AAAS

Science

24 JULY 2025

PROTECTING FISH

Marine protected areas limit illegal fishing—when rules are enforced

pp. 352, 392, & 396

Apply for the BII and Science Translational Medicine **Prize for Innovations in Women's Health**

BII Science Translational Medicine

Prize for Innovations in Women's Health

The prize seeks to recognize researchers who have developed innovative advances with translational potential to impact women's health.

BioInnovation Institute (BII) and Science Translational Medicine introduce a new annual prize to celebrate scientists whose research has advanced translation and innovation to address high unmet medical needs in the underserved fields of women's health.

The winner will have their essay published in Science Translational Medicine, will receive a cash reward of $25,000, and be invited to BII headquarters in Copenhagen, Denmark, for the grand award ceremony.

Contributions within any area of women's health will be considered. These include:

- Female specific conditions;
- Reproductive health;
- Conditions that affect females and males differently.

Deadline
November 1, 2025
bii.dk/community/prizes

BII BioInnovation Institute

Science Translational Medicine

CONTENTS

24 JULY 2025 | VOLUME 389 | ISSUE 6758

338

South Africa's Cradle of Humankind region hosts the world's richest hominin fossil deposits.

EDITORIAL

NEWS

FEATURES

COMMENTARY

POLICY FORUM

PERSPECTIVES

Science AAAS

CHEN TIANQIAO & CHRISSY INSTITUTE

CO-HOST ETH zürich

AI & MENTAL HEALTH

SEPTEMBER 8–9, 2025 | ZURICH

Mental health is a vast and growing worldwide problem. In addition to the suffering experienced by individuals and their affected family members it also has an enormous economic impact. Because it can affect people at a young age they will often need treatment for many years and during that period they will not be able to participate to their full potential in society.

REGISTER TODAY!

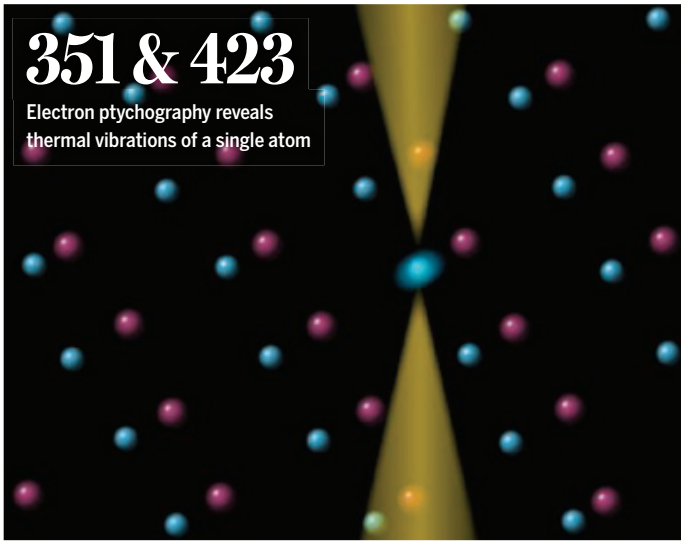


IMAGE: YICHAO ZHANG AND PINSHANE HUANG

ON THE COVER

Marine protected areas (MPAs) can protect fish and maintain healthy ecosystems, such as that of the Galapagos Islands, shown here, where a sea lion hunts salema fish. As demonstrated by two related papers in this issue, however, such areas are only effective against illegal fishing when established limitations are enforced. See pages 352, 392, and 396.
Photo: Enric Sala/National Geographic

ON THE PODCAST

Robots that eat other robots, and an ancient hot spot of early human relatives

WORKING LIFE

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 or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news, commentary, and book 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.** *Science* (ISSN 0036-8075) is published weekly on Thursday, except 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. Copyright © 2025 by the American Association for the Advancement of Science. The title *Science* is a registered trademark of the AAAS. Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $2865; Foreign postage extra: Air assist delivery: $135. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. 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:** $15 each plus shipping and handling available from backissues.science.org; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act can be obtained through the Copyright Clearance Center (CCC), www.copyright.com. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Science AAAS

EDITOR-IN-CHIEF **Holden Thorp**, hthorp@aaas.org

EXECUTIVE EDITOR Valda Vinson

EDITORS, RESEARCH Sacha Vignieri, Jake S. Yeston **EDITOR, COMMENTARY** Lisa D. Chong

DEPUTY EXECUTIVE EDITOR Lauren Kmec

DEPUTY EDITORS Stella M. Hurtley (UK), Phillip D. Szuromi **SENIOR EDITORS** Caroline Ash (UK), Michael A. Funk, Angela Hessler, Di Jiang, Priscilla N. Kelly, Marc S. Lavine (Canada), Sarah Lempriere (UK), Mattia Maroso, Yevgeniya Nusinovich, Ian S. Osborne (UK), L. Bryan Ray, H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Yury V. Suleymanov, Valerie B. Thompson, Brad Wible **ASSOCIATE EDITORS** Jack Huang, Sumin Jin, Bianca Lopez, Sarah Ross (UK), Madeleine Seale (UK), Corinne Simonti, Ekeoma Uzogara **SENIOR LETTERS EDITOR** Jennifer Sills **NEWSLETTER EDITOR** Christie Wilcox **NEWSLETTER INTERN** Jasmin Galvan **RESEARCH & DATA ANALYST** Jessica L. Slater **LEAD CONTENT PRODUCTION EDITORS** Chris Filiatreau, Harry Jach Sr. **CONTENT PRODUCTION EDITOR** Amelia Beyna **CONTENT PRODUCTION EDITORS** Anne Abraham, Robert French, Julia Haber-Katris, Nida Masiulis, Abigail Shashikanth, Suzanne M. White **SENIOR PROGRAM ASSOCIATE** Maryrose Madrid **EDITORIAL MANAGER** Joi S. Granger **EDITORIAL ASSOCIATES** Aneera Dobbins, Lisa Johnson, Jerry Richardson, Anita Wynn **SENIOR EDITORIAL COORDINATORS** Alexander Kief, Ronmel Navas, Isabel Schnaidt, Alice Whaley (UK), Brian White **EDITORIAL COORDINATORS** Samuel Bates, Clair Goodhead (UK), Kat Kirkman, Daniel Young **ADMINISTRATIVE COORDINATOR** Karalee P. Rogers **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI OFFICE MANAGER** Carly Hayward (UK) **ASI SR. OFFICE ADMINISTRATORS** Simon Brignell (UK), Jessica Waldock (UK) **COMMUNICATIONS DIRECTOR** Meagan Phelan **DEPUTY DIRECTOR** Matthew Wright **SENIOR WRITERS** Walter Beckwith, Joseph Cariz, Abigail Eisenstadt **WRITER** Mahathi Ramaswamy **SENIOR COMMUNICATIONS ASSOCIATES** Zachary Graber, Sarah Woods **COMMUNICATIONS ASSOCIATES** Kiara Brooks, Haley Riley, Mackenzie Williams

NEWS EDITOR Tim Appenzeller

NEWS MANAGING EDITOR John Travis **INTERNATIONAL EDITOR** David Malakoff **DEPUTY NEWS EDITORS** Rachel Bernstein, Martin Enserink, David Grimm, Eric Hand, Michael Price, Kelly Servick, Matt Warren (Europe) **SENIOR CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis **ASSOCIATE EDITORS** Michael Greshko, Katie Langin **NEWS REPORTERS** Jeffrey Brainard, Adrian Cho, Phie Jacobs, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **CONSULTING EDITOR** Elizabeth Culotta **CONTRIBUTING CORRESPONDENTS** Vaishnavi Chandrashekhar, Dan Charles, Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Virginia Morell, Dennis Normile (Tokyo), Catherine Offord, Cathleen O'Grady, Elisabeth Pain (Careers), Charles Piller, Zack Savitsky, Richard Stone (Senior Asia Correspondent), Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **INTERNS** Nazeefa Ahmed, Annika Inampudi **COPY EDITORS** Julia Cole (Senior Copy Editor), Hannah Knighton, Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

CREATIVE DIRECTOR Beth Rakouskas

DESIGN MANAGING EDITOR Chrystal Smith **GRAPHICS MANAGING EDITOR** Chris Bickel **PHOTOGRAPHY MANAGING EDITOR** Emily Petersen **MULTIMEDIA MANAGING PRODUCER** Kevin McLean **DIGITAL DIRECTOR** Kara Estelle-Powers **DESIGN EDITOR** Marcy Atarod **DESIGNER** Noelle Jessup **SENIOR SCIENTIFIC ILLUSTRATOR** Noelle Burgess **SCIENTIFIC ILLUSTRATORS** Austin Fisher, Kellie Holoski, Ashley Mastin **SENIOR GRAPHICS EDITOR** Monica Hersher **GRAPHICS EDITOR** Veronica Penney **SENIOR PHOTO EDITOR** Charles Borst **PHOTO EDITOR** Elizabeth Billman **SENIOR PODCAST PRODUCER** Sarah Crespi **SENIOR VIDEO PRODUCER** Meagan Cantwell **SOCIAL MEDIA STRATEGIST** Jessica Hubbard **SOCIAL MEDIA PRODUCER** Sabrina Jenkins **WEB DESIGNER** Jennie Pajerowski

CHIEF EXECUTIVE OFFICER AND EXECUTIVE PUBLISHER

Sudip Parikh

PUBLISHER, SCIENCE FAMILY OF JOURNALS Bill Moran

DIRECTOR, BUSINESS OPERATIONS & ANALYSIS Eric Knott **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR MANAGER, BUSINESS ANALYSIS** Cory Lipman **BUSINESS ANALYSTS** Kurt Ennis, Maggie Clark, Isacco Fusi **BUSINESS OPERATIONS ADMINISTRATOR** Taylor Fisher **DIGITAL SPECIALIST** Marissa Zuckerman **SENIOR PRODUCTION MANAGER** Jason Hillman **SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS** Marcus Spiegler **CONTENT OPERATIONS MANAGER** Rebecca Doshi **PUBLISHING PLATFORM MANAGER** Jessica Loayza **PUBLISHING SYSTEMS SPECIALIST, PROJECT COORDINATOR** Jacob Hedrick **SENIOR PRODUCTION SPECIALIST** Kristin Wowk **PRODUCTION SPECIALISTS** Kelsey Cartelli, Audrey Diggs **SPECIAL PROJECTS ASSOCIATE** Shantel Agnew

MARKETING DIRECTOR Sharice Collins **ASSOCIATE DIRECTOR, MARKETING** Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **ASSOCIATE DIRECTOR, MARKETING SYSTEMS & STRATEGY** Aimee Aponte **SENIOR MARKETING MANAGER** Shawana Arnold **MARKETING MANAGER** Ashley Evans **MARKETING ASSOCIATES** Hugues Beaulieu, Ashley Hylton, Lorena Chirinos Rodriguez, Jenna Voris **MARKETING ASSISTANT** Courtney Ford **SENIOR DESIGNER** Kim Huynh

DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING Erika Gebel Berg **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SENIOR TRAFFIC ASSOCIATE** Christine Hall

DIRECTOR, PRODUCT MANAGEMENT Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **ASSOCIATE DIRECTOR, PUBLISHING INTELLIGENCE** Rasmus Andersen **SR. PRODUCT ASSOCIATE** Robert Koepke **PRODUCT ASSOCIATES** Caroline Breul, Anne Mason

ASSOCIATE DIRECTOR, INSTITUTIONAL LICENSING MARKETING Kess Knight **ASSOCIATE DIRECTOR, INSTITUTIONAL LICENSING SALES** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Nazim Mohammedi, Claudia Paulsen-Young **SENIOR MANAGER, INSTITUTIONAL LICENSING OPERATIONS** Judy Lillibridge **MANAGER, RENEWAL & RETENTION** Lana Guz **SYSTEMS & OPERATIONS ANALYST** Ben Teincuff **FULFILLMENT ANALYST** Aminta Reyes

ASSOCIATE DIRECTOR, INTERNATIONAL Roger Goncalves **ASSOCIATE DIRECTOR, US ADVERTISING** Stephanie O'Connor **US MID WEST, MID ATLANTIC AND SOUTH EAST SALES MANAGER** Chris Hoag **DIRECTOR, OUTREACH AND STRATEGIC PARTNERSHIPS, ASIA** Shoupeng Liu **SALES REP, ROW** Sarah Lelarge **SALES ADMIN ASSISTANT, ROW** Victoria Glasbey **DIRECTOR OF GLOBAL COLLABORATION AND ACADEMIC PUBLISHING RELATIONS, ASIA** Xiaoying Chu **ASSOCIATE DIRECTOR, INTERNATIONAL COLLABORATION** Grace Yao **SALES MANAGER** Danny Zhao **MARKETING MANAGER** Kilo Lan **ASCA CORPORATION, JAPAN** Rie Rambelli (Tokyo), Miyuki Tani (Osaka)

DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS Emilie David **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **LICENSING ASSOCIATE** Virginia Warren **RIGHTS AND LICENSING COORDINATOR** Dana James **CONTRACT SUPPORT SPECIALIST** Michael Wheeler

EDITORIAL
science_editors@aaas.org

NEWS
science_news@aaas.org

INFORMATION FOR AUTHORS
science.org/authors/
science-information-authors

REPRINTS AND PERMISSIONS
science.org/help/
reprints-and-permissions

MULTIMEDIA CONTACTS
SciencePodcast@aaas.org
ScienceVideo@aaas.org

MEDIA CONTACTS
scipak@aaas.org

PRODUCT ADVERTISING & CUSTOM PUBLISHING
advertising.science.org
science_advertising@aaas.org

CLASSIFIED ADVERTISING
advertising.science.org/
science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE
employers.sciencecareers.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL SUBSCRIPTIONS
science.org/subscriptions

MEMBER BENEFITS
aaas.org/membership/
benefits

INSTITUTIONAL SALES AND SITE LICENSES
science.org/librarian

Eighty years at the nuclear brink

Ernest J. Moniz

Eighty years ago, the atomic bombings of Hiroshima and Nagasaki ushered in the nuclear age with weapons of previously unimagined destructive power. Since then, an arms race led to a global inventory of 70,000 nuclear weapons, followed by a period of arms control that brought the number down to almost 12,000 today. Despite this tremendous achievement, nuclear arms still figure prominently in global crises. The world is in desperate need of both short-term nuclear risk reduction and a longer-term vision for enduring peace and stability.

Today's nuclear weapons—some of which are 100 times more powerful than those used by the United States against Japan at the end of World War II—could inflict catastrophic damage on civilization. Given threatening geopolitical trends, including regional instability and conflicts, the risk that nuclear weapons will be used again is widely viewed as higher than at any time since the 1962 Cuban Missile Crisis. Among the nine nuclear weapons states, a majority—the United States, Russia, Israel, India, Pakistan, and North Korea—have been involved in armed conflicts over the past year, with the attendant risk of escalation to nuclear warfare. Recent military strikes on Iran by the United States were tied to its ambiguous position as a threshold nuclear state—a country with technical capability and resources to quickly produce nuclear weapons. This rising threat level has been further punctuated by Russia crossing a line with nuclear saber rattling at a non-nuclear weapon state, Ukraine.

Today's nuclear weapons... could inflict catastrophic damage on civilization.

Against this backdrop, world leaders need a two-pronged approach to move back from the nuclear brink: near-term steps that minimize the risk of nuclear weapons use, and a long-term fundamental rethinking of strategic relationships that preserve stability and of the role of nuclear weapons toward this end.

Near-term risk reduction steps should include reopening talks between the United States and Russia to extend the New Strategic Arms Reduction Treaty (New START)—the last remaining limitation on the two largest stockpiles (each country has over 5000 nuclear warheads)—before it expires in February 2026, until a long-term replacement is negotiated. New START allows each country to inspect the other's nuclear arsenal multiple times each year. Also critical is ensuring, through diplomacy, that once and for all, Iran's nuclear program will be used only for peaceful applications, such as nuclear energy development. There must also be an effort by all nuclear states to ensure that their weapons are as safe as possible from accidental use through "fail-safe" reviews to mitigate risk. Following the lead of the United States, each country would conduct a unilateral risk assessment, both in its own interest of avoiding accidental use and for the collective good of avoiding missteps into a nuclear exchange.

As for the longer term, leaders must work creatively to develop the architecture for a new era of strategic stability. Nuclear nonproliferation remains central to global security but may need to reflect post–Cold War geopolitical realities. For example, how the US continues its commitment to defend its allies is central for discouraging new nuclear weapons programs. Meanwhile, China's move to develop a nuclear weapons capability similar to those of the US and Russia requires a fundamental reinvention of arms control. The Cold War paradigm for managing nuclear deterrence, which was based on bilateral agreements between the US and the Soviet Union to preserve approximate numerical nuclear weapons equality, is not viable with three major players.

Going beyond the three great powers, the rise of regional nuclear powers, such as North Korea, India, and Pakistan, requires new approaches to containing risks in the context of regional security challenges. How can each of these smaller nuclear nations ensure its own safety without setting off a regional arms race? Old assumptions that nuclear deterrence is based on the actions of rational actors and that security can be achieved through the threat of mutual annihilation no longer apply, especially in light of the persistent threat of terrorism and the heightened role of non-state actors in the 21st century.

One of the victims of Hiroshima was a girl named Sadako Sasaki. Two years old in 1945, she succumbed to radiation-induced leukemia 10 years later. From her hospital bed, she tried to fold 1000 origami cranes, an act that according to Japanese tradition reflected her wish to live. This August, the world can honor Sadako and the "hibakusha" survivors of Hiroshima and Nagasaki by working to advance the vision of nuclear risk reduction and the ultimate elimination of nuclear weapons. Perhaps the simple act of folding a crane for peace—by scientists, cultural and faith leaders, policymakers, students, and others around the globe—can mark a new path away from nuclear conflict. □

Ernest J. Moniz is the chief executive officer of the Nuclear Threat Initiative, Washington, DC, USA and an emeritus professor of physics at the Massachusetts Institute of Technology, Cambridge, MA, USA. moniz@nti.org

10.1126/science.aea6515

SCIENTIFIC PUBLISHING

Science retracts 'arsenic life' paper despite protests

Study authors and others question belated decision, whereas some critics of work are pleased **CATHERINE OFFORD**

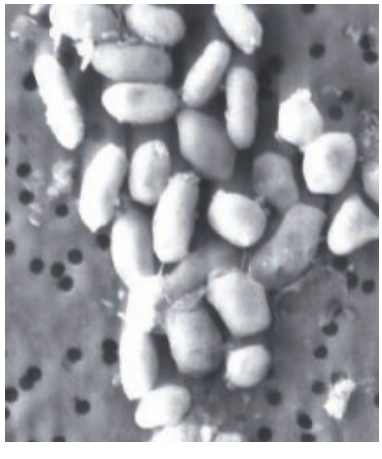
California's Mono Lake was the source of a bacterium (below) that a research team claimed incorporated arsenic into its molecules.

The claim was striking, the criticisms ferocious: A 2010 *Science* paper reported that a microbe found in the inhospitable waters of Mono Lake in California could, unlike other known life forms, use arsenic to grow. Now, on p. 357, the journal has retracted the work. But instead of closing the books on one of its most infamous papers, the decision has stirred the old controversy—and sparked new debate on when the scientific record should formally be "corrected."

Several long-standing critics of the research applaud the retraction. But the study's authors have protested the decision and the journal editors' justification, which includes an accusation of flawed data. Others question the retraction's timing, 15 years after the paper's publication but just months after *The New York Times* profiled the first author's return to research following the scandal.

