famously provides a description for the steady states of

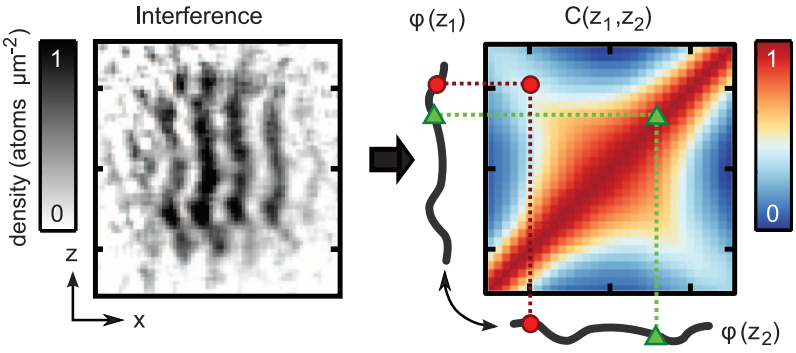



**Fig. 1. Experimental concept.** A non-equilibrium system is prepared by splitting a 1D Bose gas into two halves. After an evolution time $t$, matter-wave interference is used to extract the local relative phase profile $\varphi(z)$ between the two halves. This is accomplished by measuring the local position of the fluctuating interference fringes. Subsequently, the phase profile is used to calculate the two-point correlation function $C(z_1, z_2) \sim \langle \exp[i\varphi(z_1) - i\varphi(z_2)] \rangle$ as a function of all possible coordinates $z_1$ and $z_2$ along the length of the measured phase profile. For $z_1 = z_2$ (e.g., green triangle on the diagonal), $C(z_1, z_2) = 1$. Coordinates with $z_1 = -z_2$ (e.g., red point) are located symmetrically around the center of the system and are found on the antidiagonal of the correlation function.

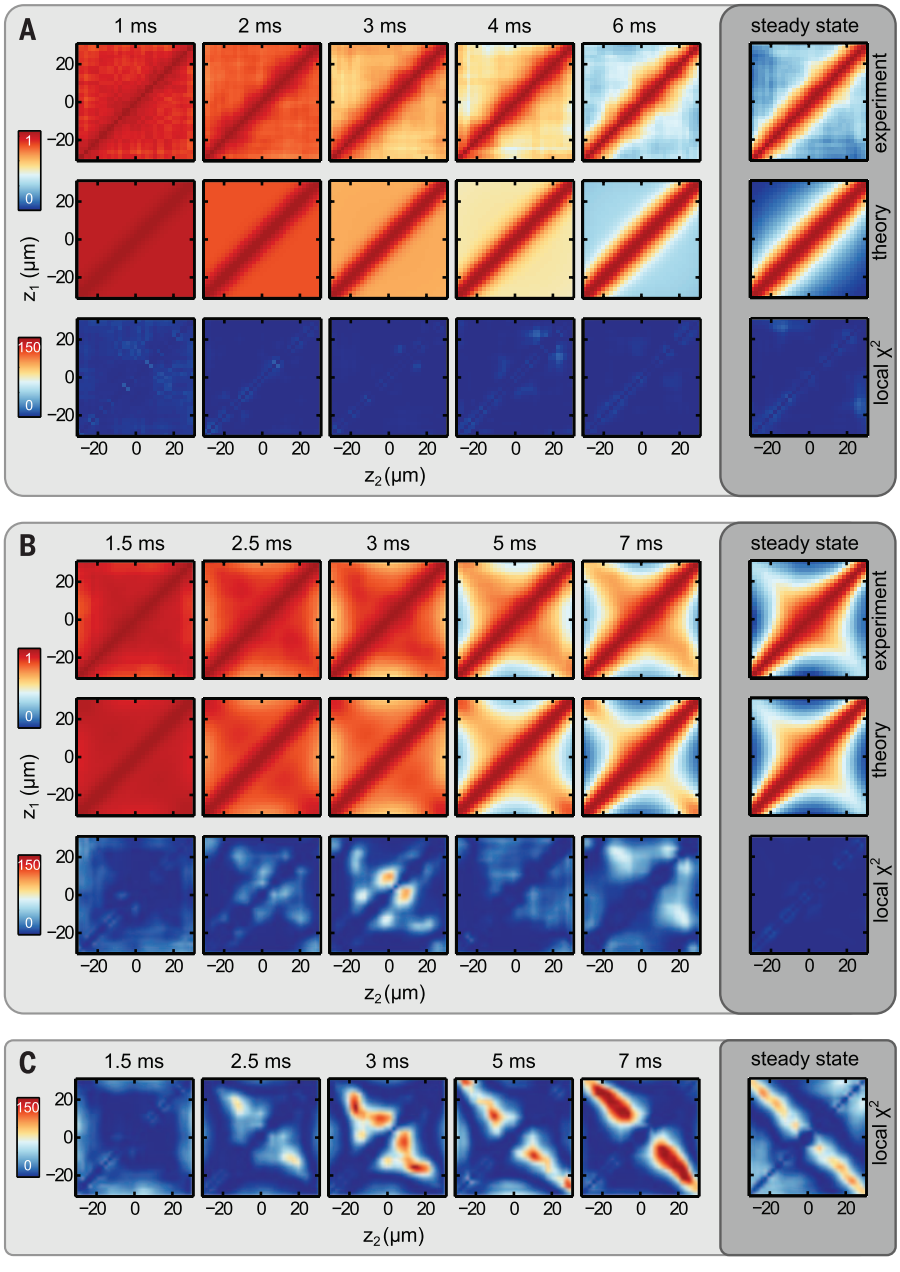



**Fig. 2. Two-point phase correlation functions $C(z_1, z_2)$ for increasing evolution time.** (**A** and **B**) Different initial states were prepared using two different splitting protocols. Both states show a characteristic maximum on the diagonal and a decay of correlations away from the diagonal. We used a $\chi^2$ analysis to quantify the agreement of our theoretical model and the experiments. The steady state and the dynamics in (A) can be well described by a single temperature $T_{\mathrm{eff}}$. (**C**) The single-temperature model fails for the steady state and the dynamics in (B), which require more temperatures to explain additional correlations on the antidiagonal (see text). The observation of different temperatures in the same steady state constitutes our observation of a GGE. The center of the system is located at $z_1 = z_2 = 0$; color marks the amount of correlations between 0 and 1 and the local $\chi^2$ contribution between 0 and 150. The uncertainty of the correlation functions is estimated via bootstrapping over approximately 150 experimental realizations (*25*).

integrable systems, which have as many independent conserved quantities as they have degrees of freedom (*9*, *11*, *14*). Even if quantities are only approximately conserved, the GGE description was formally shown to be valid for extended time scales (*16*). The GGE has also been suggested as a description for many-body localized states (*17*). Numerical evidence for the emergence of a GGE has been provided for many systems, but a direct experimental observation has been lacking.

Here, we experimentally studied the relaxation of a trapped one-dimensional (1D) Bose gas. Our system is a close realization of the Lieb-Liniger model describing a homogeneous gas of 1D bosons with contact interactions, which is one of the prototypical examples of an integrable system (*18*, *19*). In the thermodynamic limit, its exact Bethe Ansatz solutions imply an infinite number of conserved quantities, which provides a memory of an initial non-equilibrium state, forcing the gas to relax to a GGE. Recent results have also shown that trapped 1D Bose gases can be approximately integrable over very long time scales, enabling the detailed investigation of integrable dynamics (*6*, *7*, *10*, *20*, *21*). To demonstrate the emergence of a GGE, we prepared such a 1D Bose gas in different initial non-equilibrium states and observed how they each relaxed to steady states that maximize entropy according to the initial values of the conserved quantities.

The experiments started with a phase-fluctuating 1D Bose gas (*22*) of $^{87}$Rb atoms that was prepared and trapped using an atom chip (*23*). We initialized the non-equilibrium dynamics by transversally splitting this single 1D gas coherently into two nominally identical 1D systems, each containing half of the atoms on average. Information about the total system was extracted using matter-wave interferometry between the two halves (*6*, *20*, *21*, *24*). This enabled the time-resolved measurement of individual two-point and higher-order $N$-point phase correlation functions,

$$C(z_1, z_2, ..., z_N) \sim \langle \Psi_1(z_1)\Psi_2^\dagger(z_1)\Psi_1^\dagger(z_2)\Psi_2(z_2)...\Psi_1^\dagger(z_N)\Psi_2(z_N)\rangle \sim \langle \exp[i\varphi(z_1) - i\varphi(z_2) + ... - i\varphi(z_N)]\rangle \quad (2)$$

where $z_1$, $z_2$, ..., $z_N$ are $N$ coordinates along the length of the system, and $\varphi(z)$ is the relative phase between the two halves (*25*). As described below, these correlation functions reveal detailed information about the dynamics and the steady states of the system.

We start with the two-point correlation function $C(z_1, z_2) \sim \langle \exp[i\varphi(z_1) - i\varphi(z_2)]\rangle$. Previously, this correlation function was studied in regions where the system is approximately translationally invariant (*21*, *26*), that is, $C(z_1, z_2) \equiv C(z_1 - z_2)$. Here, more comprehensive information about generic many-body states is obtained by mapping the full correlation function $C(z_1, z_2)$ for any combination of the coordinates $z_1$ and $z_2$ (Fig. 1).

Our observations after a typical splitting, which is fast relative to the dynamics of the system and therefore realizes a quench (*25*), are summarized in Fig. 2A. As every point in the system is perfectly correlated with itself, the correlation functions exhibit a maximum on the diagonal $z_1 = z_2$ for all times. Away from the diagonal, the system shows a light cone–like decay of correlations (*21*) leading to a steady state. From a theoretical point of view, the emergence of this steady state is a consequence of prethermalization (*6*, *27*–*31*), which in the present case can be described as the dephasing of phononic excitations (*29*, *31*–*33*). The occupation numbers $n_m$ of these excitations are the conserved quantities of the corresponding integrable model (*25*).

With a knowledge of the conserved quantities, we can directly calculate the Lagrange multipliers $\lambda_m$ for the GGE. In terms of the excitation energies $\varepsilon_m$ they can be written as $\lambda_m = \beta_m\varepsilon_m$, which defines an effective temperature $1/\beta_m$ for every excitation mode.

For the steady state illustrated in Fig. 2A, the proportionality factor $\beta_m$ can be well described by $\beta_m \approx \beta_{eff} = 1/k_B T_{eff}$ (where $k_B$ is the Boltzmann constant). A fit yields $k_B T_{eff} = 0.50$ ($\pm 0.01$) $\times$ $\mu$, which is independent of $m$ and in very good agreement with theory (*6*, *33*). Here, $\mu$ denotes the chemical potential in each half of the system. Although it is a GGE in principle, for our experiment (which observes the relative phase between the two halves of the system) it becomes formally equivalent to the usual Gibbs ensemble with a single temperature $T_{eff}$ (*25*).

To obtain direct experimental signatures of a genuine GGE, we modified the initial state so that it exhibited different temperatures for different excitation modes. This was accomplished by changing the ramp that splits the initial gas into two halves (*25*). The results are shown in Fig. 2B. In addition to the maximum of correlations on the diagonal, a pronounced second maximum on the antidiagonal can be seen. This corresponds to enhanced correlations of the points $z_1 = -z_2$, which are located symmetrically around the center of the system. These correlations are a direct consequence of an increased population of quasi-particle modes that are even, and a decreased population of quasi-particle modes that are odd, under a mirror reflection with respect to the longitudinal trap center. Consequently, the observations can be described to a first approximation by the above theoretical model but with different temperatures—that is, with $\beta_{2m} = 1/[k_B(T_{eff} + \Delta T)]$ for the even modes and $\beta_{2m-1} = 1/[k_B(T_{eff} - \Delta T)]$ for the odd modes, respectively. Fitting the experimental data of the steady state with this model, we find $k_B T_{eff} = 0.64$ ($\pm 0.01$) $\times$ $\mu$, $\Delta T = 0.48$ ($\pm 0.01$) $\times$ $T_{eff}$, and a reduced $\chi^2 \approx 6$.

More detailed insights and a more accurate description of the experimental data can be gained by fitting the steady state with the individual mode occupations as free parameters. The results yield a reduced $\chi^2$ close to 1 and thus

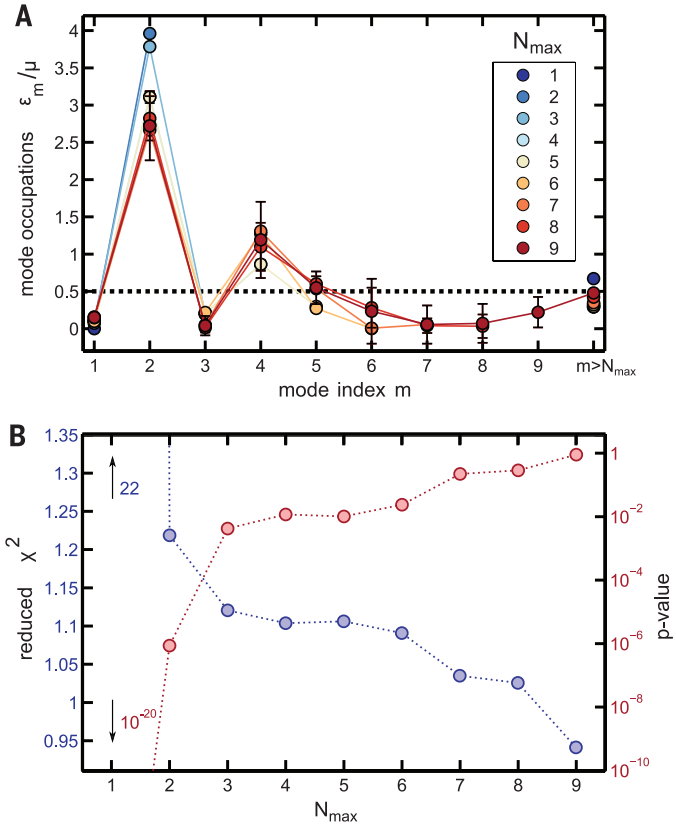



**Fig. 3. Statistical analysis of the fitting procedure.** (**A**) Occupation numbers $n_m$ (in units of $\varepsilon_m/\mu$) of the quasi-particle modes with index $m$, for different fitting procedures applied to the data from Fig. 2B. The color of the points encodes results where modes up to $m = N_{max}$ are fitted freely, whereas all higher modes are fit with the same occupation number. The plot reveals that the occupation of the lowest even modes is increased, and that of the lowest odd modes is decreased, relative to the single-temperature state from Fig. 2A (dashed line). (**B**) Scaling of the reduced $\chi^2$ value, and of the corresponding $P$ value with $N_{max}$. The occupation numbers of the lowest nine modes and a single occupation number for all higher modes are sufficient to describe the experimental data to very good accuracy.

a very good description of the experimental data (see Fig. 2B). As expected from our intuitive two-temperature model, the fitting confirms that the occupation of even modes is strongly enhanced, whereas the occupation of odd modes is reduced (see Fig. 3A). Fixing these occupation numbers extracted from the steady state, our theoretical model also describes the complete dynamics very well. This clearly demonstrates that the different populations of the modes were imprinted onto the system by the splitting quench and are conserved during the dynamics. In contrast, a simple model based on a usual Gibbs ensemble with just one temperature for all modes clearly fails to describe the data [best fit: $k_B T_{eff} = 0.38\ (\pm 0.01) \times \mu$, reduced $\chi^2 \approx 25$], as visualized in Fig. 2C.

Notably, our fitting results for the GGE exhibit strong correlations between the different even modes and the different odd modes, respectively. This demonstrates the difficulty in fully and independently determining the parameters of such complex many-body states. In fact, any full tomography of all parameters would require exponentially many measurements. The results thus clearly show the presence of a GGE with at least two, but most likely many more temperatures.

Our work raises the interesting question of how many Lagrange multipliers are needed in general to describe the steady state of a realistic integrable quantum system. In analogy to classical mechanics, where $N$ conserved quantities exist for a generic integrable system with $N$ degrees of freedom, integrability in quantum many-body systems has been proposed to be characterized by the fact that the number of independent local conserved quantities scales with the number of particles (*5*). Here we conjecture that many experimentally obtainable initial states evolve in time into steady states, which can be described to a reasonable precision by far fewer than $N$ Lagrange multipliers (*8*, *34*). This would have the appeal of a strong similarity to thermodynamics, where also only a few parameters are needed to describe the properties of a system on macroscopic scales.

To illustrate this in our specific case, we investigated how many distinct Lagrange multipliers need to be considered in the GGE to describe our data with multiple temperatures (Fig. 3). Including more and more modes in the fitting of the experimental data, we found that the reduced $\chi^2$ values decrease and settle close to unity for nine included modes, with all higher modes being fitted by one additional Lagrange multiplier. Looking at the $P$ value for the measured $\chi^2$ (*35*) shows that only a limited number of Lagrange multipliers needs to be specified to describe the observables under study to the precision of the measurement. A simple numerical estimate based on the decreasing contribution of higher modes to the measured correlation functions and the limited imaging resolution leads to approximately 10 Lagrange multipliers, which is in good agreement with our observations (*25*). Moreover, comparing this result with the single-temperature steady state discussed earlier illustrates that the complexity of the initial state plays an important role for the number of Lagrange multipliers that need to be included in a GGE.

In general, deviations of steady states from the GGE description are expected to manifest first in higher-order correlation functions. To provide further evidence for our theoretical description and the presence of a GGE, Fig. 4 shows examples of measured 4-point, 6-point, and 10-point correlation functions of the steady state. Like the 2-point correlation functions, they are in very good agreement with the theoretical model and clearly reveal the difference between the GGE and the usual Gibbs ensemble. This confirms that the description based on a GGE with the parameters extracted from the 2-point correlation functions also describes many-body observables, at least up to 10th order.

Our results clearly visualize, both experimentally and theoretically, how the unitary evolution of our quantum many-body system connects to a steady state that can be described by a classical statistical ensemble. We expect our measurements of correlation functions to high order to play an

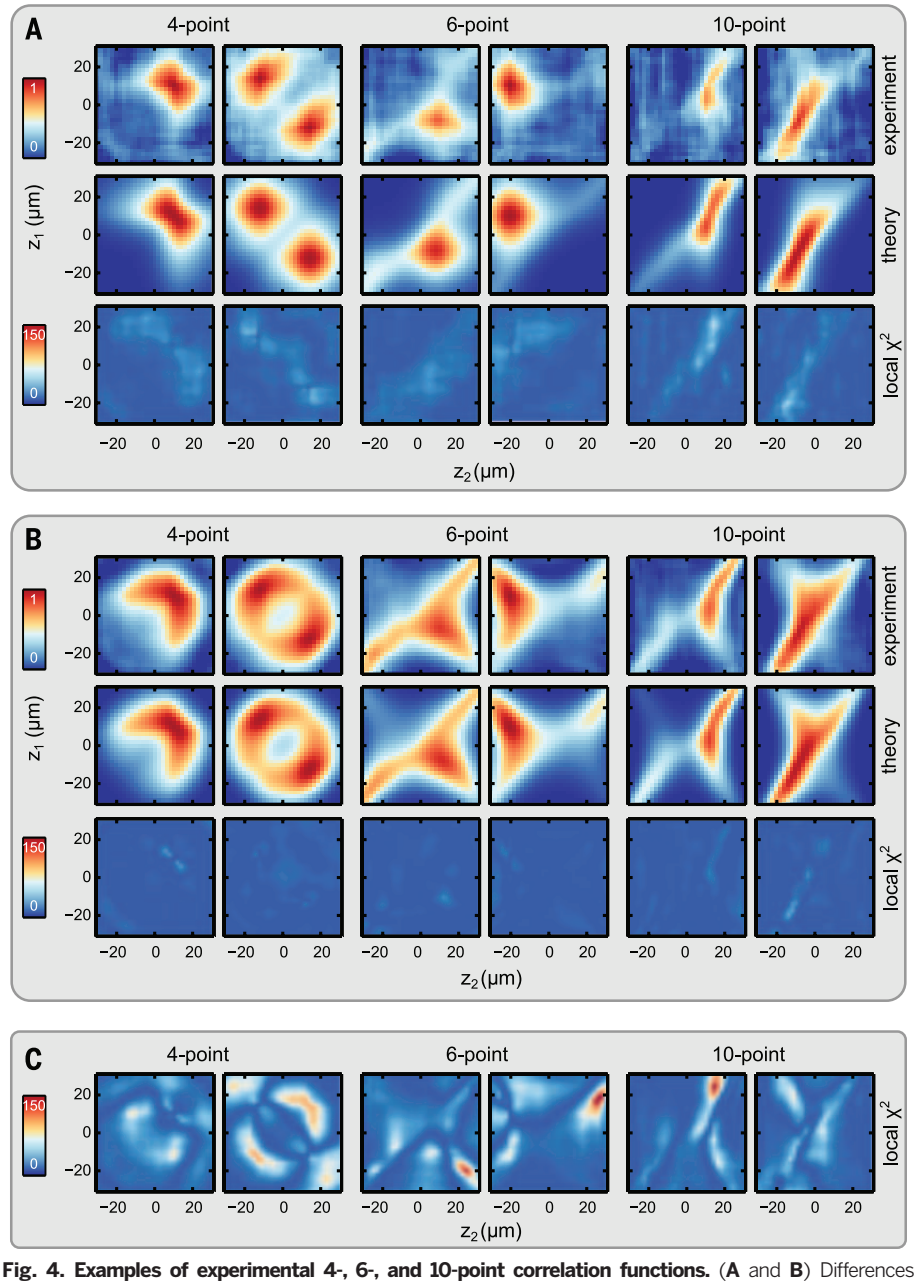



**Fig. 4. Examples of experimental 4-, 6-, and 10-point correlation functions.** (**A** and **B**) Differences between the steady state described by a single temperature (A) and the steady state described by multiple temperatures (B) are significant and can be captured by the theoretical model. (**C**) The single-temperature model cannot describe the state with multiple temperatures, which is reflected in high values of the local $\chi^2$. From left to right, we plot $C(z_1, 10, z_2, 10)$, $C(z_1, -12, z_2, 14)$, $C(z_1, 10, 10, z_2, -20, 10)$, $C(z_1, -8, 8, z_2, -24, -20)$, $C(z_1, 4, 10, z_2, -8, z_2, -22, -18, 10, -4)$, and $C(z_1, -22, -8, z_2, -22, -26, -22, z_2, -26, -24)$. All coordinates are given in μm and were chosen as representative cases for our high-dimensional data.

important role in new tomography techniques for complex quantum many-body states (*36*). Moreover, the observed tunability of the nonequilibrium states suggests that our splitting process could in the future be used to prepare states tailored for precision metrology (*37*).

## REFERENCES AND NOTES


1. C. E. Shannon, W. Weaver, *The Mathematical Theory of Communication* (Univ. of Illinois Press, Urbana, IL, 1949).
2. E. T. Jaynes, *Phys. Rev.* **106**, 620–630 (1957).
3. E. T. Jaynes, *Phys. Rev.* **108**, 171–190 (1957).
4. K. Huang, *Statistical Mechanics* (Wiley, New York, 1987).
5. A. Polkovnikov, K. Sengupta, A. Silva, M. Vengalattore, *Rev. Mod. Phys.* **83**, 863–883 (2011).
6. M. Gring *et al.*, *Science* **337**, 1318–1322 (2012).
7. J. P. Ronzheimer *et al.*, *Phys. Rev. Lett.* **110**, 205301 (2013).
8. J.-S. Caux, F. H. Essler, *Phys. Rev. Lett.* **110**, 257203 (2013).
9. P. Calabrese, F. H. L. Essler, M. Fagotti, *Phys. Rev. Lett.* **106**, 227203 (2011).
10. T. Kinoshita, T. Wenger, D. S. Weiss, *Nature* **440**, 900–903 (2006).
11. M. Rigol, V. Dunjko, V. Yurovsky, M. Olshanii, *Phys. Rev. Lett.* **98**, 050405 (2007).
12. M. Rigol, *Phys. Rev. Lett.* **103**, 100403 (2009).
13. M. A. Cazalilla, *Phys. Rev. Lett.* **97**, 156403 (2006).
14. J.-S. Caux, R. M. Konik, *Phys. Rev. Lett.* **109**, 175301 (2012).
15. M. Rigol, V. Dunjko, M. Olshanii, *Nature* **452**, 854–858 (2008).
16. M. Kollar, F. A. Wolf, M. Eckstein, *Phys. Rev. B* **84**, 054304 (2011).
17. R. Vosk, E. Altman, *Phys. Rev. Lett.* **110**, 067204 (2013).
18. E. H. Lieb, W. Liniger, *Phys. Rev.* **130**, 1605–1616 (1963).
19. V. Korepin, N. Bogoliubov, A. Izergin, *Quantum Inverse Scattering Method and Correlation Functions* (Cambridge Univ. Press, Cambridge, 1993).
20. M. Kuhnert *et al.*, *Phys. Rev. Lett.* **110**, 090405 (2013).
21. T. Langen, R. Geiger, M. Kuhnert, B. Rauer, J. Schmiedmayer, *Nat. Phys.* **9**, 640–643 (2013).
22. D. S. Petrov, G. V. Shlyapnikov, J. T. M. Walraven, *Phys. Rev. Lett.* **85**, 3745–3749 (2000).
23. J. Reichel, V. Vuletic, Eds., *Atom Chips* (Wiley-VCH, Berlin, 2011).
24. T. Schumm *et al.*, *Nat. Phys.* **1**, 57–62 (2005).
25. See supplementary materials on *Science* Online.
26. T. Betz *et al.*, *Phys. Rev. Lett.* **106**, 020407 (2011).
27. J. Berges, S. Borsányi, C. Wetterich, *Phys. Rev. Lett.* **93**, 142002 (2004).
28. T. Gasenzer, J. Berges, M. G. Schmidt, M. Seco, *Phys. Rev. A* **72**, 063604 (2005).
29. J. Berges, T. Gasenzer, *Phys. Rev. A* **76**, 033604 (2007).
30. M. Eckstein, M. Kollar, P. Werner, *Phys. Rev. Lett.* **103**, 056403 (2009).
31. T. Kitagawa, A. Imambekov, J. Schmiedmayer, E. Demler, *New J. Phys.* **13**, 073018 (2011).
32. R. Bistritzer, E. Altman, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9955–9959 (2007).
33. T. Kitagawa *et al.*, *Phys. Rev. Lett.* **104**, 255302 (2010).
34. P. Barmettler, C. Kollath, http://arxiv.org/abs/1312.5757 (2013).
35. I. G. Hughes, T. Hase, *Measurements and Their Uncertainties* (Oxford Univ. Press, Oxford, 2010).
36. R. Hübener, A. Mari, J. Eisert, *Phys. Rev. Lett.* **110**, 040401 (2013).
37. T. Berrada *et al.*, *Nat. Commun.* **4**, 2077 (2013).


## ACKNOWLEDGMENTS

We acknowledge discussions with E. Demler, E. Dalla Torre, K. Agarwal, J. Berges, M. Karl, V. Kasper, I. Bouchoule, M. Cheneau, and P. Grisins. Supported by the European Union (SIQS and ERC advanced grant QuantumRelax), the Austrian Science Fund (FWF) through the Doctoral Programme CoQuS (W1210) (B.R. and T.S.), the Lise Meitner Programme M1423 (R.G.) and project P22590-N16 (I.E.M.), Deutsche Forschungsgemeinschaft grant GA677/7,8 (T.G.), the University of Heidelberg Center for Quantum Dynamics, Helmholtz Association grant HA216/EMMI, and NSF grant PHY11-25915. T.L., T.G., and J.S. thank the Kavli Institute for Theoretical Physics, Santa Barbara, for its hospitality.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/207/suppl/DC1
Materials and Methods
Supplementary Text
References (*38*–*44*)

5 June 2014; accepted 2 March 2015
10.1126/science.1257026

BIOANALYSIS

# Mass spectrometry imaging with laser-induced postionization

**Jens Soltwisch,[1] Hans Kettling,[1,2] Simeon Vens-Cappell,[1,2] Marcel Wiegelmann,[1] Johannes Müthing,[1] Klaus Dreisewerd[1,2]***

**Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) can simultaneously record the lateral distribution of numerous biomolecules in tissue slices, but its sensitivity is restricted by limited ionization. We used a wavelength-tunable postionization laser to initiate secondary MALDI-like ionization processes in the gas phase. In this way, we could increase the ion yields for numerous lipid classes, liposoluble vitamins, and saccharides, imaged in animal and plant tissue with a 5-micrometer-wide laser spot, by up to two orders of magnitude. Critical parameters for initiation of the secondary ionization processes are pressure of the cooling gas in the ion source, laser wavelength, pulse energy, and delay between the two laser pulses. The technology could enable sensitive MALDI-MS imaging with a lateral resolution in the low micrometer range.**

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is used for the analysis of large nonvolatile biomolecules (*1*). Typically, the analytes are embedded into crystalline MALDI matrices (often aromatic carboxylic acids). Laser-induced codesorption of both constituents into the gas phase is convoluted with concomitant ionization in primary and secondary ionization processes (*2*–*4*). MALDI-MS imaging (MALDI-MSI) visualizes the lateral distribution of numerous biomolecules or that of administered drugs simultaneously by scanning matrix-coated tissue slices with a finely focused laser beam and recording the ion profiles per irradiated pixel (*5*, *6*). For

[1]Institute for Hygiene, University of Münster, Robert-Koch-Strasse 41, 48149 Münster, Germany. [2]Interdisciplinary Center for Clinical Research (IZKF), University of Münster, Domagkstrasse 3, 48149 Münster, Germany.
*Corresponding author. E-mail: klaus.dreisewerd@uni-muenster.de

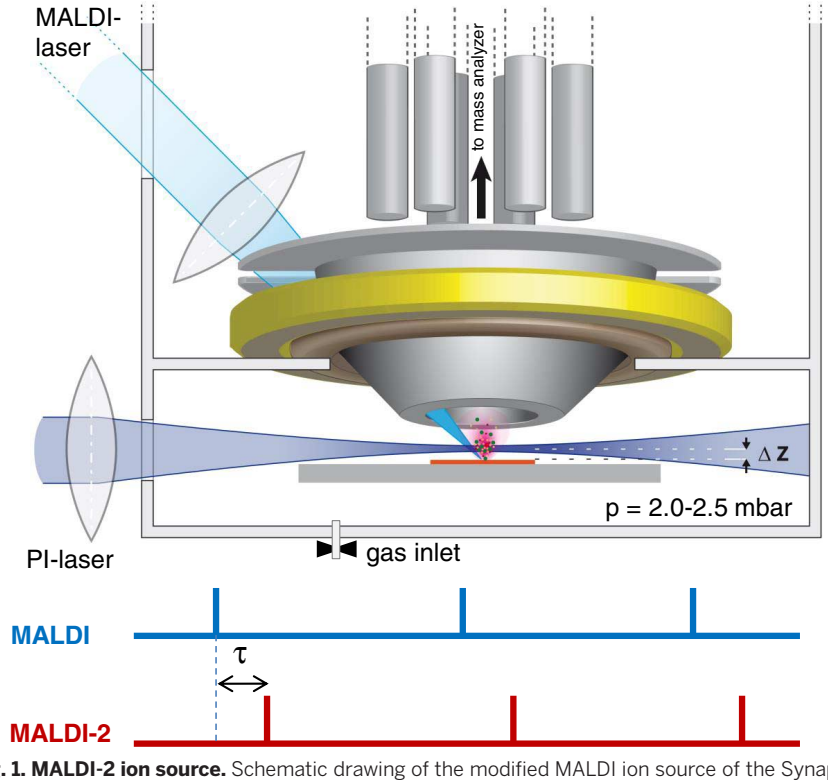



**Fig. 1. MALDI-2 ion source.** Schematic drawing of the modified MALDI ion source of the Synapt G2-S mass spectrometer with primary MALDI and PI laser beams and shielding aperture for increasing the cooling gas pressure in the region of ion generation; 2.0 to 2.5 mbar of $N_2$ were used for the MALDI-2-MSI measurements. The lower panel illustrates the laser pulse triggering sequence.

example, lipid profiles are increasingly used for visualization of malignancies (*7*, *8*) or to retrieve valuable information about the general biochemistry in tissue (*9*). A lateral resolution below 10 µm can now be achieved (*10*, *11*) that could potentially allow one to visualize the molecular content of single cells, thereby fostering numerous applications in the life sciences (*12*–*14*).

However, restricted ion yields—less than 1 out of 1000 desorbed molecules is on average ionized (*15*, *16*)—and ion suppression effects (*3*) can impede obtaining a more comprehensive picture about the tissue composition. For example, lipid mass spectra recorded in the positive ion mode are typically dominated by the signals of phosphatidylcholines (PCs), whereas other phospho- (PLs) and glycolipids (GLs) are often barely detectable, even if these are present in the tissue in high abundance (*17*). Because of the minute amount of material that is ejected per laser shot, this problem is aggravated in high-resolving imaging (*18*). To increase the ion yields, single and multiphoton laser ionization of neutral gas phase molecules (*19*, *20*), as well as postionization (PI) by means of an electrospray ionization (ESI) beam (*21*), have been developed. However, extensive analyte fragmentation limits the use of direct photoionization to small molecules, and ESI requires the use of ambient pressure ion sources. For most analyte classes, the state-of-the

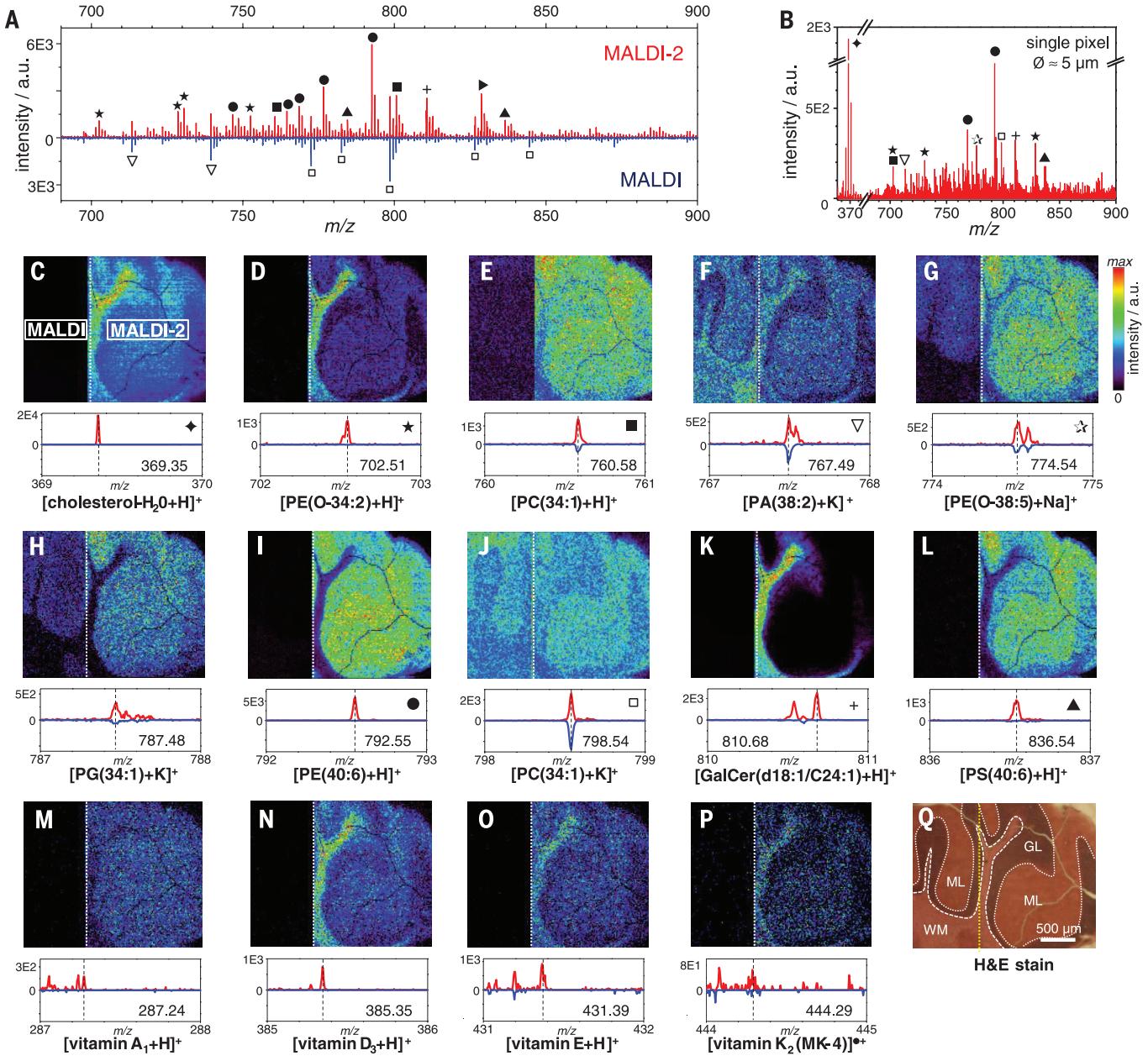



**Fig. 2. Mass spectra and MS images recorded from mouse cerebellum in the positive ion mode.** (**A**) MALDI-2 (red trace) and conventional MALDI mass spectra (blue trace) acquired in the high-resolution mode of the instrument. Each spectrum was accumulated over 144 pixels (two parallel lines, 20 laser shots/pixel); $\lambda_{PI}$ = 280 nm, $\tau$ = 9 µs, $p$ = 2.5 mbar. (**B**) Single pixel spectrum acquired in the resolution mode. (**C** to **L**) Conventional MALDI-MS (left) and MALDI-2-MS images (right) of brain lipids and (**M** to **P**) of liposoluble vitamins; the molecular identities were corroborated by tandem MS (fig. S2). Images were recorded with a pitch size (pixel-to-pixel distance) of 15 µm. (**Q**) Hematoxylin and eosin (H&E) stain; the slice was taken approximately from interaural −2.5 to −3 mm. GL: granular layer, ML: molecular layer, WM: white matter. ✦: [cholesterol – $H_2O$ + H]$^+$; ∇: [PA + Na]$^+$ (PA: phosphatidic acid) or [PA + K]$^+$; ■: [PC + H]$^+$; □: [PC + Na]$^+$ or [PC + K]$^+$; ●: [PE + H]$^+$; ★: [PE-O + H]$^+$; ☆: [PE-O + Na]$^+$ or [PE-O + K]$^+$; ▲: [PS + H]$^+$; ▸: [PC-O + H]$^+$; +: [GalCer + H]$^+$.

art sensitivity achieved by classical MALDI-MSI performed under vacuum is not met with these methods.

Here we introduce a PI strategy, called MALDI-2, that initiates secondary MALDI-like ionization processes in the gas phase. In MALDI-2, the beam of a pulsed ultraviolet (UV) laser intercepts the expanding particle plume in an $N_2$ cooling gas environment (*22*, *23*) (Fig. 1), which contrasts with previous photoionization studies where classical high-vacuum ion sources ($p \leq 10^{-6}$ mbar) were implemented (*20*, *24*). An effective diameter of ~5 μm of the primary laser beam (fig. S1) was achieved by beam shaping (*23*) and by mounting the focusing lens inside the MALDI ion source (Fig. 1) (*25*).

Two positive ion mode mass spectra that we recorded from mouse cerebellum with and without the PI laser are compared in Fig. 2A and table S1. For this experiment, we chose 2,5-dihydroxybenzoic acid (DHB) as a classical MALDI matrix and prepared the samples using a sublimation and recrystallization protocol, to obtain a particular uniform microcrystalline coating (fig. S1) (*25*). We observed several classes of membrane lipids that are difficult to image by conventional MALDI-MSI of PC-rich tissue [e.g., cholesterol, phosphatidylethanolamine (PE), plasmalogens (PE-O), phosphatidylserine (PS), and neutral GLs such as galactosylceramide (GalCer)] in the MALDI-2 spectra with $[M + H]^+$ signal intensities that were up to two orders of magnitude higher. The signal intensities of alkali metal adducts increased more moderately by a factor of up to 5, depending on the lipid class, or remained essentially unchanged for PCs. Based on exact mass, we could tentatively assign about two to three times as many lipid species in the mass-to-charge ratio (*m/z*) range of 670 to 850. The benefit of the high signal intensities that we obtained from a single 5 μm-wide pixel (Fig. 2B) for MS imaging is demonstrated in Fig. 2, C to L, with the example of selected PLs and GalCer analyzed from mouse cerebellum. The distributions of liposoluble vitamins $A_1$ (retinol), $D_3$ (cholecalciferol), E (α-tocopherol), and $K_2$ (menachinon-4) could only be visualized with MALDI-2 (Fig. 2, M to P). These compounds exhibit a high extinction coefficient at 280 nm and are therefore detected both as $[M + H]^+$ and resonantly photoionized $M^{\bullet+}$ species.

Next we asked if measurements in the negative ion mode would also benefit from the post-ionization step; typically, positive and negative ion mode measurements provide valuable complementary information about the overall lipid composition in tissue. We obtained up to one to two orders of magnitude higher $[M - H]^-$ signal intensities for several lipid classes (e.g., for PA, plasmalogens PE-P and PE-O, PE, and PS) if desorbed from a norharmane matrix (Fig. 3 and table S2). This enabled MS imaging with high signal contrast for about five times as many lipid species. Only phosphatidylinositols (PIs; Fig. 3H) and sulfatides (STs; Fig. 3I), compounds that are already detected sensitively by regular MALDI-MSI (*23*), display a more moderate signal increase or no notable gain (STs).

We also tested a few more tissue types and analyte classes. For example, seminiferous tubule structures of rat testis could be visualized via their triacylglycerol (TG) content (fig. S5). Atypical for conventional MALDI, these polar lipids were detected as protonated molecules. MALDI-2-MSI of plant (apple) tissue produced strongly enhanced signals of mono- and oligosaccharides and of polyphenolic glycosides (fig. S6). In contrast to lipids and glycans, MALDI-2 of peptides has so far not lead to a signal enhancement, possibly because of different ionization mechanisms and plume expansion dynamics (*1*, *2*, *4*).