And some condemned what they said was an unnecessarily harsh measure in a case with no suggestion of research misconduct. "I don't see this as serving any purpose," says NASA scientist Mary Voytek, whose agency funded the work and who spoke at an early press conference about the findings. "It doesn't feel constructive."

The retracted study concerned GFAJ-1, a bacterium collected in Mono Lake's arsenic-filled waters by Felisa Wolfe-Simon, then a researcher at the NASA Astrobiology Institute and the U.S. Geological Survey. Citing experiments on the microbe, she and colleagues suggested it used the normally toxic element in place of phosphorus to make DNA and proteins. A NASA release claimed the discovery "broadens our thinking about the possibility of life on other planets, and begs a rewrite of biology textbooks." But the work swiftly attracted challenges by other researchers, and attempts to replicate it failed.



In a blog post (https://scim.ag/ArsenicRetraction) accompanying this week's retraction notice, *Science*'s current Editor-in-Chief Holden Thorp and Valda Vinson, executive editor of the *Science* family of journals, emphasize there is no suggestion of foul play in the GFAJ-1 paper. Instead, pointing to subsequent commentary and research that suggest some of the paper's findings stem from contamination, not arsenic use by bacteria, they write: "*Science* believes that the key conclusion of the paper is based on flawed data."

PHOTOS: (TOP TO BOTTOM) HENRY BORTMAN/ZUMA; NASA

Speaking with *Science*'s News team, which operates independently from its research arm, study co-author and Arizona State University geochemist Ariel Anbar says the team disputes that assessment and has already addressed the referenced criticisms. "We stand by the data," he adds.

After controversy sparked by the paper's online publication delayed its appearance in print, *Science* ultimately packaged the work in 2011 with eight "technical comments" challenging the paper's conclusions and proposing alternative explanations. The following year, it published two papers suggesting GFAJ-1's growth did require phosphate like other known life and that although the microbe could tolerate arsenic, it didn't incorporate the element into DNA at significant levels. To many, the matter was settled, although some still wanted a retraction and the authors have repeatedly defended their findings. They argue that the 2012 papers didn't faithfully reproduce the original study.

Thorp, who became editor-in-chief of *Science* in 2019, notes that the question of retracting the study kept arising. After *The New York Times* contacted *Science* about the saga late last year, there was "a realization that we needed to go ahead and decide what we're going to do," he says. Otherwise, "It's just going to come around again."

Retractions by *Science* and other journals have become more common for papers blighted by major errors, not just the more serious charge of misconduct, Thorp says. The rise of artificial intelligence (AI)-powered literature searches makes this formal act especially important for correcting the scientific record, he argues. "Until you put the word 'retracted' in the title, you don't trigger all of the [AI] tools."

Science accepts responsibility for flawed peer-review and editorial decisions on the paper, he and Vinson note in their post. The journal should have sought reviewers with a wider range of expertise, Thorp says. It also should have condemned the attacks on Wolfe-Simon that appeared on social media and elsewhere, he adds. (Bruce Alberts, *Science*'s editor-in-chief from 2008 to 2013, did not respond to a request for comment.)

In recent months, *Science* has engaged with the authors and the Committee on Publication Ethics (COPE), which produces guidance and advises on specific cases. COPE's criteria for retraction have long included major errors that undermine a paper's findings, not just fabrication and falsification. "Retractions are a mechanism to correct the published record and should not be seen as punitive," says COPE Chair Nancy Chescheir in an emailed statement.

In their published response to the retraction notice, which the team saw in advance and—unlike the editors' blog post—doesn't cite data flaws, Anbar and colleagues say their work "could have been written and discussed more carefully." But they say disputes about papers' conclusions are "a normal part of the process of science." "We entirely reject the assertion of a major flaw," Anbar says.

Now-retired University of British Columbia microbiologist Rosie Redfield, an early critic of the original study, is among those pleased, however. She called the retraction "especially important because the paper's incorrect conclusion was so dramatic and received so much publicity."

Richard Sever, chief science and strategy officer at openRxiv, which oversees preprint servers bioRxiv and medRxiv, found *Science*'s move was reasonable, but questioned why it hadn't been done before. Ivan Oransky, co-founder of the blog Retraction Watch, expressed a similar sentiment: "Does it require a major news outlet calling for a journal to do the right thing?" he says.

Although Oransky commends Thorp's proactive approach to engaging with problematic papers, he and several others note that if *Science* were to apply the same criteria to all its publications, it would likely have to retract many more. Voytek, now exobiology branch chief at NASA's Ames Research Center, instead questions whether errors—which she says haven't been proved here—should compel retractions at all, as there are less stigmatizing procedures, such as errata, for such situations.

Thorp denies the decision marks a new approach, saying: "We've retracted plenty of papers for similar kinds of errors over the years." There will be no sweeping review of *Science*'s back issues, either, though the journal continues to investigate readers' concerns about published research on a case-by-case basis, and tries to give a "straight answer" about decisions, he says. "This is a paper we've been asked about over and over and over again. … It took longer than it should, but we're giving a response now." □

IN OTHER NEWS

PROTESTS SPREAD TO NSF, NASA

Some 149 employees of the National Science Foundation this week denounced steps taken by President Donald Trump's administration that they say have politicized federal science and are making the country less competitive. The NSF petition, which calls out proposed budget cuts, staff firings, and "a covert and ideologically driven secondary review process by unqualified political appointees," came a day after a similar Voyager Declaration from 287 current and former NASA scientists. The latter decried "rapid and wasteful changes" at the space agency and was addressed to NASA's interim head. The NSF petition was sent to Representative Zoe Lofgren (CA), the top Democrat on the House of Representatives science committee, who has repeatedly accused Trump of waging an assault on science. Both follow protest letters last month signed by hundreds of scientists and staff at the National Institutes of Health (NIH) and the Environmental Protection Agency. —*Science* News staff

RUSSIA TO VET INTERNATIONAL PARTNERSHIPS Russian research institutions will soon be obliged to report all scientific collaborations with foreign citizens to the country's security services. Under the new law, set to come into effect on 1 September, organizations will need to register both existing and future international agreements on a state database, where they will be vetted by the Russian Federal Security Service. If a project fails to meet the agency's criteria—which have not been disclosed—then it will be terminated. The government says the law aims to prevent the unauthorized transfer of scientific results outside of the country. But some scientists interviewed by *Science* say the changes will further isolate Russian researchers from the global community. —*Science* News staff

NIH CAPS PROPOSALS TO SIX PER YEAR

Scientists hoping to obtain NIH funding will soon be limited to submitting six applications per calendar year, according to a notice the agency released last week. Although the policy is designed to prevent researchers from overwhelming NIH's grant-review system with large numbers of artificial intelligence–generated proposals, some critics worry it may discourage collaboration and hinder scientists who are already struggling to obtain federal grants amid NIH funding freezes and as the Trump administration seeks to cut the agency's annual budget by more than one-third. —Phie Jacobs

PHOTO: PHIL MOORE/AFP VIA GETTY IMAGES

CLIMATE CHANGE

‘This is delicate.’ Chemically altering swamps could help curb methane emissions

Rising levels of potent greenhouse gas prompt a new kind of geoengineering **PAUL VOOSEN**

Emissions from vast tropical swamplands in the Sudd, in South Sudan, have helped drive global methane levels to new highs.

With their sluggish, oxygen-starved waters and buffets of organic muck, wetlands are ideal habitats for anaerobic microbes that produce methane, a potent greenhouse gas. In a menacing feedback, a warming climate seems to be turbocharging that microbial activity, helping drive a global surge in methane—which leads to more warming. To interrupt it, some researchers are exploring an exotic idea: hacking the swamps.

Starting this year, an effort called Feedback Research and Action on Methane Emissions (FRAMES) plans to treat wetland microbial ecosystems with chemicals that could curb their methane releases, first in the lab and later in small field trials. The researchers stress that this new kind of geoengineering could have unforeseen consequences; understanding them is a primary reason for the work. “This is delicate,” says Brian Buma, who co-leads the Climate Innovation Initiative at the Environmental Defense Fund, the U.S. nonprofit backing the effort. “We’re in uncharted waters.”

Methane levels have risen by nearly 10% in the past 2 decades, and the gas is thought to be responsible for one-third of overall global warming. Researchers know the main source of the rising methane levels isn’t the oil and gas industry, because atmospheric methane has become more enriched in carbon-12, a light isotope favored by biology.

The growth of cattle ranching and landfills is certainly playing a role, but in recent years, researchers have realized that emissions from wetlands are also rising. Global warming is shifting rainfall patterns, expanding swamps in some regions and thawing permafrost to create new ones. The warming also favors methanogens: Heat speeds up their metabolisms, and warmer waters hold less dissolved oxygen, making it easier for the anaerobic organisms to dominate ecologically.

Last year, a team of U.S. Department of Energy scientists found that methane emissions from wetlands across northern Eurasia and North America increased 9% from 2002 to 2021. And by carefully examining shifts in the seasonal ebb and flow of methane at long-term monitoring sites, a study published in May in *Nature* found that global wetland emissions have been increasing since the 1980s. Both studies tied the trend to global warming. “It’s inescapable that this is all happening,” says Paul Palmer, an atmospheric chemist at the University of Edinburgh.

The best way to stanch the methane emissions would be to stop the wetlands from warming in the first place. That means burning less fossil fuels, says Emily Ury, an ecologist at Colgate University who published a review last year with Buma laying out ideas for how the wetland feedback could be mitigated. But absent such action, it’s worth at least studying artificial interventions, she says.

One well-known way to curb wetland methane comes from an earlier environmental crisis. Last century, coal-fired power plants spewed huge amounts of sulfur into the air, creating acid rain. These emissions damaged human and ecosystem health alike. But the acid rain also caused a notable decline in the methane produced by swamps. The extra sulfate in the system, researchers found, allowed sulfate-consuming bacteria to outcompete methanogens, starving them out.

Today that lesson plays out in the rice paddies of East Asia, where farmers have unintentionally reduced methane emissions by adding sulfate-rich gypsum to alleviate salt stresses on the plants. Other additives, such as nitrate, iron, and biochar—a form of charcoal used as a soil conditioner—are also promising.

As part of FRAMES, Ury and Meredith Holgerson, an ecologist at Cornell University, are establishing an array of soil cores—"wetlands in a jar"—to test these options. Wetlands are so variable that even if a treatment is successful on one type, it might totally fail on another, Ury says. "We want to do these experiments in a really reproducible way."

Field trials of sulfate could also come as soon as this year. Irena Creed, a hydrologist at the University of Toronto, says one of her partners, the conservation organization Ducks Unlimited, is eager to see whether it can stanch methane emissions from restored wetlands it oversees in Canada. Such a trial would involve a baseline study of the chemistry and microbiology of the swamp and its greenhouse gas emissions, followed by careful monitoring after the sulfate treatment. Creed also works on wetlands in the cities and farmland of the Prairie Pothole region of western Canada, and hopes to start trials there in the next 18 months.

When these chemicals are deployed, the emphasis will be on monitoring for potential negative consequences to the ecosystem, Creed says. For example, sulfates have been shown to boost the rate at which microbes convert mercury from pollution into methylmercury, a toxin that can bioaccumulate in aquatic species. "We can't just apply [an additive] for 12 hours and move on," Creed says.

There are other ways to reduce a swamp's methane emissions, notes Danielle Potocek, a biogeochemist at Spark Climate Solutions, a small nonprofit focused on methane mitigation. You can dig a channel to reconnect a coastal marsh to the sulfate-rich ocean. Or you could raise a swamp's water levels, helping prevent methane from escaping to the atmosphere. Vegetation type is also critical: Hollow sedges and grasses act like straws for methane, speeding up its escape. Aquatic ferns, meanwhile, are known to promote oxygenation in wetland soils, reducing methanogen populations.

As with many geoengineering proposals, "scale is the problem," Potocek adds. Regularly adding huge amounts of chemicals to wetlands would be costly and logistically infeasible, and no one wants to tamper with the vast unmanaged wetlands in Africa or South America, where much of the methane feedback could play out.

Perhaps a small way to start would be to incorporate the ideas into wetland restoration, a surging field worldwide, Buma says. Even if just a few projects prioritized methane, he says, "you could get a huge amount of additional climate benefit." □

IN FOCUS

These insects have a static superpower

Sap-sipping bugs called treehoppers come in an extraordinary range of shapes, distinguished by variations in a structure on their back called a helmet. But the appendage isn't just for show: It helps these insects detect electric fields, researchers reported this week in the *Proceedings of the National Academy of Sciences*. Scientists at the Natural History Museum Berlin studied 11 species in Costa Rica and found that an electrode repelled them. Their helmets' sharp points and long edges evidently foster a buildup of static electricity, which then bends tiny hairs. That electrical sense may help treehoppers tune into the fields generated by other insects. The researchers found that predatory wasps carry higher electrical charges than harmless bees, which suggests treehoppers may use electric fields to help distinguish friend from foe. —Erik Stokstad

PHOTO: ROBERT OELMAN/GETTY IMAGES

IN OTHER NEWS

BIOTECH HALTS GENE THERAPY The biotech company Sarepta Therapeutics has agreed to a request from the U.S. Food and Drug Administration (FDA) to stop distributing its gene therapy for Duchenne muscular dystrophy (DMD) following three patient deaths from liver failure. The therapy, marketed as Elevidys, uses a modified virus to carry into muscle cells a shortened gene for the protein dystrophin, which DMD patients lack. FDA approved the treatment in 2023 despite limited evidence that it slowed the progression of DMD. But two teenage boys who received Elevidys recently died, and last week the agency demanded the halt from Sarepta after it publicly disclosed the death of a man who got a similar, experimental gene therapy from the firm for a different muscle disease. Sarepta in June halted Elevidys use in wheelchair-bound patients but initially refused to do the same for younger boys who can still walk, noting no new safety issues in this group. It changed course earlier this week. —Jocelyn Kaiser

WETLAND RESTORATION EFFORTS SINK Efforts to stanch the loss of Louisiana's coastal wetlands—estimated to be sinking into the Gulf of Mexico at the rate of 10 city blocks' worth of land per day—were dealt a severe blow last week. The state announced it was canceling a $3 billion project aimed at creating 54 square kilometers of land by diverting muddy Mississippi River water into nearby Barataria Bay. Although the Mid-Barataria Sediment Diversion project had once enjoyed bipartisan support, Republican Jeff Landry, a vocal critic, became Louisiana governor in 2024. The state agency in charge of the effort cited costs and concern over impacts on local shrimp and oyster fisheries as reasons for scuppering it. The agency said it would now pursue smaller land restoration projects in the area. But an alliance of conservation groups dismissed that approach as shown to be ineffective and condemned the state's change of heart as "a complete abandonment of science-driven decision-making." —Warren Cornwall

PHOTO: KAI-YANG LIN/ASIAA

ASTRONOMY

All-sky radio telescopes ditch the dish

Arrays of simple antennas hunt for nearby fast radio bursts

DENNIS NORMILE

Think of a radio telescope, and you may picture a massive dish, tens or even hundreds of meters across. That's the classic design, which gathers radio waves from space and focuses them onto an antenna above the dish, creating a field of view that resembles a narrow beam.

The Bustling Universe Radio Survey Telescope in Taiwan (BURSTT), a humble-looking device nestled among the trees at the Fushan Botanical Garden south of Taipei, does away with the dish. Instead, it packs 256 antennas, shaped like miniature pine trees, into a so-called phased array to catch radio waves directly. Computers digitally combine the signals from the antennas, synthesizing multiple beams that can view roughly half of the visible sky at once. The array has no moving parts, but the beams can be steered electronically to view other parts of the sky. "There's no common nomenclature for this gizmo yet," says University of California, Berkeley radio astronomer Dan Werthimer. "I call them all-sky, all-the-time telescopes."

BURSTT, which is expected to begin its sky surveys in the coming months, showcases the hottest new technology in radio astronomy. Three similar arrays are due to come online by the end of the year in China, the Netherlands, and California (see table, p. 333). "I see this kind of array design becoming central to radio astronomy in the coming years and decades," says Liam Connor, an astrophysicist at the Center for Astrophysics | Harvard & Smithsonian who is co-principal investigator for the Coherent All-Sky Monitor (CASM), an array being deployed at the Owens Valley Radio Observatory in California.

The downside of the approach is that, without a dish, the instrument's sensitivity drops. The arrays will likely detect only the brightest radio sources in the Milky Way and neighboring galaxies. Researchers are embracing that limitation, however. They are using the all-sky telescopes to hunt for fast radio bursts (FRBs), which are orders of magnitude more luminous than any other radio sources in the sky.

Since FRBs were first discovered in 2007, astronomers have identi-

Taiwan's BURSTT array uses 256 radio antennas to monitor half the sky at once.

fied a thousand or so of the mysterious, fleeting blasts. Theories suggest they might be generated by black holes, supernovae, or spinning, magnetized, stellar remnants called magnetars. But, with one exception, known FRBs are so distant that even the biggest telescopes can trace them only to their host galaxies and not to specific astrophysical objects. And because of their narrow field of view, those telescopes are rarely pointed in the right direction at the right time to catch events closer to Earth. Devices like BURSTT will cast a wide net to capture these nearby FRBs. "Finding [FRBs] in our Galaxy means we can then point a bunch of other telescopes in that direction and confirm what produces them," says Jessica Dempsey, director of ASTRON, the Netherlands Institute for Radio Astronomy, which plans to turn on its Arthropod array this month.

A major impetus for the new telescopes came in 2020, when astronomers captured the first and only confirmed Milky Way FRB and connected it to a magnetar called SGR 1935+2154. Astronomers had already been enlarging the fields of view of traditional radio telescopes by replacing the single antennas that lie at the dishes' foci with phased arrays. The Milky Way event suggested that by deploying the antenna arrays as direct detectors, observers could capture many more. "It was a game changer," says Ron Ekers, a retired radio astronomer at Australia's Commonwealth Scientific and Industrial Research Organisation.

Vikram Ravi, a radio astronomer at the California Institute of Technology and a CASM co–principal investigator, says the new telescopes "are very complementary" to existing FRB-hunting radio telescopes that probe deep space. Comparing nearby and distant events could, for example, reveal different FRB rates that depend on the galaxy type, offering clues to the conditions that give rise to FRBs and hints to how they evolve over cosmic time.

A bigger haul of FRBs could enable astronomers to use them as cosmic probes. Electrons in the wispy gases of the intergalactic medium distort radio waves, and the effect is greatest at low frequencies. By measuring those distortions in distant FRBs, astronomers can "weigh" the matter the signals have passed through on their way to Earth (*Science*, 19 June, p. 1250). Similarly, catching FRBs from the nearby universe will allow scientists to determine how much ordinary matter the Milky Way contains, which is "a major open question in astrophysics," Connor says.

The new dishless telescopes could also watch for signals from the colliding black holes and other cosmic cataclysms that generate gravitational waves. At present, gravitational wave observatories send out an alert when they detect an event, triggering a scramble to observe its source with other telescopes. All-sky monitors, in contrast, might capture radio signals from gravitational wave events as they occur. "We may not find anything, or it may be spectacular," Ekers says.

The search for extraterrestrial intelligence (SETI) might benefit as well. SETI hunters typically look for a continuous signal, as might be emitted by a transmitter, Werthimer says. But if the signal is like an intermittent lighthouse beam, "you're never going to find it with a telescope that looks at a little piece of the sky at a time," he says. "You really want one of these very wide field, all-sky [monitors]."