To identify optimal PI conditions, we performed a multiparameter study (Fig. 4). This revealed that the MALDI-2 ionization efficiency depends critically on a few key parameters, namely, the buffer gas pressure *p*, delay between the laser pulses τ, wavelength of the PI laser $\lambda_{PI}$, and its pulse energy $E_{PI}$. Optimal ion yields were found in distinct (matrix-dependent) parameter ranges. In particular, an elevated pressure was required (Fig. 4A and fig. S7). At $p = 0.5$ mbar, where only a low signal increase was found, the derived mean average velocity of neutral DHB molecules of about 500 $ms^{-1}$ (Fig. 4C) corresponds well to

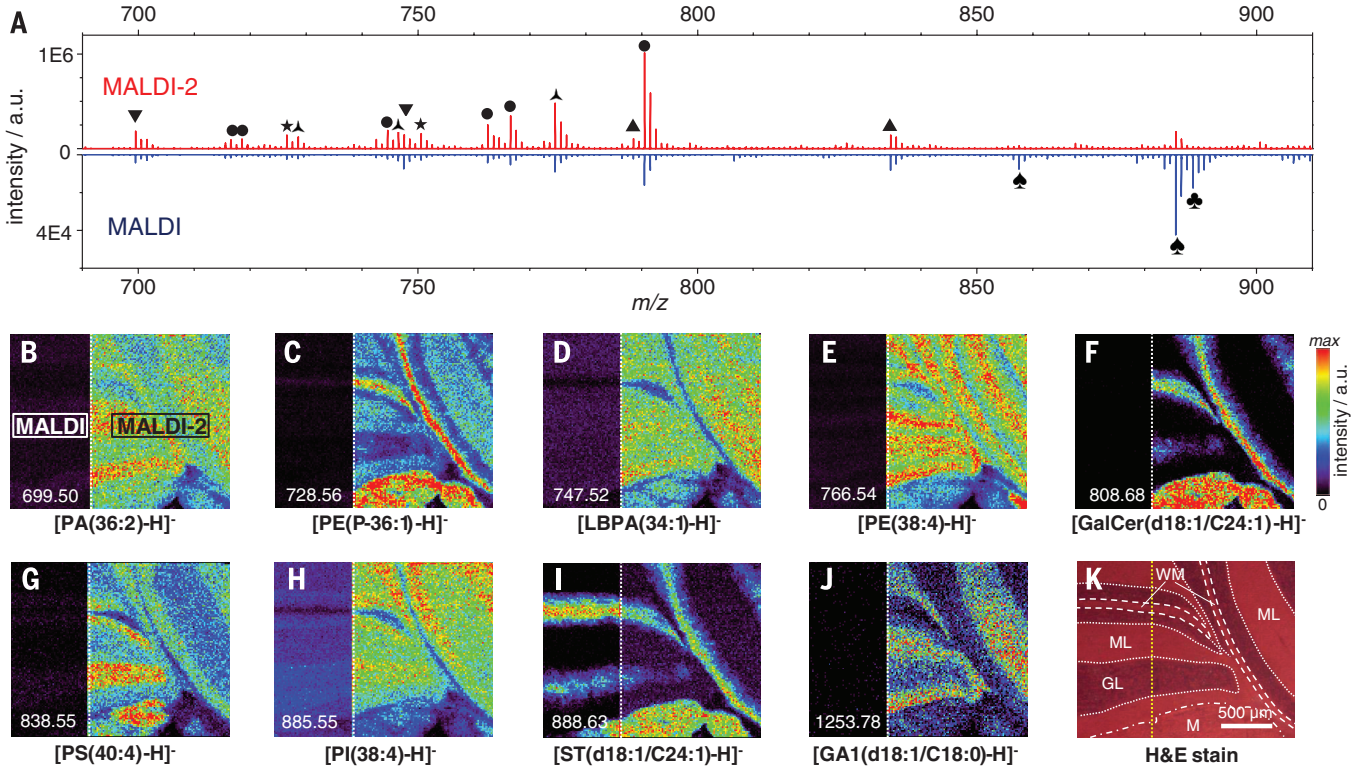



**Fig. 3. Mass spectra and MS images recorded from mouse cerebellum in the negative ion mode.** (**A**) Mass spectra acquired in the high-resolution mode from 300 adjacent pixels, each; $\lambda_{PI}$ = 260 nm, τ = 10 μs, $p$ = 2.0 mbar (compare fig. S3 for a single pixel spectrum). (**B** to **J**) MALDI- and MALDI-2-MS images of selected compounds (compare fig. S4 for tandem MS data); pitch size: 15 μm. (**K**) H&E stain; the slice was taken approximately from interaural −3.7 to −4.2 mm and comprises also a small part of the medulla (M). ▼: $[PA - H]^-$; ●: $[PE - H]^-$; ★: $[PE\text{-}O - H]^-$; ⅄: $[PE\text{-}P - H]^-$; ▲: $[PS - H]^{--}$; ♠: $[PI - H]^{--}$; ♣: $[ST - H]^{--}$.

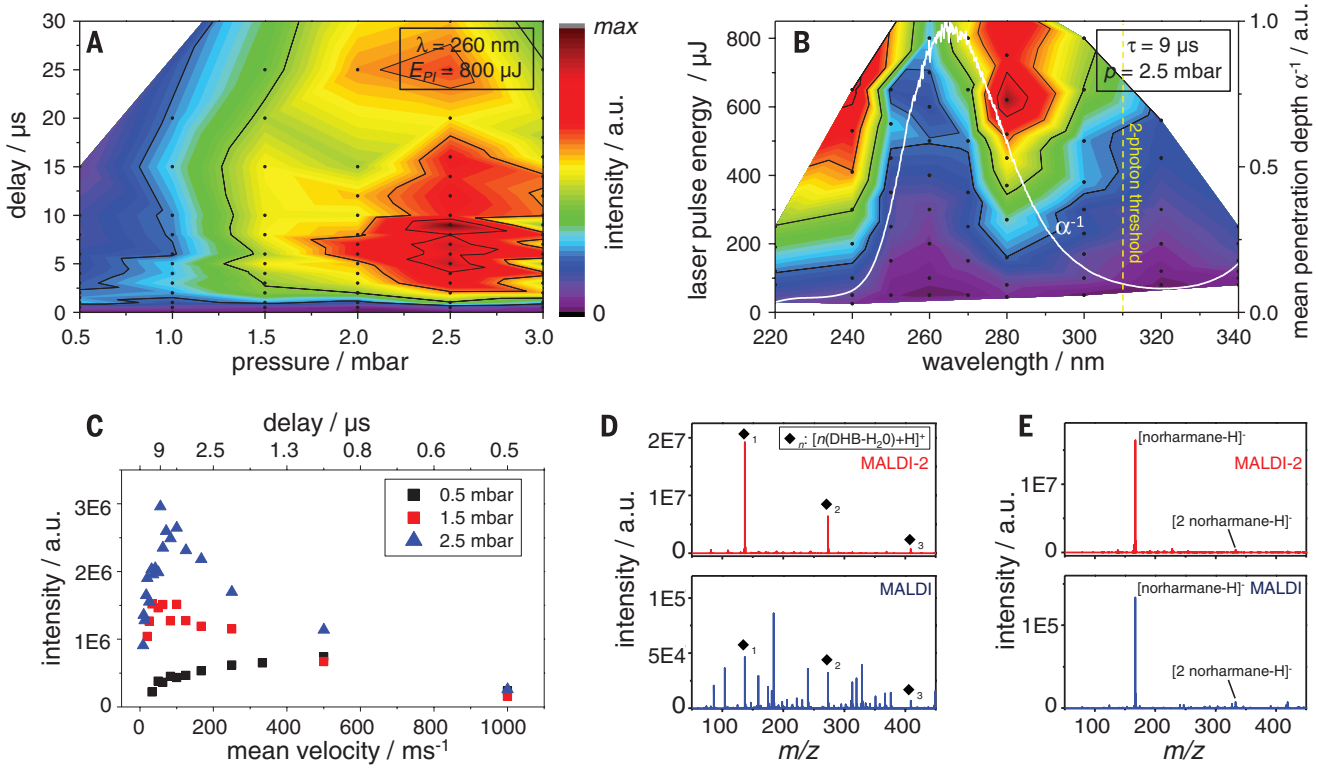



**Fig. 4. Parameters affecting the MALDI-2 ion yields.** (**A**) Heat map illustrating the effect of the cooling gas pressure $p$ and laser pulse delay $\tau$ on the signal intensities of protonated PE molecules. Each data point (black dots) was recorded by applying 600 laser pulses onto approximately 50 positions on DHB-coated homogenized liver tissue. (**B**) Effect of the PI laser wavelength $\lambda$ and pulse energy $E_{PI}$. Solid white line: solution phase mean penetration depth ($\alpha^{-1}$) of DHB; dashed yellow line: two-photon ionization threshold. (**C**) MALDI-2 signal intensities of protonated PE as a function of $\tau$ for three different gas pressures; average mean velocities are calculated as $v = \Delta z/\tau$ with $\Delta z$ = 0.5 mm. (**D** and **E**) Matrix mass ranges in the positive (DHB matrix) and negative ion mode (norharmane), respectively.

values that were previously determined for classical high-vacuum MALDI conditions (*2*). However, at elevated pressures (1.5 and 2.5 mbar), the molecules were effectively slowed down to mean velocities below 100 $ms^{-1}$. The absence of entrainment in a cooling gas flow and subsequent gas-phase reactions probably account for the inability of previous MALDI-PI studies to produce protonated matrix molecules and higher molecular weight compounds (*20*).

Apart from a small hypsochromic shift, reflecting the difference between solution- and gas-phase absorption spectra (*26*), the MALDI-2 ion yields of lipids that are nonabsorbing at the PI laser wavelength followed the optical absorption characteristics of the matrix. Particularly high $[M + H]^+$ signals were obtained for wavelengths below the two-photon ionization threshold of DHB of ~310 nm (*27*) (Fig. 4, B and C). The same behavior was found for the total ion count (TIC) (fig. S7). In contrast, for alkali metal adducts of the lipids, the wavelength characteristics more closely reflected the overall absorption profile of the matrix. Similar to conventional MALDI-MS (*28*), high signals are here found also above 310 nm (fig. S7D). Possibly, gaseous $Na^+$ and $K^+$ ions were generated by the photodissociation of neutral matrix clusters (*29*), which gave rise to additional $[M + Na]^+$ and $[M + K]^+$ species.

We hypothesize that the mechanisms underlying ionization by proton transfer in MALDI-2 involve resonant two-photon ionization of the matrix (m) by the PI laser (giving rise to $m^{+\bullet}$ ions and free $e^-$), succeeding collisions with neutral matrix molecules (leading to the generation of protonated or deprotonated matrix), and proton transfer to or from neutral analyte molecules (M) in subsequent collisions to yield the observed $[M + H]^+$ or $[M - H]^-$ products. With regard to the analyte ionization step, similar mechanisms are discussed for conventional MALDI (*3*). These assumptions are supported by the detection of high abundances of $[m - H_2O + H]^+$ ions of the DHB matrix and oligomers thereof, and of $[m - H]^-$ ions of norharmane and a few oligomers thereof, respectively (Fig. 4, D and E, and table S3). Also, the observed adduct formation between some analytes (e.g., PE) and these matrix-derived ions supports the notion of ample gas-phase reactions (fig. S8). Notably, the soft nature of the ionization is accompanied by a low degree of analyte fragmentation (fig. S9A).

Next to the liposoluble vitamins (fig. S9, C and D), we observed the direct two-photon ionization of analyte molecules, giving rise to additional $M^{\bullet+}$ and $M^{\bullet-}$ species, also for sialic acid–containing gangliosides. The direct absorption of the laser light by the weakly bound sialic acid residue(s) and succeeding bond cleavage probably led to the high levels of observed asialoganglioside ions (Fig. 3J, fig. S9B, and table S2). So far, we have observed the strong PI effect only with laser spot sizes in the low 10- to sub–10-µm range. This could be caused by the confinement of the particle plume or by an insufficient overlap of the PI laser beam with a more extended particle plume.

Given the high image contrast that was achieved from only 5-µm-wide pixels, we postulate that MALDI-2 will in the future enable MS imaging with even higher lateral resolution in the 1- to 2-µm range. This could provide analytical possibilities that at present are mainly in the domain of optical microscopy techniques. Additionally, the more comprehensive picture that is obtained about the molecular composition in tissue could substantially improve the confidence level of methods that use altered lipid profiles as markers for disease states.

## REFERENCES AND NOTES


1. F. Hillenkamp, J. Peter-Katalinić, Eds., *MALDI MS—A Practical Guide to Instrumentation, Methods and Applications* (Wiley, Weinheim, Germany, ed. 2, 2013).
2. K. Dreisewerd, *Chem. Rev.* **103**, 395–426 (2003).
3. R. Knochenmuss, R. Zenobi, *Chem. Rev.* **103**, 441–452 (2003).
4. T. W. Jaskolla, M. Karas, *J. Am. Soc. Mass Spectrom.* **22**, 976–988 (2011).
5. M. Stoeckli, P. Chaurand, D. E. Hallahan, R. M. Caprioli, *Nat. Med.* **7**, 493–496 (2001).
6. K. Chughtai, R. M. A. Heeren, *Chem. Rev.* **110**, 3237–3277 (2010).
7. N. Goto-Inoue, T. Hayasaka, N. Zaima, M. Setou, *Biochim. Biophys. Acta* **1811**, 961–969 (2011).
8. S. Guo, Y. Wang, D. Zhou, Z. Li, *Sci. Rep.* **4**, 5959 (2014).

9. K. A. Zemski *et al.*, *Chem. Rev.* **111**, 6491–6512 (2011).
10. A. Römpp *et al.*, *Angew. Chem. Int. Ed.* **49**, 3834–3838 (2010).
11. A. Zavalin *et al.*, *J. Mass Spectrom.* **47**, 1473–1481 (2012).
12. J. L. Norris, R. M. Caprioli, *Chem. Rev.* **113**, 2309–2342 (2013).
13. E. J. Lanni, S. S. Rubakhin, J. V. Sweedler, *J. Proteomics* **75**, 5036–5051 (2012).
14. B. Spengler, *Anal. Chem.* **87**, 64–82 (2015).
15. R. Knochenmuss, L. V. Zhigilei, *Anal. Bioanal. Chem.* **402**, 2511–2519 (2012).
16. K. Dreisewerd, M. Schürenberg, M. Karas, F. Hillenkamp, *Int. J. Mass Spectrom. Ion Process.* **141**, 127–148 (1995).
17. J. Schiller *et al.*, *Prog. Lipid Res.* **43**, 449–488 (2004).
18. K. Dreisewerd, *Anal. Bioanal. Chem.* **406**, 2261–2278 (2014).
19. L. Hanley, R. Zimmermann, *Anal. Chem.* **81**, 4174–4182 (2009).
20. C. H. Becker, K. J. Wu, *J. Am. Soc. Mass Spectrom.* **6**, 883–888 (1995).
21. P. Nemes, A. Vertes, *Anal. Chem.* **79**, 8098–8106 (2007).
22. J. Soltwisch, J. Souady, S. Berkenkamp, K. Dreisewerd, *Anal. Chem.* **81**, 2921–2934 (2009).
23. H. Kettling *et al.*, *Anal. Chem.* **86**, 7798–7805 (2014).
24. A. Rohlfing, A. Leisner, F. Hillenkamp, K. Dreisewerd, *J. Chem. Phys. C* **114**, 5367–5381 (2010).
25. Further experimental details can be found in the supplementary online materials.
26. J. Y. Zhang, D. S. Nagra, L. Li, *Anal. Chem.* **63**, 2995–2999 (1991).
27. Q. Lin, R. Knochenmuss, *Rapid Commun. Mass Spectrom.* **15**, 1422–1426 (2001).
28. J. Soltwisch, T. W. Jaskolla, F. Hillenkamp, M. Karas, K. Dreisewerd, *Anal. Chem.* **84**, 6567–6576 (2012).
29. A. N. Krutchinsky, B. T. Chait, *J. Am. Soc. Mass Spectrom.* **13**, 129–134 (2002).



## ACKNOWLEDGMENTS

We thank Waters Corp. and GWU Lasertechnik for technical support; F. Spieker and J. Haier for animal tissue; J. Klingauf for scanning electron microscopy images; W. Kramer for Fig. 1; and S. Albrecht, H. Karch, T. Kuhlmann, G. Pohlentz, and J. Y. Yew for helpful discussions. Financial support by the German Science Foundation (grants DR416/8-2 and DR416/9-1 to K.D., SO976/2-1 to J.S., GRK 1409 and MU845/4-2 to J.M.) and the Interdisciplinary Center for Clinical Research (IZKF) Münster (Z03 to K.D.) is gratefully acknowledged. We dedicate this work to the memory of Dr. Franz Hillenkamp in recognition of his lifetime achievements in the field of MALDI mass spectrometry.


## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/211/suppl/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S3
References (*30*–*39*)


17 October 2014; accepted 23 February 2015
Published online 5 March 2015;
10.1126/science.aaa1051


QUANTUM COMPUTING

# Quantum versus classical annealing of Ising spin glasses


**Bettina Heim,[1] Troels F. Rønnow,[1] Sergei V. Isakov,[2] Matthias Troyer[1]***



**Quantum annealers use quantum fluctuations to escape local minima and find low-energy configurations of a physical system. Strong evidence for superiority of quantum annealing (QA) has come from comparing QA implemented through quantum Monte Carlo (QMC) simulations to classical annealing. Motivated by recent experiments, we revisit the question of when quantum speedup may be expected. Although a better scaling is seen for QA in two-dimensional Ising spin glasses, this advantage is due to time discretization artifacts and measurements that are not possible on a physical quantum annealer. Simulations in the physically relevant continuous time limit, on the other hand, do not show superiority. Our results imply that care must be taken when using QMC simulations to assess the potential for quantum speedup.**


With the first archaeological records dating back more than 6000 years (*1*), thermal annealing is likely to be the oldest optimization method in human history. First heating a material and then letting it cool down slowly can relieve internal stresses and allow the material to achieve a lower-energy state. Inspired by thermal annealing, the simulated annealing (SA) algorithm (*2*) was proposed to find the ground states of combinatorial optimization problems—in particular, Ising spin glasses described by the Hamiltonian

$$H_c = -\sum_{i<j} J_{ij} s_i s_j - \sum_i h_i s_i \qquad (1)$$

Here, the $N$ spins $s_i$ can take the values ±1. Spins $s_i$ and $s_j$ on lattice sites $i$ and $j$ are coupled by the Ising term $J_{ij}$, and the $h_i$ are local fields. Nonconvex optimization problems—such as finding the ground state of this Ising spin glass (*3*), job scheduling (*4*), circuit minimization (*5*), and chain optimization (*6*)—find applications in many areas. These problems are all nondeterministic polynomially complete (*7*), which implies polynomial time mapping from one problem to the other. Thus, any method to efficiently find solutions to the Ising spin glass problem would provide an efficient way of solving other important problems.

Applying the Metropolis algorithm (*8*), Kirkpatrick *et al.* demonstrated that using SA—i.e., simulating the process of cooling Ising spin glasses in a Monte Carlo simulation—is an excellent method to minimize $H_c$ (*2*). Starting in a random state at high temperature, the system is slowly cooled. Thermal excitations allow the escape from local minima and relaxation into a low-energy state with energy equal or near that of the ground state $E_0$ (*9*). We will refer to the difference between the final energy $E$ and $E_0$ as the residual energy $E_{res} = E - E_0$. Quantum annealing (QA) (*10*–*14*) uses quantum tunneling instead of thermal excitations to escape from local minima, which can be advantageous in systems with tall but narrow barriers, which are easier to tunnel through than to thermally climb over. To perform QA of Ising spin glasses, an additional kinetic term is added, usually by applying a transverse magnetic field. The time-dependent Hamiltonian of QA is given by

$$H_q = -\sum_{i<j} J_{ij}\sigma_i^z\sigma_j^z - \sum_i h_i\sigma_i^z - \Gamma(t)\sum_i \sigma_i^x \qquad (2)$$

where $\sigma_i^z$ and $\sigma_i^x$ are Pauli $z$ and $x$ operators, respectively. The transverse field $\Gamma(t)$ is initially much larger then the couplings, $\Gamma(0) \gg |J_{ij}|, |h_i|$, and the spins start out aligned in the $x$ direction. $\Gamma(t)$ is slowly reduced to zero such that at the end of the annealing process, we recover the Hamiltonian of the initial Ising spin glass problem. On a perfectly coherent quantum device, this algorithm (*13*) would find the ground state of the spin glass in question with high probability, provided that the annealing time $t_a$ is sufficiently long to stay adiabatically in the ground

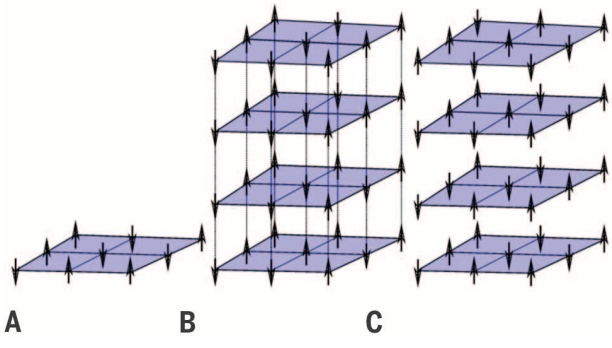



**Fig. 1. Illustration of computational resources used in SA and DT-SQA.** (**A**) A lattice of classical spins used in SA. (**B**) DT-SQA maps the transverse field Ising model to a classical system representing imaginary time paths of the quantum spins in an additional imaginary time direction. We show an example of $M = 4$ time slices with discrete time steps $\Delta\tau = \beta/M$. (**C**) With similar computational effort, we can perform SA on $M$ independent replicas.


[1]Theoretische Physik, ETH Zurich, 8093 Zurich, Switzerland. [2]Google, Brandschenkestrasse 110, 8002 Zurich, Switzerland.
*Corresponding author. E-mail: troyer@phys.ethz.ch

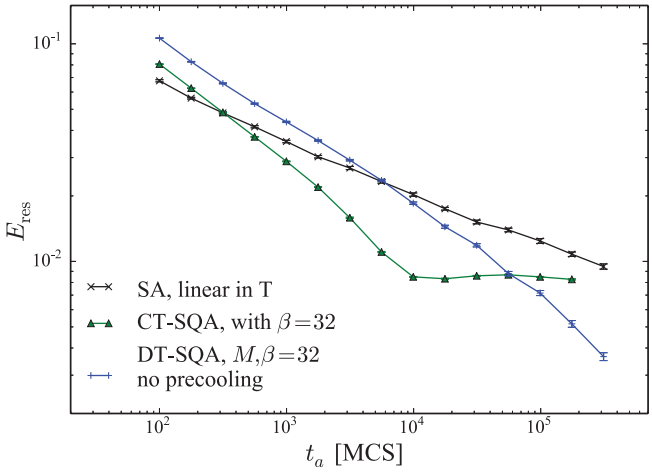



**Fig. 2. Residual energy $E_{res}$ dependence as a function of annealing time $t_a$.** Graph shows the residual energy for SA, DT-SQA, and CT-SQA for the square-lattice Ising spin glass instance of (*19*) with 6400 spins. The annealing time is in units of MCS, corresponding to one attempted update per spin. $E_{res}$ and error bars (indicating 1-σ statistical error) are obtained by averaging over 40 annealing runs.

state (*15*, *16*). QA can also be performed at nonzero temperature—for example, on spin glass materials (*17*) or in programmable devices (*18*).

Although the simulation of unitary time evolution scales exponentially with $N$, a variant of QA can be efficiently performed on a classical computer using stochastic dynamics in a quantum Monte Carlo (QMC) simulation (*19*, *20*). In this simulated quantum annealing (SQA) algorithm, the partition function of the quantum Ising model in a transverse field is mapped to that of a classical Ising model in one higher dimension corresponding to the imaginary time direction (*21*), as shown in Fig. 1. Details of the algorithms are discussed in the supplementary materials (*22*).

Strong evidence for QA being superior to classical annealing comes from a comparison of the performance of SQA and SA (*14*, *19*, *20*). Upon increasing the annealing time, the residual energy in a two-dimensional (2D) Ising spin glass was seen to drop faster in SQA than in SA, indicating that quantum tunneling may be advantageous in finding low-energy states. However, recent studies of the performance of D-Wave devices (D-Wave Systems, Burnaby, Canada), which are designed to use superconducting circuits to realize QA, failed to demonstrate indications of quantum speedup (*23*). Furthermore, in contrast to the work in (*19*, *20*), no advantage of SQA over SA was observed.

To understand these contradictory findings, we first confirm the results of Santoro *et al.* (*19*) by observing better scaling of SQA compared with SA in Fig. 2 [and in more detail in fig. S1 in (*22*)]. However, these simulations were performed with a small number of time slices $M$ and a corresponding large time step $\Delta\tau = \beta/M = 1$, which we refer to as a discrete time SQA (DT-SQA) simulation. Here, $\beta = 1/k_BT$ ($k_B$, Boltzmann's constant; $T$, temperature) is the inverse temperature. Discrete time steps incur time discretization errors of order $\mathcal{O}(\beta^3/M^2)$. To obtain accurate thermal averages for the quantum system, one needs to extrapolate results to $\Delta\tau \to 0$ or perform a continuous time simulation (CT-SQA) that works directly in the limit $\Delta\tau \to 0$ (*24*).

Repeating the simulations using CT-SQA, we see an entirely different behavior. Although the performance is improved for short $t_a$, the residual energy saturates for longer $t_a$ at a level higher than that reached by SA. Whereas the time discretization error in DT-SQA does not affect its use as a classical optimization algorithm, it does not reflect a physical quantum system, for which the continuous time limit is relevant. Hence, the circumstances under which SQA outperforms SA depend on whether we use SQA as a quantum-inspired classical algorithm or as simulation of a physical system.

Understanding the role of time discretization in SQA is important both to estimate the performance of experimental quantum annealers as well as to tune SQA as a classical optimization algorithm. To achieve this, we went beyond the single instance of (*19*) and studied 1000 random spin glass instances on an 80-by-80 square lattice with periodic boundary conditions. We use the same distribution of uniform couplings $J_{ij} \in (-2, 2)$ and $h_i = 0$. The exact ground-state energies $E_0$ are obtained using the spin glass server (*25*).

In Fig. 3A, we show $E_{res}$ as a function of $t_a$ for various $M$. As expected, for $M \to \infty$, DT-SQA converges toward the continuous time limit. We confirm the behavior already indicated in Fig. 2: Although the initial scaling is better in CT-SQA, lower residual energies are reached using a large time step. Comparing $M = 16$ and 64, a lower residual energy of $2 \cdot 10^{-3}$ is found for $M = 16$ compared with $5 \cdot 10^{-3}$ for $M = 64$, despite the computational effort being four times smaller.

Analyzing the residual energies as a function of temperature with a constant number of time slices in Fig. 3B leads to a similar observation. For $\beta < 20$, the DT-SQA results match well with the CT-SQA results [shown in fig. S4 in (*22*)], indicating that $M = 64$ is sufficient to converge to the continuous time limit. At lower temperatures, deviations from CT-SQA are seen, and the larger $\Delta\tau = \beta/M$ allows us to find states with lower energy than in CT-SQA.

When using SQA as a classical optimization algorithm, we can search the final configuration for the time slice with the lowest energy instead of averaging over all time slices. However, we also have to take into consideration the increased computational effort of QMC simulations compared with SA and multiply the number of Monte Carlo steps (MCS) by $M$ for DT-SQA and by β for CT-SQA. Plotting the residual energy as a function of total computational effort in Fig. 4A, we find that with suitable chosen β and $M$, DT-SQA outperforms SA, agreeing with (*19*).

To use SQA as a classical optimization algorithm, it is thus advantageous to use a small $\Delta\tau$ for short $t_a$, because the continuous time limit has a more rapid initial decrease of $E_{res}$. When annealing for longer $t_a$, a lower temperature and larger $\Delta\tau$ are preferred to reach lower asymptotic residual energies. To reach the lowest energies, rather large $\Delta\tau$ values of order unity are needed, where the system consists of few moderately coupled individual replicas instead of a more tightly coupled continuous path of configurations.

Simulated quantum annealing simulations can escape a local minimum through path-integral configurations corresponding to tunneling events. These configurations spend only short (imaginary) times in high-energy configurations on the barrier, thus avoiding the full cost of the barrier. Although the Monte Carlo dynamics in SQA is different from unitary dynamics of QA, they both use a form of quantum tunneling to escape local minima and are thus similarly sensitive to the nature of barriers. Hence, SQA may indicate whether a physical QA device is expected to show an advantage over SA. A good correlation of SQA [and a mean-field version thereof (*26*)] has been seen for spin glass problems on a D-Wave One device (*27*).

To use SQA to predict where QA might outperform SA, we have to take the physical continuous time limit and use either CT-SQA or DT-SQA with large enough $M$. Figure 4B shows the slightly higher residual energy obtained this way. For short $t_a$ up to $t_a = 10^5$ MCS, SQA still outperforms SA when choosing an appropriate temperature, but the asymptotic scaling is better for SA.

In all of our simulations, the residual energy saturates at some point in SQA but not in SA, indicating that the simulations consistently get stuck in local minima for SQA. This may be understood by the more deterministic dynamics of QA, which prefers a subset of low-energy states (*28*). Repeating SQA, one may thus get consistently stuck in similar local minima. SA, on the other hand, starts in a random state at high temperatures and therefore explores the configuration space more evenly. The more deterministic nature of SQA also explains the counterintuitive result that for some choices of parameters (see Fig. 4), the $E_{res}$ may increase when annealing more slowly. Similar to the work in (*29*), perturbing a quantum annealer (for example, by annealing faster) can excite a system out of a local minimum and thus help to ultimately uncover a better solution.

Reinvestigating evidence for QA outperforming classical annealing for spin glass instances, we thus find that the performance advantage previously observed for path-integral QMC annealing compared with SA for 2D spin glasses (*19*, *20*) is

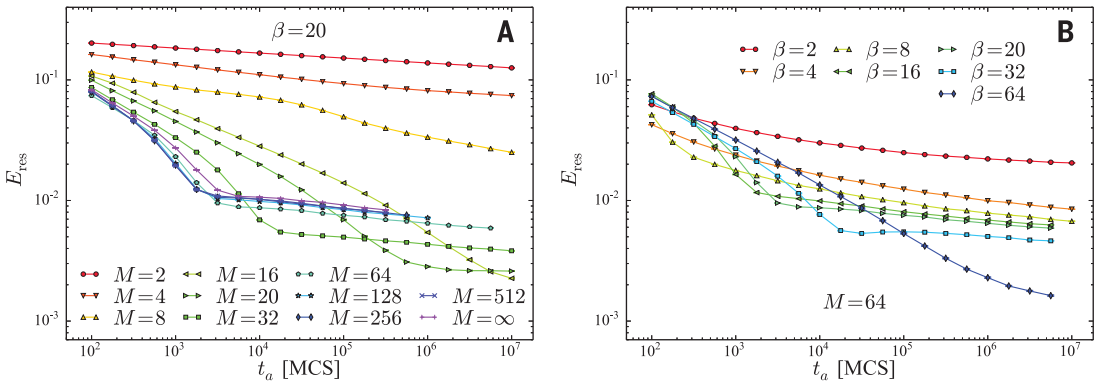



**Fig. 3. Convergence and temperature dependence of SQA.** (**A**) Convergence of DT-SQA toward the continuous time limit $M = \infty$ obtained by CT-SQA. Shown is the average of 1000 different disorder realizations annealed at an inverse temperature of $\beta = 20$. The lowest-energy configuration along the imaginary time axis was taken to calculate the residual energy. (**B**) Temperature dependence of DT-SQA with a constant number of $M = 64$ time slices. Lowering the temperature increases the time step $\Delta\tau = \beta/M$. This reduces the initial drop in energy but allows us to ultimately find a final configuration with lower energy.

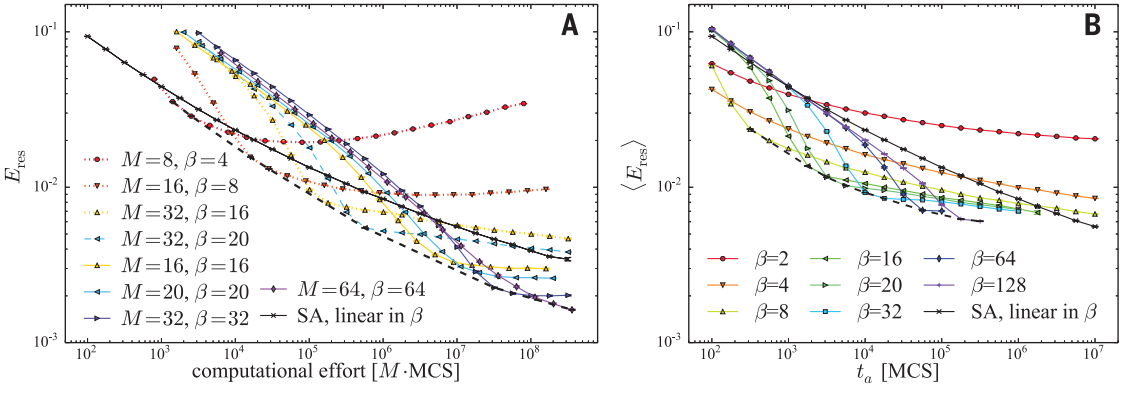



**Fig. 4. Performance of SQA as a classical optimizer versus a physical system.** (**A**) SQA as a classical optimizer: When choosing a suitable annealing temperature and time step, DT-SQA scales better than SA, consistent with the results of (*19*). (**B**) SQA as a simulation of a physical system, using large enough $M$ to be converged to the continuous time limit. Here we average the final energy over imaginary time instead of picking the lowest-energy configuration.

due to large imaginary time steps in the path integral and choosing the lowest energy over all time slices. When taking the limit of continuous time and measuring the average energy, the advantage vanishes. These results are also consistent with recent arguments that 2D spin glasses should not see any quantum speedup in QA (*30*). It will now be important to explore whether 3D spin glasses or spin glasses with long-range couplings, where barriers are taller than in 2D, show superiority for QA. There, and for problem instances derived from hard application problems, it will once more be essential to investigate both discrete and continuous time simulations. CT-QMC simulations can estimate the potential of QA to outperform SA, but ultimately the unitary dynamics of QA will need to help it outperform DT-SQA, which is by itself an efficient classical optimization algorithm. Although the answer to this question will have to wait for experiments on improved QA devices, SQA simulations performed in the correct way can guide the design of these future experiments.

## REFERENCES AND NOTES


1. M. Radivojević, T. Rehren, J. Kuzmanović-Cvetković, M. Jovanović, J. P. Northover, *Antiquity* **87**, 1030–1045 (2013).
2. S. Kirkpatrick, C. D. Gelatt Jr., M. P. Vecchi, *Science* **220**, 671–680 (1983).
3. F. Barahona, *J. Phys. Math. Gen.* **15**, 3241–3253 (1982).
4. R. Graham, *Bell Syst. Tech. J.* **45**, 1563–1581 (1966).
5. D. E. Knuth, *The Art of Computer Programming, Volume 4, Fascicle 1: Bitwise Tricks & Techniques; Binary Decision Diagrams* (Addison-Wesley Professional, Boston, ed. 12, 2009).
6. L.-P. Wong, M. Y.-H. Low, C. S. Chong, *Int. J. Artif. Intell. Tools* **19**, 305–334 (2010).
7. S. A. Cook, in *Proceedings of the Third Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 1971), pp. 151–158.
8. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **21**, 1087–1092 (1953).
9. D. Bertsimas, J. Tsitsiklis, *Stat. Sci.* **8**, 10–15 (1993).
10. P. Ray, B. K. Chakrabarti, A. Chakrabarti, *Phys. Rev. B* **39**, 11828–11832 (1989).
11. A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, J. D. Doll, *Chem. Phys. Lett.* **219**, 343–348 (1994).
12. T. Kadowaki, H. Nishimori, *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **58**, 5355–5363 (1998).
13. E. Farhi *et al.*, *Science* **292**, 472–475 (2001).
14. A. Das, B. K. Chakrabarti, *Rev. Mod. Phys.* **80**, 1061–1081 (2008).
15. L. Landau, *Phys. Z. Sowjetunion* **2**, 46–51 (1932).
16. C. Zener, *Proc. R. Soc. London Ser. A* **137**, 696–702 (1932).
17. J. Brooke, D. Bitko, T. F. Rosenbaum, G. Aeppli, *Science* **284**, 779–781 (1999).
18. M. W. Johnson *et al.*, *Nature* **473**, 194–198 (2011).
19. G. E. Santoro, R. Martonák, E. Tosatti, R. Car, *Science* **295**, 2427–2430 (2002).
20. R. Martoňák, G. E. Santoro, E. Tosatti, *Phys. Rev. B* **66**, 094203 (2002).
21. M. Suzuki, *Prog. Theor. Phys.* **56**, 1454–1469 (1976).
22. See the supplementary materials on *Science* Online.
23. T. F. Rønnow *et al.*, *Science* **345**, 420–424 (2014).
24. H. Rieger, N. Kawashima, *Eur. Phys. J. B* **9**, 233–236 (1999).
25. www.informatik.uni-koeln.de/spinglass/.
26. S. W. Shin, G. Smith, J. A. Smolin, U. Vazirani, http://arxiv.org/abs/1401.7087 (2014).
27. S. Boixo *et al.*, *Nat. Phys.* **10**, 218–224 (2014).
28. Y. Matsuda, H. Nishimori, H. G. Katzgraber, *New J. Phys.* **11**, 073021 (2009).
29. E. Crosson, E. Farhi, C. Y.-Y. Lin, H.-H. Lin, P. Shor, http://arxiv.org/abs/1401.7320 (2014).
30. H. G. Katzgraber, F. Hamze, R. S. Andrist, *Phys. Rev. X* **4**, 021008 (2014).


## ACKNOWLEDGMENTS

We thank H. G. Katzgraber, G. Santoro, and I. Zintchenko for useful discussions and G. Santoro for providing the spin glass instance used in (*19*). This work was supported by the Swiss National Science Foundation through the National Competence Center in Research QSIT and by the European Research Council (ERC) through ERC Advanced Grant SIMCOFE. M.T. acknowledges the hospitality of the Aspen Center for Physics, supported by NSF grant 1066293. The spin glass server (*25*) was used to obtain the ground states for our problem instances. The original data used to create the figures can be obtained from the corresponding author.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/215/suppl/DC1
Materials and Methods
Fig. S1 to S4
References (*31*–*34*)

3 December 2014; accepted 26 February 2015
Published online 12 March 2015;
10.1126/science.aaa4170

MARS ATMOSPHERE

# Strong water isotopic anomalies in the martian atmosphere: Probing current and ancient reservoirs

G. L. Villanueva,[1,2]* M. J. Mumma,[1] R. E. Novak,[3] H. U. Käufl,[4] P. Hartogh,[5] T. Encrenaz,[6] A. Tokunaga,[7] A. Khayat,[7] M. D. Smith[1]

**We measured maps of atmospheric water ($H_2O$) and its deuterated form (HDO) across the martian globe, showing strong isotopic anomalies and a significant high deuterium/hydrogen (D/H) enrichment indicative of great water loss. The maps sample the evolution of sublimation from the north polar cap, revealing that the released water has a representative D/H value enriched by a factor of about 7 relative to Earth's ocean [Vienna standard mean ocean water (VSMOW)]. Certain basins and orographic depressions show even higher enrichment, whereas high-altitude regions show much lower values (1 to 3 VSMOW). Our atmospheric maps indicate that water ice in the polar reservoirs is enriched in deuterium to at least 8 VSMOW, which would mean that early Mars (4.5 billion years ago) had a global equivalent water layer at least 137 meters deep.**

We report maps of water isotopologues ($H_2O$ and its deuterated form HDO) across the martian globe, which show strong isotopic anomalies that are difficult to reconcile with simple fractionation processes. These maps address fundamental unknowns such as the current representative ratio of D/H in water and how much water was lost over the geological life of the planet. Isotopic ratios are among the most valuable indicators for the loss of volatiles from an atmosphere. Deuterium fractionation also reveals information about the cycle of water on the planet and informs us of its stability on short- and long-term scales. The vapor pressures of HDO and $H_2O$ differ substantially near the freezing point, making the condensation/sublimation cycle of the isotopologues sensitive to local temperatures and saturation levels and to the presence of aerosol condensation nuclei.

Although many maps of $H_2O$ do exist [e.g., (*1*, *2*)], and some of HDO [e.g., (*3*, *4*)], two-dimensional (2D) maps of atmospheric D/H enrichment on Mars do not. Apart from preliminary 2D maps presented at conferences (*5*), previous mapping attempts to measure D/H provided only 1D maps of HDO and $H_2O$ based on nearly simultaneous measurements at ground-based infrared (IR) observatories [e.g., (*6*–*9*)] or 1D HDO maps (ground-based) and quasisimultaneous $H_2O$ measurements taken from Mars orbit by the Planetary Fourier Spectrometer long-wavelength channel (*10*). Local isotopic ratios such as derived by Curiosity/ Mars Science Laboratory (MSL) (*11*) are representative of that specific location and time and may not represent the actual D/H of atmospheric water on Mars (see more below), even though reports of hemispherically averaged D/H in water indicated a similar value of 5 to 6 times the D/H of Earth's oceans [Vienna standard mean ocean water (VSMOW)] (*12*, *13*). Importantly, such isolated (in time and in space) measurements of D/H in the atmosphere were typically—but incorrectly—assumed to be representative of the bulk atmosphere. Spatially resolved measurements of D/H at different times of day and seasons are necessary to disentangle local from global phenomena. Such maps reveal the true isotopic ratio of current water reservoirs, with implications for the global loss of water over geologic time, and may also assist in the identification of new sources of water on Mars.

Currently, the martian atmosphere and surface form an arid and highly inhospitable environment (*14*, *15*). The upper layers of soil were heavily eroded by long-term aeolian activity, and soil chemistry was strongly modified through exposure to energetic photons and particles that penetrated owing to inadequate magnetospheric and atmospheric shielding. However, there is ample evidence that ancient Mars was wet and likely hosted habitable conditions (*16*, *17*), particularly during the Noachian age (3.6 to 4.5 billion years ago), probably leading to the formation of rich subsurface aqueous reservoirs. Moreover, the presence of extensive volcanism likely gave rise to widespread hydrothermal activity and to the formation of diverse chemical environments. Measurements of epithermal neutron fluxes obtained by Mars Odyssey (*18*) suggest the presence of important near-surface hydrogen concentrations on Mars, and the polar layered deposits (PLD) are estimated to contain ~21 m global equivalent layer (GEL) of water (*19*, *20*). The atmosphere acts as a buffer between the exosphere (at the boundary with outer space) and the main reservoirs of H, C, and O (e.g., regolith and polar caps), with atmospheric isotopic/abundance ratios providing key diagnostics for quantifying the exchange among these environments.

The maps reported herein are based on observations of isotopic water made with high-resolution spectrometers [CRIRES (cryogenic infrared echelle spectrograph), NIRSPEC (near-infrared spectrograph), and CSHELL (cryogenic echelle spectrograph)] at powerful ground-based observatories [VLT (Very Large Telescope), Keck, and IRTF (InfraRed Telescope Facility), respectively] [supplementary materials section 1 (SM-1)]. This program is founded on initial observations performed using NASA's Kuiper Airborne Observatory and the Kitt Peak National Observatory in 1989 (*13*), and later with IRTF (*6*, *7*). Specifically we targeted the $\nu_1$ band of HDO near 2720 $cm^{-1}$ (3.7 μm) and the $2\nu_2$ band of $H_2O$ near 2990 $cm^{-1}$ (3.3 μm) (see Fig. 1). Spectral lines of these bands on Mars are observable through our atmosphere when measured from high-altitude observatories at moderately high Doppler shifts (>11 km $s^{-1}$), i.e., when Mars's lines are displaced sufficiently far from the cores of their counterpart telluric absorbing lines. Observations of HDO are additionally favored by the large D/H enrichment present in the martian atmosphere and the strong D/H depletion in terrestrial air at the high altitudes of these observatories (*9*), which leads to high telluric transmittances. We obtained the greatest sensitivities to HDO and $H_2O$ on Mars by performing observations at times of maximum Doppler shifts, typically several months before or after Mars' closest approach to Earth (every ~2 years).

Mapping was achieved by stepping the spectrometer's entrance slit across the planet and sampling the planet's spectrum at intervals along the slit for each step, leading to full 2D (east-west and north-south) coverage across the observable disk (see details in SM-1). The distance between slit positions was set in accord with observational image quality ("seeing"), guiding precision, and slit width. For a typical Mars diameter of 8″, a slit-width of 0.2″ and a "seeing" of 0.6″, a typical map would consist of 11 slit positions stepped across the planet, with a minimum of 9 positions for less favorable conditions. For each slit position, abundance measurements are obtained at intervals of 0.6″ along the slit. The maps are of relatively high spatial resolution (~500 km near sub-Earth positions), and, importantly, each map was obtained within a short time interval (1 to 2 hours), thereby providing hemiglobal snapshots in time that are critical for investigating the effects of local phenomena (e.g., orographic clouds and planet-scale waves) and transient processes (e.g., polar releases and diurnal effects).

The data reported here were collected over several years and seasons on Mars from March 2008 until January 2014. From this database, we made a localized D/H measurement over the Viking 1 landing site (*9*) and obtained a comprehensive search for organics in the martian atmosphere (*15*). Here, we present results from the best data sets targeting D/H, obtained at times

[1]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. [2]Catholic University of America, Washington, DC 20064, USA. [3]Iona College, New Rochelle, NY 10801, USA. [4]European Southern Observatory, Munich, Germany. [5]Max Planck Institute for Solar System Research, Katlenburg-Lindau 37191, Germany. [6]Observatoire de Paris-Meudon, Meudon 92195, France. [7]University of Hawaii-Manoa, Honolulu, HI 96822, USA.
*Corresponding author. E-mail: geronimo.villanueva@nasa.gov

of low telluric water, high Doppler shift, and maximum spatial coverage (CRIRES, 8–9 September 2009; CSHELL, 25 March 2008; NIRSPEC, 24 January 2014; CRIRES 29–30 January 2014) (see Fig. 2 and fig. S1). The data span seasons from late northern winter to late northern spring on Mars (areocentric longitude or season $L_s$ = 335°, 50°, 80°, and 83°) and sample the critical interval when the northern polar cap sublimates and replenishes the atmosphere with water.

The highest spectral resolution is achieved with the narrow slit (0.2″) of CRIRES, leading to a resolving power of λ/δλ ~100,000, whereas the cross-dispersed feature of NIRSPEC provides the broadest spectral coverage at a high resolving power of λ/δλ ~40,000. Such broad spectral coverage allows us to sample HDO and $H_2O$ with a single NIRSPEC setting, whereas measurements with CRIRES and CSHELL require mappings at two spectral settings (3.7 μm and 3.3 μm). Each setting also samples lines of $CO_2$ that we use to establish the total atmospheric column and to quantify extinction and scattering effects by water ice and dust aerosols (SM-2). Abundance ratios and column densities for $H_2O$ and HDO at each footprint are then corrected for local (spatial and spectral) variations in the comeasured $CO_2$ column, removing an important source of systematic error that would otherwise affect isotopic measurements. We derive molecular column densities at each footprint by comparing the residual spectrum to a synthetic Mars spectrum (color traces in Fig. 1), itself affected by the monochromatic telluric transmittance for each spectral line at its Mars Doppler-shifted spectral position. For this process, we employ the efficient and robust Levenberg-Marquardt curve-fitting method, which also provides reliable error estimates as derived from the measured Jacobians (SM-1). We typically obtain high signal-to-noise ratios of ~70 for the HDO columns with CRIRES, but the precision of the D/H ratio is defined mainly by the $H_2O$ columns. The median accuracy of the D/H measurements is 0.2, 0.5, 0.8, and 0.4 VSMOW for the January 2014 CRIRES, January 2014 NIRSPEC, September 2009 CRIRES, and March 2008 CSHELL observations, respectively.

The $H_2O$ and HDO disk maps (Fig. 2) reveal strong local anisotropies and seasonal variability. The slow replenishing of water vapor in the northern hemisphere as the polar cap sublimates during northern spring is quite noticeable, in particular when compared to the baseline measurements in late northern winter ($L_s$ 335°). Variability of $H_2O$ with latitude and season is apparent (Fig. 2) and is generally consistent with previous spacecraft measurements of the water cycle (*21*, *22*). HDO maps superficially resemble those of $H_2O$ in showing strong variability, but their direct comparison reveals strong differences that are most easily seen in the maps of the ratio of D/H enrichment in water vapor, relative to Earth's oceans (VSMOW).

The maps of D/H enrichment show a remarkably different evolution and structure compared with total water, with an apparent relationship to topography and atmospheric temperature. Low D/H values are seen at high altitudes, whereas high D/H values are seen in basins and orographic depressions. Very low D/H [1 to 3 VSMOW] appears in the winter hemisphere, but the spring hemisphere shows much higher values. The vapor pressure isotope effect, which produces an isotopic fractionation at condensation (e.g., cloud formation and frost/ground fog formation), could explain some of this latitudinal variability, yet the observed localized (in time and space) anisotropies are certainly higher than was predicted by current atmospheric models (*23*, *24*). For instance, preferential condensation of HDO could perhaps explain the strong D/H depletion seen in clouds over Elysium Mons on 24 January 2014 and similarly over the Tharsis district as observed in September 2009 and January 2014. Interestingly, strong clouds are observed over this region by the Mars Color Imager camera onboard the Mars Reconnaissance Orbiter (*25*). As discussed in (*24*), Rayleigh distillation would provide the strongest isotopic fractionation, yet (on Mars) competition between condensation and sedimentation may lead to certain nonequilibrium fractions, while kinetic processes may enhance or diminish the efficiency of fractionations on Mars as is observed on Earth (*26*).

Importantly, our maps reveal notably higher deuterium enrichment than was found in globally averaged observations that indicated D/H values of 5.0 ± 0.2 VSMOW (*13*) and 5.8 ± 2.6 VSMOW (*12*). This difference is expected because

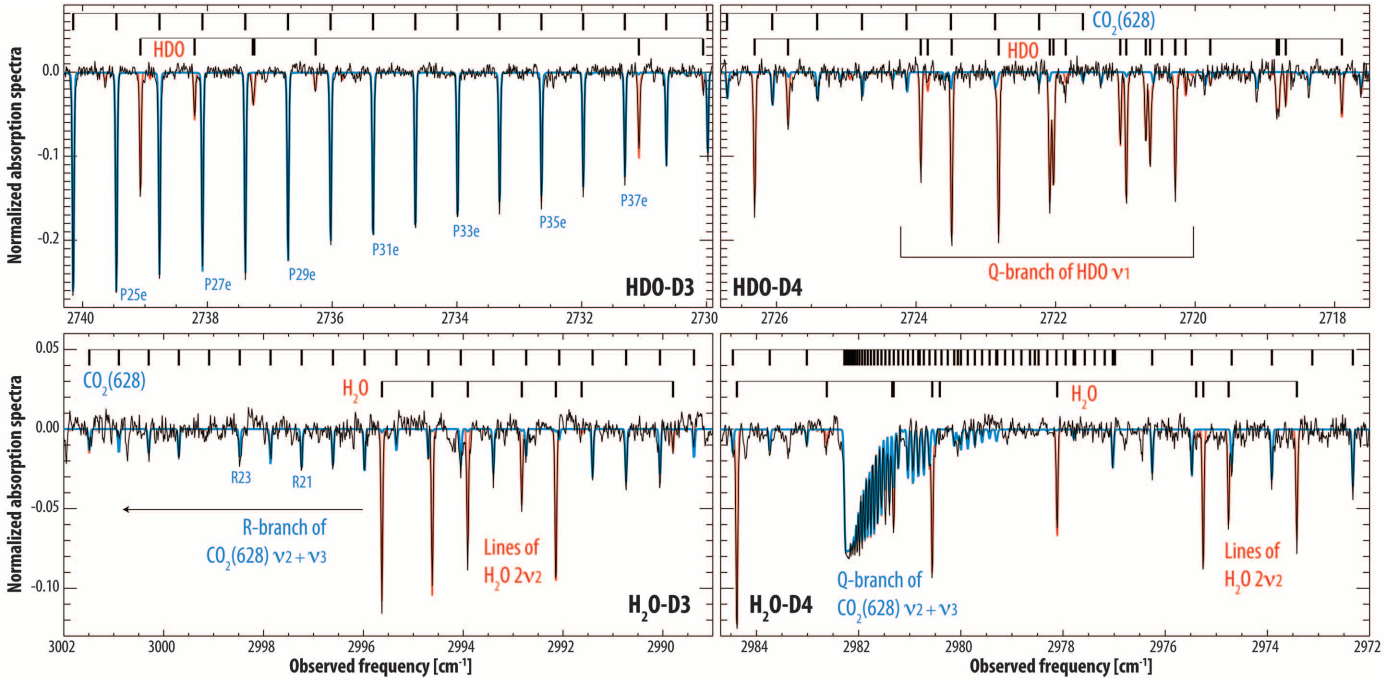



**Fig. 1. Mars high-resolution (λ/δλ ~ 100,000) residual spectra of HDO, $H_2O$, and $CO_2$ acquired with two settings of CRIRES at VLT on 29 January 2014 ($L_s$ 83°).** The data shown were extracted for a field-of-view of 0.2″ (slit width) × 0.256″ (3 pixels along the slit) over Mars's northern hemisphere (latitude, 57° to 69°; longitude, 106° to 133E°). Spectra from detectors D3 and D4 of each CRIRES setting are shown; detectors D1 and D2 reveal additional lines of these species (SM-1). Residual spectral deviation is 0.006 per pixel (1σ). Each martian residual spectrum was derived by subtracting synthetic spectra of the comeasured telluric and solar absorption features (*15*), after normalizing to the martian continuum. Synthetic models for HDO and $H_2O$ (red trace) and for $CO_2$ (blue trace) are in good agreement with the measured spectra.

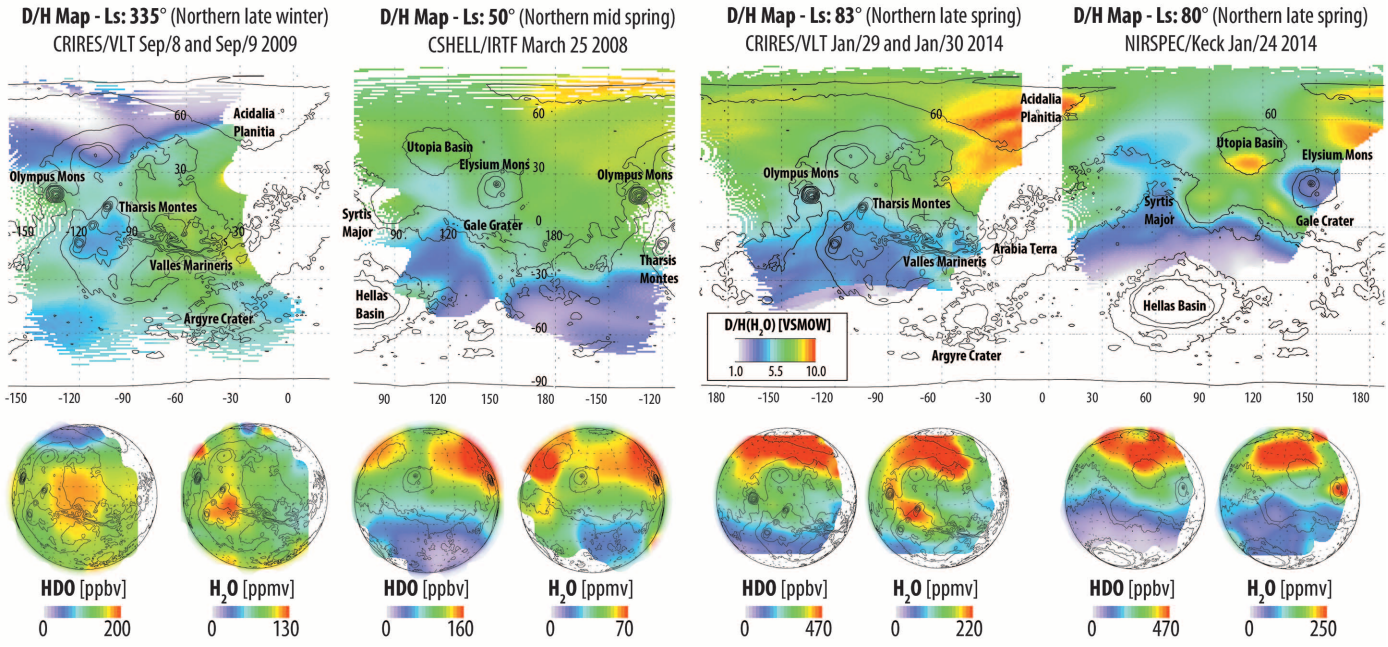



**Fig. 2. Maps of HDO and $H_2O$ and their ratio on Mars, obtained at four different seasons (from late northern winter to late northern spring).** D/H maps (upper row) were obtained by calculating the ratio of the measured abundance of HDO and $H_2O$ extracted from maps of the individual species (lower row) and are presented relative to the D/H value in Earth's ocean water (VSMOW). The ratio maps (upper row) show the progressive enrichment of D/H in the northern hemisphere as the polar cap sublimes during northern spring. The isotopologue maps (HDO and $H_2O$) also reveal strong local enhancements and variability but with significant differences between them, associated with global and local climatology. In particular, low D/H values are observed at regions of low temperature and/or high altitude, with high values observed at orographic depressions (e.g., Acidalia Planitia and Utopia Basin).

full disk measurements reflect the mean of diverse regions with low and high D/H, as revealed by our maps (Fig. 2). Moreover, isolated measurements from a discrete region may reflect local climatological effects and so may not represent the deuterium enrichment of the bulk atmosphere. For instance, Gale Crater (location of the MSL Curiosity rover) is at the boundary of a strong gradient between two regions of D/H that range from 1 to 7 VSMOW, where local and seasonal effects may affect retrievals of D/H ($L_s$ 50) (Fig. 2). Our value at Gale Crater (6 VSMOW; $L_s$ 50 and 80) is in good agreement with that derived with the MSL Tunable Laser Spectrometer instrument (6 ± 1 VSMOW) in late 2012 (*11*). However, our maps reveal that in regions and seasons where both isotopologues are expected to be fully gaseous (equatorial and mid-latitudes during mid and late spring, $L_s$ 50, 83, 80), the atmosphere shows high D/H (~7 VSMOW), with some regions showing strong enrichments reaching 9 to 10 VSMOW.

Temperature is the main parameter governing isotopic ratios in an atmosphere, but other factors such as relative humidity (condensation level), presence of dust particles (condensation nuclei), and dynamical/transport processes (e.g., expansive cooling due to vertical lifting) affect the rate at which condensation occurs for each isotopologue. These factors affect the observable gas phase D/H ratio (see D/H correlations with temperature and water column in SM-3). Models for the seasonal behavior of the D/H ratio on Mars incorporate some of these effects. Using a 3D general circulation model that included fractionation effects, Montmessin *et al.* (*24*) predicted hemispheric variability with a peak atmospheric D/H value 15% lower than the value assumed for the polar caps. In general, our results are in agreement with the predicted level of variability with latitude, but the strong local anisotropies observed across the planet will require a more realistic model to account for several climatological processes acting on the isotopologues. Importantly, our typical atmospheric value (7 VSMOW) implies that the permanent polar caps contain a D/H of 8 VSMOW, following the modeling results in (*24*).

The great obliquity variations experienced by Mars at million-year intervals (*27*) should have caused all major ice reservoirs to vaporize and reform repeatedly, causing refreshed mixing of water from the different repositories at regular intervals. If so, all near-surface and polar reservoirs of water on Mars should share a relatively common D/H ratio. Because we also observe even higher D/H values (above 8 VSMOW) in certain regions, such mixing would suggest that the current reservoirs of water on Mars contain a higher D/H than initially thought and would consequently require greater loss of water over the planet's lifetime.

Multiple reservoirs have been proposed to account for the current inventory of water on Mars, ranging from the observable PLD (*19*, *20*), to ice-rich regolith at mid-latitudes (*28*, *29*), near-surface reservoirs at high latitudes (*30*), and subsurface reservoirs as implied by gamma ray and neutron observations (*31*). As summarized by Kurokawa *et al.* (*32*), the PLD can be defined as the minimum estimate of the current inventory of Mars. The PLD would correspond to $1.4 \pm 0.3 \times 10^{18}$ kg of water [or 10 ± 2 m GEL of water for the north polar region (*19*)] and $1.6 \pm 0.2 \times 10^{18}$ kg (11 ± 1.4 m GEL) for the south polar region (*20*), totaling 21 ± 3.4 m for the PLD reservoir. Recent estimates suggest that the PLD contain 17 to 21 m GEL, with 25 to 29 m GEL across all directly measured reservoirs (*33*). The initial reservoir of water on Mars ($M_p$) can be estimated from the current inventory of water on Mars ($M_c$) and its isotopic ratio ($I_c$) when knowledge of the ancient isotopic ratio ($I_p$) exists, together with the fractionation escape rate ($f$) of the isotopologues, using the relation equation $M_p/M_c = (I_c/I_p)^{[1/(1-f)]}$. From Hubble Space Telescope measurements of D and H Lyman-α emissions in the upper atmosphere of Mars, Krasnopolsky *et al.* (*34*) derived a fractionation rate of 0.02. Usui *et al.* (*35*) measured a water D/H of 1.275 VSMOW in melt inclusions (estimated to be 4.5 billion years old) within the Mars meteorite Yamato 980459. Taking this value as $I_p$ and 21 m GEL as $M_c$, our estimate of 8 VSMOW (for the PLD) as $I_c$ implies that Mars had at least 137 m GEL of water 4.5 billion years ago (see Fig. 3).

This value could be defined as a lower bound for the original GEL because it is based on the minimum current water reservoir (PLD) and neglects the young Sun's high extreme ultraviolet fluxes (*36*), which would have led to the unfractionated escape of H and D. If we assume

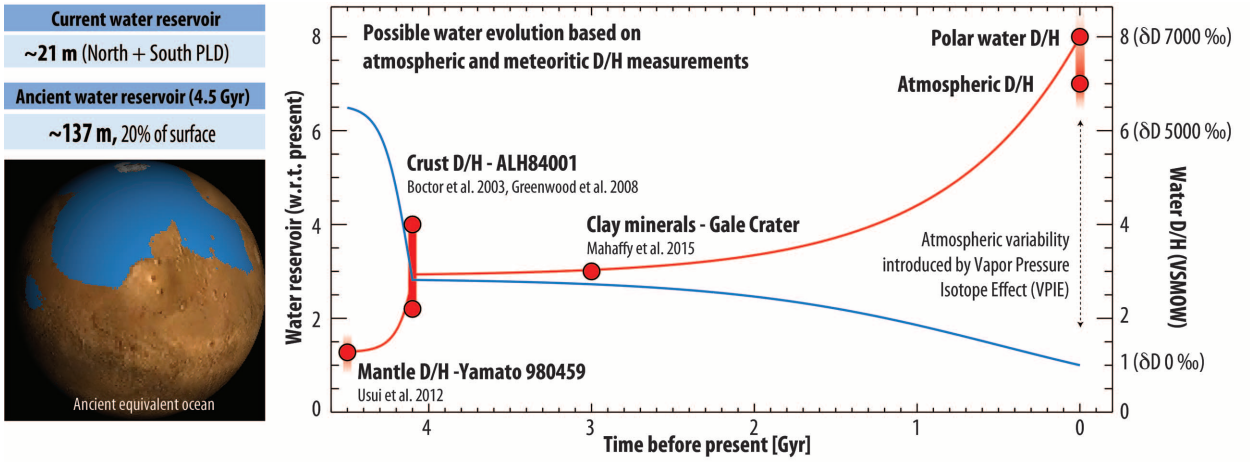



**Fig. 3. Isotopic enrichment as evidence for global loss of water on Mars.** After correcting for local climatological fractionation of the measured D/H ratio (Fig. 2), the current ratio for D/H in atmospheric water on Mars is at least 7 VSMOW, implying a D/H ratio of 8 VSMOW in the north polar reservoir (red curve and right axis). Assuming a fractionation factor *f* of 0.02, the D/H ratios obtained from water in Mars meteorites (Yamato 980459, 4.5 billion years old) imply that Mars's initial water reservoir was larger than the current water available on Mars by a factor of at least 6.5 (blue curve and left axis). When considering the current PLD content of 21 m of water, this would imply that at least 137 m GEL of water was present on Mars 4.5 billion years ago, covering 20% of the planet's surface.

that all currently measured water reservoirs (25 to 29 m), including the atmosphere, share the representative 7 VSMOW enrichment, then the ancient estimate is 142 to 165 m, further establishing the 137-m GEL value as a lower bound. Considering Mars's current topography (*37*), 137 m GEL water would have covered up to 20% of the planet's surface (*38*). Meteoritic records dated to 4.1 billion years ago (*39*, *40*) and clay minerals in Gale Crater (*41*) indicate a much higher D/H than in Yamato 980459, implying that the loss of water from Mars occurred in stages (*32*), with a substantial amount of water being lost in the first 0.5 billion years. Geomorphological records on Mars do indicate a wetter past (up to >2000 m GEL), yet—as summarized by Carr and Head (*38*) and more recently confirmed by subsurface observations with the Mars Advanced Radar for Subsurface and Ionosphere Sounding (*42*)—the best estimate is provided by the Vastitas Borealis Formation, which implies $2.3 \times 10^7$ km$^3$ of water (156 m GEL). This estimate is in relatively good agreement with our value (137 m GEL) as inferred from our estimate of D/H enrichment (8 VSMOW) in the PLD. The difference between the two estimates would also mean that ~20 m GEL of water are currently "missing" and could be stored in other proposed water reservoirs (e.g., deep aquifers).

Our D/H maps highlight the importance of obtaining isotopic measurements on Mars that are both spatially and temporally resolved in order to separate climatological from evolutionary effects. Such investigation ultimately leads to a more accurate estimate of the actual D/H of water reservoirs on Mars and improves both the estimate of water loss over geologic time and estimates for the "missing" water that might reside in undiscovered reservoirs. More realistic estimates on current and ancient fractionation rates (e.g., from NASA's Mars Atmosphere and Volatile Evolution Mission) and a larger sample of ancient D/H values derived from authentic (meteorites) and from drilled and/or returned samples would better constrain the water inventory on Mars, current and past.

## REFERENCES AND NOTES


1. M. D. Smith, *Icarus* **167**, 148–165 (2004).
2. M. A. Tschimmel *et al.*, *Icarus* **195**, 557–575 (2008).
3. R. E. Novak, M. J. Mumma, M. A. DiSanti, N. Dello Russo, K. Magee-Sauer, *Icarus* **158**, 14–23 (2002).
4. T. Fouchet *et al.*, *Planet. Space Sci.* **59**, 683–690 (2011).
5. G. L. Villanueva *et al.*, *The Mars Atmosphere: Modeling and Observations Conference*, Williamsburg, VA, 10–13 November 2008, 9101 (Lunar and Planetary Institute, Houston, TX, 2008).
6. M. J. Mumma *et al.*, *Sixth International Conference on Mars*, Pasadena, CA, 20–25 July 2003, 3186 (Lunar and Planetary Institute, Houston, TX, 2003).
7. R. E. Novak, M. J. Mumma, G. L. Villanueva, B. P. Bonev, M. A. Disanti, *Seventh International Conference on Mars*, Pasadena, CA, 9–13 July 2007, 3283 (Lunar and Planetary Institute, Houston, TX, 2007).
8. R. E. Novak, M. J. Mumma, G. L. Villanueva, *Planet. Space Sci.* **59**, 163–168 (2011).
9. G. L. Villanueva *et al.*, *J. Quant. Spectrosc. Radiat. Transf.* **113**, 202–220 (2012).
10. S. Aoki *et al.*, *European Planetary Science Congress 2013*, London, 8–13 September 2013, **8**, EPSC2013-191 (European Planetary Science Congress, Paris, 2013).
11. C. R. Webster *et al.*, *Science* **341**, 260–263 (2013).
12. T. Owen, J. P. Maillard, C. de Bergh, B. L. Lutz, *Science* **240**, 1767 (1988).
13. G. L. Bjoraker, M. J. Mumma, H. P. Larson, *Bull. Am. Astron. Soc.* **21**, 991 (1989).
14. J. P. Grotzinger, *Science* **341**, 1475 (2013).
15. G. L. Villanueva *et al.*, *Icarus* **223**, 11–27 (2013).
16. J.-P. Bibring *et al.*, *Science* **312**, 400–404 (2006).
17. M. H. Carr, *J. Geophys. Res. Planets* **104** (E9), 21897–21910 (1999).
18. W. C. Feldman *et al.*, *J. Geophys. Res. Planets* **109**, E09006 (2004).
19. M. T. Zuber *et al.*, *Science* **282**, 2053–2060 (1998).
20. J. J. Plaut *et al.*, *Science* **316**, 92–95 (2007).
21. M. D. Smith, *J. Geophys. Res. Planets* **107**, 5115 (2002).
22. T. Fouchet *et al.*, *Icarus* **190**, 32–49 (2007).
23. T. Fouchet, E. Lellouch, *Icarus* **144**, 114–123 (2000).
24. F. Montmessin, T. Fouchet, F. Forget, *J. Geophys. Res. Planets* **110**, E03006 (2005).
25. M. C. Malin, B. A. Cantor, M. R. Wu, L. M. Saper, MRO MARCI Weather Reports (2014) (available at www.msss.com/msss_images/subject/weather_reports.html).
26. J. Jouzel, L. Merlivat, *J. Geophys. Res. Atmos.* **89** (D7), 11749 (1984).
27. J. Laskar *et al.*, *Icarus* **170**, 343–364 (2004).
28. J. B. Murray *et al.*, *Nature* **434**, 352–356 (2005).
29. D. P. Page, M. R. Balme, M. M. Grady, *Icarus* **203**, 376–389 (2009).
30. P. H. Smith *et al.*, *Science* **325**, 58–61 (2009).
31. W. V. Boynton *et al.*, *Science* **297**, 81–85 (2002).
32. H. Kurokawa *et al.*, *Earth Planet. Sci. Lett.* **394**, 179–185 (2014).
33. J. Lasue *et al.*, *Space Sci. Rev.* **174**, 155–212 (2013).
34. V. A. Krasnopolsky, M. J. Mumma, G. R. Gladstone, *Science* **280**, 1576–1580 (1998).
35. T. Usui, C. M. O. Alexander, J. Wang, J. I. Simon, J. H. Jones, *43rd Lunar and Planetary Science Conference*, The Woodlands, TX, 19–23 March 2012, 1341 (Lunar and Planetary Institute, Houston, TX, 2012).
36. H. Lammer, *Springer Briefs in Astronomy* **174**, 113–154 (2012).
37. D. E. Smith *et al.*, *J. Geophys. Res. Planets* **106** (E10), 23689 (2001).
38. M. H. Carr, J. W. Head, *J. Geophys. Res. Planets* **108**, 5042 (2003).
39. N. Z. Boctor, C. M. O. D. Alexander, J. Wang, E. Hauri, *Geochim. Cosmochim. Acta* **67**, 3971–3989 (2003).
40. J. P. Greenwood, S. Itoh, N. Sakamoto, E. P. Vicenzi, H. Yurimoto, *Geophys. Res. Lett.* **35**, L05203 (2008).
41. P. R. Mahaffy *et al.*, *Science* **347**, 412–414 (2015).
42. J. Mouginot, A. Pommerol, P. Beck, W. Kofman, S. M. Clifford, *Geophys. Res. Lett.* **39**, L02202 (2012).


## ACKNOWLEDGMENTS

We thank the staff of the Very Large Telescope (runs 83.C-0538 and 92.C-0436), the NASA InfraRed Telescope Facility, and the W. M. Keck Observatory for their exceptional support throughout our long Mars observing Programs. G.L.V. acknowledges support from NASA's Planetary Astronomy Program (08 PAST08 0034) and NASA's Planetary Atmospheres Program (08 PATM08 0031). NASA's Planetary Astronomy Program (RTOP 344 32 07) and NASA's Astrobiology Program (RTOP 344 53 51) supported M.J.M. and G.L.V. NSF Research in Undergraduate Institutions supported R.E.N. through grant AST 0805540. This work was also supported by a NASA Keck PI Data Award. The authors recognize and acknowledge the very important cultural role and reverence that the summit of MaunaKea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/218/suppl/DC1
Supplementary Text SM-1 to SM-3
Figs. S1 to S8
References (*43*–*48*)

24 November 2014; accepted 6 February 2015
Published online 5 March 2015;
10.1126/science.aaa3630

OCEANOGRAPHY

# Eddy-driven subduction exports particulate organic carbon from the spring bloom

Melissa M. Omand,[1]* Eric A. D'Asaro,[2] Craig M. Lee,[2] Mary Jane Perry,[3] Nathan Briggs,[3] Ivona Cetinić,[3] Amala Mahadevan[1]†

**The export of particulate organic carbon (POC) from the surface ocean to depth is traditionally ascribed to sinking. Here, we show that a dynamic eddying flow field subducts surface water with high concentrations of nonsinking POC. Autonomous observations made by gliders during the North Atlantic spring bloom reveal anomalous features at depths of 100 to 350 meters with elevated POC, chlorophyll, oxygen, and temperature-salinity characteristics of surface water. High-resolution modeling reveals that during the spring transition, intrusions of POC-rich surface water descend as coherent, 1- to 10-kilometer–scale filamentous features, often along the perimeter of eddies. Such a submesoscale eddy-driven flux of POC is unresolved in global carbon cycle models but can contribute as much as half of the total springtime export of POC from the highly productive subpolar oceans.**

The biological carbon pump (*1*), coupled with physical processes, makes the North Atlantic ocean a perennial and substantial sink of atmospheric $CO_2$ (*2*, *3*). Seasonally, strong winter-time cooling and winds enhance $CO_2$ solubility (*4*) and gas transfer, drive convective and turbulent circulations deepening the mixed layer, sequester carbon (*5*), and entrain nutrients into the mixed layer. With the onset of spring, increasing light, a shoaling mixed layer (*6*), enhanced stratification (*7*), and suppressed turbulent convection (*8*) lead to exceptionally high rates of phytoplankton production, resulting in a bloom. A critical piece of the carbon cycle is the ensuing export of organic carbon from the surface productive layer to the mesopelagic (100 to 1000 m). Traditionally, carbon-flux studies in the region have focused on export by sinking particles (*9*). Models have been used to infer the contribution of oceanic subduction to carbon export, most often described in terms of large-scale water-mass transformation and advection over seasonal time scales (*10*–*12*). Here, using observations complemented by a high-resolution physical-biological process model, we present evidence that export also occurs through localized subduction of small, nonsinking particulate carbon (Fig. 1A). The intensification of fronts within the eddying flow field, through convergence, frontogenesis, and wind-forcing, generates downward velocities (*13*, *14*) along sloping isopycnal surfaces that create subsurface intrusions of the surface water (*15*, *16*). The resulting eddy-driven subduction at scales of order 1 to 10 km leads to episodic injections of water from the surface mixed layer (*17*, *18*) (Fig. 1A) and net transport of particulate organic carbon (POC) to depth (*19*, *20*).

During the North Atlantic Bloom study of 2008 (NAB08), we carried out intensive observations of the onset (*7*) and progression of the spring bloom. The observations were made within the water-following reference frame provided by a Lagrangian mixed-layer float, sampled spatially by four autonomous Seaglider robots at distances of 1 to 50 km from the float. The gliders measured chlorophyll fluorescence (Chl F), optical backscatter (a proxy for POC), oxygen, salinity, temperature, and pressure. Analysis was performed on 774 glider profiles (to 1000 m) obtained over 25 days starting (at yearday 120; yearday 1 is 1 January 2008) a few days after the onset of the bloom. The backscatter and Chl F records were filtered to remove spikes caused by large, sinking aggregates (*21*), likely containing diatom spores (*22*), and the remaining signal was attributed to small POC with negligible sinking rates. The co-occurrence of POC and Chl F in ratios typical of phytoplankton suggests that a good fraction of these particulates were live phytoplankton cells.

In 9.2% of the analyzed glider profiles, we observed distinct subsurface maxima [see the supplementary materials, section S1.3 (SM S1.3)] in POC, Chl F, and oxygen co-occurring with anomalous temperature-salinity properties ("spice") (fig. S2) characteristic of the surface mixed layer (Fig. 1, B and C). These lateral intrusions of water (*27*) extending downward from the surface occurred at depths ranging from 100 to 350 m, well below the mixed layer and euphotic zone where phytoplankton can grow. In fact, 63 (of 71) subsurface maxima in POC collapse onto distinct potential density surfaces (isopycnals), suggesting that these are multiple observations of the same few intruding features (Fig. 1D). The dynamic height estimated from an objective map of glider-based depth-averaged currents (SM S1.2) from this period indicates that the platforms (float and gliders) were orbiting an anticyclonic eddy (Fig. 1E). Theory and modeling suggest that such regions of enhanced vorticity are preferred locations for subduction (*15*).

The subsurface intrusion of surface water with elevated POC in the density range $\sigma_t$ = 27.43 ± 0.01 kg $m^{-3}$ (purple segments, Fig. 1, C and D) was sampled multiple times over a span of 17 days (yeardays 128 to 145) (Fig. 1E). The data suggest that this isopycnal was in the mixed layer before yearday 130, after which it remained below the mixed layer and euphotic zone, isolated from light or air-sea gas exchange for at least 17 days, which is sufficiently long for remineralization to return the organic material to dissolved inorganic form. As the bloom developed, POC increased and oxygen became increasingly supersaturated in the surface mixed layer. The apparent oxygen utilization (AOU) calculated with reference to saturation at the surface was thereby negative (open circles, Fig. 2) in the mixed layer. Within the subsurface features, the oxygen and POC decreased (filled circles, Fig. 2) and AOU became positive. We interpret the increase in mixed-layer POC and oxygen as resulting from photosynthesis, and their subsurface depletion as resulting from respiration of organic carbon. The ratio between changes in the AOU (or oxygen) and POC is ~1.5 (Fig. 2), a respiratory quotient consistent with previous observations (*23*). Assuming negligible ventilation or photosynthesis, the rates of oxygen utilization imply a remineralization rate of 0.4 mmol $O_2$ $m^{-3}$ $d^{-1}$ between yeardays 130 and 145, consistent with previous estimates in the North Atlantic (*24*).

Eddy-driven subduction was mechanistically explored with a three-dimensional process-study ocean model (*25*) coupled to a simple phytoplankton model with a light-dependent growth rate (*26*) (SM S2). The domain is 480 km in meridional extent and 96 km east-west, with periodic boundaries in the zonal direction. The model is initialized in midwinter (yearday 30), with three fronts spanning the domain; forced with winds; and allowed to evolve for 4 months (*7*). By yearday 120 (the start of the observations presented here), the modeled phytoplankton bloom is well developed. The light-dependent growth of phytoplankton results in high concentrations of POC near the surface, particularly where stratification has set in. The model shows eddies with filaments of positive and negative vorticity (Fig. 3A) that correspond with horizontal gradients in buoyancy and enhanced POC (Fig. 3B). Downward-sloping, tongue-like intrusions of water rich in POC extend from the surface into the interior. Vertical profiles through these features (Fig. 3C) are remarkably similar to the glider profiles. Approximately 10% of the model profiles exhibit subsurface maxima, similar in prevalence to the observations. Projected onto a horizontal surface at 200 m, profiles with

[1]Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. [2]Applied Ocean Physics Laboratory, University of Washington, Seattle, WA 98195, USA. [3]Darling Marine Center, University of Maine, Walpole, ME 04573, USA. *Present address: Graduate School of Oceanography, University of Rhode Island, Narragansett, RI 02882, USA. †Corresponding author. E-mail: amala@whoi.edu

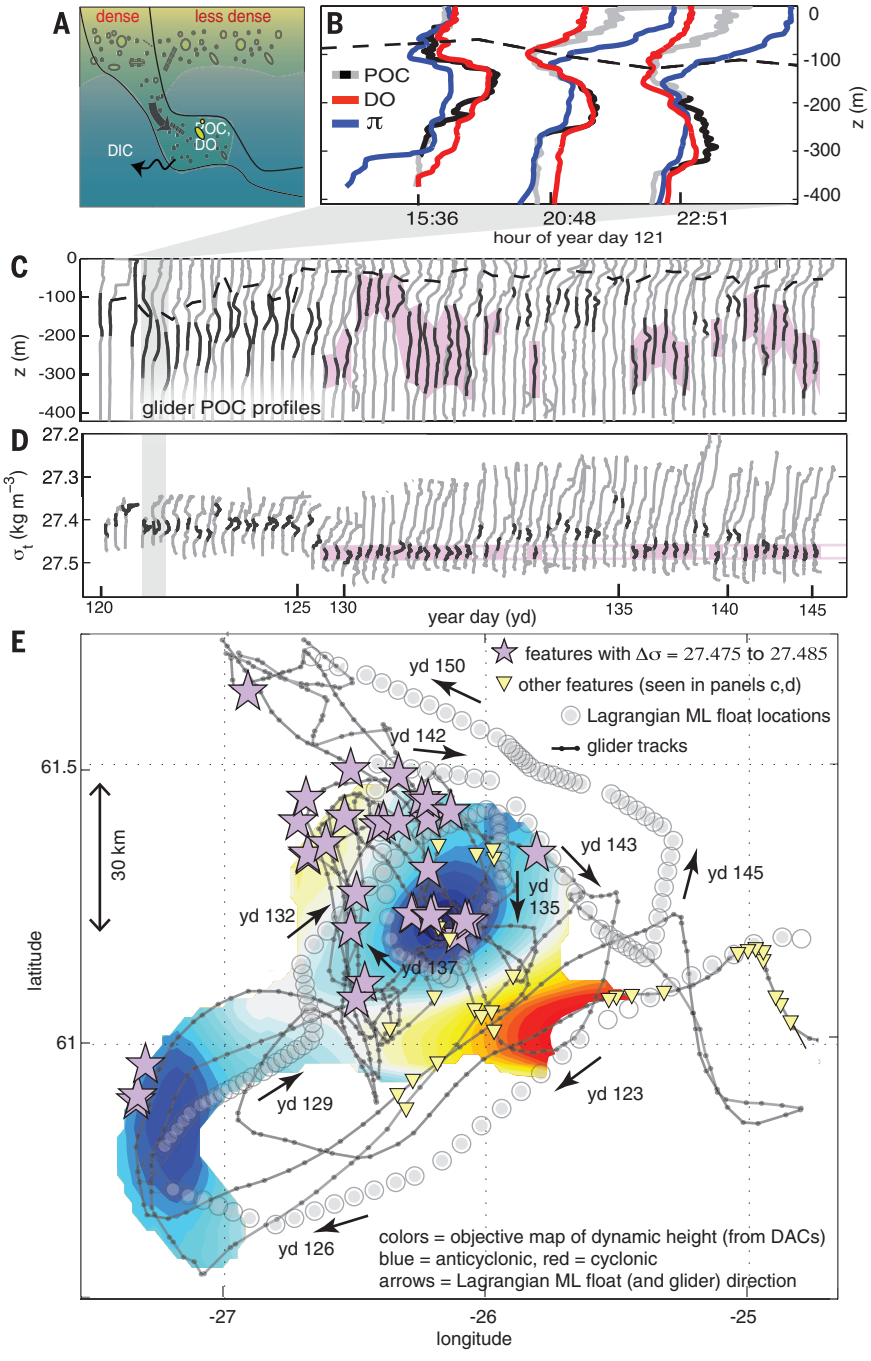



**Fig. 1. (A) Eddy-driven subduction transports POC- and oxygen-rich surface water along tilted isopycnals.** (**B**) Examples of Seaglider profiles of POC (gray, range 71 mg $m^{-3}$), dissolved oxygen (DO) (red, range 25 mmol $m^{-3}$), and spice ($\pi$, blue, range 0.17 kg $m^{-3}$) from yearday 121 show features with elevated concentrations below the mixed layer (depth defined by $\Delta\sigma_t$ = 0.03, dashed black line). (**C**) Thirty-day time series of 71 (of 772) POC profiles with distinct subsurface features versus depth. Black segments indicate the $\sigma_t$ span (0.005) of each feature. (**D**) POC profiles from (C) versus $\sigma_t$, with 33 of the features falling within the same water mass (defined by $\sigma_t$ = 27.48 ± 0.005 kg $m^{-3}$) indicated by purple shading. (**E**) Locations of the subsurface features (colored symbols) from (C) and (D) overlaid on the tracks of the Seagliders (gray lines) and mixed-layer float (open circles, locations shown every 3 hours). The approximate location and yearday of the mixed-layer float is indicated by arrows. Most features were observed in the vicinity of anticyclonic regions (the dynamic height is shown in colors, with anticyclonic streamlines indicated by blue colors and an overall range of 8 cm), based on objective mapping of the depth-averaged currents (DAC) from gliders.

deep local maxima in POC are clustered in coherent streaks and patches, most often in regions of subsurface negative vorticity (black points, Fig. 3, A and B). The probability distributions of depth-averaged currents, POC concentration, isopycnal depth, and isopycnal tilts are very similar in the model and observations (SM S2.1), giving us confidence that the model captures the temporal and spatial scales of both the physics and biology and can thus be used to estimate the POC flux due to eddy-driven subduction.

The intensification of fronts on the peripheries of mixed-layer eddies drives secondary ageostrophic circulations, generating vertical velocities of approximately 30 m $d^{-1}$ and ageostrophic cross-front flows of 3 to 5 cm $s^{-1}$ (*15*). This ageostrophic flow is largely along isopycnal surfaces and carries water from the surface layer, below and across the front, and delivers its contents to the stratifying interior (*19*). Similarly, the secondary circulations transport water from the base of the mixed layer to the surface. Because POC is produced in the sunlit surface layers as the mixed layer is stratifying, there is a strong vertical gradient in POC. This results in a net downward flux of POC.

The vertical flux of POC in the model is the covariance between POC anomalies, defined as $c' = POC - \langle POC \rangle$, and vertical velocity $w'$, calculated as $Flux = \langle w'c' \rangle$, where angle brackets denote the horizontal average at each depth (shown at 100 m in Fig. 4A). On average, a negative covariance (subduction of POC) results from downward transport of POC-rich water

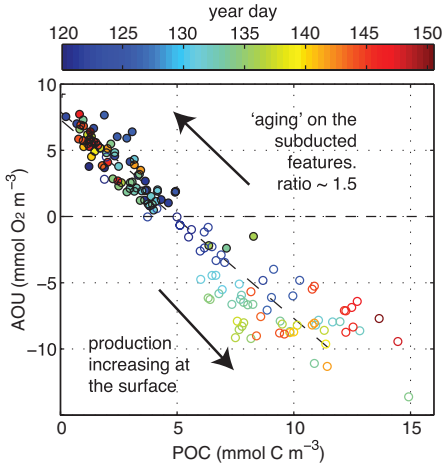



**Fig. 2. AOU versus POC averaged within the subducted features (solid circles) and over the mixed layer (open circles), colored by yearday.** The AOU increases as the POC decreases on the subducted features over time, indicative of respiration and diverging from simultaneously measured surface values. The best fit to the filled circles (dashed black line) has a slope of 1.5, consistent with a respiratory quotient arising from bacterial metabolization. AOU is defined as the difference between the saturation oxygen (at measured temperature, salinity, and surface pressure) and the measured oxygen.

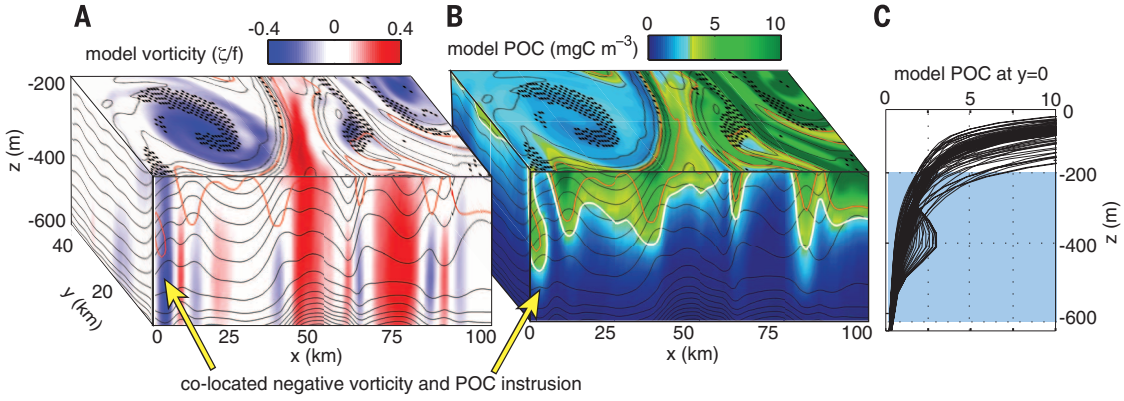



**Fig. 3. A subdomain of modeled (A) relative vorticity (ζ) normalized by planetary vorticity (*f*) and (B) POC from 200 to 600 m depth on yearday 125.** The *x*-*y* locations of subsurface features identified with the model are indicated by black points in the *z* = 200 m plane. Black lines indicate the potential density ($\sigma_t$) contours. The domain is selected to intersect a region at the periphery of an anticyclone, containing high POC and deep negative vorticity consistent with frontogenetic subduction. (**C**) Examples of 96 model POC (mgC m$^{-3}$) profiles at *y* = 0 show that roughly 10% contain subsurface features of elevated POC similar to the observations with Seagliders. The depth range shown in (A) and (B) is indicated by blue shading.

and upward transport of water relatively devoid of POC. Our model shows that $w'$, $c'$, and their covariance are coherent in filaments that are just a few kilometers in width (fig. S10) and deliver water to depths approaching the winter mixed layer. Model diagnostics suggest that only a small fraction of the subducted water is reentrained into the mixed layer, and this occurs mostly within the first 2 days. More than 75% of the water remains subducted until the end of the model simulation (a further 10 days), suggesting that it would remain below the mixed layer over the course of the summer (SM S2.2).