Tsinghua University's Li Di, who leads China's Cosmic Antennae pilot project, is expecting the unexpected. "If you go into a new parameter space, there's always something weird that pops up," he says. □

All ears

One of the hottest trends in radio astronomy is using antennas to capture a wide field of view without a dish. Four facilities are set to come online in the coming months.

NAME	LOCATION	SPONSOR	START OF SCIENCE OPERATIONS
Bustling Universe Radio Survey Telescope in Taiwan	Fushan Botanical Garden	Academia Sinica	To be determined
Cosmic Antennae	Hangzhou, China	Zhejiang Lab	August
Arthropod	Westerbork observatory	ASTRON, the Netherlands Institute for Radio Astronomy	September
Coherent All Sky Monitor	Owens Valley Radio Observatory in California	California Institute of Technology; Center for Astrophysics | Harvard & Smithsonian	Fall

RESEARCH BUDGETS

Europe plans to double flagship funding scheme

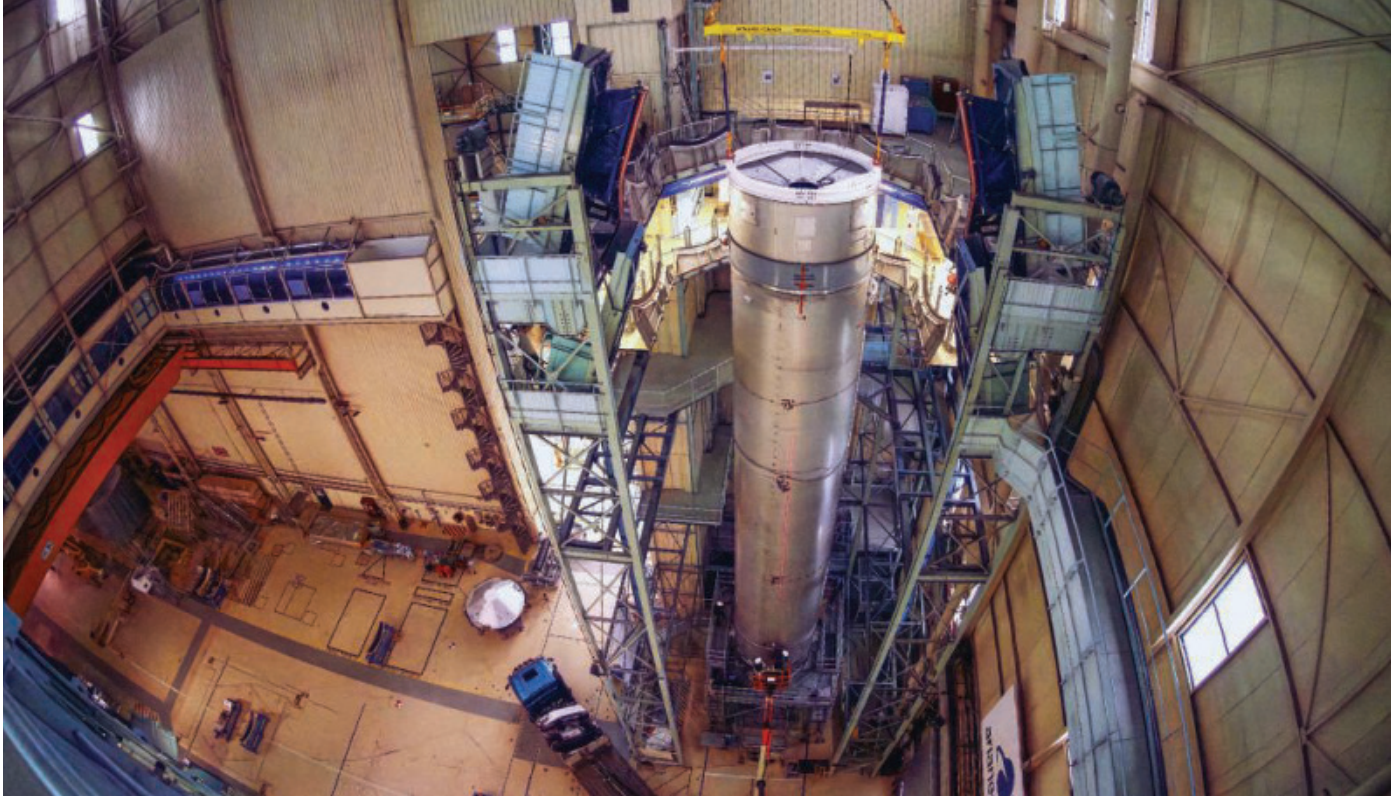
Starting in 2028, the next iteration of Horizon Europe will also focus on boosting competitiveness

CATHLEEN O'GRADY

The next iteration of Horizon Europe, the European Union's flagship research funding program, is finally starting to take shape, following months of speculation among the bloc's research community. In its long-term budget proposal last week, the European Commission revealed it plans to give the program €175 billion from 2028 to 2034—double the funds of the current 7-year program, which ends in 2027. The program will also have a renewed emphasis on funding research that will give the bloc a competitive edge, as well as "moonshot" projects in areas such as clean energy and quantum computing.

The budget boost represents a "significant increase in ambition" and a reflection of how important research is to Europe, says Martin Smith, a policy analyst at the Wellcome Trust. It arrives as, in the United States, President Donald Trump has proposed drastic cuts to federal research funding in his budget for the next fiscal year. But it's likely the sum will "be pared down over the coming couple of years of wrangling," says Smith, as the European Parliament and Council negotiate over the plans. In 2018, the Commission proposed a budget of €120 billion for the current program, which was cut by approximately 25% in later negotiations.

The money for Horizon will come from an overall €409 billion proposed for a new European Competitiveness Fund, which will fund "the strategic technologies of tomorrow,"



The Horizon Europe–funded SALTO project is developing Europe's first reusable space launch system. Space technologies are a priority for the next iteration of the program.

PHOTO: S. CORVAJA/EUROPEAN SPACE AGENCY

said European Commission President Ursula von der Leyen in last week's announcement. The plans include €131 billion for space and defense, and €51.5 billion for decarbonization. Impetus for the changes comes from a report on European competitiveness by economist Mario Draghi, who recommended the EU drastically increase its research funding levels and focus on technological advances.

But the new plans specify that Horizon will remain a distinct program. That eases concerns in the research community, says Kurt Deketelaere, secretary-general of the League of European Research Universities. When the idea of a competitiveness fund was first floated earlier this year, researchers began to worry it would absorb Horizon entirely, potentially sidelining basic research in favor of commercial R&D or work driven by political goals. Still, the proposals say Horizon will be "tightly-connected" to the competitiveness fund, without further detail. This lacks transparency, Deketelaere says. "One wonders why such a connection is needed at all."

The new Horizon program will be made up of four "pillars," one more than the current program. The first, Excellent Science, will focus on support for early-career researchers, and fund team research grants—including for basic science—through the European Research Council. The pillar will receive €44 billion in funding, up from approximately €25 billion currently. This means success rates for funding applications will go up, but "not as much as we would like," says Jan Palmowski, secretary-general of the Guild of European Research-Intensive Universities.

The second pillar, Competitiveness and Society, with €75.9 billion, will be the most tightly linked to the competitiveness fund, focusing on research of strategic benefit to the EU as well as research on "global societal challenges" such as disinformation, migration, and democracy. A third Innovation pillar will get €38.8 billion, mostly to support high-risk, high-reward research through the European Innovation Council, which draws inspiration from the Advanced Research Projects Agency and the Defense Advanced Research Projects Agency in the U.S. The new fourth pillar, with €16.2 billion, will focus on research infrastructure and supporting European countries that lag behind in research—programs that already exist, but not as a "pillar" of Horizon. The Commission also promises to use Horizon Europe to support technology-driven "moonshot" projects in areas such as fusion energy, regenerative therapies, and clean aviation.

In its proposal, the Commission says the new Horizon program will cut red tape, for which previous iterations were notorious. The changes—which include greater flexibility about the topic areas the program will fund and quicker decisions on applications—will "reduce burden and over complexity and favour speed and flexibility." Speedier grant decisions would be welcome, says Amanda Crowfoot, secretary-general of the European University Association, because "delays can cause promising ideas to lose momentum."

For countries that are outside the EU but hope to sign up to join the new Horizon program—a process known as "associating"—the proposal is a disappointment, as it does not change the arduous existing rules, Deketelaere says. Certain countries that have long historical relationships with Horizon—such as the United Kingdom and Switzerland—deserve a "fast-track" association procedure that could shortcut the lengthy negotiation process, he says. The proposal also does not say whether associated countries will have any say in the topics chosen for the Competitiveness pillar, Smith says.

Although there is still much to be worked out, the proposals are an achievement, Deketelaere adds, and will be a relief to many in the research community following months of "precarious" debate. Now, he says, it's up to Parliament and the Council to "improve them where necessary." □

U.S. FUNDING

New scheme for overhead payments faces doubts

FAIR Model aims to head off drastic cuts in government payments, but some say it's too costly and complex

JEFFREY MERVIS

It may seem counterintuitive. But a coalition of 10 U.S. higher education organizations hopes to block the threatened loss of billions of federal dollars with a plan that could result in the government paying out more, not less.

The recommendations from the Joint Associations Group on Indirect Costs (JAG), released last week, are meant to reframe a decadeslong debate over how to reimburse universities for the so-called indirect costs of research. Indirect cost payments are added to the costs of the actual research and cover functions ranging from maintaining libraries and animal care facilities to complying with research security rules.

President Donald Trump's administration and many members of Congress from both parties say the government is paying too much under the current system, in which reimbursement rates are negotiated between the government and individual institutions using a complicated formula. They also accuse universities of using the money for purposes unrelated to research. Four agencies—the National Institutes of Health (NIH), the National Science Foundation, the Department of Defense, and the Department of Energy—have tried to cut payments sharply by setting one low, flat rate but have been blocked, for now, by federal judges.

University administrators, in contrast, say they are being shortchanged because some costs are excluded and others capped. The new model, called Financial Accountability in Research (FAIR), is meant to achieve full cost recovery, says Kelvin Droegemeier, Trump's former science adviser and leader of JAG. But at a virtual town hall held last week to discuss the proposal, some university administrators worried the FAIR Model is both too complicated and, potentially, too costly for the government to embrace.

One target audience for JAG, which represents nearly every university, medical school, and independent research institution receiving federal funding, is the White House Office of Management and Budget. OMB is developing new regulations to implement dozens of Trump's executive orders pertaining to research, including how the government manages indirect costs. Congress could also address the issue in pending legislation that would set spending levels for the 2026 fiscal year that begins on 1 October.

In a major change to the existing method, the FAIR Model would end the periodic negotiations that set the reimbursement rate universities can charge. (A rate of 63%, not uncommon among major research institutions, turns a $1 million grant into a total award of $1.63 million, with indirect costs accounting for about 40% of the total.) Instead, universities would document exactly what they spent to support each project and include those costs in the overall grant budget.

The FAIR Model would eliminate the 26% cap on administrative costs, one of two current categories of indirect costs (facilities is the other), which university officials have long complained is too low. And most expenses now labeled as indirect costs would be renamed as "essential research performance support" and shifted into the pot for direct costs to bolster JAG's argument that indirect costs are not simply overhead.

"FAIR would make clear exactly what this money is being spent on, and that they are legitimate costs of doing research," says biomedical lobbyist Carrie Wolinetz of Lewis-Burke Associates, who works with an organization that is part of JAG.

The FAIR Model would preserve one catch-all category of indirect costs, labeled general research operations (GRO), that JAG says "are impractical to assign to a given project." At 15% of the total cost of the grant, GRO alone would generate roughly half of what the average institution now receives in indirect costs.

At the town hall last week, some participants worried the model's complexity could make it a hard sell. "We need something that policymakers—and the general public—can understand," said Patrick Clark, associate vice president for research at the University of Cincinnati.

Others noted that the FAIR Model seems likely to result in larger government payments. "I'd like to preserve what we have, instead of asking for more, which is what the new model basically does," Anshuman Razdan, vice president for research and innovation at the University of Oregon, said after the town hall. Although Droegemeier told the gathering he doesn't know whether institutions would gain or lose money, Razdan says under earlier FAIR Models Oregon would recoup up to 40% more of its costs than today.

Even so, some experts who advised JAG think the political climate is favorable for changes to the current system. "OMB is under pressure to get it done as soon as possible because of all these lawsuits," says Gil Tran, who spent 27 years at the OMB office that sets rules for research payments before joining the consulting firm Attain Partners. "But first OMB would have to do the math before deciding if that approach is consistent with the president's policies."

The cost of research extends well beyond people, materials, and lab equipment.

PHOTO: VINCENT ALBAN/REUTERS

Wolinetz, a former senior NIH administrator, predicts policymakers will welcome the added transparency the FAIR Model promises. The key, she says, will be "making the case for why FAIR is an improvement over the status quo. I feel confidence that we can do that," she adds. "But I don't want to pretend that will be an easy lift." □

Life scientists in the Boston area are exceptionally productive.

FUNDING

'Wandering scholars' analysis reveals how location drives productivity

Concentrating funding at high-powered universities can maximize output but may sacrifice broader benefits **KATIE LANGIN**

The mantra "location, location, location" isn't just about real estate. For life scientists, more than 50% of their productivity can be attributed to the institution where they work, according to a new study that tracked the publications of researchers as they moved during their career. The findings, published this month as a National Bureau of Economic Research working paper, play into an active and long-running debate over whether to prevent limited grant dollars from being concentrated at a handful of big-name universities—and shed light on the potential impact of the funding freeze the White House has imposed on Harvard University.

The study quantifies a phenomenon that's probably familiar to many academics, says University of California, Berkeley economist Carolyn Stein, who wasn't involved in the new study. It's easy to imagine that "you can pick someone up and move them to a more productive place, and it will make them more productive." Still, the magnitude of the observed effect is striking, she says. "The role of luck and path dependence in science is maybe larger than I've completely appreciated."

The new study compiled data for about 300,000 U.S.-based life scientists who published between 1945 and 2023. Researchers at Boston-area institutions had the highest productivity—publishing two or three times more papers per year in 15 journals that cover basic life science research, including *Cell*, *Nature*, and *Science*, compared with researchers in many other metropolitan areas. When a researcher moved from a less productive institution to one with higher average productivity, they became more productive as well, according to a "wandering scholars" analysis that included about 38,000 scientists who had a publication record from before and after moving between institutions.

The team wasn't able to pinpoint what institutional characteristics led to the productivity gains. Lead author Amitabh Chandra, an economist at Harvard, notes that "it could be something about resources, facilities, graduate students."

A previous study suggests access to collaborators may be a key. Published in *Science Advances* in 2022, it found that faculty who work in fields where collaboration and co-authorship are the norm were more productive at top U.S. universities in large part because they led larger lab groups. "It's this labor that increases faculty productivity, not really inherent characteristics of the faculty themselves," says Aaron Clauset, a computer scientist at the University of Colorado Boulder and senior author on the 2022 study.

Chandra and Harvard colleague Connie Xu began the new study well before President Donald Trump's administration halted all research funding to Harvard. But the findings indicate how much could be lost if funding isn't reinstated. "Entire countries produce less than what Harvard produces," notes Chandra, whose analysis found that his institution publishes 3.6% of the global output of top life science papers—the most of any single institution in the world. "When we turn off the funding to one of these large producers … the implications are colossal."

Recent political developments aside, the paper also bears on a decadeslong policy discussion about how to allocate federal grant money. Should funders aim to maximize output—resulting in a concentration of grants at relatively few elite institutions—or would it better serve the public to spread funding around more broadly?

The new study makes plain the benefits of favoring the elites: If two applicants for funding are otherwise comparable but one is at an institution with double the average research output of the other, "funders could get more than 50% more research by prioritizing a scientist at the more productive institution," Chandra and Xu write. But Chandra acknowledges it's also "perfectly valid" to base funding decisions on priorities other than research output.

Some argue against further enriching highly productive intuitions. "Those institutions are in a position to make faculty that much more productive … because they've received large amounts of research support over many, many decades," says Kirk Dombrowski, vice president for research and economic development at the University of Vermont who also serves as board chair for a nonprofit that promotes science in underfunded states. Allocating resources to them disproportionately would "reflect more of the historical inequities that have created differences," he says. A paper published in *PLOS ONE* in 2023 also suggests federal dollars go further in underfunded states, finding that their research communities publish more research articles and garner more citations per million dollars in grant funding than those in states with greater federal support.

The National Institutes of Health (NIH) and the National Science Foundation have long-running programs, called IDeA and EPSCoR, respectively, that reserve a portion of their budget for projects in states that receive the least funding through traditional funding tracks. As the prospect of major funding cuts looms, some worry about the future of such programs. But in comments last month at a Senate hearing about the upcoming fiscal year's budget, NIH Director Jayanta "Jay" Bhattacharya voiced his support. "It's absolutely vital that NIH investments are geographically dispersed," Bhattacharya told West Virginia Senator Shelley Moore Capito (R), whose state receives less than 1% of total NIH grant spending. "In my mind, it's probably … less funded than it ought to be." □

PHOTO: BOSTECHSTOCK/ALAMY

TRUMP ADMINISTRATION

EPA begins to dismantle its research office

Extensive layoffs could have wide-ranging impacts on environmental science **ERIK STOKSTAD**

In a sweeping reduction of the U.S. government's research capabilities, the Environmental Protection Agency (EPA) announced last week it will begin to dissolve its Office of Research and Development (ORD) and lay off staff. The toll is expected to include hundreds of scientists and their research on environmental hazards.

"I'm really worried that this will be a setback that has reverberations for a long, long time," says Jana Milford, an environmental engineer at the University of Colorado Boulder.

Since EPA's founding in 1970, the agency has had its own research arm. With 10 labs around the country and—at the start of President Donald Trump's second term—more than 1500 staff, ORD has historically pursued a wide range of research, including developing methods for detecting and testing chemicals, studying how toxic substances move through air and water, and informing legal standards for pollutants and cleanups. Those activities have long made ORD a target of regulated industries, who have argued its work has led to overly onerous and costly rules.

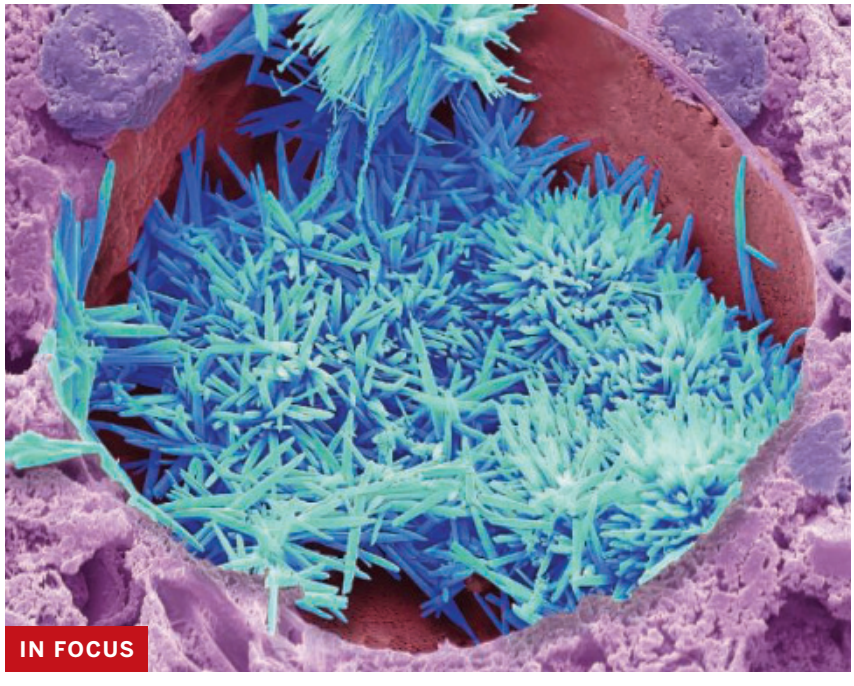
Their views have won favor in the Trump administration. The decision to eliminate ORD leaked in March, but lawsuits delayed the reorganization and layoffs. On 18 July, EPA said it will replace ORD with an Office of Applied Science and Environmental Solutions "to prioritize research and science more than ever before." The move will save $749 million, the agency says. Details remain scarce.