We parameterize the eddy-driven flux of POC $\langle w'c' \rangle$ in terms of the eddy-driven vertical flux of buoyancy $\langle w'b' \rangle = \psi_e M^2$ (*28*), where $\psi_e$ is the overturning stream function arising from mixed-layer eddies and $M^2 = |\nabla_H b|$ is the characteristic horizontal gradient of buoyancy $b \equiv -g\rho_0^{-1}(\rho-\rho_0)$ in the mixed layer. We use an existing parameterization, $\psi_e \sim C_e M^2 H^2 f^{-1}$ (*29*), where $H$ is the mixed layer depth, $f$ is the Coriolis parameter, and $C_e = 0.08$ is a scaling constant. The vertical POC flux must account for the fact that isosurfaces of POC are not aligned with the isopycnal surfaces (isosurfaces of $b$). Whereas the isopyncals are sloping, the POC gradient is largely vertical. Thus, $\langle w'c' \rangle \sim \langle w'b' \rangle \langle \frac{\partial POC}{\partial z} \rangle \langle b_z \rangle^{-1}$. We evaluate $\langle \frac{\partial POC}{\partial z} \rangle$ as $[POC]/H^*$, where [POC] is the spatial average of surface POC, and $H^*$ is the depth of the photosynthetically productive surface layer taken to be $H$ in the North Atlantic but set equal to the euphotic layer when it is deeper than $H$ (SM S3.4.1). The resulting scaling estimate for the vertical flux of POC due to eddy-driven subduction (see SM S3.3 for details) is therefore

$$Flux = \langle w'c' \rangle \sim \psi_e \Gamma [POC]/H^* \qquad (1)$$

where $\Gamma = M^2/N^2 = |\nabla b|/b_z$ represents the mean slope of the isopycnals in the restratifying region.

Spatially integrated, the export of POC by eddy-driven subduction is time dependent and varies with mixed-layer depth, vertical and lateral buoyancy gradients, and the surface concentration of POC. Using domain-averaged values of $M^2$, $N^2$, $H$, and surface POC over the upper 50 m of our model, we find that our scaling estimate (Eq. 1) captures the mean (Fig. 4A) and 88% of the variability in the domain-averaged flux at 100-m depth, calculated directly from the model as $F_{100} = \langle w'c' \rangle$ (fig. S14, B and C). The maximal flux occurs when the mixed layer is stratifying but is still deep in places, stratification is patchy, lateral buoyancy gradients exist, and the bloom has begun—i.e., [POC] is high. In our model, we observe that the POC flux $F_{100}$ increases from yearday 120 as the upper ocean becomes stratified and the bloom grows (Fig. 4A), attaining values of 170 mgC m$^{-2}$ d$^{-1}$. Later, although the near-surface POC continues increasing (fig. S13A), the POC flux due to subduction is restricted to a much shallower region as the mixed layer shoals and stratification intensifies, leading to a reduction in $F_{100}$ to roughly 50 mgC m$^{-2}$ d$^{-1}$. Eddy-induced fluxes act mainly along sloping isopycnal surfaces (*30*), and eddy-driven subduction of POC is most effective when a phytoplankton bloom occurs in conjunction with eddy-driven stratification (*7*).

Comparing our estimate of the eddy-driven POC flux with the sinking export of POC is not straightforward, because observations of sinking flux events are usually local and not integrated in space and time. Concurrent estimates from NAB08 based on nitrate, POC, and dissolved oxygen indicated that the total carbon export peaked at 984 mgC m$^{-2}$ d$^{-1}$ (*23*). Export diagnosed from $^{234}$Th disequilibria over a similar period (in May) was up to 500 mgC m$^{-2}$ d$^{-1}$ (*31*). During the short-lived diatom bloom observed within a patch of productive water, sinking export of large particulates was estimated as 514 mgC m$^{-2}$ d$^{-1}$ (*21*) at a depth of 100 m. During the Joint Global Ocean Flux Study North Atlantic Bloom Experiment campaign, the export flux of sinking particles peaked at 492 mgC m$^{-2}$ d$^{-1}$ in early May (*32*). In comparison with event-based estimates of sinking export (*21*, *32*), eddy-driven subduction of POC accounts for about 25% of the total export flux.

To arrive at an integrated estimate of the eddy-driven POC subduction globally in spring, we evaluate our scaling (Eq. 1) with the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) satellite–derived surface POC and the Monthly Isopycnal and Mixed-Layer Ocean Climatology, from which we derive mixed-layer depth $H$, vertical and lateral buoyancy gradients $N^2$ and $M^2$ during March through May for the Northern Hemisphere, and September through November for the Southern Hemisphere, respectively (SM S3.4 and fig. S15). The largest fluxes occur in regions with seasonally varying mixed layers (Fig. 4B). We compare this estimate to the total export, which includes the sinking flux of POC estimated from satellite data with a model (*33*), and our estimate of eddy-driven subduction. During the spring, our estimate suggests that eddy-driven subduction of POC contributes as much as half the total export of POC in the North Atlantic, the Kuroshio extension, and the Southern Ocean (Fig. 4C). The covariation of high POC, lateral buoyancy gradients, and relatively deep winter mixed layers provides the conditions for mixed-layer eddies, which during the early spring bloom drive a subduction of POC and oxygen, as evidenced from our observations and modeling. Such a mechanism for the export of POC, dissolved organic matter, and oxygen is unresolved in global climate models.

Previous estimates of POC export have typically fallen short of closing the budget for carbon in the mesopelagic ocean (*34*, *35*). Disaggregation of large POC by zooplankton (*36*) and the detrainment of POC-rich water from the mixed layer by seasonal stratification via the so-called mixed-layer pump (*5*) are some of the mechanisms that have been invoked to reconcile this budget and explain the export of small particles (0.2 to 20 μm) to depths up to 1000 m in the Norwegian Sea (*37*). The high spatiotemporal resolution of our observations and modeling points, however, to an alternate pathway for carbon export. Unlike the mixed-layer pump, this mechanism transports water from the most productive surface layers as they experience eddy-driven stratification and rapid growth of phytoplankton due to enhanced light exposure. The fate of the subducted carbon, and its potential to be sequestered for long periods, depends

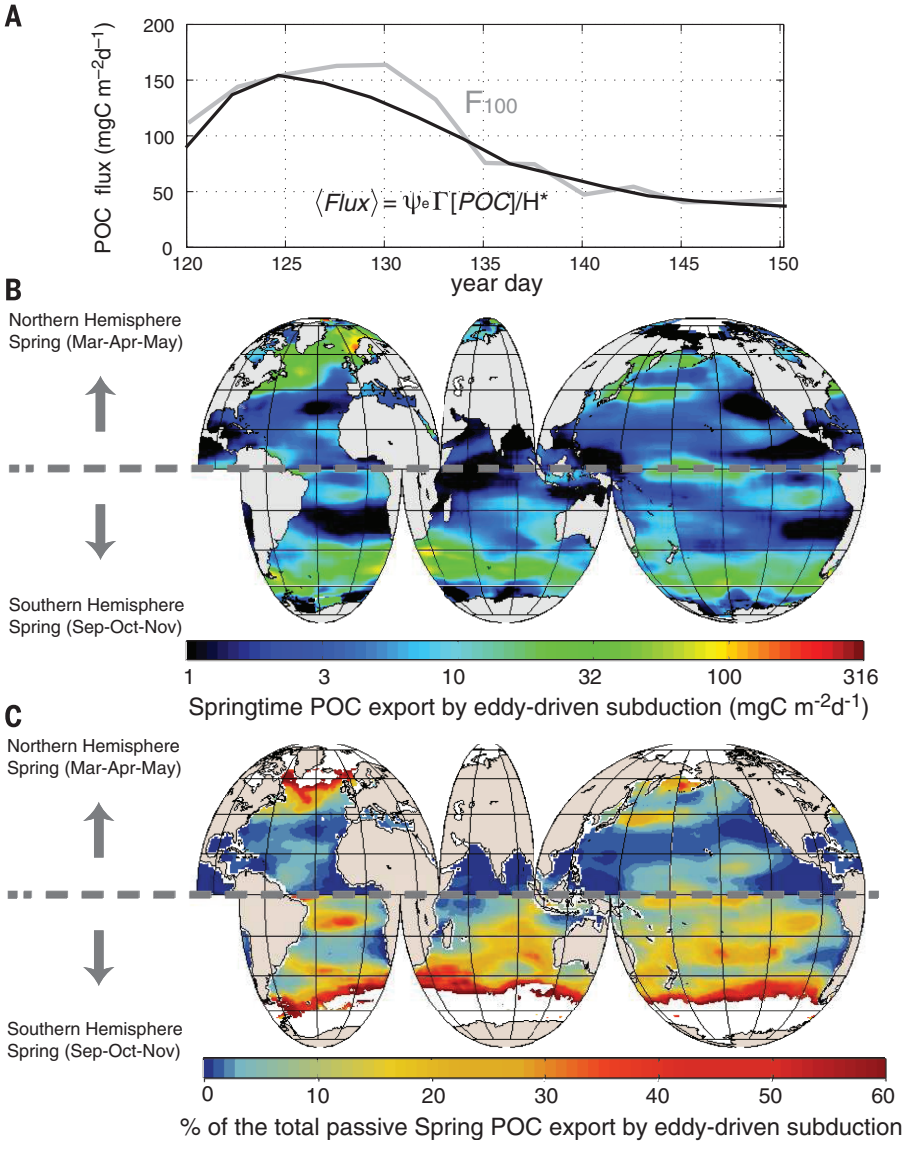



**Fig. 4. Model-based estimate of eddy-driven subduction of POC.** (**A**) The model-derived flux of POC $F_{100} = \langle w'POC' \rangle$ at 100-m depth is compared with a model-based scaling estimate $Flux = \psi_e \Gamma \frac{[POC]}{H^*}$. The two are correlated with $r^2 = 0.88$. (**B**) A global estimate of eddy-driven subduction of POC. (**C**) The contribution of eddy-driven subduction [shown in (B)] as a percentage of the total export of POC. The total POC export is defined as the sinking export (*33*) plus the eddy-driven export (fig. S18).

on several factors, including lateral advection by the large-scale circulation.

In the ocean, large cells and aggregates of organic matter are exported efficiently by sinking. But the total POC biomass is roughly evenly distributed among all size classes due to the relatively large numbers of small cells and particles that constitute a large fraction of the net primary production (*38*). Eddy-driven subduction provides an export pathway for the small size classes of POC and likely also dissolved organic carbon. (DOC) Such a subduction of small cells substantially augments the POC export arising from the sinking of large particles. Small particles and DOC provide an important source of organic carbon to the microbial food webs (*36*), and because the subduction is spatially heterogeneous, we anticipate that it stimulates hot spots of microbial activity that have implications for mesopelagic ecology.

## REFERENCES AND NOTES


1. R. Sanders *et al.*, *Prog. Oceanogr.* **129**, 200–218 (2014).
2. A. Körtzinger *et al.*, *J. Geophys. Res.* **113** (C4), C04020 (2008).
3. C. L. Sabine *et al.*, *Science* **305**, 367–371 (2004).
4. M. J. Follows, R. G. Williams, J. C. Marshall, *J. Mar. Res.* **54**, 605–630 (1996).
5. W. D. Gardner, S. P. Chung, M. J. Richardson, I. D. Walsh, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **42**, 757–775 (1995).
6. H. U. Sverdrup, *J. Cons. Int. Explor. Mer* **18**, 287–295 (1953).
7. A. Mahadevan, E. D'Asaro, C. Lee, M. J. Perry, *Science* **337**, 54–58 (2012).
8. J. Taylor, R. Ferrari, *Limnol. Oceanogr.* **56**, 2293–2307 (2011).
9. H. W. Ducklow, R. P. Harris, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **40**, 1–8 (1993).
10. J. B. Sallée, R. J. Matear, S. R. Rintoul, A. Lenton, *Nat. Geosci.* **5**, 579–584 (2012).
11. J. A. Barth, *J. Geophys. Res.* **107** (C6), 3057 (2002).
12. P. Karleskind, M. Lévy, L. Memery, *J. Geophys. Res.* **116** (C2), C02025 (2011).
13. A. Mahadevan, *Ocean Model.* **14**, 222–240 (2006).
14. A. Mahadevan, A. Tandon, *Ocean Model.* **14**, 241–256 (2006).
15. M. A. Spall, *J. Geophys. Res.* **100** (C2), 2543–2557 (1995).
16. M. Lee, A. Nurser, *J. Phys. Oceanogr.* **42**, 1762–1780 (2012).
17. J. C. McWilliams, F. Colas, M. J. Molemaker, *Geophys. Res. Lett.* **36**, L18602 (2009).
18. M. Lévy, P. Klein, A. M. Treguier, *J. Mar. Res.* **59**, 535–565 (2001).
19. R. Pollard, L. Regier, *Nature* **348**, 227–229 (1990).
20. S. Fielding *et al.*, *J. Mar. Syst.* **30**, 287–304 (2001).
21. N. Briggs *et al.*, *Deep Sea Res. Part I Oceanogr. Res. Pap.* **58**, 1031–1039 (2011).
22. T. Rynearson *et al.*, *Deep Sea Res. Part I Oceanogr. Res. Pap.* **82**, 60–71 (2013).
23. M. B. Alkire *et al.*, *Deep Sea Res. Part I Oceanogr. Res. Pap.* **64**, 157–174 (2012).
24. J. B. Palter, M. S. Lozier, R. T. Barber, *Nature* **437**, 687–692 (2005).
25. A. Mahadevan, J. Oliger, R. Street, *J. Phys. Oceanogr.* **26**, 1868–1880 (1996).
26. W. Bagniewski, K. Fennel, M. J. Perry, E. A. D'Asaro, *Biogeosciences* **8**, 1291–1307 (2011).
27. T. Johnston, O. Cheriton, J. T. Pennington, F. P. Chavez, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **56**, 246–259 (2009).
28. I. Held, T. Schneider, *J. Atmos. Sci.* **56**, 1688–1697 (1999).
29. B. Fox-Kemper, R. Ferrari, R. Hallberg, *J. Phys. Oceanogr.* **38**, 1145–1165 (2008).
30. A. Mahadevan, A. Tandon, R. Ferrari, *J. Geophys. Res.* **115** (C3), C03017 (2010).
31. P. Martin *et al.*, *Deep Sea Res. Part I Oceanogr. Res. Pap.* **58**, 338–349 (2011).
32. K. O. Buesseler, M. P. Bacon, J. K. Cochran, H. D. Livingston, *Deep-Sea Res. A, Oceanogr. Res. Pap.* **39**, 1115–1137 (1992).
33. D. A. Siegel *et al.*, *Global Biogeochem. Cycles* **28**, 181–196 (2014).
34. D. Steinberg *et al.*, *Limnol. Oceanogr.* **53**, 1327–1338 (2008).
35. A. Burd *et al.*, *Deep Sea Res. Part II Top. Stud. Oceanogr.* **57**, 1557–1571 (2010).
36. S. L. Giering *et al.*, *Nature* **507**, 480–483 (2014).
37. G. Dall'Olmo, K. A. Mork, *Geophys. Res. Lett.* **41**, 2921–2927 (2014).
38. T. L. Richardson, G. A. Jackson, *Science* **315**, 838–840 (2007).


## ACKNOWLEDGMENTS

M.M.O. and A.M. were supported by the National Science Foundation (NSF OCE 1260080, 1434788), the Office of Naval Research, and a fellowship to A.M. from the Radcliffe Institute for Advanced Study. The NAB08 field experiment was supported by NSF OCE-0628107 and OCE-0628379 and NASA NNX-08AL92G. Data from NAB08 is publicly available at the Biological and Chemical Oceanography Data Management Office: http://osprey.bcodmo.org/project.cfm?flag=view&id=102&sortby=project.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/222/suppl/DC1
Materials and Methods
Figs. S1 to S18
Table S1
References (*39*–*51*)

15 August 2014; accepted 4 March 2015
Published online 26 March 2015;
10.1126/science.1260062

PALEOCEANOGRAPHY

# Middle Miocene closure of the Central American Seaway

C. Montes,[1]* A. Cardona,[2] C. Jaramillo,[3] A. Pardo,[4] J. C. Silva,[5] V. Valencia,[6] C. Ayala,[7] L. C. Pérez-Angel,[1] L. A. Rodriguez-Parra,[1] V. Ramirez,[8] H. Niño[8]

**Uranium-lead geochronology in detrital zircons and provenance analyses in eight boreholes and two surface stratigraphic sections in the northern Andes provide insight into the time of closure of the Central American Seaway. The timing of this closure has been correlated with Plio-Pleistocene global oceanographic, atmospheric, and biotic events. We found that a uniquely Panamanian Eocene detrital zircon fingerprint is pronounced in middle Miocene fluvial and shallow marine strata cropping out in the northern Andes but is absent in underlying lower Miocene and Oligocene strata. We contend that this fingerprint demonstrates a fluvial connection, and therefore the absence of an intervening seaway, between the Panama arc and South America in middle Miocene times; the Central American Seaway had vanished by that time.**

Closure of the Central American Seaway, defined here as the deep oceanic seaway along the tectonic boundary of the South American plate and the Panama arc, is thought to have modified the salinity of the Caribbean Sea, ultimately affecting ocean circulation patterns and global climate (*1*), as well as to have triggered the Great American Biotic Interchange (*2*). However, the role of the formation of the Panamanian Isthmus in such global changes remains controversial, in part because of the difficulty of establishing a precise chronology of seaway closure (*3*). Data on the chronology of Isthmus emergence suggests that the closure not only occurred earlier than previously thought (*4*) but also may have resulted from factors other than the emergence of currently high terrain in Panama (*5*, *6*).

The Uramita suture (*7*) separates the young Panama arc to the west from the old Andean terranes to the east (Fig. 1). These are mutually exclusive geochronological domains that are ideally suited for documenting the time of detrital exchange. The young Panama magmatic arc was built on an oceanic plateau substrate (*8*) during latest Cretaceous to Eocene times [67 to 39 million years ago (Ma), with a peak around 50 Ma] (*6*, *9*), with renewed magmatic activity as young as 19 Ma (*10*) east of the Canal Basin and 10 Ma and younger west of it (*11*). To further characterize the Panama magmatic arc fingerprint, we dated a string of incompletely mapped granitic plutons along the northeastern coast of Panama and western Colombia (Fig. 1C), obtaining eight U/Pb magmatic zircon ages ranging between 59 and 42 Ma (217 U/Pb analyses; table S1). The northern Andes, in contrast, include magmatic rocks accreted during latest Cretaceous times (*8*) to a core of plutonometamorphic rock of late Precambrian (*12*–*14*) and Permo-Triassic age (*14*, *15*), and plutonic rocks of Jurassic to Cretaceous age (*16*). Middle Eocene magmatism is

[1]Universidad de los Andes, Bogotá, Colombia. [2]Universidad Nacional de Colombia, Medellín, Colombia. [3]Smithsonian Tropical Research Institute, Ciudad de Panamá, Panamá. [4]Universidad de Caldas, Manizales, Colombia. [5]University of Houston, Houston, TX 77004, USA. [6]Washington State University, Pullman, WA 99164, USA. [7]Corporación Geológica Ares, Bogotá, Colombia. [8]Ecopetrol, Bogotá, Colombia.
*Corresponding author. E-mail: cmontes@uniandes.edu.co

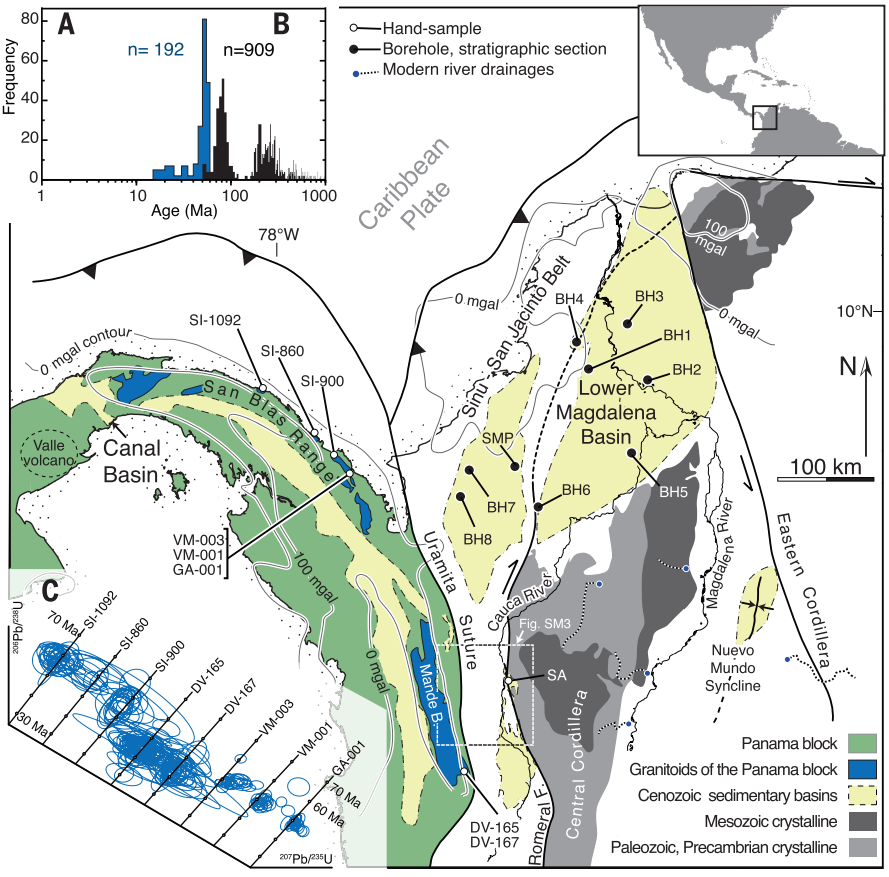



**Fig. 1. Tectonic setting of the study area and location of samples.** See Table 1. Thick lines represent major boundaries (*28*). Zero-milligal (0 mgal) contour (*29*) highlights the geodynamic continuity of the Panama arc; there are no structural breaks between the Uramita suture and the Canal Basin. (**A** and **B**) Detrital zircon ages recovered from (A) lower Miocene strata in the Canal Basin (*6*) and (B) Oligocene-Miocene strata in the Nuevo Mundo Syncline (*18*) and rivers draining the Eastern and Central cordilleras (*19*). (**C**) New U/Pb zircon ages for granitoids of the Panama arc; data point error ellipses are 68.3% confidence (see table S1).

**Table 1. Sample summary.** See fig. S1 for stratigraphic location of samples.

| Sample/ locality name | Detrital zircon sample name | Latitude (decimal degrees) | Longitude (decimal degrees) | Sample type | Age |
|---|---|---|---|---|---|
| *U/Pb in detrital zircons from boreholes* | | | | | |
| BH1 | SI-5796 | 9.4699 | –75.0015 | Cored borehole | Middle Miocene |
| | SI-6816 | 9.4699 | –75.0015 | Cored borehole | Oligocene–early Miocene |
| BH2 | SI-6635 | 9.3651 | –74.4474 | Drill cuttings | Oligocene–early Miocene |
| BH3 | SI-5900 | 9.8907 | –74.6344 | Drill cuttings | Middle Miocene |
| BH4 | SI-6878 | 9.7160 | –75.1124 | Cored borehole | Middle Miocene |
| BH5 | SI-6703 | 8.6716 | –74.5976 | Cored borehole | Oligocene–early Miocene |
| BH6 | SI-6657 | 8.1702 | –75.4761 | Drill cuttings | Middle Miocene |
| | SI-6672 | 8.1702 | –75.4761 | Drill cuttings | Oligocene–early Miocene |
| BH7 | SI-7045 | 8.5168 | –76.1256 | Cored borehole | Middle Miocene |
| BH8 | SI-6833 | 8.2680 | –76.2033 | Drill cuttings | Oligocene–early Miocene |
| *U/Pb in detrital zircons from surface stratigraphic sections* | | | | | |
| SMP | SI-6997 | 8.5488 | –75.6922 | Hand sample | Oligocene–early Miocene |
| SA | 030337-2 | 6.4513 | –75.7403 | Hand sample | Oligocene–early Miocene |
| | 030337-6 | 6.4513 | –75.7403 | Hand sample | Oligocene–early Miocene |
| | 030339 | 6.5414 | –75.7754 | Hand sample | Oligocene–early Miocene |
| | 030357 | 6.4085 | –75.7241 | Hand sample | Middle Miocene |
| | 030359 | 6.4166 | –75.7255 | Hand sample | Middle Miocene |
| | 030362 | 6.5064 | –75.8194 | Hand sample | Oligocene–early Miocene |
| *U/Pb in magmatic zircons from granitoid surface samples* | | | | | |
| DV-165 | | 5.7680 | –76.2490 | Hand sample | 43.8 ± 0.8 Ma |
| DV-167 | | 5.7709 | –76.2475 | Hand sample | 42.5 ± 1.3 Ma |
| GA-001 | | 8.5094 | –77.3371 | Hand sample | 49.5 ± 0.9 Ma |
| VM-001 | | 8.4993 | –77.3221 | Hand sample | 49.7 ± 1.6 Ma |
| VM-003 | | 8.4847 | –77.3402 | Hand sample | 49.5 ± 1.1 Ma |
| SI-860 | | 8.8370 | –77.6218 | Hand sample | 59.0 ± 1.9 Ma |
| SI-900 | | 8.6638 | –77.4321 | Hand sample | 58.6 ± 1.6 Ma |
| SI-1092 | | 9.2985 | –78.1047 | Hand sample | 58.3 ± 1.0 Ma |

absent in the northern Andes (*17*–*19*). Therefore, detrital zircons of Eocene age can be used to track the detrital contribution of the Panama arc to sedimentary basins of northwestern South America.

To track detrital contributions from the Panama arc, we sampled fluvial strata at the western flank of the Central Cordillera of Colombia (site SA in Fig. 1), following a roughly northeastern trend toward shallow marine strata (sites SMP and BH1 to BH8) in the Lower Magdalena Basin. Seven stratigraphic levels of middle Miocene age and 11 stratigraphic levels of Oligocene to early Miocene age were sampled in eight boreholes and two surface stratigraphic sections (Table 1). We obtained 18 U/Pb detrital zircon ages, as well as petrographic and heavy mineral analyses of the sedimentary rock (1654 U/Pb analyses; tables S2 to S5). All detrital zircon samples recovered from Oligocene and Miocene strata contained typical north Andean detrital signatures that included late Precambrian, Permo-Triassic, and Late Cretaceous populations. Middle Miocene strata, however, contained an additional Eocene magmatic zircon population that was absent in older strata (Fig. 2).

To establish the age of strata sampled in the Lower Magdalena Basin, we used published foraminiferal and palynological studies performed on the same boreholes where we sampled, further bracketed by our detrital zircon minimal ages (Fig. 2). In the western flank of the Central Cordillera, we relied on mapped cross-cutting relationships (*20*) of volcanic and subvolcanic rocks interbedded with and intruding fluvial, coal-bearing strata. Available geochronology (*21*) and palynological (*22*) studies along with our detrital zircon minimal ages were used to establish ages of the strata sampled (fig. S3).

Because the geochronological makeup of the northern Andes is incompletely known, we used published U/Pb detrital zircon data as proxies for the magmatic age distribution of the eastern Panama arc and the northern Andes. We used the Oligocene-Miocene strata of the Canal Basin as a proxy for the Panama arc (Fig. 1A) (*6*). For the northern Andes (Fig. 1B), we used two data sets as proxies of its geochronological makeup: Oligocene-Miocene strata of the Nuevo Mundo Syncline (*18*) and active-sediment river samples draining the Eastern and Central Cordilleras (*19*). We found that detrital zircons from basins and rivers in the northern Andes are decidedly older than those from Panama; the mean for Nuevo Mundo Syncline and active-sediment river samples is 304.1 Ma, whereas the mean for Panama is 49.7 Ma (Kolmogorov-Smirnov test, $P < 0.001$, $D = 0.98$; fig. S2, A and B). Nuevo Mundo Syncline and active-sediment river samples have an age range of 51.2 to 2675.4 Ma, with only 16 ages between 51.2 and 63.1 Ma (mean 56 Ma). Panama ages range from 17.6 to 65.1 Ma (fig. S2B).

Oligocene to middle Miocene strata sampled in the northern Andes can be separated into two age groups according to their detrital zircon populations: one containing an Eocene population and another missing it (Fig. 2). The Oligocene–early Miocene strata show an age range of 54 Ma to 3103.6 Ma. Only two of 1045 zircons have ages younger than 65.1 Ma (54 and 64 Ma). In contrast, middle Miocene sandy strata in the same sampling sites (Fig. 2A) show an age range from 13.1 Ma to 3189.9 Ma. A large number of them (103 of 609) have ages younger than 65.1 Ma, with a mean age of 36.8 Ma, slightly younger than the mean age of Panamanian detrital zircons (mean Panama = 49.7 Ma, Kolmogorov-Smirnov test, $P < 0.001$, $D = 0.53$; fig. S2).

The Eocene detrital zircon population documented in middle Miocene strata of northwestern South America (Fig. 2A) could have only been derived from the emerged Panama arc, as there are no igneous bodies of that age in the northern Andes (*18*, *19*). The magmatic roots of the Panama arc had been cooling (*5*, *6*), emerging, and eroding (*23*) since at least late Eocene times (*6*); therefore, they were available as source areas by middle Miocene times. Both fluvial coal-bearing strata (*20*, *24*) and shallow marine strata of middle Miocene age contain the Panama arc

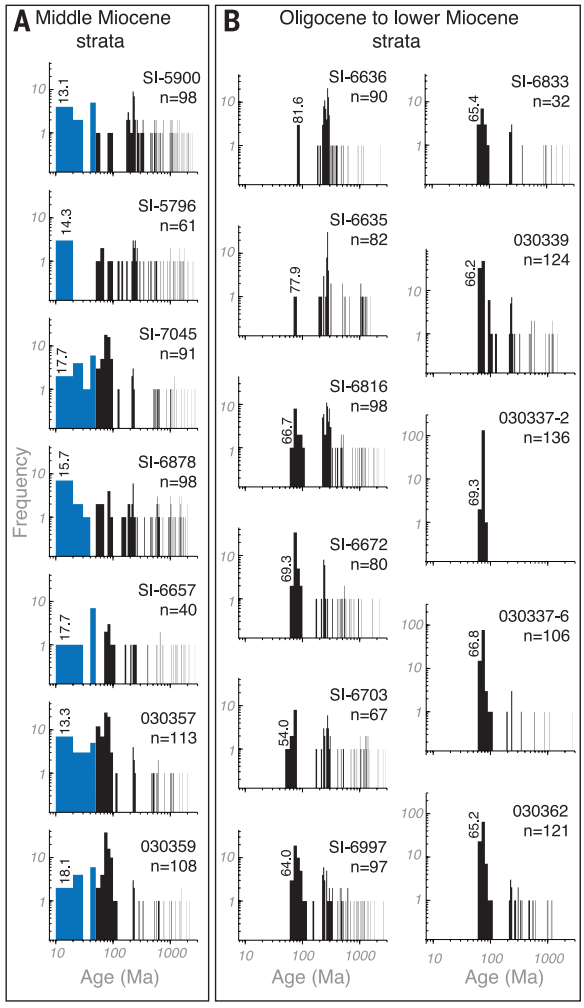



**Fig. 2. Detrital U/Pb zircon populations for all the samples recovered.** (**A**) Middle Miocene strata; (**B**) Lower Miocene and Oligocene strata of the northern Andes. Data are binned in sections of 10 million years (see tables S2 and S3). The youngest detrital zircon is labeled on top of the leftmost bin. See Table 1 and Fig. 1 for sample location.

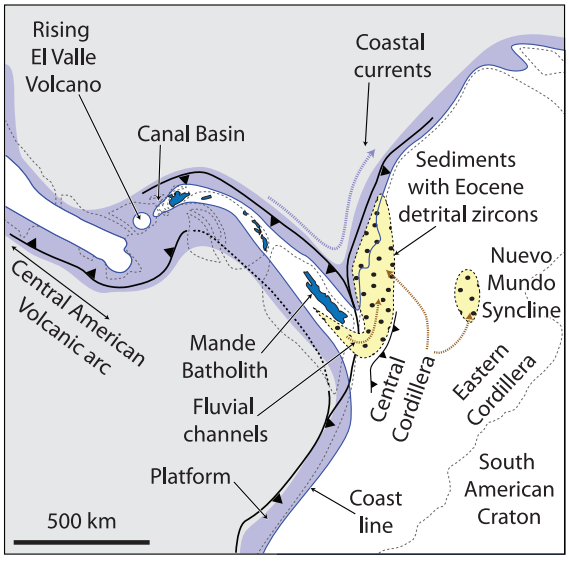



**Fig. 3. Paleogeographic reconstruction (*28*) of the Panama arc and northwestern South America during middle Miocene times (13 to 15 Ma).** The first detrital loads from Panama arrived in one of two, or both, paths to the basins of northwestern South America: (i) along coastal currents transporting detritus product of the erosion of exposed plutonic rocks along the northern coast of the Panama arc and/or (ii) along fluvial channels draining emerging ranges parallel to the length of the Isthmus. The El Valle volcano, an edifice rising from sea level starting before 10 Ma (*11*), would only allow shallow and transient seaways between 15 and 10 Ma, as the Canal basin was connected to North America by a land bridge from Oligocene until middle Miocene times (*30*).

signature (Fig. 2). This signature—in fluvial strata to the south and in shallow marine strata to the north—suggests that the Panama arc had docked and emerged and was shedding detrital material to north-bound currents parallel to the Uramita Suture, similar to today's Cauca River (Fig. 1), and to northeast-bound coastal currents (Fig. 3).

Our results imply that by middle Miocene times (13 to 15 Ma), rivers originating in the Panama arc were transporting sediment to the shallow marine basins of northern South America. This implies that (i) at least a segment of the Panama arc, including an emerged (*6*) Mande batholith and San Blas Range (Figs. 1 and 3), had already docked and (ii) the Central American Seaway was closed. Continued Caribbean-Pacific water exchange may have taken place along narrow, shallow, and transient channels that fragmented (*5*) the Isthmus west of the Canal Basin (*4*) (Fig. 3). These results support recent paleoceanographic studies (*25*, *26*) that show a decrease in the transport of deep and intermediate Pacific waters into the Caribbean by 10 to 11 Ma, probably related to a closing Central American Seaway. Recorded changes in Caribbean water salinity at ~4.2 Ma (*1*), and a delay of nearly 10 Ma in the Great American Biotic Interchange (*2*) after the first detrital loads crossed the Isthmus, could be unrelated to seaway closure and instead may be linked to Plio-Pleistocene global climatic transitions (*3*, *27*).

## REFERENCES AND NOTES


1. G. H. Haug, R. Tiedemann, R. Zahn, A. C. Ravelo, *Geology* **29**, 207–210 (2001).
2. L. G. Marshall, S. D. Webb, J. J. Sepkoski Jr., D. M. Raup, *Science* **215**, 1351–1357 (1982).
3. P. Molnar, *Paleoceanography* **23**, 15 (2008).
4. A. G. Coates, L. S. Collins, M.-P. Aubry, W. A. Berggren, *Geol. Soc. Am. Bull.* **116**, 1327–1344 (2004).
5. D. W. Farris *et al.*, *Geology* **39**, 1007–1010 (2011).
6. C. Montes *et al.*, *Geol. Soc. Am. Bull.* **124**, 780–799 (2012).
7. H. Duque-Caro, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **77**, 203–234 (1990).
8. A. C. Kerr, J. Tarney, *Geology* **33**, 269–272 (2005).
9. W. Wegner, G. Wörner, M. E. Harmon, B. R. Jicha, *Geol. Soc. Am. Bull.* **123**, 703–724 (2011).
10. S. Whattam *et al.*, *Lithos* **142–143**, 226–244 (2012).
11. P. Hidalgo, T. Vogel, T. Rooney, R. Currier, P. Layer, *Contrib. Mineral. Petrol.* **162**, 1115–1138 (2011).
12. P. A. Restrepo-Pace, J. Ruiz, G. Gehrels, M. Cosca, *Earth Planet. Sci. Lett.* **150**, 427–441 (1997).
13. U. G. Cordani, A. Cardona, D. M. Jimenez, D. Liu, A. P. Nutman, *Geol. Soc. London Spec. Publ. 246*, 329–346 (2005).
14. C. J. Vinasco, U. G. Cordani, H. Gonzalez, M. Weber, C. Pelaez, *J. S. Am. Earth Sci.* **21**, 355–371 (2006).
15. A. Cardona *et al.*, *J. S. Am. Earth Sci.* **29**, 772–783 (2010).
16. J. A. Aspden, W. J. McCourt, M. Brook, *J. Geol. Soc. London* **144**, 893–905 (1987).
17. G. Bayona *et al.*, *Earth Planet. Sci. Lett.* **331–332**, 97–111 (2012).
18. J. Nie *et al.*, *Geology* **38**, 451–454 (2010).
19. J. Nie *et al.*, *Earth Sci. Rev.* **110**, 111–126 (2012).
20. E. Grosse, *Estudio Geológico del Terciario Carbonífero de Antioquia en la parte occidental de la Cordillera Central de Colombia entre el rio Arma y Sacaojal* [Dietrich Reimer (Ernst Vohsen), Berlin, 1926].
21. G. Rodriguez, G. Zapata, *Bol. Geol.* **36**, 85–102 (2014).
22. T. Van der Hammen, *Bol. Geol.* **6**, 1–56 (1958).
23. W. P. Woodring, *U.S. Geol. Surv. Prof. Pap. 306-A*, 145 (1957).
24. J. Silva Tamayo, G. Sierra, L. Correa, *J. S. Am. Earth Sci.* **26**, 369–382 (2008).
25. A. H. Osborne *et al.*, *Paleoceanography* **29**, 715–729 (2014).
26. P. Sepulchre *et al.*, *Paleoceanography* **29**, 176–189 (2014).
27. A. M. Mestas Nuñez, P. Molnar, *Paleoceanography* **29**, 508–517 (2014).
28. C. Montes *et al.*, *J. Geophys. Res.* **117**, B04105 (2012).
29. G. K. Westbrook, in *The Geology of North America*, G. Dengo, J. E. Case, Eds. (Geological Society of America, Boulder, CO, 1990), vol. H, Plate 7.

30. M. X. Kirby, B. MacFadden, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **228**, 193–202 (2005).

**ACKNOWLEDGMENTS**

Supported by Ecopetrol-ICP "Cronología de la Deformación en las Cuencas Subandinas," Smithsonian Institution, Uniandes P12. 160422.002/001, Autoridad del Canal de Panama (ACP), the Mark Tupper Fellowship, Ricardo Perez S.A.; NSF grant EAR 0824299 and OISE, EAR, DRL 0966884, Colciencias, and the National Geographic Society. We thank N. Hoyos, D. Villagomez, A. O'Dea, C. Bustamante, O. Montenegro, and C. Ojeda. All the data reported in this manuscript are presented in the main paper and in the supplementary materials.

**SUPPLEMENTARY MATERIALS**

www.sciencemag.org/content/348/6231/226/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S3
Tables S1 to S5
References (*31*–*42*)

12 November 2014; accepted 2 March 2015
10.1126/science.aaa2815

EARTH HISTORY

# Ocean acidification and the Permo-Triassic mass extinction

**M. O. Clarkson,[1]*† S. A. Kasemann,[2] R. A. Wood,[1] T. M. Lenton,[3] S. J. Daines,[3] S. Richoz,[4] F. Ohnemueller,[2] A. Meixner,[2] S. W. Poulton,[5] E. T. Tipper[6]**

**Ocean acidification triggered by Siberian Trap volcanism was a possible kill mechanism for the Permo-Triassic Boundary mass extinction, but direct evidence for an acidification event is lacking. We present a high-resolution seawater pH record across this interval, using boron isotope data combined with a quantitative modeling approach. In the latest Permian, increased ocean alkalinity primed the Earth system with a low level of atmospheric $CO_2$ and a high ocean buffering capacity. The first phase of extinction was coincident with a slow injection of carbon into the atmosphere, and ocean pH remained stable. During the second extinction pulse, however, a rapid and large injection of carbon caused an abrupt acidification event that drove the preferential loss of heavily calcified marine biota.**

The Permo-Triassic Boundary (PTB) mass extinction, at ~252 million years ago (Ma), represents the most catastrophic loss of biodiversity in geological history and played a major role in dictating the subsequent evolution of modern ecosystems (*1*). The PTB extinction event spanned ~60,000 years (*2*) and can be resolved into two distinct marine extinction pulses (*3*). The first occurred in the latest Permian [Extinction Pulse 1 (EP1)] and was followed by an interval of temporary recovery before the second pulse (EP2), which occurred in the earliest Triassic. The direct cause of the mass extinction is widely debated, with a diverse range of overlapping mechanisms proposed, including widespread water column anoxia (*4*), euxinia (*5*), global warming (*6*), and ocean acidification (*7*).

Models of PTB ocean acidification suggest that a massive and rapid release of $CO_2$ from Siberian Trap volcanism acidified the ocean (*7*). Indirect evidence for acidification comes from the interpretation of faunal turnover records (*3*, *8*), potential dissolution surfaces (*9*), and Ca isotope data (*7*). A rapid input of carbon is also potentially recorded in the negative carbon isotope excursion (CIE) that characterizes the PTB interval (*10*, *11*). The interpretation of these records is, however, debated (*12*–*16*) and is of great importance to understanding the current threat of anthropogenically driven ocean acidification (*11*).

To test the ocean acidification hypothesis, we have constructed a proxy record of ocean pH across the PTB using the boron isotope composition of marine carbonates ($\delta^{11}B$) (*17*). We then used a carbon cycle model (supplementary text) to explore ocean carbonate chemistry and pH scenarios that are consistent with our $\delta^{11}B$ data and published records of carbon cycle disturbance and environmental conditions. Through this combined geochemical, geological, and modeling approach, we are able to produce an envelope that encompasses the most realistic range in pH, which then allows us to resolve three distinct chronological phases of carbon cycle perturbation, each with very different environmental consequences for the Late Permian–Early Triassic Earth system.

We analyzed boron and carbon isotope data from two complementary transects in a shallow marine, open-water carbonate succession from the United Arab Emirates (U.A.E.), where depositional facies and stable carbon isotope ratio ($\delta^{13}C$) are well constrained (*18*). During the PTB interval, the U.A.E. formed an expansive carbonate platform that remained connected to the central Neo-Tethyan Ocean (Fig. 1A) (*18*). Conodont stratigraphy and the distinct $\delta^{13}C$ curve are used to constrain the age model (*17*).

The PTB in the Tethys is characterized by two negative $\delta^{13}C$ excursions interrupted by a short-term positive event (*10*). There is no consensus as to the cause of this "rebound" event and so we instead focus on the broader $\delta^{13}C$ trend. Our $\delta^{13}C$ transect (Fig. 1B) starts in the Changhsingian (Late Permian) with a gradual decreasing trend, interrupted by the first negative shift in $\delta^{13}C$ at EP1 (at 53 m, ~251.96 Ma) (Figs. 1B and 2). This is followed by the minor positive rebound event (at 54 m, ~251.95 Ma) (Figs. 1B and 2) before the minima of the second phase of the negative CIE (58 to 60 m, ~251.92 Ma) (Figs. 1B and 2) that marks the PTB itself. After the CIE minimum, $\delta^{13}C$ gradually increases to ~1.8 per mil (‰) and remains relatively stable during the earliest Triassic and across EP2.

Our boron isotope record shows a different pattern to the carbon isotope excursion. The boron isotope ratio ($\delta^{11}B$) is persistently low (Fig. 1C) at the start of our record during the late-Changhsingian, with an average of 10.9 ± 0.9‰ (1σ). This is in agreement with $\delta^{11}B$ values (average of 10.6 ± 0.6‰, 1σ) reported for early-Permian brachiopods (*19*). Further up the section (at ~40 m, ~252.04 Ma) (Fig. 1C), there is a stepped increase in $\delta^{11}B$ to 15.3 ± 0.8‰ (propagated uncertainty, 2σf) and by implication an increase in ocean pH of ~0.4 to 0.5 (Fig. 2). $\delta^{11}B$ values then remain relatively stable, scattering around 14.7 ± 1.0‰ (1σ) and implying variations within 0.1 to 0.2 pH, into the Early Griesbachian (Early Triassic) and hence across EP1 and the period of carbon cycle disturbance (Figs. 1 and 2).

After the $\delta^{13}C$ increase and stabilization (at ~85 m, ~251.88 Ma) (Fig. 1), $\delta^{11}B$ begins to decrease rapidly to 8.2 ± 1.2‰ (2σf), implying a sharp drop in pH of ~0.6 to 0.7. The $\delta^{11}B$ minimum is coincident with the interval identified as EP2. This ocean acidification event is short-lived (~10,000 years), and $\delta^{11}B$ values quickly recover toward the more alkaline values evident during EP1 (average of ~14‰).

The initial rise in ocean pH of ~0.4 to 0.5 units during the Late Permian (Fig. 2) suggests a large increase in carbonate alkalinity (*20*). We are able to simulate the observed rise in $\delta^{11}B$ and pH through different model combinations of increasing silicate weathering, increased pyrite deposition (*21*), an increase in carbonate weathering, and a decrease in shallow marine carbonate depositional area (supplementary text). Both silicate weathering and pyrite deposition result in a large drop in partial pressure of $CO_2$ ($P_{CO_2}$) (and temperature) for a given increase in pH and saturation state (Ω). There is no evidence for a large drop in $P_{CO_2}$, and independent proxy data

[1]School of Geosciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3FE, UK. [2]Faculty of Geosciences and MARUM–Center for Marine Environmental Sciences, University of Bremen, 28334 Bremen, Germany. [3]College of Life and Environmental Sciences, University of Exeter, Laver Building, North Parks Road, Exeter EX4 4QE, UK. [4]Institute of Earth Sciences, NAWI Graz, University of Graz, Heinrichstraße 26, 8010 Graz, Austria. [5]School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK. [6]Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.
***Corresponding author. E-mail: matthew.clarkson@otago.ac.nz**
†Present address: Department of Chemistry, University of Otago, Union Street, Dunedin, 9016, Post Office Box 56, New Zealand.

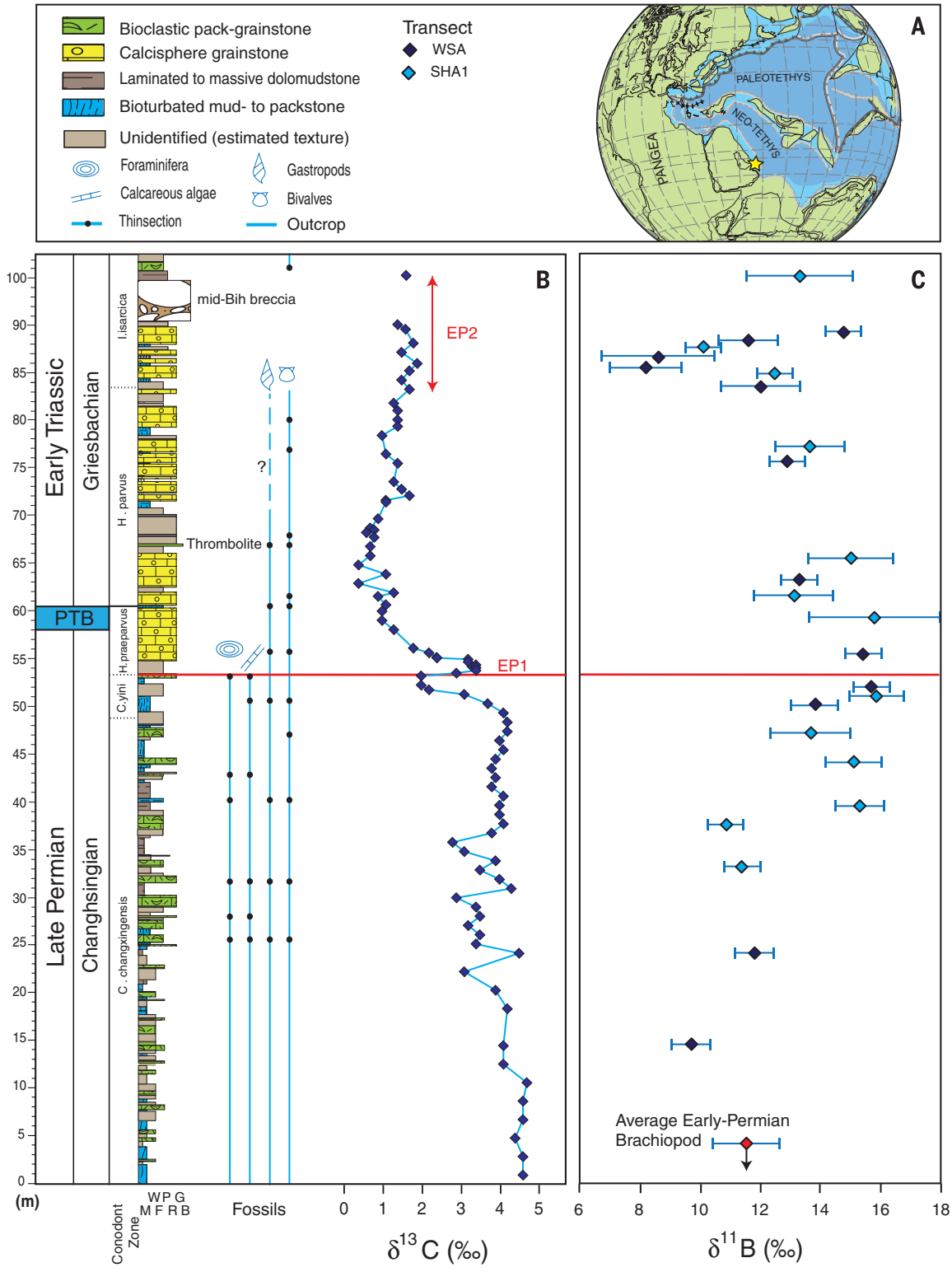



**Fig. 1. Site locality and high-resolution carbon and boron isotope data.** (**A**) Paleogeographic reconstruction for the Late Permian showing the studied section Wadi Bih, in the Musandam Mountains of U.A.E., that formed an extensive carbonate platform in the Neo-Tethyan Ocean. [Modified from (*35*).] (**B**) Shallow water $\delta^{13}C$ record (*18*). (**C**) Boron isotope ($\delta^{11}B$) record (propagated uncertainty given as $2\sigma f$) and average Early Permian brachiopod value ($n$ = 5 samples) (*19*). Lithology, biota, and transect key are provided in (A). Only *Hindeodus parvus* has been found so far in this section (*18*), and the conodont zones with dashed lines are identified from the $\delta^{13}C$ record and regional stratigraphy (*36*–*38*).

indicate only a minor temperature decrease of a few degrees celsius during the Changhsingian (*22*), suggesting that these mechanisms alone cannot explain the pH increase (fig. S5). Conversely, an increase in carbonate input or a reduction in rates of carbonate deposition both result in increases in Ω, with a greater impact on pH per unit decrease in $P_{CO_2}$ and temperature (fig. S6).

A decrease in carbonate sedimentation is consistent with the decrease in depositional shelf area that occurred because of the second-order regression of the Late Permian (*23*). With the added expansion of anoxia into shelf environments (*24*), this would effectively create both bottom-up and top-down pressures to reduce the area of potential carbonate sedimentation.

Sea-level fall also exposed carbonates to weathering (*23*), which would have further augmented the alkalinity influx. The pH increase event supports the $CO_2$*Lo* initialization scenario [$CO_2$ ~3 present atmospheric levels (PAL), pH ~8, $\delta^{11}B_{SW}$ ~34‰] (supplementary text) because the simulated $CO_2$ and temperature decrease is much reduced and therefore is more consistent with

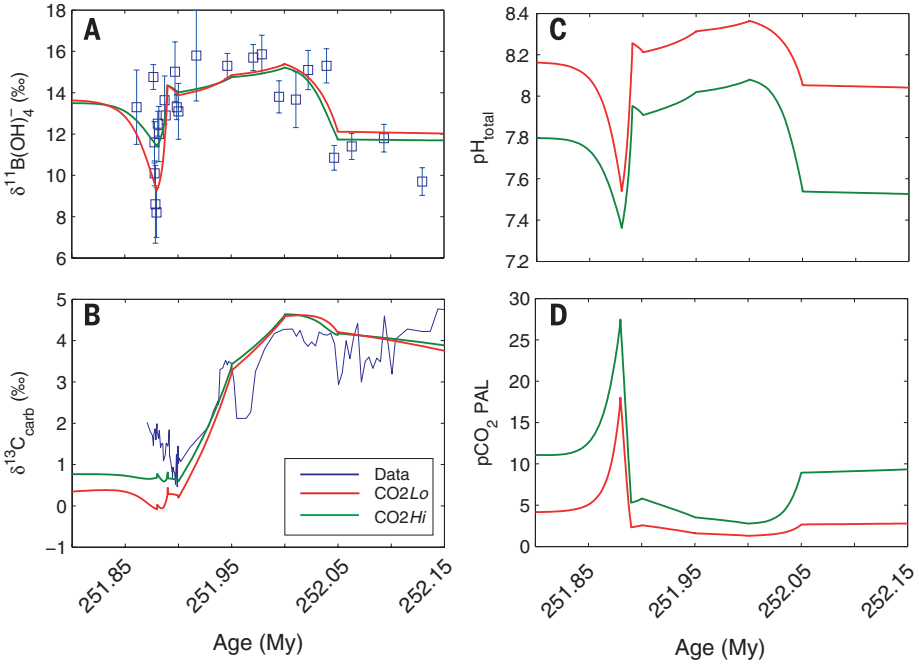



**Fig. 2. Model results of carbon cycle parameters for high- and low-$CO_2$ end-member scenarios.** (**A**) Model-reproduced $\delta^{11}B$ versus data. (**B**) Modeled $\delta^{13}C$ versus data. (**C**) Modeled pH envelope incorporating uncertainty of seawater B isotope composition ($\delta^{11}B_{SW}$) and dynamic temperatures. (**D**) Calculated atmospheric $CO_2$.

independent proxy data (*22*), as compared with $CO_2$*Hi* ($CO_2$ ~10 PAL, pH ~7.5, $\delta^{11}B_{SW}$ ~36.8‰) (Fig. 2D).

Before EP1, $\delta^{13}C_{carb}$ values began to decrease before reaching the minimum of the globally recognized negative CIE at the PTB (Fig. 1). At this time, both $\delta^{11}B$ and ocean pH remained stable. Hypotheses to explain the negative CIE require the input of isotopically light carbon, such as from volcanism (*14*, *25*), with the assimilation of very light organic carbon from the surrounding host rock (*26*), methane destabilization (*27*), collapse of the biological pump (*15*), and/or a decrease in the burial of terrestrial carbon (*16*). We can simulate the observed drop in $\delta^{13}C$, while remaining within the uncertainty of the $\delta^{11}B$ data (Fig. 2), by combining a cessation of terrestrial carbon burial with a relatively slow (50,000 years) carbon injection from any of the above sources (fig. S8). A small source of methane ($3.2 \times 10^{17}$ mol C with $\delta^{13}C = -50‰$) gives the least change in $\delta^{11}B$ and pH, whereas either a larger source of organic carbon (~$6.5 \times 10^{17}$ mol C with $\delta^{13}C = -25‰$) or a mixture of mantle and lighter carbon sources (~$1.3 \times 10^{18}$ mol C with $\delta^{13}C = -12.5‰$) are still within the measured uncertainty in $\delta^{11}B$.

This relatively slow addition of carbon minimizes the tendency for a transient decline in surface ocean pH in an ocean that was already primed with a high Ω and hence high buffering capacity from the Late Permian. The global presence of microbial and abiotic carbonate fabrics after EP1 (*28*) is indicative that this high Ω was maintained across the CIE. The carbon injection triggers an increase in $P_{CO_2}$, temperature, and silicate weathering, creating an additional counterbalancing alkalinity flux, which is consistent with independent proxy data (*6*). The alkalinity source may have been further increased through soil loss (*29*), the emplacement of easily weathered Siberian Trap basalt, or the impact of acid rain (*30*), which would have increased weathering efficiency.

The negative $\delta^{11}B_{carb}$ excursion at 251.88 Ma represents a calculated pH decrease of up to 0.7 pH. This pH decrease coincides with the second pulse of the extinction (Fig. 1), which preferentially affected the heavily calcifying, physiologically unbuffered, and sessile organisms (*3*). This was also accompanied by the temporary loss of abiotic and microbial carbonates throughout the Tethys (*31*, *32*), suggesting a coeval decrease in Ω. To overwhelm the buffering capacity of the ocean and decrease pH in this way requires a second, more abrupt injection of carbon into the atmosphere, yet remarkably, the acidification event occurs after the decline in $\delta^{13}C$, when $\delta^{13}C$ has rebounded somewhat and is essentially stable (Fig. 1).

Unlike the first carbon injection, the lack of change in $\delta^{13}C$ at this time rules out very $^{13}C$-depleted carbon sources because no counterbalancing strongly $^{13}C$-enriched source exists. Instead, it requires a carbon source near ~0‰. A plausible scenario for this is the decarbonation of overlying carbonate host rock, into which the Siberian Traps intruded (*26*), or the direct assimilation of carbonates and evaporites into the melt (*33*). Host carbonates would have had $\delta^{13}C$ ~ +2 to 4‰, which when mixed with mantle carbon (~−5‰) potentially produces a source near 0‰. We can simulate the sharp drop in pH and stable $\delta^{13}C$ values (Fig. 2) through a large and rapid carbon release of $2 \times 10^{18}$ mol C over 10,000 years (fig. S8). This is undoubtedly a massive injection of 24,000 PgC at a rapid rate of 2.4 PgC/year, but it is physically plausible given existing estimates of the volume of carbonate host sediments subject to contact metamorphism and postulated mechanisms of carbon release (supplementary text). This second rapid carbon release produces a sharp rise in $P_{CO_2}$ to ~20 PAL and warming of ~15°C, which is consistent with the observation of peak temperatures after EP1 (*22*). Initialization of the carbon cycle model under $CO_2$*Hi* cannot generate the magnitude of $\delta^{11}B$ drop (Fig. 2A) because the nonlinear relation between pH and $\delta^{11}B$ fractionation sets a lower limit of $\delta^{11}B$ at ~10‰ in this case (fig. S3). Thus, low initial $CO_2$ of ~3 PAL in the Late Permian ($CO_2$*Lo*) is more consistent with our data.

An acidification event of ~10,000 years is consistent with the modeled time scale required to replenish the ocean with alkalinity, as carbonate deposition is reduced and weathering is increased under higher $P_{CO_2}$ and global temperatures. Increased silicate weathering rates drive further $CO_2$ drawdown, resulting in stabilization (Fig. 2D). High global temperature (*6*) and increased silicate weathering are consistent with a sudden increase in both $^{87}Sr/^{86}Sr$ (*34*) and sedimentation rates (*29*) in the Griesbachian.

The PTB was a time of extreme environmental change, and our combined data and modeling approach falsifies several of the mechanisms currently proposed. Although the coincident stresses of anoxia, increasing temperature, and ecosystem restructuring were important during this interval, the $\delta^{11}B$ record strongly suggests that widespread ocean acidification was not a factor in the first phase of the mass extinction but did drive the second pulse. The carbon release required to drive the observed acidification event must have occurred at a rate comparable with the current anthropogenic perturbation but exceeds it in expected magnitude. Specifically, the required model perturbation of 24,000 PgC exceeds the ~5000 PgC of conventional fossil fuels and is at the upper end of the range of estimates of unconventional fossil fuels (such as methane hydrates). We show that such a rapid and large release of carbon is critical to causing the combined synchronous decrease in both pH and saturation state that defines an ocean acidification event (*11*).

## REFERENCES AND NOTES


1. D. H. Erwin, *Nature* **367**, 231–236 (1994).
2. S. D. Burgess, S. Bowring, S. Z. Shen, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3316–3321 (2014).
3. H. J. Song, P. B. Wignall, J. A. Tong, Y. Hongfu, *Nat. Geosci.* **6**, 52–56 (2012).
4. P. B. Wignall, R. J. Twitchett, *Science* **272**, 1155–1158 (1996).
5. K. Grice *et al.*, *Science* **307**, 706–709 (2005).
6. Y. Sun *et al.*, *Science* **338**, 366–370 (2012).
7. J. L. Payne *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8543–8548 (2010).
8. A. H. Knoll, R. K. Barnbach, J. L. Payne, S. Pruss, W. W. Fischer, *Earth Planet. Sci. Lett.* **256**, 295–313 (2007).
9. J. L. Payne *et al.*, *Geol. Soc. Am. Bull.* **119**, 771–784 (2007).
10. C. Korte, H. W. Kozur, *J. Asian Earth Sci.* **39**, 215–235 (2010).

11. B. Hönisch *et al.*, *Science* **335**, 1058–1063 (2012).
12. C. L. Blättler, H. C. Jenkyns, L. M. Reynard, G. M. Henderson, *Earth Planet. Sci. Lett.* **309**, 77–88 (2011).
13. P. B. Wignall, S. Kershaw, P. Y. Collin, S. Crasquin-Soleau, *Geol. Soc. Am. Bull.* **121**, 954–956 (2009).
14. R. A. Berner, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4172–4177 (2002).
15. M. R. Rampino, K. Caldeira, *Terra Nova* **17**, 554–559 (2005).
16. W. S. Broecker, S. Peacock, *Global Biogeochem. Cycles* **13**, 1167–1172 (1999).
17. Materials and methods are available on *Science* Online
18. M. O. Clarkson *et al.*, *Gondwana Res.* **24**, 233–242 (2013).
19. M. M. Joachimski, L. Simon, R. van Geldern, C. Lecuyer, *Geochim. Cosmochim. Acta* **69**, 4035–4044 (2005).
20. The alternative way to drive an increase in pH would be through a removal of carbon; however, this would be evident in the $\delta^{13}C$ record, so we can rule it out.
21. Bacterial sulfate reduction (BSR) is a net source of alkalinity if the generated $H_2S$ is buried as pyrite. Pyrite deposition is seen widely in certain settings during the Late Permian to PTB. Further information is provided in the supplementary materials.
22. M. M. Joachimski *et al.*, *Geology* **40**, 195–198 (2012).
23. H. Yin *et al.*, *Earth Sci. Rev.* **137**, 19–33 (2014).
24. P. B. Wignall, R. J. Twitchett, in *Catastrophic Events and Mass Extinctions: Impacts and Beyond*, C. Koeberl, K. G. MacLeod, Eds. (Geological Society of America, Boulder, CO, 2002), pp. 395–413.
25. C. Korte *et al.*, *J. Asian Earth Sci.* **37**, 293–311 (2010).
26. H. Svensen *et al.*, *Earth Planet. Sci. Lett.* **277**, 490–500 (2009).
27. E. S. Krull, G. J. Retallack, *Geol. Soc. Am. Bull.* **112**, 1459–1472 (2000).
28. A. D. Woods, *Earth Sci. Rev.* **137**, 6–18 (2013).
29. T. J. Algeo, Z. Q. Chen, M. L. Fraiser, R. J. Twitchett, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **308**, 1–11 (2011).
30. B. A. Black, J. F. Lamarque, C. A. Shields, L. T. Elkins-Tanton, J. T. Kiehl, *Geology* **42**, 67–70 (2014).
31. A. Baud, S. Richoz, S. Pruss, *Global Planet. Change* **55**, 81–89 (2007).
32. S. Richoz *et al.*, *J. Asian Earth Sci.* **39**, 236–253 (2010).
33. B. A. Black, L. T. Elkins-Tanton, M. C. Rowe, I. U. Peate, *Earth Planet. Sci. Lett.* **317**, 363–373 (2012).
34. C. Korte *et al.*, *Int. J. Earth Sci.* **93**, 565–581 (2004).
35. G. M. Stampfli, G. D. Borel, *Earth Planet. Sci. Lett.* **196**, 17–33 (2002).
36. S. Z. Shen *et al.*, *Earth Planet. Sci. Lett.* **375**, 156–165 (2013).
37. B. Koehrer *et al.*, *GeoArabia* **15**, 91–156 (2010).
38. F. Maurer, R. Martini, R. Rettori, H. Hillgartner, S. Cirilli, *GeoArabia* **14**, 125–158 (2009).



### ACKNOWLEDGMENTS

M.O.C. acknowledges funding from the Edinburgh University Principal's Career Development Scholarship, the International Centre for Carbonate Reservoirs, and The Marsden Fund (U001314). R.A.W., T.M.L., and S.W.P. acknowledge support from the Natural Environment Research Council through the "Co-evolution of Life and the Planet" scheme (NE/I005978). T.M.L. and S.J.D. were supported by the Leverhulme Trust (RPG-2013-106). S.A.K. and A.M. acknowledge support from the German Research Foundation (Deutsche Forschungsgemeinschaft) Major Research Instrumentation Program INST 144/307-1. This is a contribution to IGCP 572, with S.R. sponsored for fieldwork by the Austrian National Committee (Austrian Academy of Sciences) for the International Geoscience Programme (IGCP). We are grateful to R. Newton and A. Thomas for helpful discussions, L. Krystyn for field assistance, F. Maurer for discussions on stratigraphy and providing photomicrographs, and B. Mills for assisting with model studies. Data are available online in the supplementary materials and at www.pangaea.de.


### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/229/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S10
References (*39*–*98*)

8 October 2014; accepted 4 March 2015
10.1126/science.aaa0193

COMETARY FORMATION

# Molecular nitrogen in comet 67P/Churyumov-Gerasimenko indicates a low formation temperature

**M. Rubin,[1]* K. Altwegg,[1,2] H. Balsiger,[1] A. Bar-Nun,[3] J.-J. Berthelier,[4] A. Bieler,[1,5] P. Bochsler,[1] C. Briois,[6] U. Calmonte,[1] M. Combi,[5] J. De Keyser,[7] F. Dhooghe,[7] P. Eberhardt,[1]† B. Fiethe,[8] S. A. Fuselier,[9] S. Gasc,[1] T. I. Gombosi,[5] K. C. Hansen,[5] M. Hässig,[1,9] A. Jäckel,[1] E. Kopp,[1] A. Korth,[10] L. Le Roy,[2] U. Mall,[10] B. Marty,[11] O. Mousis,[12] T. Owen,[13] H. Rème,[14,15] T. Sémon,[1] C.-Y. Tzou,[1] J. H. Waite,[9] P. Wurz[1]**

**Molecular nitrogen ($N_2$) is thought to have been the most abundant form of nitrogen in the protosolar nebula. It is the main N-bearing molecule in the atmospheres of Pluto and Triton and probably the main nitrogen reservoir from which the giant planets formed. Yet in comets, often considered the most primitive bodies in the solar system, $N_2$ has not been detected. Here we report the direct in situ measurement of $N_2$ in the Jupiter family comet 67P/Churyumov-Gerasimenko, made by the Rosetta Orbiter Spectrometer for Ion and Neutral Analysis mass spectrometer aboard the Rosetta spacecraft. A $N_2$/CO ratio of $(5.70 \pm 0.66) \times 10^{-3}$ ($2\sigma$ standard deviation of the sampled mean) corresponds to depletion by a factor of ~25.4 ± 8.9 as compared to the protosolar value. This depletion suggests that cometary grains formed at low-temperature conditions below ~30 kelvin.**

Thermochemical models of the protosolar nebula (PSN) suggest that molecular nitrogen ($N_2$) was the principal nitrogen species during the disk phase (*1*) and that the nitrogen present in the giant planets was accreted in this form (*2*). Moreover, Pluto and Triton, which are both expected to have formed in the same region of the PSN as Jupiter family comets (JFCs), have $N_2$-dominated atmospheres and surface deposits of $N_2$ ice (*3*, *4*). This molecule has never been firmly detected in comets; however, CN, HCN, NH, $NH_2$, and $NH_3$ among others have been observed spectroscopically (*5*, *6*). The abundance of $N_2$ in comets is therefore a key to understanding the conditions in which they formed.

Condensation or trapping of $N_2$ in ice occurs at similar thermodynamic conditions as those needed for CO in the PSN (*7*, *8*). This requires very low PSN temperatures and implies that the detection of $N_2$ in comets and its abundance ratio with respect to CO would put strong constraints on comet formation conditions (*7*, *8*). Ground-based spectroscopic observations of the $N_2^+$ band in the near ultraviolet are very difficult because of the presence of telluric $N_2^+$ and other cometary emission lines. Searches conducted with high-resolution spectra of comets 122P/de Vico, C/1995 O1 (Hale-Bopp), and 153P/2002 C1 (Ikeya-Zhang) have been unsuccessful and yielded upper limits of $10^{-5}$ to $10^{-4}$ for the $N_2^+/CO^+$ ratio (*9*, *10*). Only one $N_2^+$ detection in C/2002 VQ94 (LINEAR) from ground-based observations is convincing, because the comet was at sufficient distance from the Sun to prevent terrestrial twilight $N_2^+$ contamination (*11*). The in situ measurements made by Giotto in 1P/Halley were inconclusive, because the resolution of the mass spectrometers aboard the spacecraft (*12*) was insufficient to separate the nearly identical masses of $N_2$ and CO during the 1P/Halley encounter, and only an upper limit could be derived for the relative production rates $[Q(N_2)/Q(CO) \leq 0.1]$ (*13*).

Here we report the direct in situ measurement of the $N_2$/CO ratio by the Rosetta Orbiter Spectrometer for Ion and Neutral Analysis (ROSINA)

[1]Physikalisches Institut, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland. [2]Center for Space and Habitability, University of Bern, Sidlerstrasse. 5, CH-3012 Bern, Switzerland. [3]Department of Geoscience, Tel-Aviv University, Ramat-Aviv, Tel-Aviv, Israel. [4]Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS)/Institute Pierre Simon Laplace–CNRS-UPMC-UVSQ, 4 Avenue de Neptune F-94100, Saint-Maur, France. [5]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, 2455 Hayward, Ann Arbor, MI 48109, USA. [6]Laboratoire de Physique et Chimie de l'Environnement et de l'Espace (LPC2E), UMR 6115 CNRS–Université d'Orléans, Orléans, France. [7]Belgian Institute for Space Aeronomy, Belgisch Instituut voor Ruimte-Aeronomie–Institut d'Aéronomie Spatiale de Belgique (BIRA-IASB), Ringlaan 3, B-1180 Brussels, Belgium. [8]Institute of Computer and Network Engineering, Technische Universität Braunschweig, Hans-Sommer-Straße 66, D-38106 Braunschweig, Germany. [9]Department of Space Science, Southwest Research Institute, 6220 Culebra Road, San Antonio, TX 78228, USA. [10]Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany. [11]Centre de Recherches Pétrographiques et Géochimiques (CRPG)–CNRS, Université de Lorraine, 15 rue Notre Dame des Pauvres, Bôite Postale 20, 54501 Vandoeuvre lès Nancy, France. [12]Aix Marseille Université, CNRS, Laboratoire d'Astrophysique de Marseille UMR 7326, 13388, Marseille, France. [13]Institute for Astronomy, University of Hawaii, Honolulu, HI 96822, USA. [14]Université de Toulouse; UPS-OMP; Institut de Recherche en Astrophysique et Planétologie (IRAP), Toulouse, France. [15]CNRS; IRAP; 9 Avenue du Colonel Roche, Boîte Postale 44346, F-31028 Toulouse Cedex 4, France.
***Corresponding author. E-mail: martin.rubin@space.unibe.ch**
†Deceased.

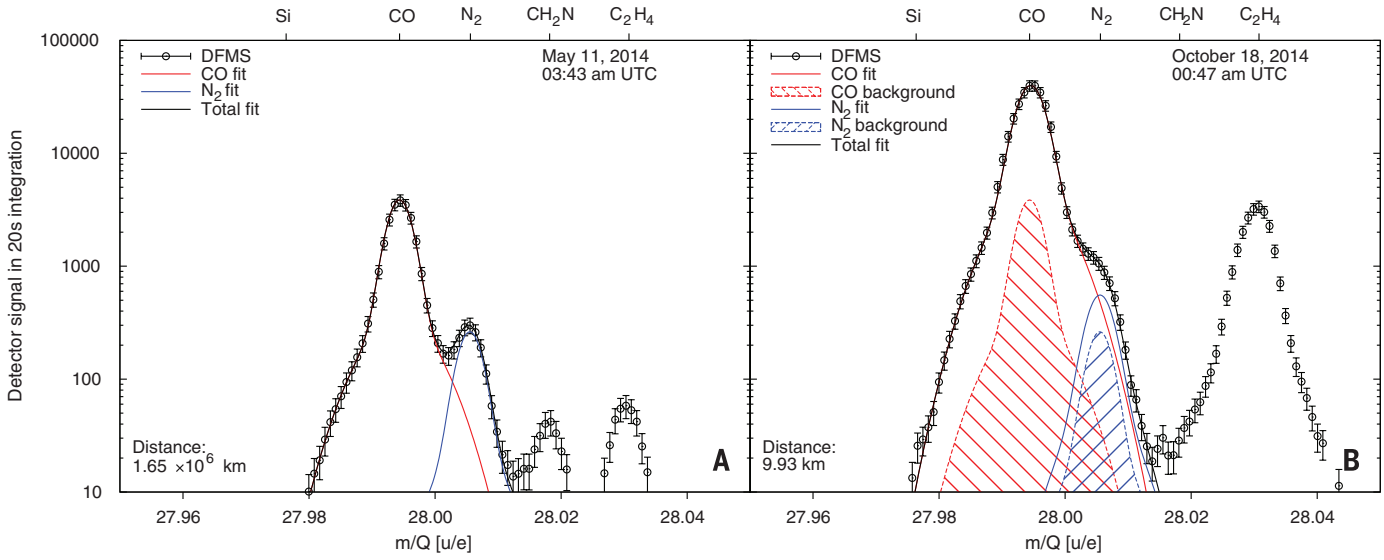



**Fig. 1. Mass per charge 28 u/*e* spectra before (A) and after (B) entering the coma of 67P, including statistical and 10% pixel gain error.** (A) was obtained in May 2014 and (B) is a representative spectrum from October containing the sum of the cometary parents and fragments and the spacecraft background signals.

in the JFC 67P/Churyumov-Gerasimenko (hereafter 67P). ROSINA is the mass spectrometer suite on the European Space Agency's Rosetta spacecraft (*14*) and measures the gas density and composition at the location of the spacecraft (*15*). The Double Focusing Mass Spectrometer (DFMS) has a high mass resolution of $m/\Delta m$ about 3000 at the 1% level (corresponding to ~9000 half peak width at the 50% level) at atomic mass per unit of charge 28 u/*e*, allowing the separation of $N_2$ from CO ($\Delta m$ = 0.011 u) by numerical peak fitting. Neutral gas is ionized by electron impact and then deflected through an electrostatic, then magnetic, filter onto a position-sensitive microchannel plate (MCP) detector. The peak shape of a single species on the MCP is well known, and therefore numerical fitting can distinguish overlapping contributions from different atoms and molecules (see the supplementary materials).

Starting on 5 August 2014, ROSINA observed the cometary gas flux rise above the spacecraft background signal for the major species, including $H_2O$, CO, and $CO_2$. For $N_2$, which has a higher relative spacecraft background, the cometary signal became apparent a few days later. The spacecraft background signal (*16*) for both species, CO and $N_2$, was derived at different times before detecting the coma and shown to be temporally quite stable. $N_2$ and CO were both observed in the Rosetta spacecraft background mass spectra, e.g., on 11 May 2014, while the spacecraft was still at a distance of $1.65 \times 10^6$ km from the comet (Fig. 1A). A comparable $N_2$ background was measured on 1 August 2014, at almost 800 km from the nucleus before the cometary signal became apparent. Another mass spectrum, representative of the measurements within a distance of 10 km from the nucleus, was obtained on 18 October 2014 (Fig. 1B) and includes both cometary and spacecraft background signal. The indicated background was subsequently removed, leaving only cometary CO and $N_2$. Furthermore, CO from dissociative electron-impact ionization of cometary $CO_2$ inside DFMS' ion source was removed (a 7 to 36% reduction), and the signal was corrected for the instrument alignment with respect to the comet (supplementary materials).

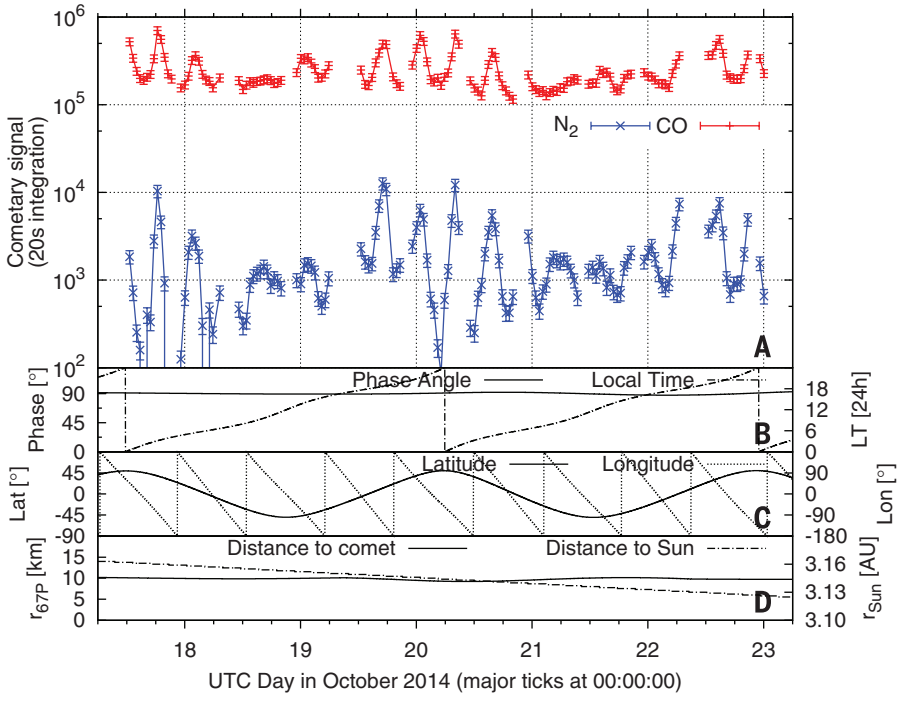



**Fig. 2. Cometary parent CO and $N_2$ signal during 17 to 23 October 2014.** (**A**) The error bars are associated with the accuracy of the fit, background subtraction, detector gain, and statistical error. Gaps in the data indicate times when ROSINA was off due to thruster operations. The sections below show phase angle and local time (**B**), latitude and longitude of the subspacecraft point (**C**) in the Cheops coordinate system, and the distances of Rosetta to the comet ($r_{67P}$) and the comet to the Sun ($r_{Sun}$) (**D**). The summer hemisphere is at positive latitudes.

This procedure was carried out for 138 spectra over two terminator orbits of the Rosetta spacecraft from 17 to 23 October 2014. Clear diurnal variations in the cometary signal of both species associated with the 12.4-hour rotation period of the comet have been observed (Fig. 2A).

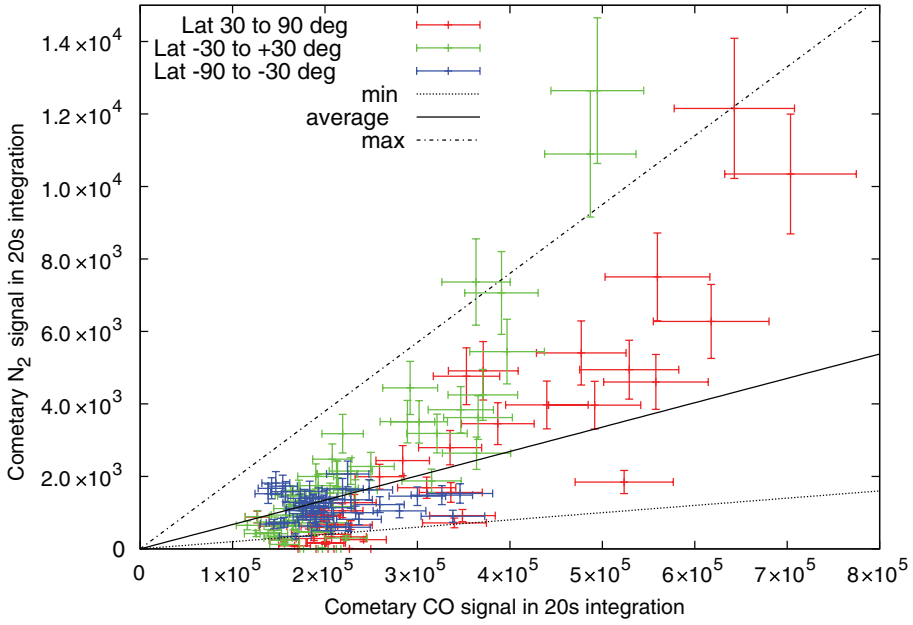



**Fig. 3. Cometary parent $N_2$ versus CO signal.** The min and max lines bracket most measurements. To derive the $N_2$/CO ratio, the detector signal ratio in the plot has to be divided by the differential sensitivity of 1.175. The average $N_2$/CO ratio of $5.70 \times 10^{-3}$ is given by the solid black line; the min and the max lines show the observed variation and correspond to ratios of $1.7 \times 10^{-3}$ and $1.6 \times 10^{-2}$.

The signal is to first order correlated to the comet's cross-section exposed to the Sun and the relative position of Rosetta (Fig. 2, B to D). The resulting mean $N_2$/CO ratio of $(5.70 \pm 0.66) \times 10^{-3}$ in the observed time period corresponds to the mean ratio of each individual measurement and includes the 2σ SD of the sampled mean. Higher outgassing is found at positive latitudes corresponding to the summer hemisphere. Over the sunlit hemisphere, the CO/$H_2O$ ratio varies between 0.1 to 0.3 (*17*), which is in agreement with variations observed at other comets (*6*). Because these measurements were achieved at a heliocentric distance of 3.1 astronomical units (AU), the water production rate may increase relative to both CO and $N_2$ as the comet approaches the Sun. We therefore expect the $N_2$/CO ratio to be more representative of the $N_2$ content in the coma than the $N_2/H_2O$ ratio. The $N_2$/CO ratio exhibits a strong variation depending on the position of Rosetta above the surface of the comet nucleus between 0.17 to 1.6% (Fig. 3). There are also hints of a nonlinear relationship between $N_2$ and CO, further indicating that thermal processes in the upper layer of the nucleus and/or surface inhomogeneities might influence the measured $N_2$/CO ratio in the coma.

With a protosolar ratio N/C of 0.29 ± 0.10 (*18*) and assuming to first order that all of N and C were in the form of $N_2$ and CO in the PSN (*1*), we derived an $N_2$/CO ratio of 0.145 ± 0.048 in the PSN gas phase. The comparison with the $N_2$/CO measurement performed in the near coma of 67P shows that the cometary $N_2$/CO ratio is depleted by a factor of about 25.4 ± 8.9 as compared to the value derived from protosolar N and C abundances. This depletion of $N_2$ relative to CO in comet 67P may be a consequence of how cometary ice formed. According to one model, comets agglomerated from pristine amorphous water ice grains originating from the interstellar medium (ISM) (*19*). In this case, the low $N_2$/CO ratio in 67P is the result of inefficient trapping of $N_2$ in amorphous water ice as compared to CO. This possibility is supported by laboratory experiments in which a mixture of water vapor with $N_2$ and CO was directed onto a cold plate in the 24 to 30 K temperature range (*7*). In these experiments, gases initially trapped in growing amorphous ice were later released when ice warmed up, and the evolved gases were measured by mass spectrometry. At 24 K, the depletion factor for the $N_2$/CO ratio was found to be ~19, a value within the range of the one observed in 67P of 25.4 ± 8.9. This yields a lower limit for the temperature experienced by the grains agglomerated by 67P, because the $N_2$/CO ratio in amorphous ice would increase at temperatures lower than 24 K because of increasing efficiency of $N_2$ trapping.

An alternative interpretation of the low $N_2$ abundance is that 67P agglomerated from grains consisting of clathrates, which are icelike crystalline solids formed by cages of water molecules that contain small nonpolar molecules (*20*). This hypothesis is based on models showing that the vaporization distance of ISM ices could have been as high as about 30 AU from the Sun when they entered the PSN (*21*). With time, the decrease of the gas temperature and pressure allowed water to condense at ~140 to 150 K in the form of crystalline ice, leaving negligible water in the gas phase to condense at low temperatures where amorphous ice is expected to form (*22*). Depending on the nature of the entrapped species, clathrates formed from preexisting crystalline water ice when the PSN temperature was lower than about 80 K, provided that the slow kinetics of the process was balanced by sufficient formation time (*8*). As in the case of trapping in amorphous ice, experiments and models suggest that $N_2$ is poorly trapped in clathrate cages, because of its small size (*8*, *23*–*25*). In particular, statistical thermodynamics models (*26*) used to compute the composition of clathrates formed from a protosolar-composition gas in the PSN show that an $N_2$/CO ratio in the comet's nucleus is consistent with the measured value in the coma if the nucleus agglomerated from grains formed in the 26 to 56 K temperature range (*8*).

Both interpretations are consistent with the idea that 67P agglomerated from grains formed at about 30 K or below. However, the measured $N_2$/CO ratio may reflect in whole or in part the comet's post-formation evolution. A possibility is that 67P agglomerated from grains formed at a lower temperature (around 20 K) in the PSN, favoring the trapping of much more $N_2$ in its building blocks, in a way consistent with the known compositions of the atmospheres and surfaces of Pluto and Triton (*3*, *4*). This possibility would be consistent with an inferred Kuiper Belt origin for 67P and its high D/H ratio (*27*). In these conditions, 67P could have been initially $N_2$-rich but subsequent post-accretion heating due to the radiogenic decay of nuclides and/or thermal cycles during its transit from the Kuiper Belt and its subsequent history in a short period orbit could have been sufficient to trigger the outgassing of $N_2$ (*8*). A scenario such as this may explain how initial nitrogen-rich cometesimals similar to Triton and Pluto evolved into nitrogen-depleted comets.