Demolishing ORD could have widespread implications, researchers say. It could, for example, delay or prevent development of new methods for assessing the risks of mixtures of chemicals, which can differ from those of individual substances, says Allen Burton, an environmental scientist at the University of Michigan. "It's a blow to the environment and wildlife, birds, fish, and humans to not be on the cutting edge."

ORD also maintains toxicology databases that are widely used by researchers, state and local governments, industry, and nonprofit groups, notes Joel Tickner, an environmental health scientist at the University of Massachusetts Lowell. Both Tickner and Milford fear EPA will no longer keep the databases up to date, such as by adding newly identified contaminants. Similarly, ORD's atmospheric and hydrologic models are used by researchers around the world to understand the sources and impacts of pollutants such as particulate matter and ozone, Milford notes.

Industry trade groups have been especially critical of ORD's Integrated Risk Information System assessments, which EPA and state agencies rely on to restrict the use of toxic chemicals and guide cleanup decisions. "This joust is nothing new, but this is total victory for the polluters right now," says Thomas Burke, an epidemiologist at Johns Hopkins University and former EPA science adviser.

More broadly, Milford and other researchers say ORD has been essential in building bridges between various forms of environmental science—a function that might get lost when its remaining scientists are scattered across EPA. ORD has also brought together academic and industry scientists in ways that improve environmental and toxicological research, Tickner says. The science prospers when it's insulated from political priorities, they add, which could change with ORD's demise. "What does the new EPA look like?" Tickner asks. "Very politically driven." □



IN FOCUS

CHOLESTEROL LIKE YOU'VE NEVER SEEN IT BEFORE An electron microscope image of a fat-laden liver cell took home this year's top Wellcome Photography Prize in the Marvels of Scientific and Medical Imaging category. Using a beam of electrons to zoom in on a 12-micrometer section, scientific photographer Steve Gschmeissner revealed the crystalized cholesterol embedded there, colorizing the image to make the cholesterol shards a vibrant blue against the purple folds of the liver cell. —Nazeefa Ahmed

PHOTO: STEVE GSCHMEISSNER

Homo *Australopithecus* *Paranthropus*

THE RIDDLE OF COEXISTENCE

Two million years ago, three kinds of hominins walked the same valley in South Africa. Did they meet—and compete?

ANN GIBBONS, at Kromdraai in South Africa

One morning in April 2014, José Braga squatted at the bottom of an open pit, cleaning a wall of red sediments with a trowel. Long ago, these rocks had formed the floor of a cave, and in 1938 they had yielded a spectacular skull of an early member of the human family, or hominin. But Braga had been scouring the sediments without luck for 12 years. He was considering throwing in his trowel and going off to search for fossils in Mongolia instead.

Then, a small, bright object fell from the wall above, bounced off his thigh, and landed in the dirt beside him. "I couldn't believe what I was seeing: a well-preserved hominin tooth!" recalls Braga, a paleoanthropologist at the University of Toulouse.

A few months later, Braga's team excavated a piece of a baby's upper jaw from the wall of the pit. The fallen molar fit perfectly into the jaw. Together, the tooth and jaw solidified the specimen's identity as an early member of our own genus, *Homo*.

It was a timely offering from the fossil gods here at Kromdraai, a roofless dolomite cave site in the rolling hills of a region of South Africa officially known as the Cradle of Humankind. This 470-square-kilometer UNESCO World Heritage Site is home to the largest known concentration of human ancestral remains anywhere in the world.

The very next year, Braga's team found another baby's jawbone. The two infants' remains had lain less than 30 centimeters apart for about 2 million years, but the new one was from a very different species: a baby *Paranthropus*, a short, robust hominin with massive molars and jaws. And an as-

Hominins who might have seen each other 2 million years ago in South Africa (left to right): *Homo erectus* (reconstruction based on an adult female), *Australopithecus sediba* (young male from Malapa), and *Paranthropus robustus* (adult male from Drimolen).

yet-unpublished skull found in 2019, just a few meters away, in sediments likely to be a bit older, is different again: It may belong to a third hominin genus, *Australopithecus*, a group of upright-walking apes with brains slightly larger than those of chimps.

That makes three kinds of hominins, including a species that was our direct ancestor, living in the same swampy valley roughly 2 million years ago. "Imagine the different hominins walking around … *Paranthropus* climbing the rock walls like baboons," Braga says.

As Braga showed visitors around in May, only one kind of hominin was visible in the valley: *Homo sapiens* day trippers from Johannesburg, jogging up Sterkfontein Road or cycling in colorful Lycra. But fossils from Kromdraai and other sites in the cradle and in East Africa show today's humans are the lone surviving branch of a remarkably bushy family tree, whose members once coexisted across the continent.

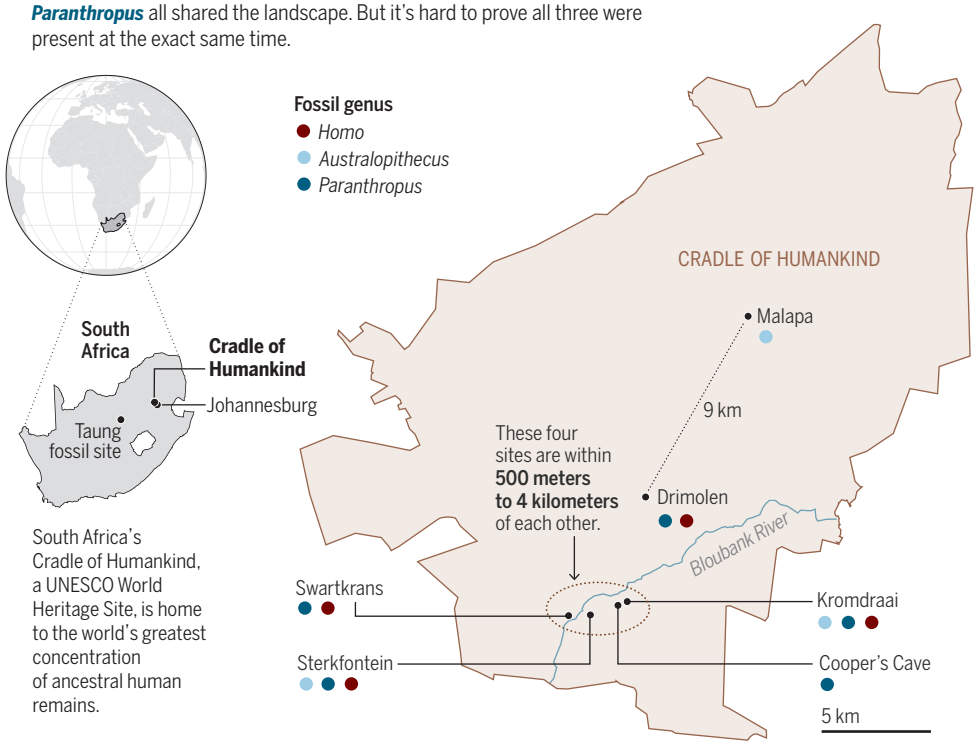
This view of the past has come into focus over the past decade as researchers have reopened and redated old sites and found new ones. They have discovered at least a half-dozen species of hominins that visited these caves from roughly 3 million years to 230,000 years ago, and have had particular success zooming in on a layer of time about 2 million years ago, when numerous hominins apparently thrived in the cradle. "Multiple sites in a small area suggest a diversity of hominins were all there, close in time," says Carol Ward, a paleoanthropologist at the University of Missouri who works in East Africa.

The cradle was prime real estate back then, says paleoanthropologist Kevin Hatala of Chatham University. The hominins walked a lush river valley with plentiful water, game, and fruiting trees, as well as caves for shelter. Hatala, who studies ancient footprints in Kenya that show two hominins overlapped there, says that in southern Africa, "There seems to be a couple of hundred thousand years when [multiple] species seem to be totally fine with sharing the landscape. There seems to be enough resources."

With new tools, including better dating and the analysis of proteins and isotopes in fossil teeth, researchers are comparing what these hominins ate, how they moved, how they grew, and how they adapted in different ways in shared territory.

Hominin hot spot

A selection of 2-million-year-old archaeological sites in the Cradle of Humankind shows that our own genus, ***Homo***; ***Australopithecus***; and a more robust hominin called ***Paranthropus*** all shared the landscape. But it's hard to prove all three were present at the exact same time.



GRAPHIC: M. HERSHER AND V. PENNEY/*SCIENCE*

The fossils' close proximity, in the same cave or within a short walk, suggests these creatures might have met, or at least been aware of one another. "They were both on this landscape for such an extensive period of time, there's no way they didn't interact with each other," says paleoanthropologist Stephanie Edwards Baker of the University of Johannesburg (UJ). She has found *Paranthropus* and early *Homo* in the same layers at nearby Drimolen cave with geochronologist Andy Herries of La Trobe University. In 2020, they proposed in *Science* (3 April, p. 47) that the region was a meeting ground for both genera as well as *Australopithecus*.

Other researchers are more cautious: "Just because I find you and an Egyptian in the same layer doesn't mean you saw each other," says paleoanthropologist Zeresenay Alemseged of the University of Chicago, who has found key fossils in Ethiopia. Given that one stratigraphic layer can contain tens of thousands of years of deposits, it's "impossible to prove" these species overlapped within any individual's lifetime, he says, Still, he acknowledges, "If indeed they were contemporaneous, [the cradle] is a good laboratory to show how hominins can coexist."

IN A CLIMATE-CONTROLLED vault at the University of the Witwatersrand (Wits) in Johannesburg, a tiny skull sits behind bulletproof glass: the Taung child, the first hominin fossil ever found in Africa, in 1924. Representing a new species called *Australopithecus africanus*, or southern ape of Africa, this skull launched paleoanthropology on the continent. Right beside it, a coffinlike glass case holds an astonishingly complete *Australopithecus* skeleton nicknamed Little Foot, its small foot bones splayed on blue velvet. Nearby lie two partial skeletons of *A. sediba*, another *Australopithecus* species from 2 million years ago. Wooden and plastic trays on shelves hold at least 15 skeletons of the much younger *H. naledi*, from the Rising Star cave 3 kilometers west of Kromdraai. This treasure trove holds much, much more, including 3500 hominin fossils collected over a century.

Some fossils were gathered by Scottish doctor and paleontologist Robert Broom in the 1930s. Broom went first to Sterkfontein cave, just 3 kilometers over a grassy hill from Kromdraai, and found adult *A. africanus*. He also bought a fossil palate from a quarry foreman for a few pounds and tracked its discovery back to Kromdraai Farm. There, in 1938, he discovered a partial

The site of Kromdraai in South Africa has yielded fossil treasure for decades, including the remains of three kinds of hominins.

skull and teeth of a new kind of robust "ape-man," now recognized as the type specimen of *P. robustus*, which was not our direct ancestor. (*Paranthropus* means "beside man.")

Broom and others realized that millions of years ago, the cradle was a magnet for many types of animals, including hominins, drawn to its reliable water sources. Rainwater has slowly dissolved the calcium carbonate in the region's bedrock of dolomite—a form of limestone—creating a porous karst landscape with bountiful springs, sinkholes, caves, and groundwater. "Water security was higher in the cradle" than on the interior grassy plateau, says Tebogo Makhubela, a geochemist at UJ. Hominins following the water found verdant river valleys lined with trees bearing figs and other fruit.

The cradle's karst landscape helped protect the fossil bounty: The calcium carbonate–rich water seeped into the bones, helping harden and preserve them for millions of years. "It gives us access to very detailed aspects of the anatomy of these creatures," Braga says.

For decades, however, the spotlight was on East Africa, where the Leakey family and many other researchers found spectacular fossils starting in the 1950s. Those discoveries include another species of *Paranthropus*, the oldest known *Homo*, and the famous partial skeleton of *A. afarensis* known as Lucy, as well as the world's oldest stone tools. The finds go back 6 million years, and many have been reliably dated from the decay of radioactive elements in the layers of volcanic ash that lace the sediments.

Researchers unearthed this jaw of a *Paranthropus robustus* infant at Kromdraai.

Given the long, well-dated record in East Africa, many researchers have thought that region was the scene of most evolutionary transitions, including from *Australopithecus* to *Homo*. Some saw the South African hominins as evolutionary dead ends stuck in a poorly dated cul-de-sac in Africa's southern tip.

But the discovery of Little Foot at Sterkfontein and other fossils in the 1990s spurred a new generation of researchers, including *National Geographic* explorer Lee Berger, then at Wits, who drew attention to the potential of South Africa.

In 2002, Braga and Wits paleoanthropologists Francis Thackeray and Bernhard Zipfel revisited Kromdraai. They identified an old test pit from the vintage Coca-Cola cans earlier generations of excavators had left behind, and dug deeper. After finding the *Paranthropus* baby molar and jaw, they won permission from the South African heritage authorities to move full-steam ahead. They scalped a flat area the size of a tennis court, removing the top meter of grass and rocky soil to expose an underground cave whose breccia roof had eroded away.

Since then the site has yielded

PHOTOS: KROMDRAAI RESEARCH PROJECT

Paleoanthropologist José Braga (center) and colleagues Shadrack Mofele (left) and Clopus Seshoene work a pit at Kromdraai.

PHOTO: KROMDRAAI RESEARCH PROJECT

fossils from at least 10 juveniles and nine adults living roughly 2 million years ago, opening a window on the diversity of ancient hominins. Braga, a scarf wrapped stylishly around his neck to ward off the early winter chill, gave visitors a tour in May. He pointed out a shallow depression below a boulder of gray breccia. Here, in 2018, his team found the unpublished skull that resembles *Australopithecus* and is likely a little older than 2 million years.

Yet the specimen looks different from two stunning skeletons of another *Australopithecus*, *A. sediba*, found in 2008 by Berger and his then-9-year-old son at Malapa, less than 20 kilometers away from Kromdraai. Dated to about 1.95 million years ago, the skeletons suggest *Australopithecus* not only coexisted with *Paranthropus* and *Homo*, but also had multiple species of its own, all living at more or less the same time.

The diversity among hominins also shows up below the species level. Researchers have identified only one species of *Paranthropus* in South Africa, but Herries says that species was quite variable in its anatomy, with more robust forms at Kromdraai and perhaps other subspecies at other caves. A groundbreaking protein study published in *Science* in May (29 May, p. 969) confirms that the variability extends to proteins. A team led by evolutionary biologist Enrico Cappellini of the University of Copenhagen deciphered the amino acid sequences in 2-million-year-old tooth enamel from four *P. robustus* individuals from one site. The enamelin protein varied by one amino acid in two individuals—enough to suggest *Paranthropus* may have had subspecies.

Multiple genera appear again just 9 kilometers up Sterkfontein Road at Drimolen, where Baker, Herries, and students at their annual field school are excavating. The students discovered pieces that fit together to form a skullcap of a juvenile member of *H. erectus* ("erect man"), a direct ancestor of our species. Just 10 centimeters away in the very same layer, the team found teeth of *P. robustus*. "They were alongside each other," Herries says. They have since found the most complete cranium of *P. robustus* known and many other hominin fossils, plus more than 100 bone tools, though they can't be sure who made them.

The discovery of early *Homo* and *Paranthropus* at multiple sites "has reinvigorated the fact that South Africa is not to be ignored. … South Africa was being viewed as a backwater; well, it ain't a backwater," says Stony Brook University paleoanthropologist Fred Grine.

But he notes that only precise dating can prove the hominins actually overlapped. Researchers have struggled to date many South African cave sites, where the stratigraphy is notoriously complex, because cave roofs collapse and animals or ground water disturb layers. But geologists have made great strides by refining and combining dating methods. University of Cape Town geochemist Robyn Pickering, for example, has spent the past decade refining radiometric dating of fossils from flowstones, deposits formed when water flows along a cave's walls or floors. By measuring the ratio of uranium to lead isotopes in flowstones above and below fossils, Pickering can date their deposition, "like the volcanic ash beds in east Africa," she says.

Within the broad window set by the radiometric dating, Herries can apply paleomagnetics and identify which reversal of Earth's magnetic field is recorded in the rock, yielding a precise

Geochemist Tebogo Makhubela (right) and Ph.D. student Tariye Onyeogu examine a new cave site in UNESCO's Cradle of Humankind region in South Africa.

PHOTO: ROBBIE SHONE/*NATIONAL GEOGRAPHIC*

date. At Drimolen, he concludes the fossils date to between 1.95 million and 2.04 million years ago. Those dates make *H. erectus* fossils at Drimolen the world's oldest, a shade older than a piece of a skull from Kenya dating to about 1.9 million years ago.

At other caves including Malapa and Swartkrans, which is only about 2 kilometers from Sterkfontein, Pickering has found six narrow windows of time when flowstones encapsulated fossils, including the same period 1.8 million to 2 million years ago. "That means all the fossils at those sites were [roughly] contemporaneous," she says.

The dating does leave wiggle room of tens of thousands of years. At Kromdraai and Drimolen, for example, the dates can't resolve whether the fossils of *Homo* and *Paranthropus* were deposited "within a week or 90,000 years" of each other, Baker says. Still, because three neighboring sites—Kromdraai, Drimolen, and Swartkrans—contain at least two types of hominins (see map, p. 339), she argues the hominins likely saw one another and recognized they were different folk.

Researchers are still working to refine the dates. Herries, Braga, and Zipfel are working together to redate Kromdraai, for example. For a decade, Herries and Braga disagreed about the age of Kromdraai's hominins. But the two made up in 2019, over eggs at a restaurant near Wits. "I paid for his breakfast," Herries recalls. Braga adds: "We realized at that point that we'd been arguing at cross purposes and we were not that far apart."

They agreed collaboration benefited both teams, and their joint field notes are revealing new patterns in the ancient hominins' lives.

ALL THREE GENERA of hominins inhabiting the cradle 2 million years ago walked upright; had short, likely hairy bodies with relatively small brains; and ventured along riverine forests near the caves. Stone and bone tools found near both *Paranthropus* and early *Homo* remains suggest both hominins may have used tools. Also, females of all three species had smaller bodies, suggesting social structures in which a few males led groups of females and infants, says paleoanthropologist David Strait of Washington University in St. Louis, who works at Drimolen. But there are key differences.

A remarkably complete, articulated *P. robustus* hip, thigh, and shin bone found in 2019 at Swartkrans shows adult females of this species were sturdy but very short, about the size of a baboon. The fossil's discoverer, UJ paleoanthropologist Matt Caruana (husband of Stephanie Baker), estimates they stood 1 meter tall and weighed about 27 kilograms.