Because $N_2$ trapped in 67P is presumably PSN gas, its $^{14}N/^{15}N$ ratio should be about 441, the value found in Jupiter and the solar wind (*28*). This is much higher than values measured in other cometary N-bearing species such as $NH_3$ and HCN (~130) (*5*). Thus, depending on the proportions of $N_2$ relative to other N-bearing species, the terrestrial $^{14}N/^{15}N$ ratio of 272 could possibly be cometary in origin, given an appropriate mix of the different nitrogen species in the comets that contributed to terrestrial volatiles (e.g., ~50% $N_2$ and ~50% $NH_3$ or HCN). Our initial ROSINA measurement for $N_2$/CO of 0.57% may be compared with $NH_3$/CO of 6% and HCN/CO of ~2% in the Oort cloud comet Hale-Bopp (*6*). The production rates of volatiles relative to water vary from one comet to another, but their values normalized to CO remain close to those measured in Hale-Bopp (*6*). If 67P is a typical JFC, then the ROSINA value for $N_2$/CO implies that the amount of $N_2$ reaching the surface of a solid body in the inner solar system from a JFC impact was almost 15 times less than the amounts of $NH_3$, HCN, and certain organic compounds (*6*). This comparison suggests that JFC comets were not the main source of Earth's nitrogen.

## REFERENCES AND NOTES

1. B. J. Fegley, R. G. Prinn, in *The Formation and Evolution of Planetary Systems*, H. A. Weaver *et al.*, Eds. (Univ. of Arizona Press, Tucson, AZ, 1989), pp. 171–205.
2. O. Mousis *et al.*, *Planet. Space Sci.* **104**, 29–47 (2014).
3. D. P. Cruikshank *et al.*, *Science* **261**, 742–745 (1993).
4. T. C. Owen *et al.*, *Science* **261**, 745–748 (1993).
5. P. Rousselot *et al.*, *Astrophys. J.* **780**, L17 (2014).
6. D. Bockelée-Morvan *et al.*, in *Comets II*, M. Festou, H. U. Keller, H. A. Weaver, Eds. (Univ. of Arizona Press, Tucson, AZ, 2004), pp. 391–423.
7. A. Bar-Nun, G. Notesco, T. Owen, *Icarus* **190**, 655–659 (2007).
8. O. Mousis *et al.*, *Astrophys. J.* **757**, 146 (2012).
9. A. L. Cochran, W. D. Cochran, E. S. Barker, *Icarus* **146**, 583–593 (2000).
10. A. L. Cochran, *Astrophys. J.* **576**, L165–L168 (2002).
11. P. P. Korsun, P. Rousselot, I. V. Kulyk, V. L. Afanasiev, O. V. Ivanova, *Icarus* **232**, 88–96 (2014).
12. D. Krankowsky *et al.*, *Nature* **321**, 326–329 (1986).
13. P. Eberhardt *et al.*, *Astron. Astrophys.* **187**, 481–484 (1987).
14. K.-H. Glassmeier, H. Boehnhardt, D. Koschny, E. Kührt, I. Richter, *Space Sci. Rev.* **128**, 1–21 (2007).
15. H. Balsiger *et al.*, *Space Sci. Rev.* **128**, 745–801 (2007).
16. B. Schläppi *et al.*, *J. Geophys. Res. Space Phys.* **115**, A12313 (2010).
17. M. Hässig *et al.*, *Science* **347**, aaa0276 (2015).
18. K. Lodders, H. Palme, H. P. Gail, in *The Solar System*, J. E. Trümper, Ed. (Springer-Verlag, Berlin Heidelberg, 2009), vol. 4B.
19. T. Owen, A. Bar-Nun, *Nature* **361**, 693–694 (1993).
20. O. Mousis *et al.*, *Icarus* **148**, 513–525 (2000).
21. K. M. Chick, P. Cassen, *Astrophys. J.* **477**, 398–409 (1997).
22. A. Kouchi, T. Yamamoto, T. Kozasa, T. Kuroda, J. M. Greenberg, *Astron. Astrophys.* **290**, 1009–1018 (1994).
23. J.-M. Herri, E. Chassefière, *Planet. Space Sci.* **73**, 376–386 (2012).
24. D. E. Sloan, C. Koh, *Clathrate Hydrates of Natural Gases* (CRC/Taylor & Franis, Boca Raton, FL, ed. 3, 2007).
25. N. Iro, D. Gautier, F. Hersant, D. Bockelée-Morvan, J. I. Lunine, *Icarus* **161**, 511–532 (2003).
26. J. I. Lunine, D. J. Stevenson, *Astrophys. J.* **58**, 493–531 (1985).
27. K. Altwegg *et al.*, *Science* **347**, 1261952 (2015).
28. B. Marty, M. Chaussidon, R. C. Wiens, A. J. G. Jurewicz, D. S. Burnett, *Science* **332**, 1533–1536 (2011).



## ACKNOWLEDGMENTS

The authors thank the following institutions and agencies, which supported this work: Work at the University of Bern was funded by the State of Bern, the Swiss National Science Foundation, and the European Space Agency PRODEX Program. Work at the Max Planck Institute for Solar System Research was funded by the Max-Planck Society and Bundesministerium für Wirtschaft und Energie under contract 50QP1302. Work at the Southwest Research Institute was supported by subcontract no. 1496541 from the Jet Propulsion Laboratory (JPL). Work at BIRA-IASB was supported by the Belgian Science Policy Office via PRODEX/ROSINA PEA 90020. This work has been carried out thanks to the support of the A*MIDEX project (no. ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French government program, managed by the French National Research Agency (ANR). This work was supported by CNES grants at IRAP; LATMOS; LPC2E; Univers, Transport, Interfaces, Nanostructures, Atmosphère et Environnement, Molécules (UTINAM); and CRPG and by the European Research Council (grant no. 267255 to B.M.). A.B.-N. thanks the Ministry of Science and the Israel Space agency. Work at the University of Michigan was funded by NASA under contract JPL-1266313. Work by J.H.W. at the Southwest Research Institute was funded by NASA JPL subcontract NAS703001TONMO710889. The results from ROSINA would not be possible without the work of the many engineers, technicians, and scientists involved in the mission, in the Rosetta spacecraft, and in the ROSINA instrument team over the past 20 years, whose contributions are gratefully acknowledged. We thank herewith the work of the whole European Space Agency (ESA) Rosetta team. Rosetta is an ESA mission with contributions from its member states and NASA. All ROSINA data are available on request until they are released to the Planetary Science Archive of ESA and the Planetary Data System archive of NASA.


## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/232/suppl/DC1
Materials and Methods

12 January 2015; accepted 3 March 2015
Published online 19 March 2015;
10.1126/science.aaa6100

HUMAN GENETICS

# Common variants spanning *PLK4* are associated with mitotic-origin aneuploidy in human embryos

**Rajiv C. McCoy,[1] Zachary Demko,[2] Allison Ryan,[2] Milena Banjevic,[2] Matthew Hill,[2] Styrmir Sigurjonsson,[2] Matthew Rabinowitz,[2] Hunter B. Fraser,[1] Dmitri A. Petrov[1]**

**Aneuploidy, the inheritance of an atypical chromosome complement, is common in early human development and is the primary cause of pregnancy loss. By screening day-3 embryos during in vitro fertilization cycles, we identified an association between aneuploidy of putative mitotic origin and linked genetic variants on chromosome 4 of maternal genomes. This associated region contains a candidate gene, *Polo-like kinase 4* (*PLK4*), that plays a well-characterized role in centriole duplication and has the ability to alter mitotic fidelity upon minor dysregulation. Mothers with the high-risk genotypes contributed fewer embryos for testing at day 5, suggesting that their embryos are less likely to survive to blastocyst formation. The associated region coincides with a signature of a selective sweep in ancient humans, suggesting that the causal variant was either the target of selection or hitchhiked to substantial frequency.**

Deviation from a balanced chromosome complement, a phenomenon known as aneuploidy, is common in early human embryos and often leads to embryonic mortality (*1*). Approximately 75% of embryos are at least partially aneuploid by day 3 because of prevalent errors of both meiotic and postzygotic origin (*2*, *3*), and this proportion increases with maternal age (*1*). The propensity to produce aneuploid embryos varies substantially, however, even among mothers of a similar age (*4*). We therefore hypothesized that variation in parents' genomes may explain variation in aneuploidy incidence. We tested this hypothesis by performing a genome-wide association study of aneuploidy risk among patients undergoing preimplantation genetic screening (PGS) of embryos collected from in vitro fertilization (IVF) cycles.

Embryo DNA (single-cell day-3 blastomere biopsies or multicell day-5 trophectoderm biopsies) and parent DNA were genotyped on a single-nucleotide polymorphism (SNP) microarray (*5*). The Parental Support algorithm (*6*) was then applied to determine the chromosome-level ploidy status of each embryo sample. This algorithm overcomes high rates of allelic dropout and other quality limitations of whole-genome amplification by supplementing these data with high-quality genotypes from parental chromosomes. The copy number of each embryonic chromosome can then be inferred by comparing microarray channel intensities from DNA amplified from the embryo biopsy to those expected given the parental genotypes at each marker. Combining these fine-scale observations across large chromosomal windows facilitates the detection of particular forms of aneuploidy and the assignment of copy number variations to specific parental homologs (*6*).

[1]Department of Biology, Stanford University, Stanford, CA, USA. [2]Natera, Inc., San Carlos, CA, USA.

Previous validation has been performed for individual blastomeres (*6*), so it is unknown how accuracy would be affected in the face of chromosomal mosaicism that could potentially affect multicell trophectoderm biopsies. We therefore performed an association study on 2362 unrelated mothers (1956 IVF patients and 406 oocyte donors) and 2360 unrelated fathers meeting genotype quality-control thresholds (*5*) and from whom at least one day-3 biopsy was obtained, with the blastomere providing a high-confidence result (a total of 20,798 blastomeres). We then separately analyzed the additional 15,388 trophectoderm biopsies to gain insight into selection occurring before this developmental stage.

We first tested for associations between the rates of errors of putative maternal meiotic origin (fig. S1) (*5*) and maternal genotypes, identifying no association achieving genome-wide significance (logistic GLM, *P*-value threshold = $5 \times 10^{-8}$). We next tested for associations between the rates of errors of putative mitotic origin and parental genotypes. The first mitotic divisions of the developing embryo take place under the control of maternal gene products provided to the oocyte (*7*) and are substantially error-prone (*2*, *3*). We hypothesized that variation in maternal gene products may thus contribute to variation in rates of postzygotic error among embryos from different mothers. To encode the mitotic error phenotype, we designated all blastomeres with aneuploidies affecting a paternal chromosome copy (excluding paternal trisomies of putative meiotic origin) as cases, and all other blastomere samples as controls (Fig. 1A). Because aneuploidy has been estimated to affect fewer than 5% of sperm (*8*) and because paternal meiotic trisomies were detected for fewer than 1% of the blastomeres in our data, this set of aneuploid cases should be nearly exclusively mitotic in origin.

The 5438 putative mitotic-origin aneuploidies were predominantly characterized by a distinct error profile involving multiple chromosome losses (Fig. 1, B and C), and their incidence was not associated with maternal age (Fig. 1D). This excess of chromosome losses is consistent with previous studies that identified anaphase lag as the primary mechanism contributing to mosaicism in preimplantation embryos (*9*, *10*). Anaphase lag refers to the delayed migration of a chromosome during anaphase, so that the lagging chromosome fails to be incorporated into the reforming nucleus, resulting in chromosome loss with no corresponding chromosome gain (Fig. 1A). This error commonly arises as a consequence of merotelic kinetochore attachment: the attachment of a single kinetochore to microtubules emanating from both spindle poles (*11*). Merotelic attachment can in turn occur because of the presence of extra centrosomes or other centrosome abberations (*12*, *13*).

From our genome-wide analysis, we identified a peak on chromosome 4, regions q28.1 to q28.2, of maternal genomes associated with this mitotic-error phenotype (Fig. 2, C to E). The SNP rs2305957 was most strongly associated, with the minor allele conferring a significantly increased rate of mitotic error [logistic GLM, (regression coefficient) $\beta = 0.218$, standard error (SE) = 0.0270, $P = 8.68 \times 10^{-16}$]. The minor allele is present in diverse human populations at frequencies of 20 to 45% (fig. S2) (*14*). We observed no significant associations between paternal genotype and the same mitotic-error phenotype (logistic GLM, $P = 0.389$), which demonstrates that population stratification did not drive the significant association with maternal genotype (Fig. 2, A and B) (*5*). We also found that the association was robust when separately tested for mothers of European and East Asian ancestries (Table 1 and fig. S3).

The observed effect was characterized by means of 24.6, 27.0, and 31.7% of blastomeres affected with paternal-chromosome aneuploidies for the GG, AG, and AA maternal genotypic classes, respectively (Fig. 3A), and was consistent across age classes (Fig. 3D). The effect size from individual blastomeres may underestimate the overall effect on aneuploidy, because diploid blastomeres will be sampled by chance from some diploid-aneuploid mosaics. The frequencies of the three genotypes were not significantly different between mothers and fathers [$\chi^2(2, N = 9,418) = 1.17$, $P = 0.557$] or between egg donors and non-donors [$\chi^2(2, N = 4,712) = 2.49$, $P = 0.288$], which together suggest that this set of IVF patients was not enriched in the mitotic-error–associated genotypes.

For validation, genotypes from 34 additional unrelated mothers, representing new cases since the initial database pull, were tested for association with the same phenotype. Despite the small sample size ($N_{patients} = 34$, $N_{blastomeres} = 283$), the association was replicated, with 25.3, 35.7, and 51.3% of blastomeres with errors affecting paternal chromosomes among the three respective maternal genotypic classes (logistic GLM, $\beta = 0.589$, SE = 0.219, $P = 0.0112$; Fig. 3B).

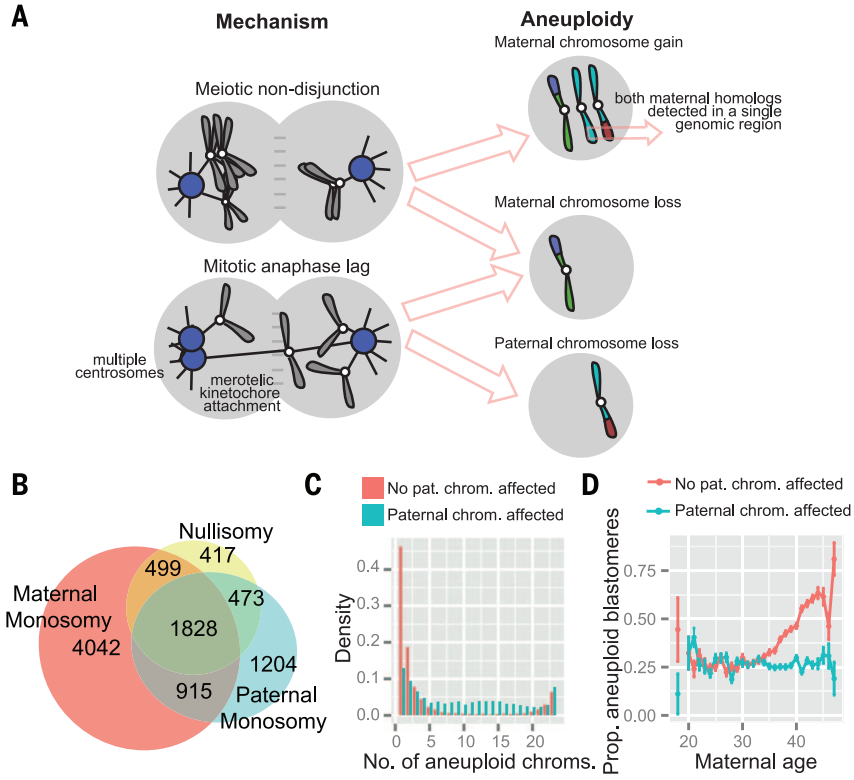



**Fig. 1. Mitotic-error phenotypes.** (**A**) Two mechanisms that frequently contribute to aneuploidy are depicted: maternal meiotic nondisjunction and mitotic anaphase lag. (**B**) Aneuploidies in which at least one paternal chromosome is affected are likely to be mitotic in origin and include an excess of chromosome losses as compared to chromosome gains, consistent with the signature of anaphase lag. Paternal chromosome loss (paternal monosomy) commonly co-occurs with other forms of chromosome loss, including maternal monosomy and nullisomy. (**C**) Blastomeres with aneuploidies affecting at least one paternal chromosome (blue; putative mitotic-origin aneuploidies) often contain multiple aneuploid chromosomes, in contrast to aneuploid blastomeres in which no paternal chromosome copies are affected (red; predominantly meiotic-origin aneuploidies). Heights of bars indicate densities (i.e., relative frequencies). (**D**) Aneuploidies in which at least one paternal chromosome copy is affected do not increase in frequency with increasing maternal age, whereas maternal aneuploidies increase in frequency beginning in the mid-30s. Error bars indicate SEs of proportions.

Highlighting its importance, genotype at rs2305957 was also a significant predictor of overall aneuploidy (logistic GLM, $\beta = 0.139$, SE = 0.0271, $P = 3.05 \times 10^{-7}$; Fig. 3E), especially for complex aneuploidies affecting more than two chromosomes (logistic GLM, $\beta = 0.234$, SE = 0.0329, $P = 1.72 \times 10^{-12}$; fig. S4). Means of 65.2, 68.3, and 71.4% of blastomeres per case were determined to be aneuploid for mothers with the GG, AG, and AA genotypes, respectively. This 6.2% difference in the proportion of aneuploid blastomeres between the two homozygous maternal genotype classes is roughly equivalent to the average effect of 1.8 years of age for mothers ≥ 35 years old (fig. S5).

Given that the association in our study was driven by complex aneuploidies affecting many chromosomes and that complex and mosaic aneuploidies are more likely to be inviable (*15*), we hypothesized that the arrest of aneuploid embryos would bias the genotypic ratios at associated SNPs for 15,388 embryos sampled at the day-5 blastocyst stage from 2998 unrelated mothers. Patients with the mitotic-error–associated genotypes at rs2305957 contributed significantly fewer trophectoderm biopsies for testing (Poisson GLM, $\beta = -0.0619$, SE = 0.0204, $P = 0.00247$; Fig. 3C), consistent with an increased proportion of inviable aneuploidies. Together these findings suggest that the mitotic-error association may affect fertility in such a way that it may take longer, on average, for women with the associated genotypes to achieve successful pregnancies.

In order to characterize the extent of the associated region, we performed genotype imputation for a subset of 1332 patients of European ancestry (*5*). The associated haplotype lies in a region of low recombination and spans over 600 Kbp of chromosome 4, regions q28.1 to q28.2 (Fig. 2E), including the genes *INTU*, *SLC25A31*, *HSPA4L*, *PLK4*, *MFSD8*, *LARP1B*, and *PGRMC2*. Although none of these candidates can yet be ruled out, we focused on *PLK4* on the basis of its well-characterized role as the master regulator of centriole duplication, a key component of the centrosome cycle (*16*, *17*). In addition, it was

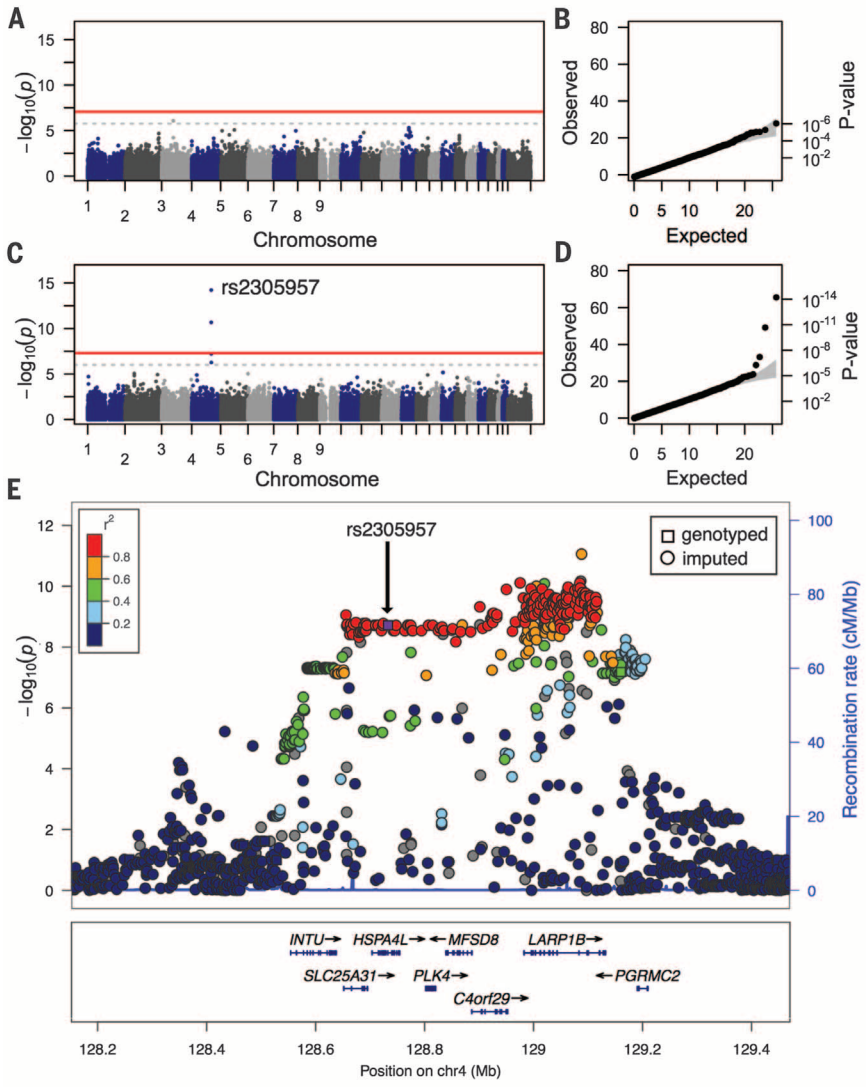



**Fig. 2. Association test for aneuploidy.** (A) to (D) Manhattan and QQ plots depicting *P* values of association tests of each genotyped SNP versus the rate of aneuploidy affecting paternal chromosomes (a proxy for mitotic aneuploidy). *P* values are corrected using the genomic control method (*5*). (**A**) Results for association with paternal genotypes, a negative control. (**B**) QQ plot of the distribution of observed *P* values versus those expected under the null. (**C**) Association with maternal genotypes, with rs2305957 highlighted as the most significant genotyped SNP. (**D**) QQ plots of *P* values. For (A) and (C), the red lines represent a standard genome-wide cutoff of $5 \times 10^{-8}$, whereas the gray dotted lines represent a less stringent *P* value of $1 \times 10^{-6}$. For (B) and (D), the gray shaded regions indicate probability bounds. (**E**) Regional association plot for mothers of European ancestry, inferred by comparison to reference populations (fig. S1). rs2305957 is indicated (purple point below arrow), whereas the colors of other variants represent linkage disequilibrium with rs2305957 (*5*).

recently demonstrated that *PLK4* is essential for mediating bipolar spindle formation during the first cell divisions in mouse embryos, which take place in the absence of centrioles (*18*).

Due in part to the observation that centrosome aberrations and aneuploidies are common in human cancers, the role of *PLK4* and its orthologs in mediating the centrosome cycle has been investigated in several model systems. *PLK4* is a tightly regulated, low-abundance kinase with a short half-life (*19*). Overexpression of *PLK4* results in centriole overduplication, thereby increasing the frequency of multipolar spindle formation and subsequent anaphase lag (*12*). Reduced expression of *PLK4* results in centriole loss (*17*), which also leads to multipolar spindle formation, as well as the formation of monopolar spindles. Both up- and down-regulation of *PLK4* therefore have the potential to induce chromosome instability, and altered *PLK4* expression is commonly observed in several forms of cancer, which is consistent with a tumor-suppressor function (*20*, *21*).

Along with hundreds of variants upstream and downstream of *PLK4*, the associated region contains two nonsynonymous SNPs within the *PLK4* coding sequence: rs3811740 (S232T) and rs17012739 (E830D), the former occurring in the protein's kinase domain and the latter occurring in the crypto Polo-box domain (*22*). Neither site exhibits strong conservation over deep evolutionary time, and both SNPs were predicted to be benign on the basis of sequence conservation, amino acid similarity, and mapping to three-dimensional protein structure (*5*).

Prompted by the observation that the minor allele of rs2305957 is derived and segregates at intermediate frequencies in diverse human populations, yet is absent from Neandertal and Denisovan genomes (*5*), we investigated whether the region showed evidence of positive selection in humans. Unfortunately, classic frequency spectrum–based tests have sensitivity over the order of $N_e$ generations, capturing only relatively recent human evolutionary history (~10,000 generations). We thus examined results of the selection scan from (*23*), which has resolution to detect signatures of ancient selective sweeps in the human lineage by identifying regions of aligned Neandertal genomes that are deficient in high-frequency human derived alleles. The mitotic-error–associated region identified in our study is among the 212 previously identified regions displaying such a signature (*23*). This finding suggests that either this seemingly deleterious allele hitchhiked with a linked adaptive variant or that the causal variant was adaptive in a context that is not currently understood.

The fact that the haplotype bearing the derived allele did not sweep to fixation and is present at similar frequencies across human populations is consistent with the action of long-term balancing selection. We speculate that the mitotic-error phenotype may be maintained by conferring both a deleterious effect on maternal fecundity and a possible beneficial effect of obscured paternity via a reduction in the probability of successful pregnancy per intercourse. This hypothesis is based on the fact that humans possess a suite of traits (such as concealed ovulation and constant receptivity) that obscure paternity and may have evolved to increase paternal investment in offspring (*24*). Such a scenario could result in balancing selection by rewarding evolutionary "free riders" who do not possess the risk allele—and thus do not suffer fecundity costs—but benefit from paternity confusion in the population as a whole.

Mitotic fidelity is affected by variation in maternal gene products controlling the initial cell divisions of preimplantation embryos. This finding is important in the context of IVF, where the selection of euploid embryos may improve the success rate of implantation and ongoing pregnancy (*25*). More broadly, factors influencing variation in rates of aneuploidy may help explain variation in fertility status among the general population. Fewer than ~30% of conceptions result in successful pregnancy, mostly due to high rates of inviable aneuploidy in early development (*26*). By altering this rate, the associated locus described in our study may influence the average time required to achieve successful pregnancy, which could be especially important for couples with already-reduced fertility. The

**Table 1. Association of SNP rs2305957 with the rate of putative mitotic-origin aneuploidy.** Sample size, β, SE, odds ratio (OR), genomic inflation factor (λ), and *P* values are reported. CI, confidence interval; NA, not applicable. The upper row gives results of an association test of all female patients, including those falling outside of the European and East Asian principal components boundaries. The middle two rows control for potential population stratification by separating analyses of female patients with a high proportion of European or East Asian ancestry, respectively.

| | Sample size | | β | SE | OR (95% CI) | Uncorrected | | Genomic control |
|---|---|---|---|---|---|---|---|---|
| | Patients | Embryos | | | | λ | *P* | *P* |
| Discovery | 2362 | 20,798 | 0.218 | 0.0270 | 1.244 (1.179–1.311) | 1.059 | $8.68 \times 10^{-16}$ | $5.99 \times 10^{-15}$ |
| Europe | 1332 | 11,861 | 0.214 | 0.0353 | 1.238 (1.155–1.327) | 1.066 | $1.91 \times 10^{-9}$ | $6.67 \times 10^{-9}$ |
| East Asia | 259 | 2222 | 0.280 | 0.0788 | 1.323 (1.133–1.543) | 1.088 | $4.58 \times 10^{-4}$ | $8.51 \times 10^{-4}$ |
| Validation | 34 | 283 | 0.589 | 0.219 | 1.802 (1.173–2.768) | NA | 0.0112 | NA |

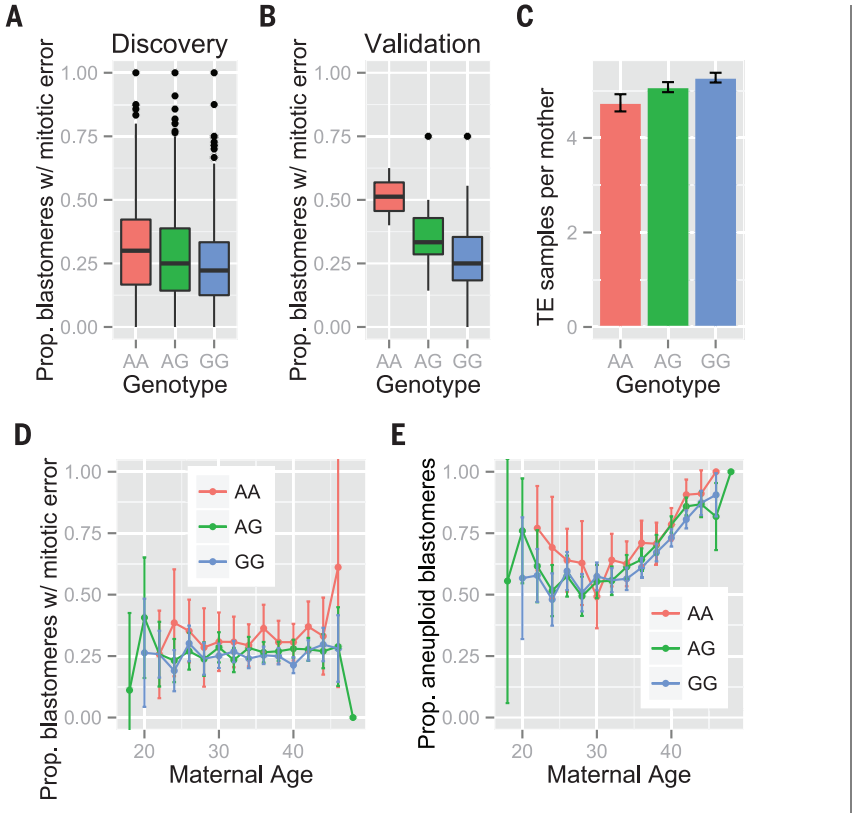



**Fig. 3. Effect of genotype on mitotic-error–related phenotypes.** For box plots, we restricted figures to include only mothers for whom >2 embryos were tested. (**A**) The proportion of blastomeres per mother with an error affecting a paternal chromosome (a proxy for mitotic aneuploidy) stratified by maternal genotype at rs2305957 for the discovery sample ($N_{patients}$ = 2362, $N_{embryos}$ = 20,798; $P = 8.68 \times 10^{-16}$). (**B**) The same phenotype as in (A), replicated in the validation sample ($N_{patients}$ = 34, $N_{embryos}$ = 283; $P$ = 0.0112). (**C**) Mean number of day-5 trophectoderm biopsies per mother, stratified by genotype at rs2305957 (Poisson GLM, $P$ = 0.00247). Error bars represent the SE. (**D**) The mean proportion of blastomeres with an aneuploidy affecting a paternal chromosome versus maternal age, stratified by genotype at rs2305957. Error bars represent the SE of the proportion. (**E**) The mean proportion of aneuploid blastomeres versus maternal age, stratified by genotype at rs2305957. Error bars represent the SE of the proportion.

identification of genetic variation influencing rates of aneuploidy is an important step in the understanding of aneuploidy risk and may assist the future development of diagnostic or therapeutic technologies targeting certain forms of infertility.

## REFERENCES AND NOTES


1. T. Hassold, P. Hunt, *Nat. Rev. Genet.* **2**, 280–291 (2001).
2. L. Voullaire, H. Slater, R. Williamson, L. Wilton, *Hum. Genet.* **106**, 210–217 (2000).
3. D. Wells, J. D. Delhanty, *Mol. Hum. Reprod.* **6**, 1055–1062 (2000).
4. J. D. Delhanty, J. C. Harper, A. Ao, A. H. Handyside, R. M. Winston, *Hum. Genet.* **99**, 755–760 (1997).
5. Materials and methods are available as supplementary materials on *Science* Online.
6. D. S. Johnson *et al.*, *Hum. Reprod.* **25**, 1066–1075 (2010).
7. W. Tadros, H. D. Lipshitz, *Development* **136**, 3033–3042 (2009).
8. C. Templado, F. Vidal, A. Estop, *Cytogenet. Genome Res.* **133**, 91–99 (2011).
9. E. Coonen *et al.*, *Hum. Reprod.* **19**, 316–324 (2004).
10. D. D. Daphnis *et al.*, *Hum. Reprod.* **20**, 129–137 (2005).
11. J. Gregan, S. Polakova, L. Zhang, I. M. Tolić-Nørrelykke, D. Cimini, *Trends Cell Biol.* **21**, 374–381 (2011).
12. N. J. Ganem, S. A. Godinho, D. Pellman, *Nature* **460**, 278–282 (2009).
13. A. J. Holland, D. W. Cleveland, *Nat. Rev. Mol. Cell Biol.* **10**, 478–487 (2009).
14. G. R. Abecasis *et al.*, *Nature* **467**, 1061–1073 (2010).
15. M. Vega, A. Breborowicz, E. L. Moshier, P. G. McGovern, M. D. Keltz, *Fertil. Steril.* **102**, 394–398 (2014).
16. R. Habedanck, Y.-D. Stierhof, C. J. Wilkinson, E. A. Nigg, *Nat. Cell Biol.* **7**, 1140–1146 (2005).
17. M. Bettencourt-Dias *et al.*, *Curr. Biol.* **15**, 2199–2207 (2005).
18. P. A. Coelho *et al.*, *Dev. Cell* **27**, 586–597 (2013).
19. E. N. Fırat-Karalar, T. Stearns, *Philos. Trans. R. Soc. London Ser. B* **369**, 20130460 (2014).
20. M. A. Ko *et al.*, *Nat. Genet.* **37**, 883–888 (2005).
21. C. O. Rosario *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6888–6893 (2010).
22. J. E. Sillibourne, M. Bornens, *Cell Div.* **5**, 25 (2010).
23. R. E. Green *et al.*, *Science* **328**, 710–722 (2010).
24. R. D. Alexander, K. M. Noonan, in *Evolutionary Biology and Human Social Organization*, N. A. Chagnon, W. G. Irons, Eds. (Duxbury Press, North Scituate, MA, 1979), pp. 436–453.
25. R. T. Scott Jr. *et al.*, *Fertil. Steril.* **100**, 697–703 (2013).
26. A. J. Wilcox *et al.*, *N. Engl. J. Med.* **319**, 189–194 (1988).


## ACKNOWLEDGMENTS


Thanks to C. Hin and J. Layne for computing support; E. Sharon for advice regarding analyses; R. Taylor and other members of the Petrov lab for helpful input; J. Sage, J. Arand, T. Stearns, and O. Cormier for advice regarding functional annotation; and C. Boggs for comments on the manuscript. D.A.P. devised the project. D.A.P., H.B.F., and Z.D. provided guidance throughout the project. R.C.M. performed analyses and wrote the manuscript. Z.D., A.R., M.B., M.H., S.S., and M.R. helped generate the data and design the algorithm by which aneuploidies are detected. All authors read the manuscript and provided comments. D.A.P. has received stock options in Natera, Inc., as consulting fees. Z.D., A.R., M.B., M.H., S.S., and M.R. are full-time employees of and hold stock or options to hold stock in Natera, Inc. De-identified aneuploidy outcome data for all blastomeres are uploaded as supplemental material, as are genome-wide association study summary statistics. Patient genotype data for the associated region are archived at Natera, which will cooperate with qualified researchers attempting to replicate the findings of this study, conditional on institutional research board approval and adherence to a materials transfer agreement. Stanford University filed a provisional patent related to this work with the U.S. Patent and Trademark Office on 14 November 2014 (USSN 62/080,251). This work was partly supported by grants R01 GM100366, R01 GM097415, and R01 GM089926 to D.A.P. Correspondence regarding the data should be addressed to Z.D. (zdemko@natera.com); other correspondence should be addressed to R.C.M. (rmccoy@stanford.edu) or D.A.P. (dpetrov@stanford.edu).


## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/235/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S5
Table S1
References (*27*–*47*)
Additional Data Files


19 November 2014; accepted 23 February 2015
10.1126/science.aaa3337

PROTEOSTASIS

# Preventing proteostasis diseases by selective inhibition of a phosphatase regulatory subunit

Indrajit Das,[1] Agnieszka Krzyzosiak,[1] Kim Schneider,[1] Lawrence Wrabetz,[2]* Maurizio D'Antonio,[2] Nicholas Barry,[1] Anna Sigurdardottir,[1] Anne Bertolotti[1]†

**Protein phosphorylation regulates virtually all biological processes. Although protein kinases are popular drug targets, targeting protein phosphatases remains a challenge. Here, we describe Sephin1 (selective inhibitor of a holophosphatase), a small molecule that safely and selectively inhibited a regulatory subunit of protein phosphatase 1 in vivo. Sephin1 selectively bound and inhibited the stress-induced PPP1R15A, but not the related and constitutive PPP1R15B, to prolong the benefit of an adaptive phospho-signaling pathway, protecting cells from otherwise lethal protein misfolding stress. In vivo, Sephin1 safely prevented the motor, morphological, and molecular defects of two otherwise unrelated protein-misfolding diseases in mice, Charcot-Marie-Tooth 1B, and amyotrophic lateral sclerosis. Thus, regulatory subunits of phosphatases are drug targets, a property exploited here to safely prevent two protein misfolding diseases.**

A first line of defense against the accumulation of misfolded proteins in the endoplasmic reticulum (ER) consists of phosphorylating the α subunit of eukaryotic translation initiation factor 2 (eIF2α) on $Ser^{51}$ to decrease protein synthesis, an adaptive stress response essential for survival (*1*–*3*). Guanabenz (GBZ) can prolong this adaptive response by selectively binding and inhibiting the regulatory subunit of the stress-induced eIF2α phosphatase composed of PPP1R15A and PP1c (*4*). GBZ spares the constitutive eIF2α phosphatase PPP1R15B-PP1c, avoiding persistent eIF2α phosphorylation, which would be lethal (*5*). As a result, GBZ increases the availability of chaperones to misfolded proteins and consequently rescues cells from proteostasis collapse (*4*). Correcting proteostasis defects could, in theory, benefit a broad range of diseases characterized by the accumulation of misfolded proteins (*6*). However, identifying a therapeutically valuable approach for progressive diseases represents a double challenge: achieving efficacy without adverse effects. In vivo, GBZ cannot be used to selectively inhibit PPP1R15A because it is a centrally active hypotensive drug with nanomolar affinity for the α2-adrenergic receptor (*7*).

Searching for PPP1R15A inhibitors devoid of α2-adrenergic activity, we synthesized GBZ derivatives and identified Sephin1 (selective inhibitor of a holophosphatase) (Fig. 1A). Like GBZ (*4*), Sephin1 specifically bound a recombinant fragment of PPP1R15A (amino acids 325 to 636) but not the highly related PPP1R15B (amino acids 340 to 698) (Fig. 1B). In cells, Sephin1 selectively disrupted the PPP1R15A-PP1c complex but spared the related PPP1R15B-PP1c complex (Fig. 1C). As a result, Sephin1 prolonged eIF2α phosphorylation after stress (Fig. 1D), delaying translation recovery (Fig. 1E). Consequently, Sephin1 attenuated expression of stress genes such as CHOP, a pro-apoptotic protein (Fig. 1D and fig. S1), because stress gene expression requires translation recovery (*8*, *9*). Activating transcription factor 4 (ATF4) is selectively translated upon stress when eIF2α is phosphorylated (*10*), and this was

[1]Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge, CB2 0QH, UK. [2]Division of Genetics and Cell Biology, San Raffaele Scientific Institute, 20132 Milan, Italy.
*Present address: Hunter James Kelly Research Institute (HJKRI), University at Buffalo School of Medicine and Biomedical Sciences, 701 Ellicott Street, Buffalo, NY 14203, USA. †Corresponding author. E-mail: aberto@mrc-lmb.cam.ac.uk

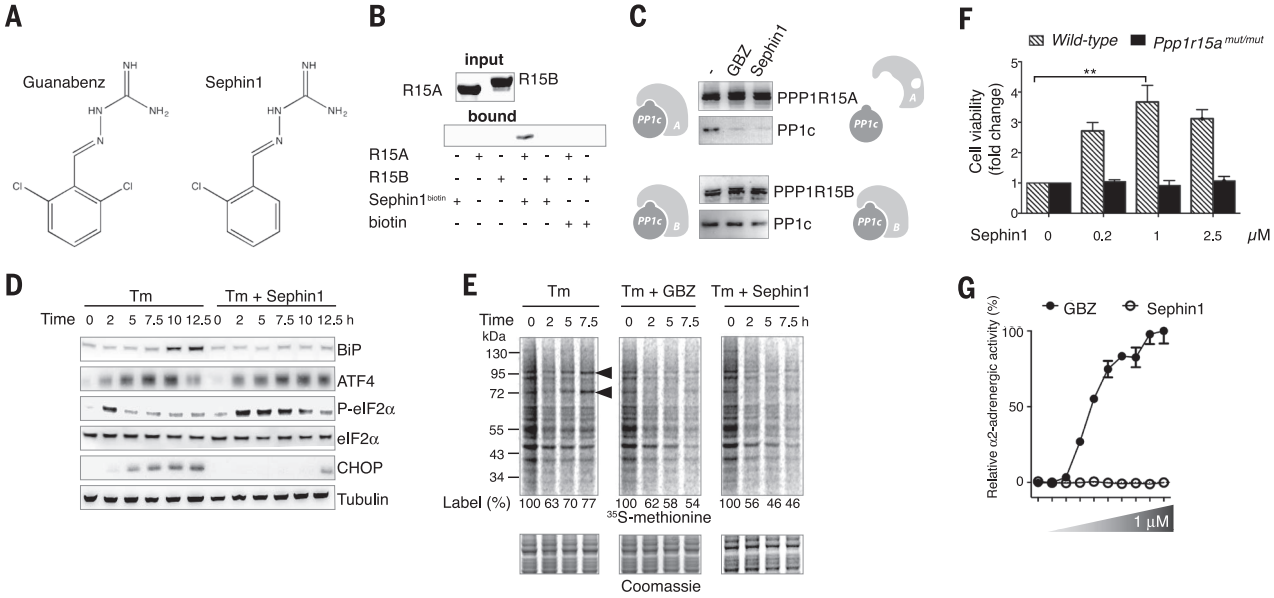



**Fig. 1. Sephin1 is a selective inhibitor of PPP1R15A.** (**A**) Structures of GBZ and Sephin1. (**B**) Coomassie-stained gel showing recombinant $PPP1R15A^{325-636}$ and $PPP1R15B^{340-698}$ (input). Biotinylated Sephin1 selectively captured PPP1R15A on neutravidin beads (bound). (**C**) Immunoprecipitations of PPP1R15 complexes from cells treated with vehicle, 50 μM GBZ, or Sephin1 for 6 hours analyzed with immunoblotting. (**D**) Immunoblots of the indicated proteins in lysates of HeLa cells treated with 2.5 μg/ml tunicamycin (Tm) in the presence or absence of 50 μM Sephin1 for the indicated time. (**E**) Newly synthesized proteins labeled with $^{35}$S-methionine in HeLa cells treated with Tm or with or without 50 μM GBZ or Sephin1 and revealed with autoradiography. (Bottom) Coomassie-stained gel. (**F**) Dose-dependent protection by Sephin1 of wild-type but not *Ppp1r15a* mutant (*mut/mut*) cells from Tm (2.5 μg/ml). Data are means ± SEM ($n$ = 4 replicates). $^{**}P \leq 0.001$. (**G**) Adrenergic activity of GBZ and Sephin1 in cells expressing recombinant human adrenergic α2A receptor. Representative results of at least three independent experiments are shown in each panel.

prolonged in Sephin1-treated cells (Fig. 1D). In the absence of stress, Sephin1 did not affect eIF2α signaling (fig. S2), which was expected because PPP1R15A is only expressed upon stress (*9*). Confirming the selectivity of Sephin1 for PPP1R15A, Sephin1 did not inhibit the catalytic subunit PP1c (fig. S3). Selective inhibition of PPP1R15A by Sephin1 protected cells from cytotoxic ER stress (Fig. 1F and fig. S4), but this was abolished in cells lacking a functional allele of *Ppp1r15a* (Fig. 1F). Thus, all the cytoprotective activity of Sephin1 in ER-stressed cells resulted from selective inhibition of PPP1R15A. Sephin1 lacked any measurable α2-adrenergic agonist activity in a cell-based assay, in contrast to GBZ (Fig. 1G). Thus, Sephin1 is a selective PPP1R15A inhibitor.