That's a shade smaller than females from other diminutive hominins such as Lucy's species, *A. afarensis*. But *Paranthropus* had a robust head, with a crest to attach its powerful chewing muscles, and a brain rather larger than a male chimp's, at 410 to 530 cubic centimeters. "If you saw *Paranthropus* walking around, the first thing you'd notice was its enormous freaking jaws and disproportionately huge head," Strait says.

In contrast, when early *Homo* bared its teeth, it revealed jaws and molars smaller than *Paranthropus*, as well as a flatter midface. Researchers also believe it was taller and had body proportions more like modern humans. By 1.8 million years ago, fossils from Dmanisi in the Republic of Georgia show *H. erectus* stood 1.4 to 1.6 meters tall and weighed 41 to 64 kilograms, with a brain size of 545 to 775 cubic centimeters.

As for *A. africanus*, it resembled Lucy's species in East Africa, with a brain slightly bigger than Lucy's or a chimp's but one-third smaller than that of early *Homo*. Shoulder and hand bones suggest *A. africanus* still

climbed trees, as did *Paranthropus*. Although its hands could have made tools, none have so far been found in the same layers with *A. africanus* or *A. sediba*.

"One of our biggest questions is how the heck do we have three different hominins all living within what we can comfortably say would be the average daily range of one chimpanzee group?" says paleoanthropologist Gary Schwartz of Arizona State University, who studies infant fossils. "Whatever *Paranthropus* and *Homo* were doing differently—whatever they were eating or moving on the landscape—it demonstrates the ability of hominins to adapt into different kinds of ecological niches."

The coexistence among these upright apes is "super-exciting, because it tells us that there's something going on behaviorally that would allow these species to be in the same place at the same time," says paleoanthropologist Ashley Hammond of the American Museum of Natural History. "There is some mechanism for niche separation."

But the nature of that mechanism remains elusive. Given the differences in facial anatomy, many researchers have long bet a key difference was diet. With its big teeth and chewing muscles, they have thought *P. robustus* specialized in hard plant foods, such as nuts, woodland leaves, and tubers. Early *Homo*, with its flatter face and smaller teeth, likely often ate softer food, such as meat. Studies of chemical isotopes, which offer clues to diet, also suggest *Homo* ate more meat than the other hominins, and that *Australopithecus* was omnivorous, munching fruit, leaves, and perhaps some small animals.

Yet the evidence suggests all three hominins were adaptable: *Homo* likely also munched fruit and tubers, for example. And a study this month in *Nature Ecology & Evolution* examined strontium isotopes in the tooth enamel of more than 20 individuals of *Paranthropus* at Kromdraai and Drimolen. Researchers found the *Paranthropus* diet was not so specialized and included grasses as well as woodland plants. Perhaps the nuts and tubers were fallback foods, says geochemist Vincent Balter of the French national research agency's Geology Laboratory of Lyon.

What's beyond dispute is that one genus—ours—persisted and the others eventually died out. Fossils of infants, which are remarkably plentiful in the cradle, may hold clues to why. Their anatomy shows that distinctions among the genera developed early in life, as Braga and his colleagues reported in an analysis of 17 infants in *Nature Communications* in June. "Big, heavy faces are present almost immediately in infants of *Paranthropus*," Schwartz says.

All three hominins weaned their babies earlier than chimps do today, according to studies of calcium isotopes in the baby teeth. Chimps nurse infants for 4 to 5 years, but *Paranthropus*, *Australopithecus*, and early *Homo* weaned before age 2, the studies suggest. Early weaning allowed hominins to have babies in closer succession and grow their populations swiftly, Braga and Balter have reported. That might have helped combat high infant mortality, perhaps from predation by fierce carnivores.

Indeed, many hominin babies apparently died young. At Drimolen, Kromdraai, and Swartkrans, 30% to 40% of *Paranthropus* fossils are infants, suggesting an infant mortality rate higher than the 20% seen in chimps. But the proportion of *Homo* infants in the cradle is smaller, suggesting lower infant mortality among our ancestors. Braga speculates that by 2 million years ago, *Homo* already had a better social safety net than *Paranthropus* or *Australopithecus*. *Homo* was set apart "not simply [by] diet or locomotion, but something more profound: the capacity for increased juvenile survival," Braga says.

Researchers know that at some point our lineage evolved the social support and food provisioning that allowed infants to grow bigger brains, have longer childhoods, and learn more complex social skills than other hominins. Braga suggests this was already beginning to happen in *Homo* in the cradle 2 million years ago. "Early *Homo* didn't just adapt physically—it created a social and developmental niche that allowed more offspring to survive and thrive," he says.

The South African discoveries show our ancestors emerged from a diversity of forms unimaginable today. But they may also reveal what made those ancestors stand out from their contemporaries. The babies in the cradle, Braga says, may "offer a window into the emergence of a deeply human kind of cooperation and life history strategy." □

They were both on this landscape for such an extensive period of time, there's no way they didn't interact with each other.

Stephanie Edwards Baker, University of Johannesburg

PHOTO: ANDY I.R. HERRIES

A healthy colony of elkhorn coral, *Acropora palmata*, inhabits Blue Bay, Curaçao (September 2019).

POLICY FORUM

CONSERVATION POLICY

Proactive assisted gene flow for Caribbean corals in an era of rapid coral reef decline

Regulatory action could facilitate cross-border efforts to retain ecosystem function

Andrew C. Baker[1], Iliana B. Baums[2,3,4], Sarah W. Davies[5], Andréa G. Grottoli[6], Carly D. Kenkel[7], Sheila A. Kitchen[8], Ilsa B. Kuffner[9], Mikhail V. Matz[10], Margaret W. Miller[11], Erinn M. Muller[12], John E. Parkinson[13], Carlos Prada[14], Andrew A. Shantz[15], Ruben van Hooidonk[16], R. Scott Winters[17,18]

Coral reefs are one of the most well-documented marine ecosystems under increasing threat from climate change. Catastrophic episodes of coral bleaching and subsequent mortality caused by prolonged heat stress (*1*) highlight the need to test and implement new approaches to prevent species loss and retain ecosystem function (*2*). One of these approaches is assisted gene flow (AGF)—the managed movement of individuals or gametes between populations within species ranges to mitigate local maladaptation (*3*). AGF has recently been approved to help save elkhorn corals in Florida from local extirpation but faces chal-

PHOTO: PAUL SELVAGGIO, SECORE INTERNATIONAL

lenges for its broader application owing to static regulatory frameworks whose precautionary nature does not readily account for the high cost of inaction in response to the dynamic ecological realities of climate change (*4*, *5*). Here, we highlight how regulatory action could help safely facilitate coral AGF across international boundaries, at least in the tropical western Atlantic (Caribbean).

In Florida and the Caribbean, the marine heat wave of 2023 devastated populations of staghorn and elkhorn coral in the genus *Acropora*—historically one of the region's dominant reef-building taxa—and severely affected many other coral species. Of the 158 genets (unique individuals) of endangered elkhorn coral, *Acropora palmata*, that existed in the upper Florida Keys before 2023, only 23 remain in the wild, with the remaining surviving genets now solely under human care in land-based facilities and in-water nurseries (*6*). Given these demographic declines, the near-zero rates of natural recruitment, and long-standing predictions of local extirpation (*6*), the Genetics Working Group of the Coral Restoration Consortium recently recommended that AGF be implemented for elkhorn corals in Florida (*7*). Managed breeding using parent colonies or gametes translocated to Florida from sites across the tropical western Atlantic where sizable populations still exist will enhance the local gene pool with alleles from healthy populations from a variety of different habitats. This in turn may increase adaptive potential by supplying additional phenotypic variation on which natural selection can act while masking the effect of deleterious alleles accumulated by the Florida population, helping restart larval recruitment, and reducing the risk of demographic collapse (*8*). Targeting diverse populations for AGF that have desired general characteristics—such as heat tolerance, disease resistance, or high recruitment—is preferable to selectively breeding specific genets from the local Florida population (if surviving genets with these traits could be found) because breeding from a restricted subset of the remaining population will likely further reduce the genetic diversity of this already severely diminished population.

Recently, the Florida Fish and Wildlife Conservation Commission approved the outplanting of offspring produced by crossing elkhorn coral parents from Florida with those from Honduras, judging the latter to be a suitable local population from which prospective parents might be sourced. This is a notable regulatory advance and is the first time that any coral offspring whose parents have been sourced from different countries have been permitted for outplanting on wild reefs anywhere in the world. However, additional diversity is still critically needed to avoid extirpation (*7*). Indeed, crossing parents from nonlocal populations may further increase diversity, and earlier AGF crosses of elkhorn coral that were produced in 2018 by fertilizing eggs from Curaçao with cryopreserved sperm from Puerto Rico and Florida (*9*) are still available in land-based facilities. These crosses were not originally approved for outplanting in Florida but could be used now to further increase diversity and provide data on the fitness effects of longer-distance cross-breeding among different populations.

What are the genetic risks of introducing diversity from outside a local population? The most commonly cited argument against AGF is the risk of outbreeding depression—disruption of the repertoire of genetic variants responsible for local adaptation in the recipient population. The risks of outbreeding depression, even in species less directly affected by climate change, are generally low and genetic rescue by means of AGF has been successful in the vast majority of cases so far across a range of animal taxa (*10*), including marine organisms such as the California sea otter and giant clams. Risks for elkhorn corals are similarly low (*7*), and assessing the survivorship and performance of Florida-Curaçao crosses on Florida reefs would provide field confirmation of experimental studies supporting their fitness (*11*); if AGF colonies or native donor colonies do not survive outplanting, then the genetic risk of outplanting AGF corals to the native Florida elkhorn population would be zero. Moreover, although outbreeding depression may not manifest itself until later generations, sexual recruitment of native elkhorn corals has been entirely absent in the Florida Keys in recent years (*6*) and any new recruits, including those produced by AGF crosses, would reasonably be considered a conservation success at this point.

...modeling suggests that unless such populations receive an influx of adaptive genetic variants... they will disappear within a few decades or sooner.

Assessing the risks of outbreeding depression relative to the counterfactual (i.e., the decision to take no AGF action) assumes that the local population will remain well adapted to local conditions, an assumption that is no longer proving valid for most coral populations as ocean warming and other impacts rapidly change environmental conditions, leading to precipitous coral declines. Many coral populations are already experiencing thermal exposures beyond their local limits of tolerance, and modeling suggests that unless such populations receive an influx of adaptive genetic variants, potentially sourced from warmer locations, they will disappear within a few decades or sooner (*12*). Consequently, there is a need to begin proactive AGF in additional species now before they reach the depleted status of Florida's elkhorn corals.

In addition to genetic risks, other risks of AGF include the unintentional introduction of additional metazoan or microbial species into new reef areas. The tropical western Atlantic has a long and ongoing history of marine disease spread (e.g., *Diadema* die-off, stony coral tissue loss disease) and introduced species (e.g., Indo-Pacific lionfish and soft corals), highlighting the capacity for new, deleterious agents to spread but also demonstrating the connectivity that already exists (naturally or human-assisted) within the region (*13*). Such risks can be mitigated by raising in vitro crosses of AGF corals in ex situ closed systems, and this is the way that all such crosses have been raised to date.

Reefs in the wider Caribbean region are a suitable place to begin coral AGF efforts, not only because of these diseases and introductions, which have contributed to their chronic decline over the past half-century, but also because of their steep environmental gradients (see the figure, left panel), the relatively broad network of restoration practitioners, and the lack of widespread traditional ownership claims, which makes addressing indigenous heritage values more achievable than in regions with greater complexity. The marine heat wave of 2023 (see the figure, right panel) illustrates the urgency of AGF to also help rescue remaining genetic diversity, particularly from the warmest sites across this region, which may be especially valuable as sources of warm-adapted coral broodstock. Saving at least some corals from these areas in advance of devastating bleaching-related mortality provides an opportunity to breed these corals in support of regional AGF efforts. Indeed, it may already be too late to begin this effort: Elkhorn and staghorn corals in some of the region's warmest areas, off the south coast of Cuba (see the figure, left panel), were exposed to unprecedented heat stress during the 2023 bleaching event (see the figure, right panel) and have experienced major losses. It is not clear whether these reefs can recover through immigration of even more thermally tolerant genets from elsewhere (*12*) because these reefs are among the warmest in the region.

REGULATORY FRAMEWORKS TO SUPPORT, NOT IMPEDE

Natural resource laws (and endangered species law in particular) are precautionary in nature, having been originally designed to maintain some historical state by preventing harm and protecting relatively small management areas from external impacts. Consequently, they tend to treat any human-aided translocation with skepticism (*5*) and do not account for accelerating climate change or the inherent dynamism of these systems (*4*). Moreover, the continued application of these static regulatory frameworks without appreciating the changing pace and scale of the threat can lead to "analysis paralysis."

...continued application of these static regulatory frameworks...can lead to "analysis paralysis."

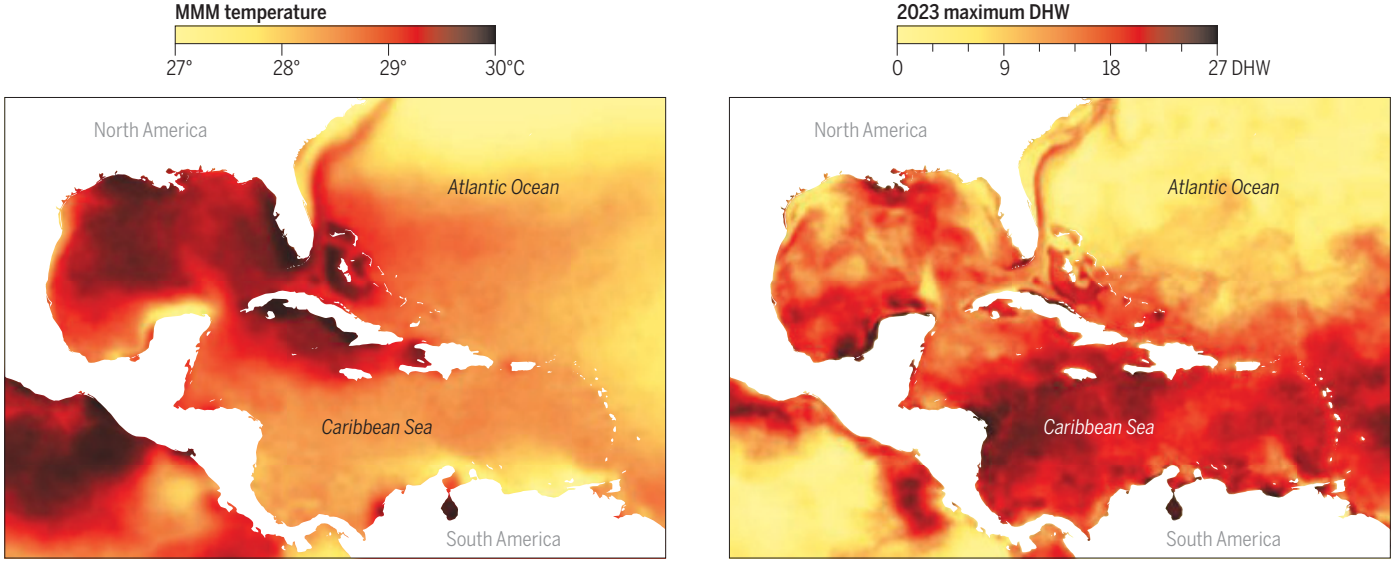
The movement of scleractinian (stony) corals between countries is regulated under the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), an agreement among 185 governments aimed at ensuring that the trade of wild animals and plants does not threaten the survival of the species in the wild. All scleractinian corals are listed under CITES Appendix II, which means that their trade must be controlled to prevent harm to the species, even if they are not presently at risk of extinction. Exporting stony corals typically requires a CITES export permit, approved by both the scientific and management authorities of the country of origin. In addition, in the US, the import or export of species listed under the Endangered Species Act can require separate permits from the US Fish and Wildlife Service, particularly if the species is listed as endangered (such as the pillar coral *Dendrogyra cylindrus*) rather than threatened (such as elkhorn coral). Depending on the country, animal exports may also require further clearances and veterinary inspections that are complicated by a lack of reliable coral health diagnostics and little coral experience among most veterinarians.

Moreover, large-scale implementation of coral translocations across the wider Caribbean is further hindered by regulatory fragmentation involving diverse national stakeholders (33 political entities, including 13 sovereign states, 12 dependencies, and 7 overseas territories), many of whom are party to the Nagoya Protocol, which outlines recommendations to ensure reciprocal and equitable benefits from the sharing of genetic (and other) resources, including Traditional Knowledge. Unlike CITES, which provides a single framework for all participating countries, the Nagoya Protocol defines genetic material as an asset of the originating country to be managed consistently with its own priorities. Transferring genetic material requires a negotiated agreement between parties, narrowly circumscribed to a specific use. Disparities in administrative infrastructure between countries can make negotiating such agreements difficult and can place undue burden on less well-resourced countries who may be asked to approve transfers to benefit corals in more well-resourced countries. Yet such agreements are essential to ensuring the long-term success of AGF efforts, and indeed their absence or abuse has led to conflict and regulatory upheaval, for example, in The Bahamas (*14*).

Despite the present challenges, regulatory frameworks would ideally support, rather than impede, resource management and species conservation. Sharing coral broodstock in regional land-based coral propagation and nursery facilities could allow at least some critical coral stocks to be safeguarded immediately, which is especially important for countries that do not have access to domestic facilities. Maintaining these stocks in centralized facilities could help preserve native genetic diversity, including locally adapted alleles, and help optimize the managed breeding of more diverse regional stocks to help species survive. These facilities could be paired with agreements and policies that help satisfy the Nagoya Protocol by allowing broodstock and/or offspring to be routinely repatriated to donor countries and providing much-needed capacity, infrastructure, and technical support, helping optimize cost-sharing and operational

Steep gradients and extreme heat in the tropical western Atlantic

Steep gradients in the maximum monthly mean (MMM) sea surface temperature in some locations indicate considerable differences in heat tolerance of local coral populations over relatively short distances (left). Maximum heat-stress accumulation during the 2023 marine heat wave is expressed as degree heating weeks (DHW), reflecting time spent above MMM-based thresholds (right). Even warm sites, such as southern Cuba, where corals have the highest thermal tolerance (characterized by high MMM) experienced heat stress that far exceeded their bleaching thresholds (>8 DHW).



Maximum DHW reflects the highest DHW recorded per pixel in 2023, as extracted from daily DHW data. See https://coralreefwatch.noaa.gov/product/5km/methodology.php and (*15*) for details on the data used.

GRAPHIC: V. PENNEY/*SCIENCE*; (DATA) NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION, HTTPS://CORALREEFWATCH.NOAA.GOV/PRODUCT/5KM

efficiency and boosting production of more climate-resilient corals for reef restoration across the region.

The following are potential actions that could be taken to optimize coral AGF implementation.

Adapt existing CITES exemptions for the noncommercial loan, donation, or exchange of live plants to also apply to corals

Corals, like plants, are clonal organisms that are routinely asexually propagated, and coral fragments can be collected without loss of that genet from the wild population. Under this exemption, corals could be exchanged among nursery facilities as live specimens akin to plant cuttings, with an exemption certificate from the originating management authority stating they are under artificial propagation (CITES Article VII, § 5). Alternatively, coral nursery facilities could qualify for a Certificate of Scientific Exchange, similar to herbariums for the exchange of live plant material (CITES Article VII, § 6). Adoption of either exemption would greatly simplify the more complex procedure described in Article IV for CITES Appendix II species and may also help alleviate problems associated with animal exports, such as veterinary inspections.

Establish regional land-based facilities, affiliated with in-water coral nurseries, to act as repositories of coral stocks from a variety of different countries

These closed-system repositories would maintain and propagate colonies of reproductive size to facilitate immediate breeding efforts. These facilities could be supplemented by regional cryobanking facilities that hold frozen material, such as sperm, for long-term preservation.

Leverage existing political networks of overseas countries and territories to facilitate immediate exchanges within the region

Transfers within some jurisdictions would not be considered international transfers and may represent an accelerated pathway to action while new international agreements are negotiated or ratified. In the wider Caribbean, this could include transfers within and between the US (specifically Florida and Texas) and its territories in the Caribbean (US Virgin Islands and Puerto Rico); between British Overseas Territories in the Caribbean (Anguilla, Bermuda, British Virgin Islands, Cayman Islands, Montserrat, Turks and Caicos Islands); between French Caribbean Overseas Departments, Regions, and Collectivities with coral reefs (Guadeloupe, Martinique, Saint Barthélemy, Saint-Martin); or within and between the Dutch Caribbean (Aruba, Curaçao, Sint Maarten) and Caribbean Netherlands (Bonaire, Sint Eustatius, and Saba). Such initiatives could provide reciprocal ecosystem, conservation, and coastal resilience benefits to these small-island developing states funded by their respective sovereign states. These benefits generally far exceed the cost of investment and would help demonstrate how future benefit-sharing among Nagoya Protocol signatories could work.

For more than 25 years, there has been growing certainty among coral reef biologists that natural rates of biological adaptation are unlikely to prevent widespread loss of corals under climate change (*1*). Given the lack of effective action in curbing carbon emissions to date, the role of long-term observer bias in dampening the recognition of ecosystem decline, and the time needed for interventions to take effect, it is prudent to take action now in anticipation of more rapid environmental change in the coming decades. Introducing additional genetic diversity to struggling populations could enhance their ability to adapt to changing conditions, thereby reducing future losses. Waiting until a population is vanishingly small before new diversity is introduced commits the population to continued loss and further endangers the ecosystem services these species provide.

Historically, the precautionary principle "take no action unless there is high certainty no harm will result" was the basis for endangered species law and was appropriate in an age of relative environmental stability (*4*). The present rate of warming and repeated occurrence of marine heat waves overwhelmingly point to further coral losses as a "not if, but when" scenario, underscoring the need for timely action (*2*). Such action could include taking proactive steps to preemptively boost genetic diversity through AGF to help buy time for local coral populations until the rate of ocean warming can be slowed (*12*). Windows of opportunity for effective large-scale implementation of AGF are closing rapidly; waiting until genetic rescue is "needed" to save coral species on the brink of extinction may well be too late. □

REFERENCES AND NOTES

1. T. P. Hughes *et al.*, *Science* **359**, 80 (2018).
2. National Academies of Sciences, Engineering, and Medicine, *A Research Review of Interventions to Increase the Persistence and Resilience of Coral Reefs* (National Academies Press, 2019).
3. S. N. Aitken, M. C. Whitlock, *Annu. Rev. Ecol. Evol. Syst.* **44**, 367 (2013).
4. J. B. Ruhl, *Environ. Law* **40**, 363 (2010).
5. A. E. Camacho, *Wash.* Univ. *Law Rev.* **92**, 849 (2015).
6. D. Williams, K. Nedimyer, A. Bright, M. Ladd, "Genotypic inventory and impact of the 2023 marine heatwave on *Acropora palmata* (elkhorn coral) populations in the Upper Florida Keys, USA: 2020–2023" (NOAA National Marine Fisheries Service Southeast Fisheries Science Center, 2024).
7. CRC Genetics Working Group, Safeguarding Florida's coral reefs: The urgency of assisted gene flow for Elkhorn coral conservation. Zenodo (2025); https://doi.org/10.5281/zenodo.14920439.
8. K. M. Quigley, L. K. Bay, M. J. H. van Oppen, *Ecol. Evol.* **9**, 11122 (2019).
9. M. Hagedorn *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2110559118 (2021).
10. J. G. Clarke, A. C. Smith, C. I. Cullingham, *Mol. Ecol.* **33**, e17532 (2024).
11. E. M. Muller *et al.*, bioRxiv 10.1101/2025.03.03.641242 (2025).
12. M. V. Matz, E. A. Treml, B. C. Haller, *Glob. Change Biol.* **26**, 3473 (2020).
13. H. A. Lessios, *Annu. Rev. Mar. Sci.* **8**, 267 (2016).
14. K. Sherman, C. Dahlgren, C. Dunn, D. Claridge, N. Higgs, *Conservation* **5**, 3 (2025).
15. Climatology (maximum monthly mean) data and maximum degree heating weeks data used in the figure were obtained from https://www.star.nesdis.noaa.gov/pub/sod/mecb/crw/data/5km/v3.1_op/climatology/nc/ and https://www.star.nesdis.noaa.gov/pub/socd/mecb/crw/data/5km/v3.1_op/nc/v1.0/daily/dhw/, respectively.

ACKNOWLEDGMENTS

We acknowledge the National Oceanographic and Atmospheric Administration (NOAA) and the National Science Foundation (NSF) for long-term support of research programs. In particular, we acknowledge an award from the Paul Allen Foundation (to I.B.B., E.M.M.) and a NOAA Ruth Gates Coral Restoration Innovation Grant (to A.C.B.) for the support of elkhorn coral AGF in the Caribbean. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US government. All authors (except R.v.H.) are members of the Genetics Working Group of the Coral Restoration Consortium; R.S.W. is chairman of the board of directors for the Coral Restoration Consortium. E.M.M., C.D.K., M.W.M., and R.S.W. serve on the US *Acropora* Recovery Implementation Team, which is a formal advisory body to the US National Marine Fisheries Service. A.C.B. and C.D.K. serve on the Intervention Risk Review Group for Australia's Reef Restoration and Adaptation Program, which serves in a quality control capacity by providing an independent assessment of intervention risk and overall risk assessment practices for the Great Barrier Reef.

10.1126/science.adx5842

[1]Department of Marine Biology and Ecology, Rosenstiel School of Marine, Atmospheric, and Earth Science, University of Miami, Miami, FL, USA. [2]Helmholtz-Institute for Functional Marine Biodiversity at the University of Oldenburg, Oldenburg, Germany. [3]Alfred Wegener Institute, Helmholtz-Centre for Polar and Marine Research, Bremerhaven, Germany. [4]Institute for Chemistry and Biology of the Marine Environment, School of Mathematics and Science, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany. [5]Department of Biology, Boston University, Boston, MA, USA. [6]School of Earth Sciences, The Ohio State University, Columbus, OH, USA. [7]Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA. [8]Department of Marine Biology, Texas A&M University at Galveston, Galveston, TX, USA. [9]US Geological Survey, St. Petersburg Coastal and Marine Science Center, St. Petersburg, FL, USA. [10]Department of Integrative Biology, University of Texas at Austin, Austin, TX, USA. [11]SECORE International, Miami, FL, USA. [12]Mote Marine Laboratory, Sarasota, FL, USA. [13]Department of Integrative Biology, University of South Florida, Tampa, FL, USA. [14]Department of Biological Sciences, University of Rhode Island, Kingston, RI, USA. [15]Cooperative Institute for Marine and Atmospheric Research, University of Hawai'i Manoa, Honolulu, HI, USA. [16]SymbioSeas, Carolina Beach, NC, USA. [17]Coral Restoration Foundation, Tavernier, FL, USA. [18]The Oxford Uehiro Centre for Practical Ethics, University of Oxford, Oxford, UK. Email: abaker@miami.edu

SIGNAL TRANSDUCTION

Harnessing insulin biology to treat diabetes and cancer

Small molecules are used to target the insulin signaling molecular machinery **Giovanni Solinas**

Class 1 phosphoinositide 3-kinases (PI3Ks) are enzymes that mediate the metabolic action of insulin and many effects of growth factors (*1*). PI3K activity is frequently increased in cancer, so the development of PI3K inhibitors is an area of intensive investigation. However, the insulin-PI3K signaling pathway is subject to feedback control, and systemic inhibition of PI3K activity leads to hyperglycemia and hyperinsulinemia, with the latter reactivating PI3K in cancer cells (*2*, *3*). On pages 402 and 409 of this issue, Terayama *et al.* (*4*) and Simanshu *et al.* (*5*), respectively, report small molecules that target the PI3Kα–rat sarcoma (RAS) complex. Terayama *et al.* describe molecular glues that stabilize the complex and display insulin-like activity, which could have applications in the treatment of diabetes. Simanshu *et al.* identify an inhibitor that disrupts the complex and has anticancer effects in preclinical models without altering blood glucose concentration.

RAS is a family of guanosine triphosphatases (GTPases) that link the activation of several membrane receptors—including the insulin receptor and the receptors for growth factors—to downstream signal transducers, such as mitogen-activated protein kinases (MAPKs) and PI3K. PI3K consists of a regulatory subunit (p85) and a catalytic subunit (p110). There are four isoforms of p110: α, β, γ, and δ (*1*). RAS potentiates PI3Kα but not PI3Kβ activity through protein-protein interaction (*6*), and insulin-PI3K signaling can still function in the absence of an active RAS (*7*). Once activated, PI3K phosphorylates the membrane lipid phosphoinositide 4,5-bisphosphate, which ultimately results in the phosphorylation and activation of the kinase AKT.

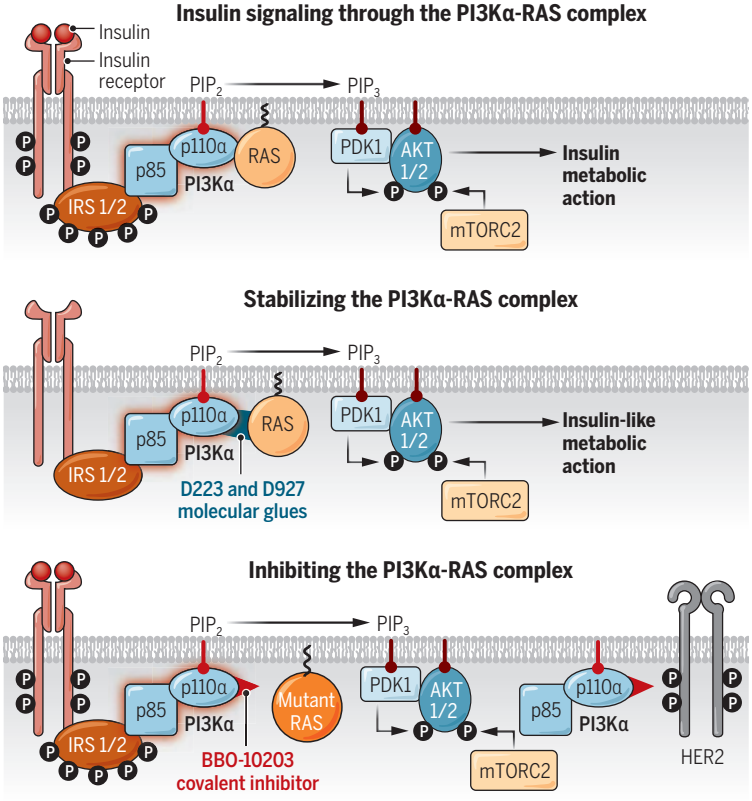
The insulin-PI3K system is best characterized in the mouse, where insulin signaling in hepatocytes and endogenous glucose production depends on the redundant activities of PI3Kα and PI3Kβ (*7*). Insulin also inhibits lipolysis and an adipocyte-derived signal that induces insulin secretion (the "adipoincretin" effect); thus insulin inhibits insulin secretion through PI3K. This effect also depends on redundant PI3Kα and PI3Kβ activities (*8*, *9*). Such redundancy provides a rationale for the success of the PI3Kα-selective inhibitor alpelisib in the treatment of certain types of breast cancer that bear mutations in the gene encoding p110α (*PIK3CA*) and PIK3CA-related overgrowth spectrum (PROS), a group of disorders characterized by tissue overgrowth caused by mosaic mutations in *PIK3CA* (*10*). Notably, a low dose of alpelisib substantially improved the condition of PROS patients without altering blood glucose (*10*). Indeed, at the hyperglycemic threshold, alpelisib is no longer selective against PI3Kβ (*11*). These findings provide a rationale for the generation of inhibitors with improved PI3Kα selectivity. Several PI3Kα-selective covalent inhibitors are in preclinical development (*12*), and one is in a clinical trial (NCT05683418).

Terayama *et al.* aimed to identify strategies to control blood glucose concentration in diabetes. They screened a library of compounds for those that had a similar effect to insulin in an in vitro model of muscle cells [i.e., induced the translocation of the glucose transporter type 4 (GLUT4) to the plasma membrane]. The authors identified two small molecules, D223 and D927, which function as molecular glues to stabilize the interaction between RAS and PI3Kα by nearly three orders of magnitude. They tested the effects of D927 in vivo and found that it showed insulin-like activity in healthy rats as well as in rat models of insulin resistance, and mouse models of type 1 and type 2 diabetes. Therefore, D223 and D927 may be lead compounds for a new class of orally available small-molecule insulin mimetics.

PIK3CA and the genes encoding RAS are the most frequently mutated genes in solid metastatic cancers, so stabilizing their interaction poses safety concerns that demand further testing. Terayama *et al.* found that D927 induced PI3K signaling in human cancer cells and promoted binding of the Gly^{12}→Asp

Targeting a multifunctional complex

Binding of insulin to the insulin receptor triggers phosphorylation events that activate PI3Kα (complex of p85 and p110α). PI3Kα then phosphorylates PIP_2 to form PIP_3, which activates AKT 1/2 through PDK1 to cause insulin's metabolic action (activation of AKT 1/2 also requires phosphorylation by mTORC2). Binding of RAS to PI3Kα potentiates its activity, but insulin-PI3Kα signaling can occur without RAS. D223 and D927 stabilize the interaction between PI3Kα and RAS, mimicking the effect of insulin. BBO-10203 prevents the binding of PI3Kα and RAS. This has an anticancer effect by preventing the actions of mutant oncogenic forms of RAS, but insulin signaling continues through PI3Kα's interaction with IRS 1/2. BBO-10203 also prevents activation of PI3Kα by HER2, which is oncogenic.



AKT 1/2, serine-threonine protein kinase; HER2, human epidermal growth factor receptor 2; IRS 1/2, insulin receptor substrate 1/2; mTORC2, mechanistic target of rapamycin complex 2; P, phospho group; PDK1, 3-phosphoinositide–dependent protein kinase 1; PI3Kα, class 1 phosphoinositide 3-kinase α; PIP_2, phosphoinositide 4,5-bisphosphate; PIP_3, phosphoinositide 3,4,5-trisphosphate; RAS, rat sarcoma.

GRAPHIC: K. HOLOSKI/*SCIENCE*

(G12D) oncogenic form of the Kristen rat sarcoma viral oncogene homolog (KRAS) protein to PI3Kα. However, insulin was also associated with tumor promotion, and D927 reduced the concentration of circulating insulin.

Terayama *et al.* used D223 to help resolve the crystal structure of the PI3Kα-KRAS complex. This revealed the amino acids involved in the binding of PI3Kα to KRAS, in particular, cysteine at position 242, which is not conserved in other class 1 PI3Ks. Simanshu *et al.* used this information to generate BBO-10203, a small-molecule inhibitor of PI3Kα that binds covalently to Cys^{242}. BBO-10203 prevented binding of PI3Kα to KRAS, Harvey rat sarcoma viral oncogene homolog (HRAS), and neurosarcoma rat sarcoma viral oncogene homolog (NRAS) without interfering with the in vitro catalytic activity of purified PI3Kα. Insulin can induce AKT phosphorylation in the absence of active RAS (*7*), and BBO-10203 had no effect on blood glucose or insulin secretion in mice.

Simanshu *et al.* tested BBO-10203 in cells containing different oncogenic mutations: It substantially reduced AKT phosphorylation in cells with the *KRAS* G12X mutation or mutations in *PIK3CA* but not in cells bearing mutations in phosphatase and tensin homolog (*PTEN*). BBO-10203 was most effective in reducing AKT phosphorylation in cells with tumor-promoting amplification of human epidermal growth factor receptor 2 (HER2). The underlying mechanism for this remains unresolved but appears to be independent of the action of the drug on the formation of PI3Kα-RAS complexes. In vivo, BBO-10203 caused tumor regression in a xenograft mouse model of esophageal squamous cell carcinoma, which bears HER2 amplification and oncogenic *KRAS* Gly^{12}→Cys (G12C) mutation. Furthermore, the drug, in combination with either trastuzumab (a monoclonal antibody targeting HER2), fulvestrant (an estrogen receptor degrader), palbociclib (an inhibitor of cyclin-dependent kinases 6 and 8), or BBO-8520 (a KRAS G12C inhibitor), achieved tumor stasis or regression in multiple xenograft tumor mouse models with different genotypes.

Terayama *et al.* and Simanshu *et al.* present two classes of compounds targeting the PI3Kα-RAS complex—one that achieves antidiabetogenic activity by stabilizing the complex and one that breaks the complex to produce an anticancer effect (see the figure). Future preclinical and clinical studies will be necessary to test the safety of D223 and D927, which are nevertheless promising leading compounds toward the generation of orally available insulin mimetics for the treatment of diabetes. BBO-10203 will soon be tested in patients with advanced solid tumors in the BREAKER-101 clinical trial (NCT06625775). In the future, it will be crucial to match the type of PI3K inhibitor with the cancer genotype. Ultimately, further basic research is required because the mechanism by which BBO-10203 inhibits PI3K activation by HER2 remains unresolved. It will also be important to resolve the feedback signals linking PI3K activity and insulin secretion. □

REFERENCES AND NOTES

1. G. Solinas, B. Becattini, in *PI3K and AKT Isoforms in Immunity*, M. Dominguez-Villar, Ed., vol. 436 of *Current Topics in Microbiology and Immunology* (Springer, 2022), pp. 311–336.
2. B. D. Hopkins *et al.*, *Nature* **560**, 499 (2018).
3. N. Vasan, L. C. Cantley, *Nat. Rev. Clin. Oncol.* **19**, 471 (2022).
4. K. Terayama *et al.*, *Science* **389**, 402 (2025).
5. D. K. Simanshu *et al.*, *Science* **389**, 409 (2025).
6. R. Fritsch *et al.*, *Cell* **153**, 1050 (2013).
7. A. Molinaro *et al.*, *Cell Metab.* **29**, 1400 (2019).
8. G. Solinas, B. Becattini, *Trends Endocrinol. Metab.* **35**, 466 (2024).
9. B. Becattini, A. Molinaro, M. Henricsson, J. Borén, G. Solinas, *Cell Rep.* **43**, 114132 (2024).
10. Q. Venot *et al.*, *Nature* **558**, 540 (2018).
11. C. Fritsch *et al.*, *Mol. Cancer Ther.* **13**, 1117 (2014).
12. L. Bissegger *et al.*, *Chem. Sci.* **15**, 20274 (2024).

ACKNOWLEDGMENTS

G.S. acknowledges support from the Swedish Research Council, the Cancerfoden, and the Diabetesfonden. G.S. is grateful to B. Becattini for critical feedback on the manuscript.