Pharmacokinetic analysis of Sephin1, administered orally at 1 or 10 mg/kg, revealed that the molecule rapidly disappeared from plasma (fig. S5) but concentrated in the nervous system, reaching concentrations 7 to 44 times higher in

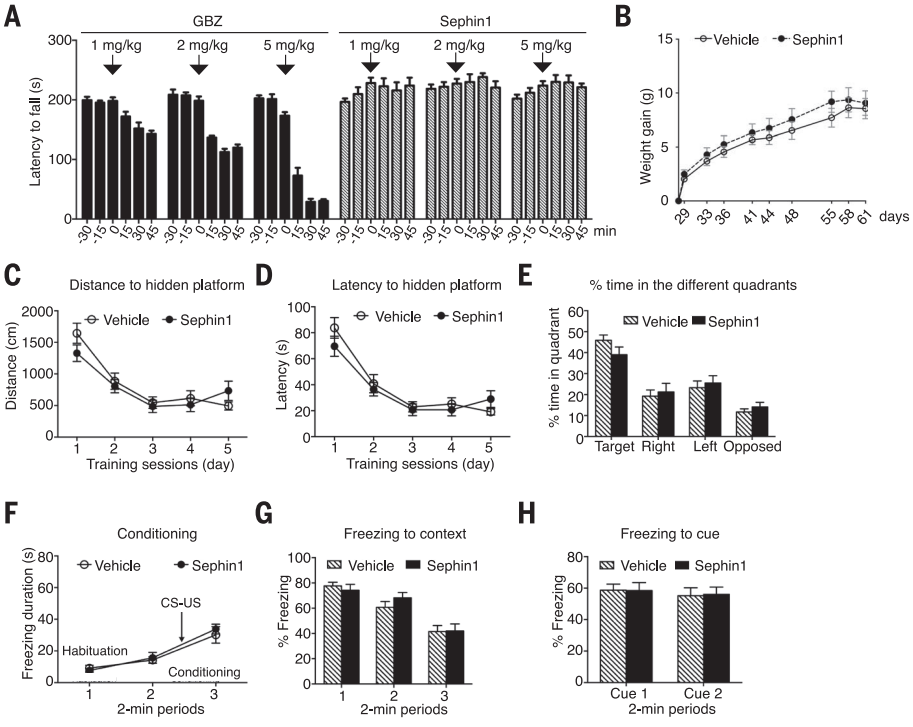



**Fig. 2. Sephin1 is devoid of adverse effects on rotarod performances, total body weight gain, or memory.** (**A**) Motor performance of mice on an accelerating rotarod before or after the indicated treatments. Data are means ± SEM. (*n* = 5 mice). (**B**) Total body weight gain of mice treated orally with Sephin1 (1 mg/kg) or vehicle twice a day, from postnatal days 28 to 61. Data are means ± SEM (*n* = 6 mice). (**C** and **D**) Distance and latency to locate a hidden platform in the Morris water maze, in five trials a day for 5 consecutive days. (**E**) Quadrant occupancy after training and removal of the platform. (**F**) Freezing response during the conditioning session, where a light/tone [conditioned stimulus (CS)] and foot shock [aversive unconditioned stimulus (US)] were applied. (**G** and **H**) Freezing responses expressed as percentage of total time (2 min) mice spent immobile during context and auditory cue testing. In (C) to (H), data are means ± SEM (*n* = 12 mice). Mice were treated with 1 mg/kg Sephin1 or vehicle twice a day for 4 weeks. No statistical differences were found between Sephin1- or vehicle-treated mice [(B) to (H)].

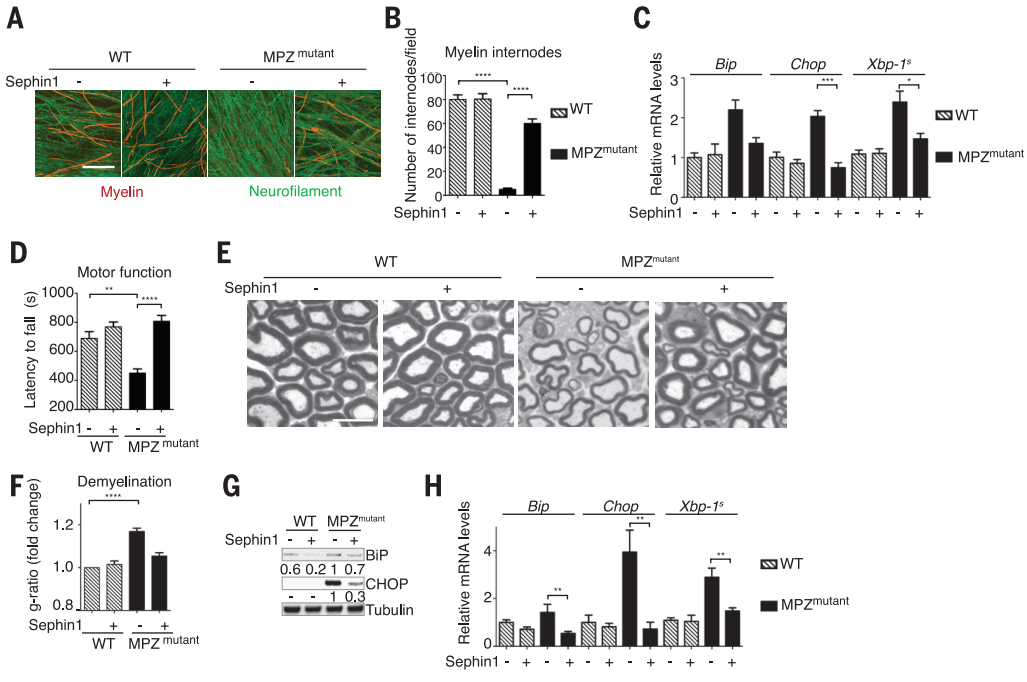



**Fig. 3. Sephin1 prevents the defects caused by misfolding-prone myelin MPZ$^{mutant}$ protein ex vivo and in vivo.** (**A** and **B**) Representative images (A) and quantification (B) of myelin internodes revealed with immuno staining for myelin basic protein (MBP, red) from cultured DRG of the indicated genotype treated for 2 weeks with vehicle or Sephin1 (100 nM). In green, neurofilament (NF). Data are means ± SEM (*n* = 4 to 6 mice). (**C**) mRNA levels [by means of quantitative polymerase chain reaction (PCR)] of the indicated ER stress markers in cultures as (A). Data are means ± SEM (*n* = 5 to 7 mice). (**D**) Motor performance of 4-month-old wild-type or MPZ$^{mutant}$ mice on a rotarod after a 3-month oral treatment with Sephin1 (1 mg/kg) or vehicle twice a day. Data are means ± SEM (*n* = 14 to 16 mice). (**E**) Myelin thickness revealed on toluidine blue–stained semithin sciatic nerve sections of 6-month-old mice of the indicated genotype after 5 months of oral Sephin1 treatment (1 mg/kg) or vehicle twice a day. Data are means ± SEM (*n* = 6 to 8 mice). (**F**) Demyelination expressed as increased *g*-ratio (ratio of axon diameter to fiber diameter) over wild-type mice. Data are means ± SEM (*n* = 5 or 6 mice). (**G**) Immunoblots on sciatic nerve lysates from mice. (**H**) mRNA levels (quantitative PCR) in sciatic nerves of 6-month-old mice as in (E). Data are means ± SEM (*n* = 5 to 7 mice). *$P \le 0.05$; **$P \le 0.01$; ***$P \le 0.001$; ****$P \le 0.0001$. Scale bars, 100 μm.

the brain and sciatic nerve (up to ~1 μM) than in the plasma (fig. S5, A and B), like GBZ does (*11*). In humans, the adrenergic agonist activity of GBZ has side effects, including drowsiness and coma, at high doses (*12*). In mice, GBZ (1 to 5 mg/kg) also exhibited side effects manifested by a rapid and dose-dependent decrease in rotarod performances (Fig. 2A). In contrast, Sephin1-treated mice (1 to 5 mg/kg) continued to run as before treatment (Fig. 2A). A chronic treatment (1 mg/kg for 1 month) with Sephin1 was also tolerable with no measurable adverse effects on body weight gain in mice (Fig. 2B). We next evaluated whether Sephin1 had adverse effects on memory because manipulations of PP1c (*13*) and eIF2α phosphorylation (*14*) affect memory. In the Morris water maze (*15*), Sephin1-treated mice showed normal spatial learning and improved their ability to locate a submerged platform as training progressed (Fig. 2, C and D) and remembered where the platform was after it had been removed (Fig. 2E). In a fear conditioning paradigm, both Sephin1- and vehicle-treated mice showed similar basal fear responses during the conditioning phase (Fig. 2F) and after (Fig. 2, G and H). Thus, Sephin1 lacks the adverse effects of GBZ in vivo and has no measurable adverse effect on general health or memory in diverse experimental paradigms.

We next examined whether Sephin1 could correct a protein misfolding disease. Deletion of serine *63* of myelin protein zero ($MPZ^{mutant}$), a transmembrane protein produced by Schwann cells in the peripheral nervous system, causes the demyelinating neuropathy Charcot-Marie-Tooth 1B (CMT1B) in humans and a similar disorder in mice. $MPZ^{mutant}$ causes CMT1B by a gain of toxic property associated with pathological signaling through CHOP and PPP1R15A (*16*, *17*). As previously reported (*16*, *17*), myelination defects were severe in dorsal root ganglia (DRG) cultures prepared from $MPZ^{mutant}$ mouse embryos (Fig. 3, A and B). A 2-week treatment with 100 nM of Sephin1 rescued myelination (Fig. 3, A and B) and decreased the expression of ER-stress genes in mutant DRG cultures (Fig. 3C). Because of the potency of Sephin1 ex vivo (Fig. 3, A to C) and its pharmacokinetic properties (fig. S5), we next treated $MPZ^{mutant}$ mice orally twice a day with 1 mg/kg of Sephin1. As reported (*18*), $MPZ^{mutant}$ mice exhibited motor defects detectable with rotarod analysis at 4 months of age (Fig. 3D), but this was completely prevented by Sephin1 (Fig. 3D). Sephin1 also rescued myelin thickness around axons in sciatic nerves (Fig. 3, E and F) and reduced the levels of ER-stress markers in $MPZ^{mutant}$ sciatic nerves (Fig. 3, G and H). Thus, without any obvious adverse effect, Sephin1 prevented the molecular, morphological, and motor defects of the $MPZ^{mutant}$ mice.

Selective inhibition of PPP1R15A could in principle ameliorate a broad range of protein misfolding diseases. Mutant and misfolding-prone superoxide dismutase 1 (SOD1) is associated with familial forms of amyotrophic lateral sclerosis (fALS), a motor neuron disease (*19*). SOD1 mutants bind to Derlin-1 on the cytosolic side of the ER membrane, blocking degradation of ER proteins and causing ER stress (*20*) with pathological PPP1R15A signaling (*21*). Transgenic mice expressing the human ALS-causing mutant $SOD1^{G93A}$ ($SOD1^{mutant}$ mice) develop a motor neuron disease that closely resembles ALS (*22*). The motor deficits in $SOD1^{mutant}$ mice were partially prevented by treatment with 1 mg/kg of Sephin1 twice a day (fig. S6). With 5 mg/kg of Sephin1 once a day, the progressive weight loss of $SOD1^{mutant}$ mice (Fig. 4A) as well as their motor deficits (Fig. 4B) were almost completely prevented, without adverse effects on weight gain or motor performance of wild-type mice (Fig. 4, A and B). The motor deficits of $SOD1^{mutant}$ mice were associated with motor neuron loss, which was prevented by Sephin1 (Fig. 4, C and D). Genetic ablation or pharmacological inhibition of PPP1R15A decreases translation rates and increases the availability of chaperones (*4*, *23*). To assess whether Sephin1 could improve proteostasis in the cytosol, we monitored aggregation of $SOD1^{mutant}$, a defining histopathological feature in ALS (*22*, *24*). Sephin1 prevented the accumulation of insoluble $SOD1^{mutant}$ (Fig. 4E) and decreased ER stress markers in transgenic spinal cords (Fig. 4F). Thus, Sephin1 prevented the molecular and organismal defects of $SOD1^{mutant}$ mice.

Here, we have shown that Sephin1 selectively inhibited PPP1R15A to prevent two otherwise unrelated protein misfolding diseases in mice. This was achieved while sparing PPP1R15B, a crucial property because the lack of both PPP1R15A and PPP1R15B is lethal in mice (*25*). Thus, PPP1R15A

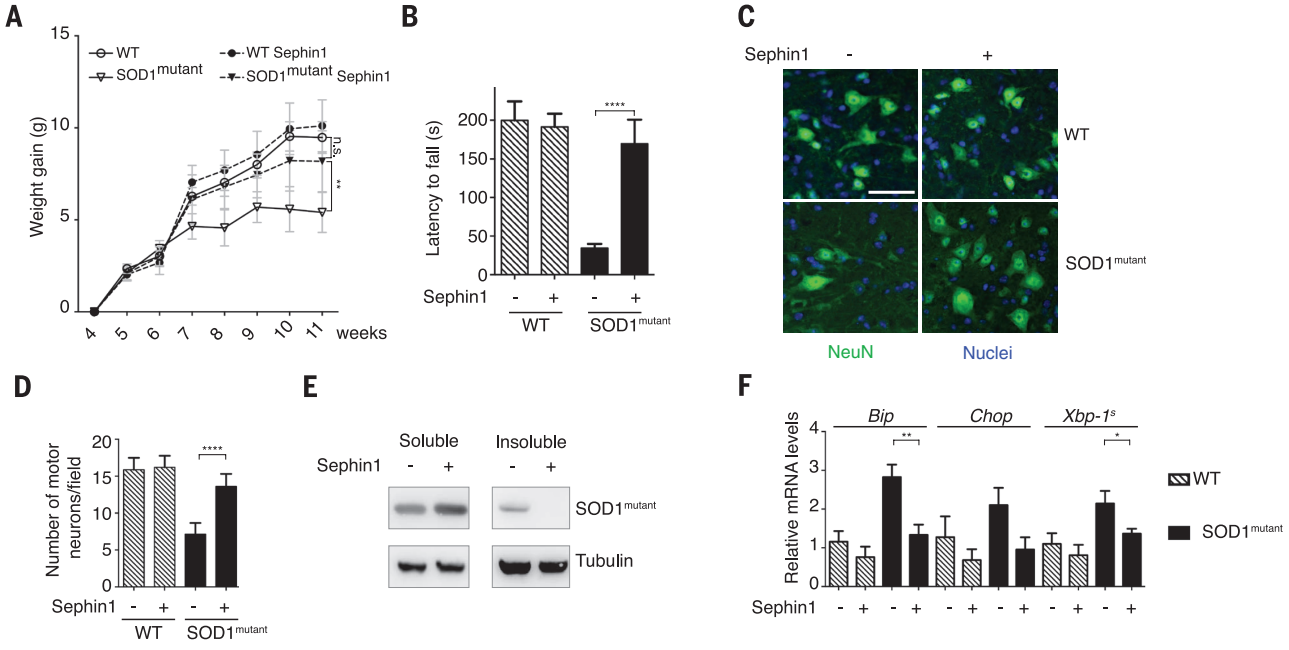



**Fig. 4. Sephin1 prevents motor deficits, motor neuron loss, and the molecular defects in $SOD1^{mutant}$ mice.** (**A**) Total body weight gain of wild-type or $SOD1^{mutant}$ mice treated orally with Sephin1 (5 mg/kg) or vehicle once a day from 4 to 11 weeks of age. Data are means ± SEM ($n$ = 4 to 6 mice). (**B**) Rotarod analysis of 110-day-old wild-type or $SOD1^{mutant}$ mice treated as in (A). Data are means ± SEM ($n$ = 4 to 6 mice). (**C** and **D**) Representative motor neuron staining (NeuN, green) and quantification (D) of sections of anterior horn of the lumbar region of spinal cord of 110-day-old mice treated as in (A). Nuclei, H33258 (blue). (**D**) Data are means ± SEM ($n$ = 6 to 8 mice). (**E**) SOD1 immunoblots on soluble and insoluble fractions of spinal cord extracts from $SOD1^{mutant}$ mice treated as in (A). (**F**) mRNA levels (quantitative PCR) in lumbar spinal cord from 4-month-old mice of indicated genotype, after treatment with Sephin1 or vehicle as in (A). Data are means ± SEM ($n$ = 4 to 6 mice). *$P \le 0.05$; **$P \le 0.01$; ***$P \le 0.001$. Scale bar, 50 μm.

inhibitors could ameliorate a broad range of diseases caused by accumulation of misfolded proteins. Because many signaling pathways operate on the same dynamic—phosphorylation in the activation phase terminated by dephosphorylation—delaying the termination phase of signaling pathways through the selective inhibition of phosphatases may be of broad relevance to safely and selectively manipulate cellular functions for therapeutic benefit.

## REFERENCES AND NOTES


1. H. P. Harding, Y. Zhang, A. Bertolotti, H. Zeng, D. Ron, *Mol. Cell* **5**, 897–904 (2000).
2. D. Scheuner *et al.*, *Mol. Cell* **7**, 1165–1176 (2001).
3. H. P. Harding *et al.*, *Mol. Cell* **11**, 619–633 (2003).
4. P. Tsaytler, H. P. Harding, D. Ron, A. Bertolotti, *Science* **332**, 91–94 (2011).
5. D. Scheuner *et al.*, *J. Biol. Chem.* **281**, 21458–21468 (2006).
6. W. E. Balch, R. I. Morimoto, A. Dillin, J. W. Kelly, *Science* **319**, 916–919 (2008).
7. B. Holmes, R. N. Brogden, R. C. Heel, T. M. Speight, G. S. Avery, *Drugs* **26**, 212–229 (1983).
8. M. A. Brostrom, X. J. Lin, C. Cade, D. Gmitter, C. O. Brostrom, *J. Biol. Chem.* **264**, 1638–1643 (1989).
9. I. Novoa *et al.*, *EMBO J.* **22**, 1180–1187 (2003).
10. H. P. Harding *et al.*, *Mol. Cell* **6**, 1099–1108 (2000).
11. R. H. Meacham *et al.*, *J. Pharmacol. Exp. Ther.* **214**, 594–598 (1980).
12. A. H. Hall, S. C. Smolinske, K. W. Kulig, B. H. Rumack, *Ann. Intern. Med.* **102**, 787–788 (1985).
13. D. Genoux *et al.*, *Nature* **418**, 970–975 (2002).
14. M. Costa-Mattioli *et al.*, *Cell* **129**, 195–206 (2007).
15. R. G. M. Morris, P. Garrud, J. N. P. Rawlins, J. O'Keefe, *Nature* **297**, 681–683 (1982).
16. M. Pennuto *et al.*, *Neuron* **57**, 393–405 (2008).
17. M. D'Antonio *et al.*, *J. Exp. Med.* **210**, 821–838 (2013).
18. L. Wrabetz *et al.*, *J. Neurosci.* **26**, 2358–2368 (2006).
19. A. Nordlund, M. Oliveberg, *HFSP J.* **2**, 354–364 (2008).
20. H. Nishitoh *et al.*, *Genes Dev.* **22**, 1451–1464 (2008).
21. L. Wang, B. Popko, R. P. Roos, *Hum. Mol. Genet.* **23**, 2629–2638 (2014).
22. M. E. Gurney *et al.*, *Science* **264**, 1772–1775 (1994).
23. S. J. Marciniak *et al.*, *Genes Dev.* **18**, 3066–3077 (2004).
24. J. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1392–1397 (2009).
25. H. P. Harding *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1832–1837 (2009).


## ACKNOWLEDGMENTS

We thank P. Tsaytler for some initial experiments on Sephin1; H. Meziane for the studies on memory; E. Fisher for SOD1$^{G93A}$ mice; E. Pettinato and C. Ferri for technical assistance; R. Roberts for discussions on CMT; A. Segonds-Pichon for statistical analysis; and members of the Bertolotti laboratory, M. Goedert, M. Hastings, and S. Munro for discussions. A.B. is an honorary fellow of the Clinical Neurosciences Department of Cambridge University. This work was supported by the Medical Research Council (UK) and the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant 309516. A.K. was supported by the European Molecular Biology Organization and Human Frontier Science Program, K.S. by the Swiss National Science Foundation, M.D. by the Italian Ministry of Health (GR-2011-02642791), and L.W. by NIH R01-NS55256. A.B. is a co-inventor on Great Britain patent WO 2014108520, covering benzylideneguanidine derivatives inhibitors of PPP1R15A. The data presented in this paper are tabulated in the main paper and the supplementary materials.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/239/suppl/DC1
Materials and Methods
Figs. S1 to S6
References (*26*–*29*)

8 December 2014; accepted 6 March 2015
10.1126/science.aaa4484

PRIMATE GENOMICS

# Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding

**Yali Xue,[1]* Javier Prado-Martinez,[2]* Peter H. Sudmant,[3]* Vagheesh Narasimhan,[1,4]* Qasim Ayub,[1] Michal Szpak,[1] Peter Frandsen,[5] Yuan Chen,[1] Bryndis Yngvadottir,[1] David N. Cooper,[6] Marc de Manuel,[2] Jessica Hernandez-Rodriguez,[2] Irene Lobon,[2] Hans R. Siegismund,[5] Luca Pagani,[1,7] Michael A. Quail,[1] Christina Hvilsom,[8] Antoine Mudakikwa,[9] Evan E. Eichler,[3,10] Michael R. Cranfield,[11] Tomas Marques-Bonet,[2,12] Chris Tyler-Smith,[1]† Aylwyn Scally[13]†**

**Mountain gorillas are an endangered great ape subspecies and a prominent focus for conservation, yet we know little about their genomic diversity and evolutionary past. We sequenced whole genomes from multiple wild individuals and compared the genomes of all four *Gorilla* subspecies. We found that the two eastern subspecies have experienced a prolonged population decline over the past 100,000 years, resulting in very low genetic diversity and an increased overall burden of deleterious variation. A further recent decline in the mountain gorilla population has led to extensive inbreeding, such that individuals are typically homozygous at 34% of their sequence, leading to the purging of severely deleterious recessive mutations from the population. We discuss the causes of their decline and the consequences for their future survival.**

Mountain gorillas came to the attention of the western world as recently as 1902 (*1*) and survive today only in two small and critically endangered populations in central Africa (*2*). Current estimates place their total population at around 800, of which just over half are in the Virunga volcanic mountain range on the borders of Rwanda, Uganda, and the Democratic Republic of Congo (Fig. 1A) (*3*, *4*). Mountain gorillas face a number of threats to their continued existence (*5*), and their survival became an urgent concern during the 1960s, when a precipitous decline in their numbers appeared to put them at grave risk of extinction (*6*). Active conservation efforts have since reversed the decline (*3*), but a loss of genetic diversity associated with small population size may still threaten their long-term viability. Infectious diseases such as Ebola, the cause of substantial recent mortality in the western lowland population (*7*), represent a particularly serious threat (*8*).

Despite extensive study of mountain gorillas in the field, few genetic analyses have been carried out, and these have been confined to mitochondrial sequences and a limited number of autosomal loci (*9*–*11*). Unlike the other great apes (*12*), mountain gorillas have not been studied on a genome-wide scale, which is key to understanding their biology, evolution, and relationship to sister taxa (Fig. 1B). It is also important in assessing their current status and in forming strategies for future conservation efforts. Phenotypic indicators of inbreeding such as syndactyly have been reported (*13*, *14*), but the full genetic impact of their decline is unknown. Additionally, the severe population bottleneck experienced by mountain gorillas provides an opportunity to study processes that may have played a recurring role in hominin evolution and extinction.

We performed whole-genome sequencing and analysis for 13 eastern gorillas, comprising seven mountain gorillas from the Virunga volcanoes region and six eastern lowland gorillas. The mountain gorillas originate from both sides of the Virunga massif, including from three groups spanning a densely populated sector on the Rwandan side, and plausibly represent genetic variation and ancestry in the Virunga population at large (*15*). Sequencing was performed to an average 26× depth (table S1), and combining these data with published gorilla genome sequences (*12*) yielded a total data set of 44 samples spanning all four gorilla subspecies (table S2).

Sequences were aligned to the gorilla reference genome (*16*), and several tools were used to call variants across samples (*15*). Comparing

[1]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK. [2]Institut de Biologia Evolutiva (CSIC/UPF), Parque de Investigación Biomédica de Barcelona (PRBB), Barcelona, Catalonia 08003, Spain. [3]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. [4]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK. [5]Department of Biology, University of Copenhagen, DK-2200 Copenhagen N, Denmark. [6]Institute of Medical Genetics, Cardiff University, Cardiff CF14 4XN, UK. [7]Department of Biological, Geological and Environmental Sciences, University of Bologna, 40134 Bologna, Italy. [8]Research and Conservation, Copenhagen Zoo, DK-2000 Frederiksberg, Denmark. [9]Rwanda Development Board, KG 9 Avenue, Kigali, Rwanda. [10]Howard Hughes Medical Institute, Seattle, WA 91895, USA. [11]Gorilla Doctors, Karen C. Drayer Wildlife Health Center, University of California, Davis, CA 95616, USA. [12]Centro Nacional de Análisis Genómico (Parc Cientific de Barcelona), Baldiri Reixac 4, 08028 Barcelona, Spain. [13]Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK.
*These authors contributed equally to this work. **†Corresponding author. E-mail: cts@sanger.ac.uk (C.T.-S.); aos21@cam.ac.uk (A.S.)**

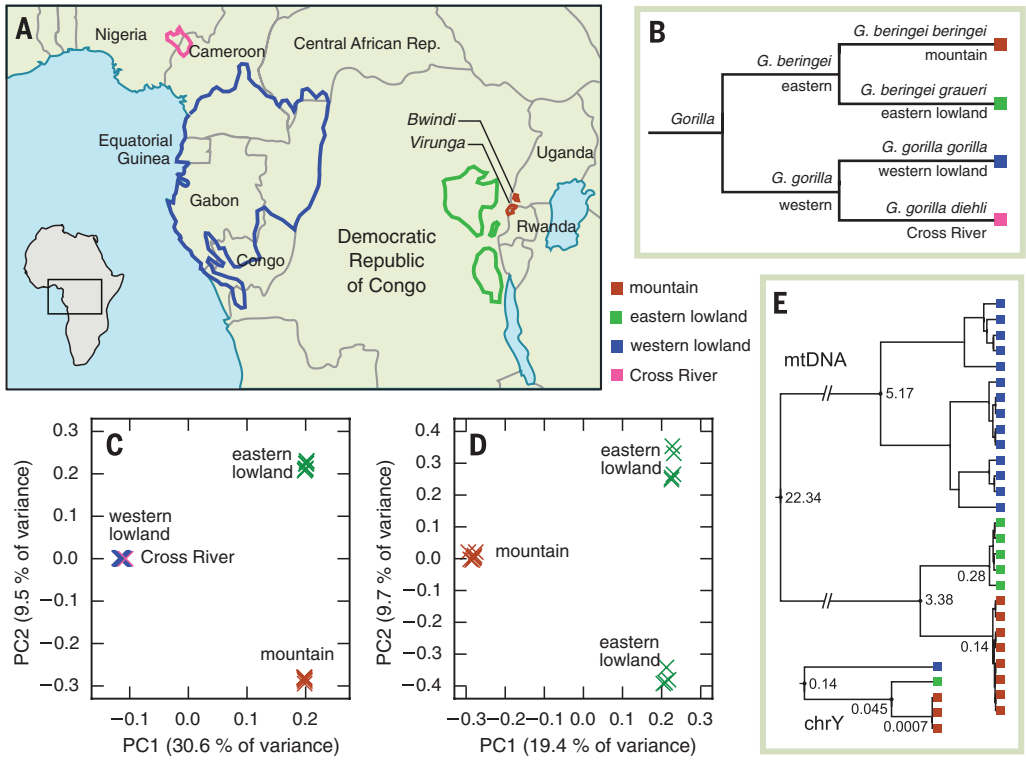



**Fig. 1. Geography, taxonomy and genetic structure of gorilla species.** (**A**) Distribution of gorilla subspecies (*2*). (**B**) *Gorilla* taxonomy. (**C**) PCA plot of SNP data for all four gorilla subspecies. (**D**) PCA plot of SNP data from mountain and eastern lowland gorilla samples only. (**E**) mtDNA and Y-chromosomal phylogenies. Node heights are in units of substitutions per base pair; each tree is drawn to a separate scale.

populations, we found that genetic diversity was lower in the eastern species than in the western by a factor of 2 to 3 (table S3), with mean autosomal heterozygosity (frequency of between-chromosomal differences) of $6.5 \times 10^{-4}$ per base pair in mountain and eastern lowland gorillas, versus $1.9 \times 10^{-3}$ per base pair in the western lowland population. These values are consistent with the small reported census population sizes of eastern gorillas relative to those in the west (*2*). A detailed analysis of variation in mitochondrial DNA (mtDNA) showed little diversity in mountain gorillas, with only three haplotypes in total differing by just one to three mutations (*15*). A similar situation was observed on the Y chromosome, with only two sites differentiating the three male mountain gorillas (*15*).

A principal components analysis (PCA) of 11,743,407 single-nucleotide polymorphisms (SNPs) across all samples (Fig. 1C) showed a hierarchical structure consistent with the accepted *Gorilla* taxonomy and the geographical distribution of these populations. In particular, the separation between eastern lowland and mountain gorilla samples confirms them as genetically distinct populations. Another PCA focusing on the eastern species only (Fig. 1D) revealed no substructure within the mountain gorillas sampled here, but did show a separation of eastern lowland gorillas into two subgroups, which may reflect structure within the subspecies as a whole or may only be a feature of the individuals we have sampled. Similar patterns are also found in a genetic structure analysis using the program ADMIXTURE (fig. S3).

Copy number variation (CNV) among gorillas was assessed with a read depth–based approach

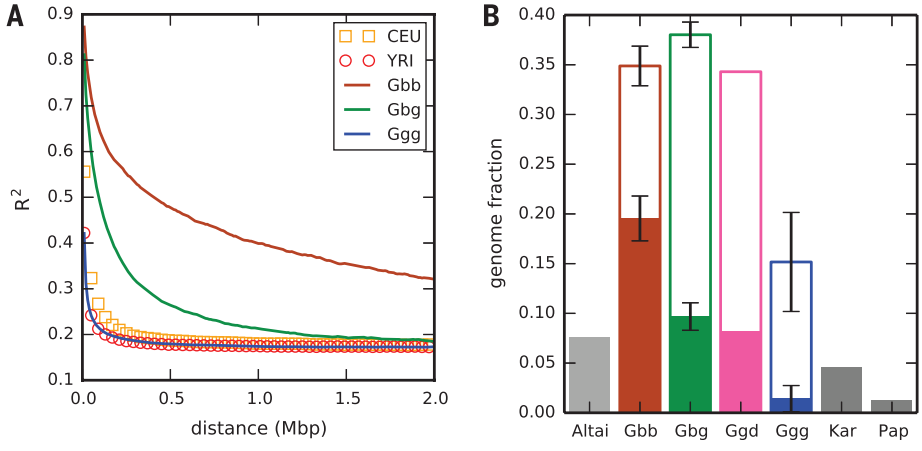



**Fig. 2. Linkage disequilibrium and homozygosity in gorillas.** (**A**) LD decay (*15*) in gorilla and human populations. Human samples are Utah residents with European ancestry (CEU) or Yoruba in Ibadan, Nigeria (YRI). (**B**) Mean per-sample genome fractions found in homozygous tracts. Open bars show total fractions for mountain (Gbb), eastern lowland (Gbg), Cross River (Ggd), and western lowland (Ggg) gorillas; solid bars show fractions in tracts of length 2.5 to 10 Mb (gorillas) or 2.5 to 10 cM in an Altai Neandertal and two human individuals [Karitiana (Kar) and Papuan (Pap)] (*19*). Error bars are ±1 SD.

(*17*), and the number of fixed loss and gain events between populations reflects the accepted taxonomy (table S5), although eastern lowland and mountain gorillas showed an excess of shared fixed deletions relative to western lowland gorillas, potentially reflecting a small ancestral population prior to their divergence. Diversity patterns at polymorphic CNVs also resembled those seen in SNPs (fig. S8) (*15*).

Accurate population monitoring and management are of central importance for conservation, and we cataloged 25,628 autosomal and 89 mtDNA ancestry-informative SNP markers (AIMs) uniquely identifying eastern lowland and mountain gorillas (as sampled here), as well as 1127 lineage-specific CNV events (data files S1 and S2) (*15*).

We found that genome-wide linkage disequilibrium (LD) varies markedly between gorilla populations (Fig. 2A). Both eastern subspecies—particularly mountain gorillas—show much more extensive LD than western lowland gorillas, in whom the decay of LD with genomic distance is similar to that in African humans (*15*). These patterns reflect differing demographic histories and suggest

a recent population bottleneck in the two eastern subspecies.

An analysis of chromosomal sequence sharing within individuals (genomic tracts of homozygosity) provided insight into recent ancestry and a measure of parental relatedness. Within mountain and eastern lowland gorilla individuals, chromosomes are typically homozygous over one-third of their length (on average 34.5% and 38.4%, respectively) (Fig. 2B), much higher than in western lowland individuals (13.8%) and exceeding even the most inbred human populations (*18*). We observed longer tracts in the eastern species, particularly in mountain gorillas, and a clear distinction in tract length distribution between eastern lowland and mountain gorillas (fig. S15) (*15*). Very long tracts (2.5 to 10 Mb) are particularly indicative of recent inbreeding, and homozygosity on this scale in mountain gorillas exceeds not only that in other gorilla populations but also that observed in the Altai Neandertal (*19*), consistent with parental relatedness equivalent to that between two half-siblings (*19*). These data suggest that mountain gorillas may have experienced several recent generations of close inbreeding. We also examined haplotype matching between individuals, finding subclustering among eastern lowland gorillas consistent with that seen in Fig. 1D; individuals in one cluster displayed a 10% higher mean level of sequence sharing than those in the other (fig. S16).

Whole-genome data also enabled us to explore the longer-term demographic history of these populations. Effective population size ($N_e$) in each of the four subspecies, as inferred by the PSMC algorithm (*20*), has varied over time (Fig. 3A), and we note the decline of $N_e$ particularly in the eastern species within the past 100,000 years. The most recent values of inferred $N_e$ are just 273 ± 54 in mountain gorillas and 290 ± 18 in eastern lowland gorillas. $N_e$, which reflects genetic diversity, is usually lower than census population size; however, in 1981 the estimated census size in the Virunga region was as low as 254 (*3*).

As judged from the date when inferred $N_e$ began to differ, the divergence of eastern and western gorillas began at least 150,000 years ago, but a more direct analysis using male X-chromosomal sequences suggests that they exchanged genetic material until around 20,000 years ago (Fig. 3B) (*15*). Given that this also coincides with a notable decline in western lowland gorilla $N_e$ (Fig. 3A), it may be that environmental changes during the Last Glacial Maximum (26,000 to 19,000 years ago), when dry savannah replaced tropical forest over much of the Congo basin (*21*), triggered a collapse in the western population and complete separation of the two species. Indeed, changes in forestation across the center of the continent over the past 150,000 years may have had a substantial influence on speciation and diversity in African apes more generally, including ancient humans (*12*).

We found no evidence for more recent east-west genetic contact, and *D*-statistic analysis provided no strong support for differences in east-west gene flow between subspecies pairs in either species (table S17).

These evolutionary differences between subspecies allowed us to investigate the relationship between demography and selection at protein-coding loci. To facilitate this, we classified coding sequence variants into three groups: synonymous, missense, and loss-of-function (LoF) (*15*). Accounting for potential reference bias, we found a relatively larger number of derived missense alleles in eastern gorillas, particularly in mountain

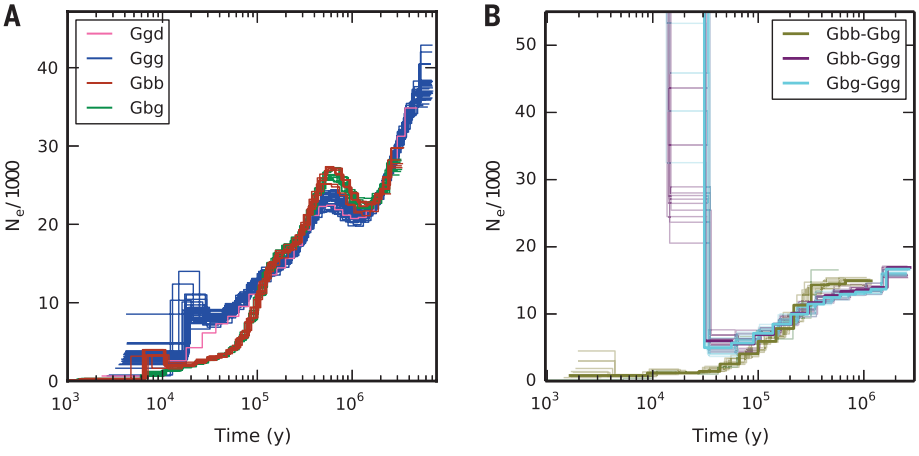



**Fig. 3. Ancestral effective population size and gene flow between gorilla populations.** (**A**) Inferred effective population size ($N_e$) history for each of the samples studied. (**B**) Cross-population $N_e$ history, based on paired male X-chromosomal sequences. Both plots are scaled using a generation time of 19.3 years and an autosomal mutation rate of $1.25 \times 10^{-8}$ per base pair per generation.

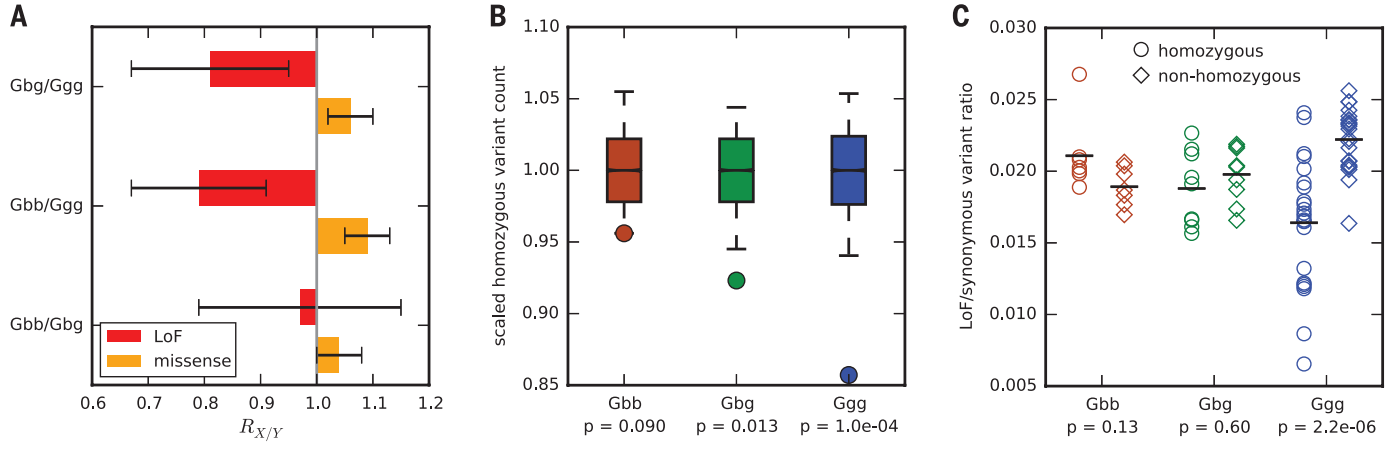



**Fig. 4. The genetic burden of missense mutations and purging of LoF mutations in eastern gorillas.** (**A**) Relative number of derived alleles at LoF (red) and missense (orange) sites that are frequent in one population and not another (*15*). Error bars represent ±2 SD. (**B**) Circles indicate the scaled number of LoF variant sites in each population where at least one sample is homozygous for the derived allele. Boxplots show distributions of the same statistic for matched samples of synonymous sites (*15*); whiskers show 5th and 95th percentiles; *P* values are the proportion of each sample distribution smaller than the corresponding LoF count. (**C**) Circles show the rate of LoF variants relative to synonymous variants in homozygous tracts for each sample; diamonds show the same ratio in nonhomozygous regions. Horizontal bars indicate population means; *P* values for each subspecies correspond to a Kolmogorov-Smirnov test for difference in distribution between homozygous and nonhomozygous regions.

gorillas, relative to western lowland gorillas (Fig. 4A). This is consistent with long-term differences in $N_e$ (Fig. 3A), because selection in a larger population is more effective at removing moderately deleterious variants. In contrast, we found that mountain gorillas have relatively fewer LoF variants than western lowland gorillas. This is again consistent with low $N_e$ and inbreeding, wherein alleles are more frequently exposed in a homozygous state, and because LoF variants are more likely to be strongly deleterious, they are less likely to persist in the population even if recessive (*22*). Further evidence for the purging of severely deleterious mutations from mountain and eastern lowland gorillas comes from the lack of any deficit of homozygous LoF genotypes in mountain gorillas (Fig. 4B) and the observation that LoF variants in both subspecies occur at the same rate in homozygous and heterozygous tracts (Fig. 4C) (*15*).

Finally, we looked at variation in genes using function inferred from homology, focusing on possible selection and adaptation in mountain gorillas. Such adaptation might be expected from the fact that mountain gorillas range over high altitudes (1500 to 4000 m), with consequences for diet, morphology, and physiology (*23*). However, we found no significant enrichment in any functional category of genes, although there are interesting examples related to nervous system morphology, immunoglobulin quantity, and red blood cell morphology (*15*). Mountain gorillas carry a significant excess of variants in genes associated with blood coagulation in humans (fig. S21), perhaps linked to high-altitude living (*15*). We also identified variants associated with cardiomyopathy, including in one deceased individual (Kaboko) in whom post mortem analysis revealed evidence of muscular hypertrophy (*15*). Cardiovascular disease has been identified as a notable cause of death in captive western lowland gorillas (*24*).

Concern about the survival of mountain gorillas is focused largely on the threat from human encroachment on their habitat (*2*, *5*), and their severe recent population decline may be responsible for the high level of inbreeding that we observe. An increased burden of deleterious mutation and low genetic diversity—including at the major histocompatibility locus, of central importance to the immune system (fig. S22)—have likely reduced their resilience to environmental change and pathogen evolution. However, the origins of this condition extend far into their history, because both eastern subspecies have experienced a long decline over tens of millennia. Indeed, the demographic histories of mountain and eastern lowland gorillas (Fig. 3A) bear unhappy resemblance to similar histories inferred from Neandertals before their disappearance (*19*).

Nonetheless, the same evidence suggests that such a fate is not inevitable: These subspecies have survived for thousands of generations at low population levels and may have developed physiological and behavioral strategies to mitigate inbreeding, such as natal dispersal and gene flow between isolated populations (*25*). The purging of strongly deleterious mutations may also be an important factor. Our findings reinforce the imperative to sustain conservation efforts that have helped to keep mountain gorillas from the brink of extinction, and provide a genomic resource for future conservation and research.