10.1126/science.adz4117

Wallenberg Laboratory, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden. Email: giovanni.solinas@wlab.gu.se

IMMUNOTHERAPY

Beyond the native repertoire

Design of T cell receptors by artificial intelligence is poised to accelerate cancer immunotherapy

Grant H. Hickok[1,2] and **Ingunn M. Stromnes**[1,2,3,4]

T cell receptors (TCRs) serve as sentinels of adaptive immunity, allowing T cells to monitor for threats by recognizing short peptides presented by major histocompatibility complex class (MHC) molecules on the cell surface. These peptide-MHC (pMHC) complexes enable T cells to identify cancerous cells bearing aberrant proteins (*1*). Leveraging TCRs for cancer immunotherapy has traditionally involved isolating T cells from patients, either as tumor-infiltrating lymphocytes or by expanding T cells from the naive repertoire against a desired epitope. However, this process is technically challenging and laborious, and naturally occurring TCRs often possess suboptimal affinity for tumor antigens. On pages 375, 380, and 386 of this issue, Householder *et al.* (*2*), Johansen *et al.* (*3*), and Liu *et al.* (*4*), respectively, report overcoming these limitations by using generative artificial intelligence (AI) to design high-specificity pMHC-binding "artificial TCRs" for more precise targeting of tumor antigens.

Powerful AI platforms...are fundamentally transforming the design of artificial T cell receptors.

To circumvent the challenges of TCR identification and facilitate the clinical application of immunotherapies based on pMHC recognition, the concept of TCR mimics has emerged. These high-affinity molecules aim to imitate the native TCR's ability to recognize pMHC complexes, including intracellular antigens previously considered "undruggable." Antibodies that recognize pMHC can be integrated into various therapeutic formats, including chimeric antigen receptors (CARs) expressed by T cells that are designed to bind a specific antigen and subsequently activate the T cell, and bispecific T cell engagers (BiTEs) that are engineered antibody-based proteins that link tumor cells and T cells, triggering T cell–mediated killing. Both approaches have shown promise in reducing tumor burden (*5*, *6*). However, broader clinical implementations of TCR mimics still face obstacles. Traditional TCR mimics often interact more strongly with the ubiquitous regions of the MHC rather than with the presented peptide, which can contribute to off-target toxicities. As an alternative, native TCRs can be "affinity-enhanced" through engineering of the TCR. Although seemingly desirable, this can lead to unpredictable recognition of healthy tissue peptides and severe toxicities (*7*). The inherent flexibility of TCR complementarity–determining regions can lead

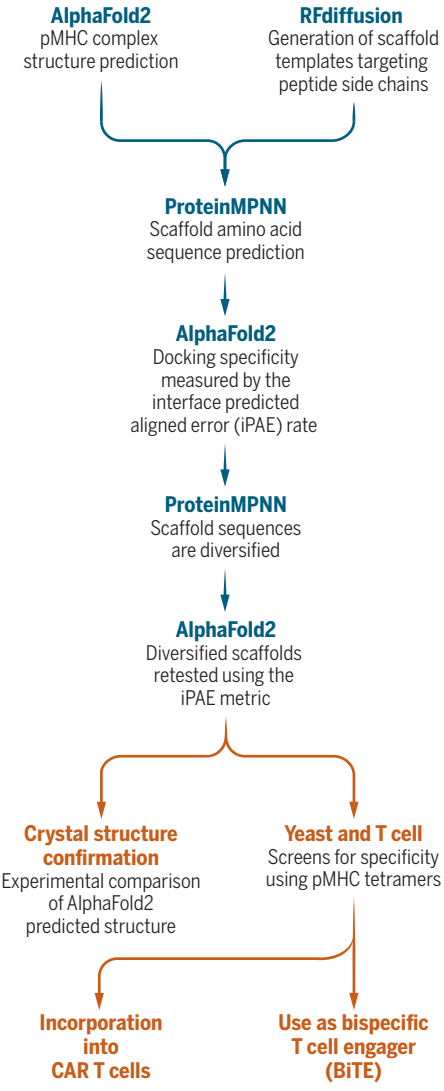
to cross-reactivity, making precise peptide specificity harder to achieve through such engineering efforts.

Powerful AI platforms, particularly generative protein models that are trained on vast datasets of protein sequences or structures, are fundamentally transforming the design of artificial TCRs. To create de novo pMHC binders that function in T cells, Householder *et al.*, Johansen *et al.*, and Liu *et al.* converged on the same core set of advanced AI tools. In this discovery scheme, the AI system AlphaFold2, which predicts the three-dimensional structure of proteins from their amino acid sequences, is consulted to predict the target pMHC structure (see the figure) (*8*). Concurrently, another model (RoseTTAFold diffusion) generates a diverse range of stable protein monomers (single-polypeptide scaffolds) with predicted structural affinity to a selected pMHC (*9*). From these proposed scaffolds, Protein Message Passing Neural Network (ProteinMPNN), a deep learning–based protein sequence design method, predicts multiple likely amino acid sequences that will generate the selected scaffold structures (*10*). AlphaFold2 is then revisited to predict how each scaffold design docks with the target pMHC, applying a ranking based on an interface predicted aligned error (iPAE) metric. More diverse sequences are generated by using ProteinMPNN and subsequently validated with pMHC binding assays using a yeast surface display followed by functional assays in T cells. Although the fundamental approach and tools are comparable across the schemes of the three studies, the exact application varied in detail and scale.

Householder *et al.* initiated their design with α-helical bundle scaffolds because of their inherent stability, structural predictability, and defined conformational constraints. As such, de novo protein design tools favor α-helical coils. They ultimately identified a single, compact, rigid mini-TCR mimic specific to the tumor antigen NY-ESO-1 peptide bound to HLA-A*02 (a common MHC allele) (*11*, *12*). Despite the availability of this pMHC complex structure in the Protein Data Base, Householder used AlphaFold2 to predict its structure to test the broader utility of the platform because most pMHCs lack known crystal structures. A comparison of experimentally determined structures revealed that the selected TCR mimic formed two tight pockets around key side-chain residues of the peptide, burying the peptide like a puzzle piece. This differed from a NY-ESO-1–specific TCR or a fragment antigen-binding region (Fab) of an antibody specific to the identical pMHC that each instead formed a single large and flexible pocket to accommodate the peptide side chains. The TCR mimic contacted 19 MHC residues, including nine shared with the TCR and Fab, and had a fivefold greater affinity [dissociation constant (K_d) = 9.5 nM] than that of the Fab (K_d = 47 nM) and a 600-fold greater affinity than that of the TCR (K_d = 5.7 μM).

Binders by design

T cell receptors (TCRs) bind to specific peptides that are bound to the major histocompatibility complex (MHC), initiating T cell activation. These peptides can be from viral proteins or tumor-associated proteins, for example. A method to rapidly synthesize small proteins that recognize specific peptide-MHC (pMHC) complexes involves **computational** and **experimental** steps. The designed binding proteins then can be engineered into chimeric antigen receptors (CARs) expressed by T cells or as bispecific T cell engagers (BiTEs) for immunotherapies. The computational programs include AlphaFold2, Protein Message Passing Neural Network (ProteinMPNN), and RoseTTAFold diffusion (RFdiffusion).



Because off-target toxicity is a concern, the authors used ProteinMPNN to computationally screen 14,363 peptides identified from a MHC motif atlas (*13*). They successfully predicted two off-target peptides, both of which were less effective at eliciting TCR mimic–mediated T cell activation, demonstrating robust peptide specificity. Householder *et al.* also generated TCR mimic–based BiTE or CAR-expressing T cells, showing antigen-specific T cell activation and cytotoxicity against a melanoma cell line bearing the pMHC. That such a high-affinity binder can elicit T cell signaling and function is promising.

Johansen *et al.* also developed a TCR mimic specific to the NY-ESO-1 peptide bound to HLA-A*02, referred to as a mini-binder (miBd). Bypassing screening in yeast, they validated specificity using pMHC binding in T cells. When incorporated into a CAR, the miBd elicited T cell cytotoxicity against cultured melanoma cells. The NY-ESO-1 peptide has protruding residues at four positions (*14*), suggesting that it represents a "best case scenario" for AI-generative approaches. However, Johansen *et al.* also formulated miBds toward a melanoma neoantigen presented on HLA-A*01:01 (*15*). The pMHC structure was modeled by using AlphaFold2, and a library of 96 miBds were screened before validation in T cells. This process revealed two miBds that displayed high specificity for the neoantigen target.

Liu *et al.* tested the generalizability of the de novo design process by generating artificial binders specific to a broad range of targets. They designed binders for 11 distinct pMHC complexes, including several predicted solely on the basis of structures from AlphaFold2 (*7*). Four different HLA complexes that present 9– or 10–nucleotide oligomer peptides derived from viral proteins, tumor-associated proteins, or cancer neoantigens were probed. For each pMHC target, Liu *et al.* assessed oligonucleotide pools encoding between 200 and 12,000 binder designs. Using a high-throughput yeast display, pMHC binding, and sequencing approach, they selected and validated binders with high target specificity and putative low off-target reactivity. Notably, a binder that recognized a specific nine–amino acid peptide presented by HLA-A*01:01 did not recognize another peptide that differed by only four residues, thus avoiding fatal cross-reactivity of an affinity-enhanced TCR (*7*). Furthermore, Liu *et al.* validated binder functionality by formulating them as CARs and testing the ability to activate human T cells. The binder-based CARs induced T cell–mediated killing of target cells presenting the cognate peptides presented by HLA-A*02. This occurred at higher efficiency compared with that of noncognate peptide or no-peptide controls, demonstrating that binders designed de novo facilitate antigen-specific T cell–mediated cell killing, a process that is vital for effective immunotherapies.

GRAPHIC: N. BURGESS/*SCIENCE*

A common thread underlying the approach taken across the three studies is the emphasis on generating binders that are highly specific to the presented peptide. This peptide-centric binding mode is crucial for achieving high specificity and reducing the risk of off-target toxicities that have plagued previous TCR engineering efforts. Notably, all three studies used experimental structural validation techniques (Householder *et al.* and Liu *et al.* used x-ray crystallography, and Johansen *et al.* used cryo–electron microscopy) to confirm that AI-predicted binding modes generally aligned with experimental structures. The discovery of minor deviations suggests that crystal structure validation still plays a qualitative role in de novo binder development. However, the need for such experimental validation may decrease as the accuracy of AI modeling increases. Further, new binders identified in the future should continue to refine AI-generative approaches with even greater precision.

The implications of these platforms are far-reaching. Johansen *et al.* estimate the development of functional pMHC binders to take less than 2 months. The hope is that de novo pMHC binders may limit cross-reactivity and overcome unpredictable safety concerns that hinder traditional TCR-based therapeutics. However, even with reliable off-target prediction methods, caution is essential when considering translation to clinical use. Elucidating how high-affinity pMHC binders affect the longevity, differentiation, and functional persistence of engineered T cells remains to be determined and will be crucial for cell-based therapies. Immunogenicity of de novo pMHC binders is another concern that may also benefit from AI-driven prediction methods. Notably, such binders could extend beyond cell therapies and include personalized targeting of autoimmune epitopes, diagnostic applications, and multispecific biologics. The AI-driven revolution underway has the potential to unlock the full therapeutic benefits of TCR-based therapeutics with unprecedented breadth and precision. □

REFERENCES AND NOTES

1. A. Peri *et al.*, *Nat. Cancer* **4**, 937 (2023).
2. K. D. Householder *et al.*, *Science* **389**, 375 (2025).
3. K. H. Johansen *et al.*, *Science* **389**, 380 (2025).
4. B. Liu *et al.*, *Science* **389**, 386 (2025).
5. M. Yarmarkovich *et al.*, *Nature* **623**, 820 (2023).
6. J. C. Hassel *et al.*, *N. Engl. J. Med.* **389**, 2256 (2023).
7. G. P. Linette *et al.*, *Blood* **122**, 863 (2013).
8. J. Jumper *et al.*, *Nature* **596**, 583 (2021).
9. J. L. Watson *et al.*, *Nature* **620**, 1089 (2023).
10. J. Dauparas *et al.*, *Science* **378**, 49 (2022).
11. J.-L. Chen *et al.*, *J. Exp. Med.* **201**, 1243 (2005).
12. G. Stewart-Jones *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5784 (2009).
13. D. M. Tadros, S. Eggenschwiler, J. Racle, D. Gfeller, *Nucleic Acids Res.* **51** (D1), D428 (2023).
14. A. I. Webb *et al.*, *J. Biol. Chem.* **279**, 23438 (2004).
15. N. P. Kristensen *et al.*, *J. Clin. Invest.* **132**, e150535 (2022).

ACKNOWLEDGMENTS

I.M.S. acknowledges support from NIH R01CA255039.

10.1126/science.adz6423

[1]Department of Microbiology and Immunology, University of Minnesota Medical School, Minneapolis, MN, USA. [2]Center for Immunology, University of Minnesota Medical School, Minneapolis, MN, USA. [3]Masonic Cancer Center, University of Minnesota Medical School, Minneapolis, MN, USA. [4]Center for Genome Engineering, University of Minnesota Medical School, Minneapolis, MN, USA. Email: ingunn@umn.edu

MATERIALS SCIENCE

Minuscule vibrations, uncovered

Computational imaging resolves atomic vibrations at picometer scale

Toma Susi

Twisted heterostructures—two atomically thin layers stacked at an angle by van der Waals forces—exhibit exotic properties beyond those of each layer alone, such as superconductivity (*1*, *2*). The interlayer interactions cause structural reconstruction of atoms within the layers, forming a Moiré superlattice. This lattice has distinctive vibrational modes, called Moiré phonons (or phasons), that affect the material's properties (*3*, *4*). However, low energy (~0.01 meV) and nanometer-scale spatial variation make the characterization of phasons with existing techniques extremely challenging. On page 423 of this issue, Zhang *et al.* (*5*) report the use of a computational imaging technique, called electron ptychography, to characterize phasons in twisted bilayers of tungsten diselenide. The findings illustrate how the extreme spatial resolution offered by electron ptychography can yield insights into a phenomenon that would be otherwise difficult to study.

Development of advanced electron microscopy techniques has enabled the imaging of individual atoms within a material. For example, in aberration-corrected scanning transmission electron microscopy, an atomically focused probe is scanned across the sample. Detecting electrons elastically (no energy transfer) scattered to specific angles uncovers information about the atomic structure of a specimen by providing contrast between elements (heavier elements scatter more and thus appear brighter) (*6*). Measuring the subsequent energy loss of inelastically scattered electrons allows evaluation of vibrational properties of the sample on the atomic scale (*7*). Although these methods can obtain images and spectra at atomic resolution, they cannot distinguish individual phasons in a material. Analyzing electron diffraction patterns recorded at each scan position has become a way to uncover further information of specimen structure and properties (*8*). In particular, electron ptychography is a computational imaging technique that reconstructs a high-resolution image of a sample from overlapping diffraction patterns. It corrects image distortions due to aberrations post-acquisition and can precisely measure distances between atoms in a twisted heterostructure close to picometer (10^{-12} m) resolution (*9*).

Two types of algorithms are available: noniterative methods, which simply reconstruct the structure over the entire scanned area, and iterative methods, which resolve intricate features by refining diffraction signals over the set of scan positions. In a conventional sample that is thicker than an atom, electrons of the probe can scatter multiple times because of the increased depth available for interaction. This makes imaging thick samples at picometer resolution challenging. The multislice algorithm divides a thick sample into multiple thin slices then simulates electron scattering from each section. The model simultaneously compares simulation to experimental data over the set of scan positions to iteratively refine what the specimen structure must be to have resulted in the recorded scattering signal. This method can not only recover depth information but also improve the reconstruction of atomic positions in a twisted heterostructure (*10*). Spatial resolution of multislice electron ptychography is only limited by smearing due to lattice vibrations (*11*). Although this technique cannot directly capture the dynamic motion of vibrating atoms because of their very high frequencies, it can effectively capture averages (snapshots) of the time-dependent dynamics.

Zhang *et al.* used electron ptychography to distinguish atomic-scale anisotropies of vibrational modes (phasons) in individual atoms of twisted bilayers of tungsten diselenide. The authors performed careful statistical analysis of the changes in the contrast of a tungsten atom in different areas of the Moiré superlattice to quantify asymmetries in solitons (networks of stacking faults in the reconstructed lattice) at the picometer scale. The measurements revealed Moiré phasons to be the most prominent thermal vibrations in low-angle twisted bilayers. The experimental measurements were in close agreement with ptychography simulations based on molecular dynamics modeling of the phasons. By contrast, standard frozen phonon models, which simulate atomic positions on the basis of different static snapshots of isotropic displacements, did not reproduce the observed asymmetry. Thus, the real-space imaging of Zhang *et al.* convincingly determine the magnitude of displacement of phasons.

A previous study reported similar anisotropy in the rattling motion of heavy barium atoms in a clathrate compound (a lattice that traps molecules or atoms) by capturing segmented detector signals using scanning transmission electron microscopy (*12*). Naturally asymmetric molecular cages in the host compound provided a highly anisotropic confinement of the trapped barium atoms (movements are more restricted in certain directions), resulting in strongly asymmetric scattering contrast in the resulting images. However, the picometer precision of Zhang *et al.* is notable, resolving the minuscule anisotropy in phasons within the lattices of twisted bilayers of tungsten diselenide. This study also demonstrates the use of electron ptychography beyond imaging static structures to examine dynamic vibration of atoms within a layered material.

The increasing availability of state-of-art direct-electron detectors has overcome limitations of measurement speed and electron dose rate. Combined with advanced computational algorithms, electron ptychography could accelerate understanding of dynamic material behaviors at atomic resolution. In addition, the atomic-scale electromagnetic potentials of the specimen leave an imprint on the scattered electrons. Thus, interesting properties such as charge transfer (*13*) and magnetic order (*14*) can be examined with the help of quantitative simulations. The next challenge is to extend the reconstruction of these more subtle properties into three dimensions. □

REFERENCES AND NOTES

1. A. K. Geim, I. V. Grigorieva, *Nature* **499**, 419 (2013).
2. Y. Cao *et al.*, *Nature* **556**, 43 (2018).
3. M. Koshino, Y.-W. Son, *Phys. Rev. B* **100**, 075416 (2019).
4. H. Ochoa, *Phys. Rev. B* **100**, 155426 (2019).
5. Y. Zhang *et al.*, *Science* **389**, 423 (2025).
6. O. L. Krivanek *et al.*, *Nature* **464**, 571 (2010).
7. F. S. Hage, G. Radtke, D. M. Kepaptsoglou, M. Lazzeri, Q. M. Ramasse, *Science* **367**, 1124 (2020).
8. C. Ophus, *Microsc. Microanal.* **25**, 563 (2019).
9. Y. Jiang *et al.*, *Nature* **559**, 343 (2018).
10. S. Gao *et al.*, *Nat. Commun.* **8**, 163 (2017).
11. Z. Chen *et al.*, *Science* **372**, 826 (2021).
12. K. Tabata *et al.*, *Small Sci.* **4**, 2300254 (2024).
13. C. Hofer, J. Madsen, T. Susi, T. J. Pennycook, *J. Microsc.* jmi.13404 (2025).
14. J. Cui, H. Sha, W. Yang, R. Yu, *Sci. Bull.* **69**, 466 (2024).