## REFERENCES AND NOTES


1. G. B. Schaller, *The Year of the Gorilla* (Univ. of Chicago Press, Chicago, 2010), p. 7.
2. M. Robbins *et al*., *Gorilla beringei* ssp. *beringei* (listing in IUCN Red List of Threatened Species, 2008); www.iucnredlist.org/details/39999/0.
3. M. Gray *et al*., *Biol. Conserv.* **158**, 230–238 (2013).
4. K. Guschanski *et al*., *Biol. Conserv.* **142**, 290–300 (2009).
5. A. H. Harcourt, *Biol. Conserv.* **75**, 165–176 (1996).
6. A. W. Weber, A. Vedder, *Biol. Conserv.* **26**, 341–366 (1983).
7. P. J. Le Gouar *et al*., *PLOS ONE* **4**, e8375 (2009).
8. IUCN, "Ebola outbreak highlights critical links between biodiversity loss and human health, says IUCN's Wildlife Health Specialist Group" (2014); www.iucn.org/?18439.
9. K. J. Garner, O. A. Ryder, *Mol. Phylogenet. Evol.* **6**, 39–48 (1996).
10. M. I. Jensen-Seaman, K. K. Kidd, *Mol. Ecol.* **10**, 2241–2247 (2001).
11. J. Roy *et al*., *Biol. Lett.* **10**, 2014.0811 (2014).
12. J. Prado-Martinez *et al*., *Nature* **499**, 471–475 (2013).
13. D. Fossey, *Gorillas in the Mist* (Harcourt Brace, New York, 1983), p. 72.
14. A. Routh, J. Sleeman, in *Proceedings of the British Veterinary Zoological Society* (Howletts and Port Lympne Wild Animal Parks, Kent, 14–15 June 1997), pp. 22–25.
15. See supplementary materials on *Science* Online.
16. A. Scally *et al*., *Nature* **483**, 169–175 (2012).
17. C. Alkan *et al*., *Nat. Genet.* **41**, 1061–1067 (2009).
18. T. J. Pemberton *et al*., *Am. J. Hum. Genet.* **91**, 275–292 (2012).
19. K. Prüfer *et al*., *Nature* **505**, 43–49 (2014).
20. H. Li, R. Durbin, *Nature* **475**, 493–496 (2011).
21. D. Anhuf *et al*., *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **239**, 510–527 (2006).
22. S. Glémin, *Evolution* **57**, 2678–2687 (2003).
23. D. P. Watts, *Am. J. Primatol.* **7**, 323–356 (1984).
24. R. McManamon, L. Lowenstine, in *Fowler's Zoo and Wild Animal Medicine* (Elsevier, St. Louis, vol. 7, 2012), pp. 408–415.
25. A. Pusey, M. Wolf, *Trends Ecol. Evol.* **11**, 201–206 (1996).


## ACKNOWLEDGMENTS


We thank the Institut Congolais pour la Conservation de la Nature and the Rwandan Development Board for the use of samples; J. Ramer, D. Zimmerman, J. F. Kinani, J. Bosco Noheri, M. Bahizi, J. P. Lukasa, J. Iyanya, M. Kabuyaya, E. Kambale, and J. Sohl of Gorilla Doctors for sample collection and processing; and M. Pollard, P. Hallast, M. Jobling, and L. Lowenstine for assistance. Supported by Royal Society grant RG130105 (A.S.), Wellcome Trust grants 098051 (Q.A., Y.C., V.N., L.P., M.A.Q., M.S., C.T.-S., Y.X., B.Y.) and 099769/Z/12/Z (V.N.), NIH grant HG002385 (E.E.E.), a European Research Council Starting Grant (260372), and Ministerio de Ciencia e Innovacion grant BFU2011-28549 (T.M.-B.). E.E.E. is an investigator of the Howard Hughes Medical Institute. Accession numbers for sequence data are given in table S1. Gorilla blood samples sequenced in this study were collected and transferred in compliance with CITES legislation and subject to a material transfer agreement between the Mountain Gorilla Veterinary Project and the Office Rwandais du Tourisme et des Parcs Nationaux. D.N.C. acknowledges support from BIOBASE GmbH through a license agreement with Cardiff University; E.E.E. is a member of the external scientific advisory board for DNAnexus Inc.


## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6231/242/suppl/DC1
Materials and Methods
Figs. S1 to S22
Tables S1 to S18
Data files S1 and S2
References (*26*–*62*)


30 November 2014; accepted 3 March 2015
10.1126/science.aaa3952

NEW ENGLAND BioLabs® Inc.

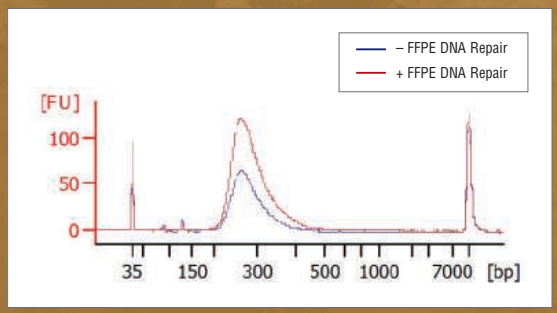
# On the mend.

## NEBNext® FFPE DNA Repair Mix from New England Biolabs®

While archiving of clinical materials as Formalin-Fixed, Paraffin-Embedded (FFPE) samples is common, it causes significant damage to the sample's DNA. As a result, such samples can be difficult to sequence. The cocktail of enzymes in the NEBNext FFPE DNA Repair Mix repairs multiple types of damage that are common among FFPE samples, thereby improving yields and overall library success rates.

Make sure your sample's DNA is on the mend before your next sequencing experiment!

Effect of FFPE DNA Repair Mix on library yields




*An example of Agilent Bioanalyzer® traces of libraries prepared from stomach tumor FFPE DNA that was treated with the FFPE DNA Repair Mix, or was untreated, before library construction. Yield improvements of 101% to 458% have been observed.*

**Visit www.neb.com/M6630 to learn more.**

# Discovery Fast Track *Challenge*

**The world is eager for new medicines.**
**This is your chance to join GSK scientists—helping to put your novel drug discovery concept on the fast track.**

Are you an academic researcher in Europe or North America? Do you have a drug discovery concept you're eager to explore?

**The Discovery Fast Track Challenge could be just what you need.**

Tell us about your idea. Our expert judges will select up to 12 researchers to win a collaboration with GSK. You'll work side-by-side with GSK scientists, discovering active compounds using our high-throughput screening capabilities and extensive compound library. In short, you and your concept will be on the fast track to success.

Submissions are being accepted March 23 - April 24.
Enter today at **gsk.com/discoveryfasttrack**

eppendorf


# Impress Yourself

**The new Eppendorf Cell Culture Consumables**

The all new product line of Eppendorf Cell Culture Consumables will truly delight your cells. Its outstanding design, reliability and purity is based on more than 50 years of experience.
Products created by experts, developed for perfectionists. Impress yourself!

> Unsurpassed quality, clarity, purity and sterility, providing reliable cell culture conditions
> Significantly improved design for more safety and consistency
> Maximum safety and confidence during storage and transportation

www.eppendorf.com/ccc

AAAS 2016 ANNUAL MEETING | WASHINGTON, DC FEBRUARY 11–15

# Global Science Engagement

Science is a global endeavor that advances when knowledge is both generated and shared. Increasingly, scientists and engineers are working both within and outside of national boundaries on local and global issues.

Challenges necessitating innovation and international scientific collaboration are abundant in food and water security, sustainable development, infectious disease and health, climate change, natural disasters, and energy. Countries with varying levels of development, education, and scientific capacity may have different goals and expectations for international scientific engagement.

What elements make international collaboration successful and sustainable? What engagement opportunities are available, and what are the responsibilities of researchers, entrepeneurs, educators, and policymakes in global scientific endeavors?

## Symposium Proposal Deadline: April 24

Share innovative and international science with the AAAS Annual Meeting community. Visit the 2016 proposal submission website for information.

**aaas.org/AM16**

### Gold Nanoparticles

Highly monodisperse, premium Gold NanoRods and Gold NanoSpheres, available in a range of different sizes and peak absorption wavelengths, are ideal for use in imaging, tumor targeting, microscopy, lateral flow assays, SERS, and several other applications. The monodispersity and minimal batch to batch variations ensure superior data quality and consistency. Detailed technical specifications are included with every product to guarantee complete transparency in surface chemistry for researchers performing their own functionalizations and conjugations. Surface coating options include a specialized silica coating in addition to commonly used coatings such as citrate, CTAB, and PEG. The silica coating provides colloidal stability, bioconjugation potential and enhanced optical properties. Silica-coated Gold NanoRods resist melting and shape distortion even when subjected to extreme heat via focused laser beams. In addition to providing thermodynamic stability, the silica-coating also facilitates better heat transfer to the surrounding fluid, thus dramatically increasing imaging signal strength.

**NanoHybrids**
For info: 512-270-8469
www.nanohybrids.net

### Laboratory Information Management System

The Genologics Clarity laboratory information management system (LIMS) has been integrated into the Illumina HiSeq-based workflows to offer next generation sequencing with the highest level of data management and methods control. RNA sequencing, target capture, whole-genome sequencing, and premade library samples are tracked by the LIMS from acquisition through QC, library prep, and Illumina sequencing. Reliable and flexible sample handling is provided for all workflows and chain of custody information is preserved. The Beckman Coulter Life Sciences' Biomek FX liquid handling and robotics platforms are integrated with the system as well, providing further sample processing and security benefits. Any number of robots and QC instruments can be brought into the process and run together, so projects can be scaled as needed. Clarity is one of the few LIMS systems that offer an upgrade option to audit to clinical standards and is approved for CFR part 11 recordkeeping.

**Beckman Coulter Genomics**
For info: 800-361-7780
www.beckmangenomics.com

### Three-Position Stage

The Three Position Stage for the VIAFLO 96 and VIAFLO 384 handheld benchtop pipettes expands available stage positions for microplates, reservoirs, and tips on a VIAFLO 96 or VIAFLO 384 from two to three thereby enhancing workflow in a wide range of applications. The Three Position Stage also features an indexing function to allow access to 384-well plates using a 96-channel pipetting head. For plate replication applications, having a tip rack in the left position, the source plate in the middle position, and the target plates in the right position of the Three Position Stage enables the user to replicate the source plate with minimal plate handling effort. Changing tips and target plates can be done without moving the source plate, eliminating the risk of spills. When undertaking compound dilutions or screening, the Three Position Stage enables accurate and rapid addition of reagent and compounds from two different sources (reservoir and plate) to one target screening plate.

**Integra Biosciences**
For info: +41-(0)-81-286-9530
www.integra-biosciences.com

### Antibodies

Two new product families, QPrEST for mass spectrometry-based protein quantification and PrEST Antigens have been designed to be used as controls in antibody assays. QPrEST represents a new category of isotope-labeled internal mass spectrometry standards, offering distinct advantages over existing products for relative and absolute protein quantification. The initial catalogue contains more than 19,000 QPrESTs targeting 13,000 human proteins. QPrEST improves the standard workflow for mass spectrometry-based proteomics. PrEST Antigens are the recombinant protein fragments used as immunogens in the generation of Triple A Polyclonals. They should be used together with the corresponding antibody for preadsorption assays, or as controls in Western blots, enzyme-linked immunosorbent assays, and protein array assays. This launch includes more than 17,000 PrEST Antigens originally used for generation of the Triple A Polyclonals and PrecisA Monoclonals.

**Atlas Antibodies**
For info: +46-8-54-59-58-50
www.atlasantibodies.com

### Flexi-Block Accessory

The new Flexi-block accessory can heat almost any sample from ambient temperature up to 200°C, including those stored in 0.2 mL, 0.5 mL, 2.0 mL, and 5.0 mL micro tubes and any diameter test-tube, vial, cuvette, or microplate. This flexibility means laboratories can achieve new levels of efficiency, as researchers are able to heat any sample, regardless of container or type, without having to change or recalibrate instruments. The Flexi-block's broad range is made possible by the incorporation of 1.5 mL diameter autoclaveable beads. Manufactured from high density ceramic, these exhibit high mechanical strength and are an excellent thermal transfer medium. This outstanding flexibility eliminates the traditional limitation of aluminium heating blocks—their specificity to a particular type of tube, multiwell plate, or glass vial. Flexi-block is unique, economically priced, and comes in two variations: double-width or a larger triple-width version for high throughput heating and incubating applications as well as sensitive analytical procedures.

**Bibby Scientific**
For info: +44-(0)-1785-812121
www.bibby-scientific.com

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.

# 2015 MRS® FALL MEETING & EXHIBIT

**November 29 – December 4, 2015 | Boston, Massachusetts**

# CALL FOR PAPERS

**Abstract Submission Opens May 18, 2015 | Abstract Submission Deadline June 18, 2015**

A Engaged Learning of Materials Science and Engineering in the 21st Century

## BIOMATERIALS AND SOFT MATERIALS

B Stretchable and Active Polymers and Composites for Electronics and Medicine
C Tough, Smart and Printable Hydrogel Materials
D Biolological and Bioinspired Materials in Photonics and Electronics—Biology, Chemistry and Physics
E Engineering and Application of Bioinspired Materials
F Biomaterials for Regenerative Engineering
G Plasma Processing and Diagnostics for Life Sciences
H Multifunctionality in Polymer-Based Materials, Gels and Interfaces
I Nanocellulose Materials and Beyond—Nanoscience, Structures, Devices and Nanomanufacturing
J Wetting and Soft Electrokinetics
K Materials Science, Technology and Devices for Cancer Modeling, Diagnosis and Treatment
L Nanofunctional Materials, Nanostructures and Nanodevices for Biomedical Applications

## NANOMATERIALS AND SYNTHESIS

M Micro- and Nanoscale Processing of Materials for Biomedical Devices
N Magnetic Nanomaterials for Biomedical and Energy Applications
O Plasmonic Nanomaterials for Energy Conversion
P Synthesis and Applications of Nanowires and Hybrid 1D-0D/2D/3D Semiconductor Nanostructures
Q Nano Carbon Materials—1D to 3D
R Harsh Environment Sensing—Functional Nanomaterials and Nanocomposites, Materials for Associated Packaging and Electrical Components and Applications

## MECHANICAL BEHAVIOR AND FAILURE OF MATERIALS

S Mechanical Behavior at the Nanoscale
T Strength and Failure at the Micro- and Nanoscale—From Fundamentals to Applications
U Microstructure Evolution and Mechanical Properties in Interface-Dominated Metallic Materials
V Gradient and Laminate Materials
W Materials under Extreme Environments (MuEE)
Y Shape Programmable Materials

## ELECTRONICS AND PHOTONICS

Z Molecularly Ordered Organic and Polymer Semiconductors—Fundamentals and Devices
AA Organic Semiconductors—Surface, Interface and Bulk Doping
BB Innovative Fabrication and Processing Methods for Organic and Hybrid Electronics
CC Organic Bioelectronics—From Biosensing Platforms to Implantable Nanodevices
DD Diamond Electronics, Sensors and Biotechnology—Fundamentals to Applications
EE Beyond Graphene—2D Materials and Their Applications
FF Integration of Functional Oxides with Semiconductors
GG Emerging Materials and Platforms for Optoelectronics
HH Optical Metamaterials—From New Plasmonic Materials to Metasurface Devices
II Phonon Transport, Interactions and Manipulations in Nanoscale Materials and Devices—Fundamentals and Applications
JJ Multiferroics and Magnetoelectrics
KK Materials and Technology for Non-Volatile Memories

## ENERGY AND SUSTAINABILITY

LL Materials and Architectures for Safe and Low-Cost Electrochemical Energy Storage Technologies
MM Advances in Flexible Devices for Energy Conversion and Storage
NN Thin-Film and Nanostructure Solar Cell Materials and Devices for Next-Generation Photovoltaics
OO Nanomaterials-Based Solar Energy Conversion
PP Materials, Interfaces and Solid Electrolytes for High Energy Density Rechargeable Batteries
QQ Catalytic Materials for Energy
RR Wide-Bandgap Materials for Energy Efficiency—Power Electronics and Solid-State Lighting
SS Progress in Thermal Energy Conversion—Thermoelectric and Thermal Energy Storage Materials and Devices

## THEORY, CHARACTERIZATION AND MODELING

TT Topology in Materials Science—Biological and Functional Nanomaterials, Metrology and Modeling
UU Frontiers in Scanning Probe Microscopy
VV *In Situ* Study of Synthesis and Transformation of Materials
WW Modeling and Theory-Driven Design of Soft Materials
XX Architected Materials—Synthesis, Characterization, Modeling and Optimal Design
YY Advanced Atomistic Algorithms in Materials Science
ZZ Material Design and Discovery via Multiscale Computational Material Science
AAA Big Data and Data Analytics for Materials Science
BBB Liquids and Glassy Soft Matter—Theoretical and Neutron Scattering Studies
CCC Integrating Experiments, Simulations and Machine Learning to Accelerate Materials Innovation
DDD Lighting the Path towards Non-Equilibrium Structure-Property Relationships in Complex Materials

*X* *Frontiers of Material Research*

**www.mrs.org/fall2015**

*The* **MRS/E-MRS Bilateral Conference on Energy** *will be comprised of the energy-related symposia at the 2015 MRS Fall Meeting.*

### Meeting Chairs

**T. John Balk** University of Kentucky
**Ram Devanathan** Pacific Northwest National Laboratory
**George G. Malliaras** Ecole des Mines de St. Etienne
**Larry A. Nagahara** National Cancer Institute
**Luisa Torsi** University of Bari "A. Moro"

### Don't Miss These Future MRS Meetings!

**2016 MRS Spring Meeting & Exhibit**
March 28 - April 1, 2016
Phoenix, Arizona

**2016 MRS Fall Meeting & Exhibit**
November 27 - December 2, 2016
Boston, Massachusetts

MRS MATERIALS RESEARCH SOCIETY®
*Advancing materials. Improving the quality of life.*
506 Keystone Drive • Warrendale, PA 15086-7573
Tel 724.779.3003 • Fax 724.779.8313
info@mrs.org • www.mrs.org

Introducing

# bio-techne®

## HIGHEST QUALITY PRODUCT PORTFOLIO


## TRUSTED GLOBAL BRANDS

R&D systems

NOVUS BIOLOGICALS™

TOCRIS

bio-techne®

LEARN MORE bio-techne.com/launch

Building Innovation Opportunities

| bio-techne.com | North America | Europe • Middle East • Africa | China | Rest of World |
|---|---|---|---|---|
| info@bio-techne.com<br>techsupport@bio-techne.com | TEL 800 343 7475 | TEL +44 (0)1235 529449 | info.cn@bio-techne.com<br>TEL +86 (21) 52380373 | bio-techne.com/find-us/distributors<br>TEL +1 612 379 2956 |

Bio-Techne is a trading name for R&D Systems


There's only one Science

# *Science* Careers Advertising

For full advertising details, go to ScienceCareers.org and click For Employers, or call one of our representatives.

**Tracy Holmes**
Worldwide Associate Director
*Science* Careers
Phone: +44 (0) 1223 326525

THE AMERICAS

E-mail: advertise@sciencecareers.org
Fax: 202 289 6742

**Tina Burks**
Phone: 202 326 6577

**Nancy Toema**
Phone: 202 326 6578

**Marci Gallun**
Sales Administrator
Phone: 202 326 6582

**Online Job Posting Questions**
Phone: 202 312 6375

EUROPE / INDIA / AUSTRALIA / NEW ZEALAND / REST OF WORLD

E-mail: ads@science-int.co.uk
Fax: +44 (0) 1223 326532

**Axel Gesatzki**
Phone: +44 (0) 1223 326529

**Sarah Lelarge**
Phone: +44 (0) 1223 326527

**Kelly Grace**
Phone: +44 (0) 1223 326528

JAPAN

**Katsuyoshi Fukamizu** (Tokyo)
E-mail: kfukamizu@aaas.org
Phone: +81 3 3219 5777

**Hiroyuki Mashiki** (Kyoto)
E-mail: hmashiki@aaas.org
Phone: +81 75 823 1109

CHINA / KOREA / SINGAPORE / TAIWAN / THAILAND

**Ruolei Wu**
Phone: +86 186 0082 9345
E-mail: rwu@aaas.org



ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

ScienceCcareers.org


For recruitment in science, there's only one Science

## What makes *Science* the best choice?

- Read and respected by 570,400 readers around the globe
- 78% of readers read *Science* more often than any other journal
- Your ad sits on specially labeled pages to draw attention to the ad
- Your ad dollars support AAAS and its programs, which strengthens the global scientific community.

## Why choose this microbiology section for your advertisement?

- Relevant ads lead off the career section with special Microbiology banner
- Bonus distribution to:

  American Society for Microbiology (ASM)
  May 30–June 2, 2015 New Orleans, LA.

## Expand your exposure. Post your print ad online to benefit from:

- Link on the job board homepage directly to microbiology jobs
- Dedicated landing page for jobs in microbiology
- Additional marketing driving relevant job seekers to the job board.


* Ads accepted until May 11 on a first-come, first-served basis.

SCIENCECAREERS.ORG

Science Careers
AAAS

To book your ad: **advertise@sciencecareers.org**

**The Americas**
202-326-6582

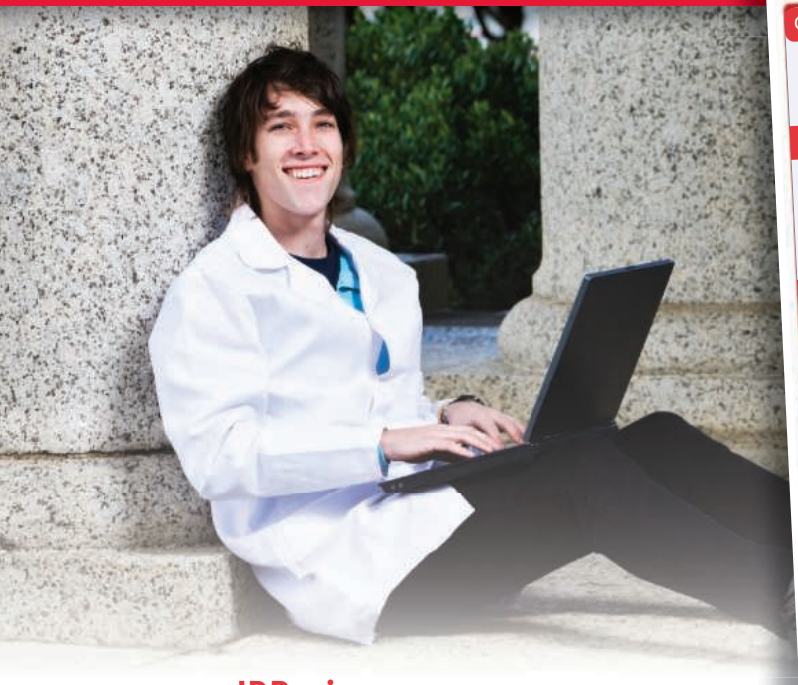
**Europe/RoW**
+44(0)1223-326500

**Japan**
+81-3-3219-5777

**China/Korea/Singapore/Taiwan**
+86-186-0082-9345

For your career in science, there's only one **Science**

# A career plan customized *for* you, *by* you.



myIDP.sciencecareers.org


Recommended by leading professional societies and endorsed by the National Institutes of Health, an individual development plan will help you prepare for a successful and satisfying scientific career.

my IDP Science Careers

In collaboration with FASEB, UCSF, and the Medical College of Wisconsin and with support from the Burroughs Wellcome Fund, AAAS and *Science* Careers present the first and only online app that helps scientists prepare their very own individual development plan.

Visit the website and start planning today!
**myIDP.sciencecareers.org**

In partnership with:

advancing the life sciences 1912-2012 FASEB Federation of American Societies for Experimental Biology

UCSF University of California San Francisco

Medical College of Wisconsin

Burroughs Wellcome Fund

杭州电子科技大学
HANGZHOU DIANZI UNIVERSITY

# Faculty Positions Available at Hangzhou Dianzi University

**About HDU:**

Hangzhou Dianzi University（HDU）, located in Hangzhou, a well-known city for its history and culture, is one of the advanced universities in Zhejiang Province, China. There are more than 28000 full-time students and over 2200 faculty and staffs in campus, 1 academician of the developing country, 5 sharing academicians, 2 young middle-aged outstanding experts, 5 distinguished experts of the "Thousand-talent Project" awarded by the Organization Department of the Central Committee of the CPC, 3 of New Century's Talent Project, 4 holders of the National Outstanding Youth Fund, 10 of the New Century's Excellent Talents Project supported by the Ministry of Education, 4 of National Outstanding Teachers and a large number of famous scholars and middle-aged experts with profound academic accomplishments. HDU has developed into a significant institution for its personnel training, scientific research and community service, and holds the lead among the universities in Zhejiang Province in every aspect as the school scale, the education level, the teaching quality and efficiency in school management.

*HDU website: http://www.hdu.edu.cn*

**Vacancies:**

Mechanical Engineering, Electric Engineering, Marine Equipment & Technology, Control Theory & Control Engineering, Biomedical Engineering, Instrument Science & Technology , Electronic Science & Technology, Computer Application Technology, Information & Communication Engineering, Economics, Business Administration, Material & Environmental Engineering, Digital Media, Artistic Designing, Journalism, Law, Mathematics, Physics, English Language and Literature and so on. (Link Recruitment of Personnel Department of HDU for further Information).

**Qualifications:**

Applicants with PHD degree from prestigious domestic and overseas universities; Professors or assistant professors/researchers of the equivalent titles, having been working at overseas prestigious universities or research institutes.

**Salary, Benefits and Support:**

Successful applicant will be offered a competitive package, including sufficient laboratory space, startup funding and competitive salary commensurate with experience, in addition to a housing allowance, and other employee benefits. HDU is recruiting overseas talents all year around. Applicants may refer to the recruitment in our website *http://renshi.hdu.edu.cn* for further information. For those exceptionally qualified talents, the salary could be negotiated in person.

**Contact:**

So appreciated for interested individuals sending curriculum vitae by e-mail to rsc@hdu.edu.cn.Or contact us by telephone 86-(0)571-86915031, fax 86-(0)571-86915029.

*Website: http://renshi.hdu.edu.cn*

# Faculty Position Available at China University of Petroleum in Qingdao

China University of Petroleum (UPC) is a national key university directly affiliated with the Ministry of Education and a member of the "211 Project" and "985 innovation platform for preponderant discipline" universities。Honored as "the cradle of scientific and technological talents for petroleum industry", UPC is an important base of training high-level talents for the petroleum and petrochemical industry and has already developed into a multi-disciplinary, well-rounded university focusing on petroleum and engineering.

UPC now has two campuses in Shandong Province, one in Qingdao and one in Dongying. Qingdao Campus, the main and new campus, is located in Qingdao which enjoys a high reputation for its charming scenery and favorable climate.

We warmly welcome scientific and technological elites all over the world to join us and build our university into a national renowned, high-level, research-oriented institute with its petroleum-related disciplines reaching the world's first-class level.

**UPC provides eligible talents with a good academic environment and excellent working and living conditions.**

1.Tenured professor of well-known overseas universities or experts and scholars who have published a thesis as the original author in "Science" or "Nature" can be recruited as a high level professor at UPC， and will be offered 2 million RMB settling-in allowance or housing subsidy and 3-8 million RMB scientific research and academic funds after the evaluation.

2.Experts or scholars who have a high academic reputation and are qualified leader in certain disciplines can be recruited as a high level or regular professor，and will be offered 1 million RMB settling-in allowance or housing subsidy and 1-3 million RMB Scientific research and academic funds after the evaluation.

3.Youth talents who qualify for China's "Youth 1000 Talents projects" can be offered 0.5 million RMB annual salary, 1.5 million RMB settling-in allowance or housing subsidy and 1-3 million RMB scientific research and academic funds.

4.Youth talents who receive the position as postdoctoral, assistant professor, associate professor or have a PhD from a famous overseas university or research institution can be offered higher treatment than similar talents in China.

***Further information is available at http:// rsc.upc.edu.cn***

***Contact us:***

Tel: 0086-532-86981808; 86981806 E-mail: teacher@upc.edu.cn

# Southwest Jiaotong University, P.R.China Anticipates Your Working Application

西南交通大学 Southwest Jiaotong University

Southwest Jiaotong University (SWJTU), founded in 1896, situates itself in Chengdu, the provincial capital of Sichuan. It is a national key multidisciplinary "211" and "985 Feature" Projects university directly under the jurisdiction of the Ministry of Education, featuring engineering and a comprehensive range of study programs and research disciplines spreading across more than 20 faculties and institutes/centers. Boasting a complete Bachelor-Master-Doctor education system with more than 2,500 members of academic staff, our school also owns 2 first-level national key disciplines, 2 supplementary first-level national key disciplines (in their establishment), 15 first-level doctoral programs, 43 first-level master programs, 75 key undergraduate programs, 10 post-doctoral stations and more than 40 key laboratories at national and provincial levels.

Our university is currently implementing the strategy of "developing and strengthening the university by introducing and cultivating talents". Therefore, we sincerely look forward to your working application.

More information available at http://www.swjtu.edu.cn/

**I. Positions and Requirements**

**A.High-level Leading Talents**

It is required that candidates be listed in national top talents programs such as ***Program of Global Experts, Top Talents of National Special Support Program, "Chang Jiang Scholars", China National Funds for Distinguished Young Scientists and National Award for Distinguished Teacher.***

Candidates are supposed to be no more than 50 years old. The limitation could be extended in the most-needed areas of disciplinary development.

Candidates who work in high-level universities/institutes and reach the above requirements are supposed to be no more than 45 years old.

**B. Young Leading Scholars**

Candidates are supposed to be listed in or qualified to apply for the following programs:

***•National Thousand Young Talents Program***

***•The Top Young Talents of National Special Support Program(Program for Supporting Top Young Talents)***

***•Science Foundation for the Excellent Youth Scholars***

Candidates should have good team spirit and leadership, outstanding academic achievements, broad academic vision and international cooperation experience and have the potential of being a leading academic researcher.

**C. Excellent Young Academic Backbones**

Candidates under 40 years old are expected to graduate from high-level universities/institutes either in China or other countries. Those who are professors, associate professors and other equal talents from high-level universities/institutes overseas could be employed as professors and associate professors as well.

**D. Excellent Doctors and Post Doctoral Fellows**

Candidates under 35 years old are supposed to be excellent academic researchers from high-level universities either in China or other countries.

**II. Treatments**

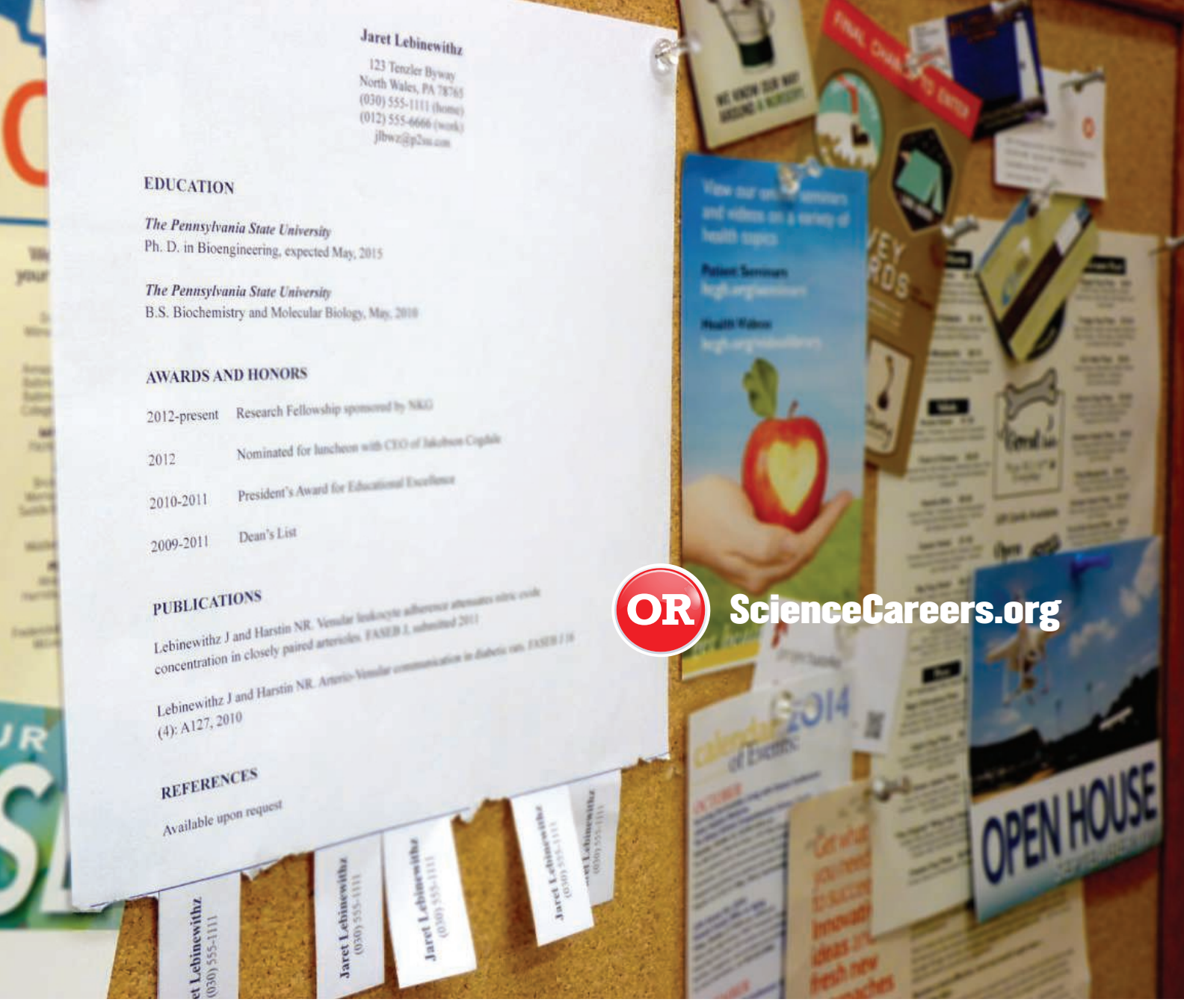
The candidates will be provided with competitive salaries and welfares that include settling-in allowance, subsidy of rental residence, start-up funds of scientific research, assistance in establishing scientific platform and research group as well as international-level training and promotion . As for outstanding returnees, we can offer further or specific treatments that can be discussed personally.

**III . Contact us:**

Contacts: Ye ZENG & Yinchuan LI Telephone number: 86-28-66366202 Email: *talent@swjtu.edu.cn*

Address: Human Resources Department of SWJTU, the western park of high-tech zone, Chengdu, Sichuan, P.R.China, 611756

**http://www.swjtu.edu.cn/**




## Learn more and conduct your job search the easy way.

- Search thousands of job postings
- Create job alerts based on your criteria
- Get career advice from our Career Forum experts
- Download career advice articles and webinars
- Complete an individual development plan at "myIDP"

Target your job search using relevant resources on **ScienceCareers.org**.

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

## NATIONAL INSTITUTE OF IMMUNOLOGY, INDIA

**Career Opportunities for Scientists**

The National Institute of Immunology, New Delhi, is a leading research Institute in India with a long-standing reputation for scientific excellence. The Institute is equipped with state-of-the-art infrastructure for pursuing research in immunology and allied sciences. The Institute also imparts vigorous long-term research training leading to a Ph.D degree.

The Institute invites applications from early as well as mid-career scientists with potential for intellectual leadership and passion for innovative research to set up independent multidisciplinary research programmes relating to the areas of (but not restricted to) immunology, virology, microbiology, structural, chemical and systems biology to address immunological problems at the expanding interface of modern biology. Applicants should possess a Ph.D degree or an M.D with four years relevant experience post-Ph.D/M.D. Those possessing Ph.D/M.D degrees with evidence of leadership with about 8 years and more of R & D experience can be considered for higher positions. Demonstrated record of scientific productivity in the form of outstanding publications in scientific journals of repute and/or internationally valid and productive patents is essential. Applicants should have a commitment to excellence and willingness to collaborate in interdisciplinary projects. Eligible & interested candidates are advised to visit the website, **http://www.nii.res.in** for detailed information.

Western University of Health Sciences

**www.westernu.edu**

### Faculty Position in Biochemistry and Discipline Leader in Microbiology

Western University of Health Sciences, a thriving center for human health care and veterinary medicine education, is growing and along with our site for the College of Osteopathic Medicine of the Pacific – Northwest (COMP-NW) in Lebanon, Oregon, we are recruiting new faculty in basic medical sciences for our campus in Pomona, California (COMP-Pomona).

The Department of Basic Medical Sciences provides the preclinical education for the College of Osteopathic Medicine, and invites applications from highly motivated and accomplished individuals for a tenure-track faculty position in Biochemistry and a Discipline Leader position in Microbiology. These are full-time, 12-month, tenure-track positions at the rank of Assistant/ Associate/ or full Professor depending upon qualifications. Successful candidates will be located at the COMP-Pomona campus. For the biochemistry position, applicants need to have a PhD with at least 2 years of postdoctoral experience, current funding or a strong potential to obtain extramural grant support for their research program, research publications, and evidence of teaching biochemistry topics. For the microbiology discipline leader position, applicants must have several years of excellent experience of teaching microbiology and virology topics, significant scholarly activity and research publications, and consistent record of extramural funding. Submit a current curriculum vitae and a cover letter describing your teaching philosophy, research activity and future goals. Please include contact information for at least three references. These positions will remain open until filled.

Nissar A. Darmani, PhD
Associate Dean for Basic Sciences and Research
Chair, Department of Basic Medical Sciences
College of Osteopathic Medicine of the Pacific
Western University of Health Sciences
309 E. Second Street, Pomona, CA 91766-1854
Email Address: ndarmani@westernu.edu

Western University of Health Sciences is an equal opportunity employer.

UNIVERSITY OF CAMBRIDGE

**www.jobs.cam.ac.uk**

## University Lecturer

Department of Physiology, Development and Neuroscience
Salary: £38,511-£48,743

Applications are invited for a research-oriented, tenure-track University Lectureship in Physiology, available from 1 November 2015 or as soon as possible thereafter. Based in central Cambridge, we are searching for an outstanding scientist, with an excellent publication record, undertaking cutting edge and fundable work in a field which complements our world class programme in physiology with a preference for comparative physiology or mammalian systems physiology research (http://www.pdn.cam.ac.uk/research/).

The successful applicant will have a PhD in a relevant subject area, will have an aptitude and enthusiasm for teaching, and be willing to contribute effectively to our undergraduate programmes in preclinical medicine, preclinical veterinary science, and natural sciences (http://www.pdn.cam.ac.uk/teaching). He/she will be expected to contribute to the design and delivery of undergraduate and graduate lecture courses and to perform other academic duties including administration, examining and assessment.

We welcome applications from all qualified candidates and we strongly encourage women to apply. Appointment will be based on merit alone.

Informal enquiries may be made to Professor Bill Harris, Head of Department (wah20@cam.ac.uk; +44 (0) 1223 766137/333772).

Appointment will be made at University Lecturer level with a probationary period of five years, with appointment to the retirement age thereafter. The pensionable salary scale starts at £38,511 to £48,743 per annum. Once an offer of employment has been accepted, the successful candidate will be required to undergo a health assessment, with a satisfactory outcome determined by the University.

**To apply online for this vacancy and to view further information about the role, please visit: http://www.jobs.cam.ac.uk/job/6620**

**Please ensure that, in the Upload section of the online application, you upload your Curriculum Vitae (CV), a research publication list which highlights your three most important published outputs, and a covering letter of no more than two pages, outlining current and future plans, indicating the fit within the research profile of the Department and how you might contribute to our teaching. If you upload any additional documents which have not been requested, we will not be able to consider these as part of your application.**

**Please quote reference PM05747 on your application and in any correspondence about this vacancy.**

**Closing Date: 30 June 2015**

The University values diversity and is committed to equality of opportunity.

The University has a responsibility to ensure that all employees are eligible to live and work in the UK.


## Download Free Career Advice Booklets!

ScienceCareers.org/booklets

**Featured Topics:**

- Networking
- Industry or Academia
- Job Searching
- Non-Bench Careers
- And More

ScienceCareers

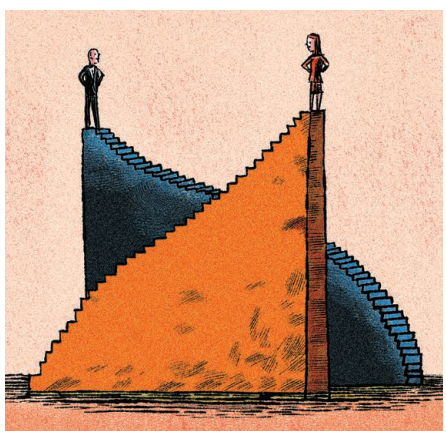
FROM THE JOURNAL SCIENCE AAAS

*By* **Maria Fadri-Moskwik**

# *Love knows no boundaries*

Because we are scientists, my husband and I have moved 13 times over the past 8 years, not always to the same place or in the same direction. Eight of those moves covered large swaths of the United States: South Carolina to Texas, Texas to Washington, Washington to Idaho. For 3 years now, we've been seeking a permanent place where we can live together and have good, long-term jobs in science—both of us. Now, those 3 years of searching and our 10 years together have come down to decisions we'll likely make in the next month: Will I follow him? Will he follow me? Will we take jobs in different places?

Things weren't supposed to end up this way. We both attended tier-one research institutions for our graduate degrees. Both he, a spatial ecologist, and I, a cell and molecular biologist, have publications in our fields. We both have generous reference letters. Years ago, we decided that I would pursue the academic track while he looked for a job with a nongovernmental organization; we hoped that would improve our odds of finding jobs in the same city. We've been married for 5 years and together for 10, but we've lived in the same city for only three. (Yes, it's a tough way to have a relationship, especially if starting a family is an eventual goal, as it is for us—but thank you, Google, for Google Chat.)



***"We've been married for 5 years and together for 10, but we've lived in the same city for only three."***

Now, my great opportunity is in Seattle, Washington. His is in Research Triangle Park, North Carolina. Both seem promising, but either could still fall through. The only certainties are that in the next 30 days, we'll commit to yet another move—at least one, somewhere, because my husband's job here in Boise is ending—and that whatever decisions we make, we'll make together. The choices are terrifying: My career over his? His over mine? Live thousands of miles apart, long term, on opposite coasts and not have a family? Break up?

It might be less terrifying if I had a clear preference for one over the other, my career over my husband (and nascent family) or vice versa. But I don't. I want them both. From graduate school and two postdocs to my first nontenure-track faculty position, I haven't lost my desire to push forward, bit by bit, the frontiers of my field—to watch with increasing precision the dance of proteins, DNA replication and repair, and chromosome segregation. When I celebrated another New Year's Eve with a trip to the core microscopy facility; when I spent my evenings last semester creating clinical problems for my students to work on during class; when, after I miscarried our first and so far only child, I returned to the lab the next day at the request of my principal investigator, I did so with the sense of cheerful self-sacrifice that comes from a deep desire to serve my institution, my students, and my science.

But my love for my husband and my hope for a family are deep and fundamental parts of who I am. We met over a plate of barbecue during my graduate training in Houston, Texas, and when he flew to Africa for a 3-year stay, we e-mailed and called frequently. He shared with me his experiences as a science teacher in Sierra Leone. I told him about my graduate work in Texas. The relationship was cemented when, after an encounter with a rabid cat forced his medical evacuation to Belgium, I flew in to be with him as he received the treatments that saved his life.

During that week and a half, I waited with him in a hotel room, looking out for signs of the disease. When none appeared, we returned to the United States and continued our interstate courtship. In time we married. He has supported me through my last year of graduate school, my first year teaching undergraduate biology, and the miscarriage of our child. Though we've lived apart for much of our marriage, neither my mind nor my heart can imagine a future without him. ■

*Maria Fadri-Moskwik is a cellular and molecular biologist and most recently a clinical assistant professor at Washington State University, Spokane. For more on life and careers, visit ScienceCareers.org. Send your story to SciCareerEditor@aaas.org.*

ILLUSTRATION: ROBERT NEUBECKER