10.1126/science.adz7157

Faculty of Physics, University of Vienna, Vienna, Austria. Email: toma.susi@univie.ac.at

MARINE CONSERVATION

A catch in ocean conservation

Satellite imagery reveals global patterns of industrial fishing across marine protected areas

Boris Worm

Humans sometimes take from Earth whatever they can, without much consideration for the future (*1*). In the ocean, for example, historical overexploitation has compromised marine biodiversity and associated services (*2*). Fortunately, many nations are now changing course by investing in the ocean's future. Recently, a United Nations Ocean Conference (UNOC 2025) saw governments commit to responsible fisheries management and increased marine protected area (MPA) coverage. However, there is a concern that persistent exploitation of MPAs may compromise their ability to safeguard valuable biodiversity assets. On pages 392 and 396 of this issue, Raynor *et al.* (*3*) and Seguin *et al.* (*4*), respectively, use cutting-edge satellite technology to show that although industrial fishing vessels are present in many protected areas worldwide, MPAs with the highest levels of protection remain largely unfished. These findings clarify the present extent of fishing pressure on protected areas and will help inform expanding ocean conservation efforts.

Approximately 8% of the world's ocean is under formal protection, with international commitments set to more than triple present coverage by 2030. Protective measures often provide long-term benefits (*5*), especially when combined with effective fisheries management (*6*). Yet these efforts can be undermined by poor regulation. For example, destructive fishing methods continue to be used in protected areas that lack appropriate safeguards (*7*). To better understand patterns of industrial fishing across protected areas, Raynor *et al.* and Seguin *et al.* applied recent advances in monitoring global fisheries from space, a rapidly emerging technology that has revolutionized the capacity to directly observe human activity at sea (*8*–*10*). By combining synthetic aperture radar (SAR) satellite imagery of coastal marine areas (*10*) with previously available automatic identification system (AIS) data (*9*), it is possible to identify the position of any fishing vessel larger than 15 m, capturing the distribution of the world's industrial fleets in unprecedented detail.

Raynor *et al.* analyzed 455 coastal MPAs that have been identified as "fully" or "highly" protected under the MPA Guide, an assessment framework that takes implemented regulations and existing management plans into account (*5*). In these fully or highly protected MPAs, all forms of industrial fishing are prohibited. The analyses focused on coastal MPAs because these overlap with available SAR imagery and because fishing may be more common in areas closer to the coastal human population. The encouraging news emerging from this study is that these MPAs are successful in excluding industrial fishing, with very few vessels seen in these waters—on average, just one detection per 20,000 km^2. However, MPAs with this level of protection are not the norm and account for less than 10% of all MPAs by number. Fortunately, some of these well-protected areas are very large,

Fishing vessels leave Tailu Central Fishing Port in China (August 2024). Industrial fishing may affect almost half of coastal protected areas worldwide, according to satellite imagery data.

PHOTO: WANG WANGWANG/VCG/GETTY IMAGES

collectively covering ~3% of the ocean and accounting for one-third of the existing global MPA area (*11*).

Seguin *et al.* quantified fishing across a larger group of 6021 coastal MPAs. Importantly, these MPAs encompassed a wide range of protective categories according to the International Union for Conservation of Nature (IUCN) management framework, which is commonly used to classify protected areas on land and in the sea. The authors detected industrial fishing vessels in approximately half (47%) of all analyzed MPAs. Furthermore, fishing effort (measured as the number of fishing hours) in the protected areas was often at a level equal to or greater than that observed in nonprotected areas nearby. The results are partially consistent with the findings of Raynor *et al.*, as fishing effort was less in more strictly protected areas (IUCN categories I and II) than in others. However, further analysis by Seguin *et al.* suggests that the presence of industrial fishing patterns in MPAs is more strongly related to the size and remoteness of the protected area, rather than to their management category per se.

It may be surprising to some that almost half of the MPAs observed by Seguin *et al.* are experiencing industrial fishing, despite the IUCN clarifying that such activity is not compatible with biodiversity conservation objectives (*12*). In the European Union, for example, previous research has shown that most MPAs are fished more heavily, on average, than surrounding unprotected areas, a conservation paradox that needs to be addressed with urgency (*7*). Likewise, Seguin *et al.* also observed some of the highest fishing vessel densities in Dutch and Belgian MPAs, as well as in China.

Collectively, these findings raise questions about the disconnect between the stated intent of MPAs to effectively protect biodiversity and the extent of industrial fishing activity in these areas. This may be partly due to national efforts to raise MPA coverage quickly, given international commitments under the Convention on Biological Diversity, and also because the ocean has been lagging far behind land in term of spatial protection. For example, less than 1% of the ocean was formally protected in the year 2000 compared with more than 10% of land, whereas by 2025 these figures have expanded to 8% of the ocean compared with 17% of land. However, such rapid expansion across the ocean came at a cost. Many MPAs have been established quickly without strong protective regulations, meaningful consultation with local stakeholders, or appropriate management capacity. In some cases, this has resulted in "paper parks" that are recognized as protected areas but do not prevent harmful activities.

These observations have important implications for ongoing efforts to expand MPA coverage to at least 30% of the ocean by 2030. Clearly, paper parks will only ever deliver suboptimal outcomes, if any at all. Yet the available data show that where proper investments are made, industrial exploitation is curtailed, and protective measures are comprehensive, long-term benefits will accrue. Most well-protected areas, both land and sea, provide meaningful and measurable benefits for biodiversity (*13*) and human well-being (*14*), with stronger protection progressively increasing benefits across both dimensions (*15*).

The overarching message is that proper investment in protected areas will pay off and that satellite technology can be one of the key tools to help ensure that such investments are kept safe. Observational capacities have continued to grow since AIS technology became widely used a decade ago (*8*), and there is rapid movement toward the ability to quantify the full extent of human activity at sea and assess the reach of marine protection. Such unprecedented transparency, combined with improved regulation, cooperation, and accountability across sectors, will help to both safeguard the ocean's future and build toward a sustainable human enterprise. □

REFERENCES AND NOTES

1. G. M. Heal, *Valuing the Future: Economic Theory and Sustainability* (Columbia Univ. Press, 1998).
2. B. Worm *et al.*, *Science* **314**, 787 (2006).
3. J. Raynor *et al.*, *Science* **389**, 392 (2025).
4. R. Seguin, F. Le Manach, R. Devillers, L. Velez, D. Mouillot, *Science* **389**, 396 (2025).
5. K. Grorud-Colvert *et al.*, *Science* **373**, eabf0861 (2021).
6. R. Hilborn *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2218 (2020).
7. M. Dureuil, K. Boerder, K. A. Burnett, R. Froese, B. Worm, *Science* **362**, 1403 (2018).
8. D. J. McCauley *et al.*, *Science* **351**, 1148 (2016).
9. D. A. Kroodsma *et al.*, *Science* **359**, 904 (2018).
10. F. S. Paolo *et al.*, *Nature* **625**, 85 (2024).
11. ProtectedSeas, Navigator data download (2025); https://navigatormap.org/data-request.
12. IUCN, Applying IUCN's global conservation standards to marine protected areas (MPAs): Delivering effective conservation action through MPAs, to secure ocean health & sustainable development. Version 1.0 (IUCN, 2018).
13. P. F. Langhammer *et al.*, *Science* **384**, 453 (2024).
14. N. Ban *et al.*, *Nat. Sustain.* **2**, 524 (2019).
15. A. J. Nowakowski *et al.*, *Nat. Sustain.* **6**, 1210 (2023).

ACKNOWLEDGMENTS

The author thanks L. Schiller for helpful discussion and comments.

10.1126/science.adz8322

Biology Department, Dalhousie University, Halifax, NS, Canada. Email: bworm@dal.ca

QUANTUM COMPUTING

Flat optics produces quantum graphs

A miniature device links multiple photon paths for bespoke entanglement

Maxim R. Shcherbakov

A photon—the smallest discrete unit (quantum) of light—is a fundamental concept in various technologies such as secure communication and quantum computing. Identical photons can "feel" each other's presence on a beam splitter, an optical device that separates a beam of light into distinct paths. This so-called quantum interference is a straightforward method for generating quantum entanglement, in which two or more photons are linked regardless of their distance. Entangling photons that travel across multiple paths is one of the primary challenges in quantum technologies. Existing devices occupy immense space for a handful of photonic qubits (quantum bits analogous to classical bits). On page 416 of this issue, Yousef *et al.* (*1*) report a metasurface—a planar array of structures with sizes smaller than the wavelength of light—that can manage photons on demand. This produces a special class of quantum states in a miniature optical device with micrometer dimensions.

In the classic 1987 study of Hong, Ou, and Mandel (*2*), two indistinguishable photons with the same wavelength and polarization arrive simultaneously at a 50:50 beam splitter, which has equal probabilities of transmitting and reflecting a photon (see the figure). Instead of parting ways, the incoming photons are either transmitted or reflected together because of quantum interference. This observation established anticorrelations between the two paths: Light detectors placed at each propagation direction will never click at the same time. Anticorrelations are an essential precursor to quantum entanglement; however, the complexity of Hong–Ou–Mandel beam-splitting apparatus grows with the number of photon paths. An updated version of the Hong–Ou–Mandel effect (*3*) uses horizontally and vertically polarized photons as input, inviting the use of polarization-sensitive metadevices (*4*) for quantum interference.

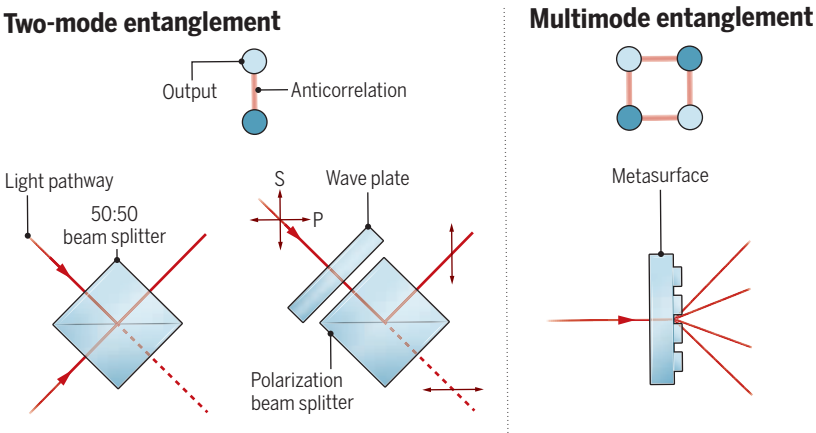
Yousef *et al.* judiciously engineered a polarization-splitting metasurface that can double the number of output channels. The incoming photons were split into four channels on a flat device with micrometer thickness and sent to four detectors. Similar to the Hong–Ou–Mandel demonstration, some detectors registered anticorrelations. By contrast, other detectors clicked simultaneously and exhibited correlations, resulting in more complex output states than those of the Hong–Ou–Mandel experiment.

Quantum interference across more than two channels has been shown in metasurfaces previously (*5, 6*). However, Yousef *et al.* took this concept further by demonstrating that the output could represent a bipartite (two-colored) graph—a mathematical construct that can visualize quantum states as colored nodes and connections. This expands the catalog of cluster states (the combination of nodes and connections), which are key resources in measurement-based quantum computation that uses a sequence of local measurements to perform computation (*7*). Many desired universal cluster states are bipartite (*8*), and some measurement-based quantum computing schemes are equivalent to measurements on bipartite graphs (*9*). Thus, the work of Yousef *et al.* suggests the potential use of metasurfaces to scale up photonic quantum computers.

The method of Yousef *et al.* is versatile. For example, adding fully programmable metasurfaces (*10*) could control output states on demand. This way, traditional photonic quantum computing schemes, such as Gaussian boson sampling (*11*) and variational eigensolvers (*12*) can be miniaturized. However, outstanding issues, such as compromised purity of output quantum states, lack of three-dimensional packaging (housing and protection), and sizable photon generation and detection devices, still need to be addressed. In addition, the quality of metasurface-generated states should be rigorously compared to those of the conventional silicon photonic devices (*13*). Nevertheless, the remarkable progress of the past decade (*14*, *15*) shows that metasurfaces have become a welcome home for photons—and a catalyst for what they will enable next. □

Entangling multiple photon modes

A beam-splitting apparatus based on the classic Hong-Ou-Mandel effect is limited by two output modes. The metasurface can double the number of outputs by splitting the incoming photons into four channels. The output states and their correlations can be graphically represented as a bipartite graph.



P, parallel polarization; S, perpendicular polarization.

GRAPHIC: A. MASTIN/*SCIENCE*

REFERENCES AND NOTES

1. K. Yousef *et al.*, *Science* **389**, 416 (2025).
2. C. K. Hong, Z. Y. Ou, L. Mandel, *Phys. Rev. Lett.* **59**, 2044 (1987).
3. T. E. Kiess, Y. H. Shih, A. V. Sergienko, C. O. Alley, *Phys. Rev. Lett.* **71**, 3893 (1993).
4. X. Yin, Z. Ye, J. Rho, Y. Wang, X. Zhang, *Science* **339**, 1405 (2013).
5. T. Stav *et al.*, *Science* **361**, 1101 (2018).
6. K. Wang *et al.*, *Science* **361**, 1104 (2018).
7. H.-J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, M. Van den Nest, *Nat. Phys.* **5**, 19 (2009).
8. H. Aschauer, W. Dür, H.-J. Briegel, *Phys. Rev. A* **71**, 012319 (2005).
9. I. D. Smith, H. P. Nautrup, H.-J. Briegel, *Phys. Rev. Lett.* **132**, 220602 (2024).
10. S. Q. Li *et al.*, *Science* **364**, 1087 (2019).
11. H.-S. Zhong *et al.*, *Science* **370**, 1460 (2020).
12. A. Peruzzo *et al.*, *Nat. Commun.* **5**, 4213 (2014).
13. J. Wang *et al.*, *Science* **360**, 285 (2018).
14. A. S. Solntsev, G. S. Agarwal, Y. S. Kivshar, *Nat. Photonics* **15**, 327 (2021).
15. T. Santiago-Cruz *et al.*, *Science* **377**, 991 (2022).

10.1126/science.adz5392

Department of Electrical Engineering and Computer Science, University of California, Irvine, CA, USA.
Email: maxim.shcherbakov@uci.edu

BOOKS ET AL.

EARTH SCIENCE

Stone and its stories

A journalist pens a love letter to geologic history **Riley Black**

The striking geology of Arizona's Antelope Canyon offers hints about our planet's past.

Strata: Stories from Deep Time
Laura Poppick
Norton, 2025. 288 pp.

Mud rarely gets the appreciation it deserves. The slurry of water, sediment, organic detritus, and microorganisms is usually treated as a nuisance—an annoying and obvious result of water suffusing eroded stone. But mud, as science writer Laura Poppick explains in a portion of the lushly descriptive *Strata*, is a wonder. It is an intersection of different earthly realms, a sedimentary state that did not always exist and may not have appeared if it were not for more-ancient interactions between the biotic and abiotic components of our planet. In Poppick's hands, something as simple as a rime of muck around the edge of a driveway is an invitation to wonder about relationships and changes that can only be understood through layers of stone stacked like storybook pages.

Geology is a challenging subject to write about. Rocks form, change, and move on timescales far longer than our human life spans allow us to comfortably comprehend. And while the present is the key to the past, as every geology student has hammered into their minds, Earth is not traveling along a stable and predictable path of progress. Earth's strata provide the chief record of what life truly was like in the distant past, before animals crawled and wiggled and walked on land, before forest fires were even possible, before molecular oxygen was an important component of our atmosphere.

"It is through these lines in stone that we can glimpse ancient iterations of this planet and gain context for the moment we're spinning through now," Poppick writes in the book's prologue. But *Strata* offers something more than just geological history. Just as the story of Earth itself is not linear, the same holds true for geological hypotheses and notions. What we know of our planet and what we still wish to know emerge from a constantly changing array of studies carried out by people whose individual perspectives, backgrounds, and interests influence what has generally accreted over the past two centuries into what we now know as "Earth science."

Covering the entirety of Earth's history would result in a book about as thick as Earth's crust. Instead, Poppick selects four main themes—air, ice, mud, and heat—that span 2.4 billion years and thread together to reach something more fundamental than a straightforward geological retelling. Each part is rich with scientific findings, visits to field sites and labs, historical sketches, and poetic swaths that Poppick deftly balances page after page. Her personal approach carries the book through immense spans of time. Considering the exposed stone of a canyon in Wyoming, Poppick writes, "I arrived at that canyon wall with a longing for a version of the planet I knew I would never see in my lifetime. A version that wasn't burning, flooding, skittering off-kilter from our own undoings."

For any lump of stone or exposed rock face, the compressed and altered sediment has stories to tell about our planet. But rock also holds stories of how colonialism and utilitarian needs have connected humanity to rock through mines and road cuts; stories of who has been fascinated by strata throughout history and who has been permitted to study it; and stories about what old rocks might tell us about our present moment—and perhaps tomorrow. *Strata* is thus ultimately less about the rocks themselves than the relationships between humans and the rocky parts of the planet we live atop. It is this focal point that allows Poppick to glide back and forth through time to retell the stories of scientific debates over the "Snowball Earth" hypothesis, which posits that the planet was once entirely covered in ice; to recount tales of attempts to address the dispossession of Indigenous peoples by mining interests; and to recall the joy she felt uncovering a chip of Jurassic dinosaur bone for the first time.

Poppick's lavish writing presents a gem on nearly every page. Throughout the book, rocks singled out for their stories have "warm buttery grains," early continental landmasses were like "cookie dough," and glacial ice comes in vibrant hues "that glow like sapphires or hard candies." And through her fondness for strata and the stories that can be found therein, readers will be left with something better than a geological crash course. They will turn the final page with a new way of looking at the interplay between the world we know and its backstory. "There's magic in the unknown, but there's also magic in the known," Poppick writes, in how "the evolution of bacteria capable of turning sunshine into sugar and a few global ice ages, together, made possible the creation of life with cells complex enough to build stems and roots and leaves and eventually brains large enough to make sense of this all."

Our history is not merely found among the hominins, mammals, and tetrapods that we count among our ancestors. Earth itself shaped humankind as we have shaped the planet in return—a relationship that requires repair but could be rehabilitated with a little more recognition of where we came from. □

10.1126/science.ady7001

The reviewer is the author of *When the Earth Was Green: Plants, Animals, and Evolution's Greatest Romance* (St. Martin's, 2025) and *The Last Days of the Dinosaurs* (St. Martin's, 2022). Email: evogeek@gmail.com

ARCHAEOLOGY

Sensing early civilizations

Armchair archaeologists summon ancient tastes, sights, and sounds to understand the past **Patrick E. McGovern**

Dinner with King Tut
Sam Kean
Little, Brown, 2025. 464 pp.

A prospective reader of *Dinner with King Tut* might be excused for thinking that it would provide a detailed description of a royal dinner served up to King Tutankhamun. The reader would be sadly mistaken. The book, which purports to introduce readers to ongoing attempts to reproduce ancient foodstuffs and technologies in order to better understand past civilizations, focuses mainly on quirky avocational experiments while giving short shrift to the exciting scholarly work being conducted in the field of experimental archaeology.

The book's chapter on Egypt, for example, describes author Sam Kean's attempt to make an ancient rendition of sourdough bread in his home kitchen on the advice of a self-proclaimed "gastro-Egyptologist." Then, taking his cue from an online recipe created by another self-described Egyptologist, Kean has a go at crumbling his unappetizing bread into a pot of boiling water, adding some herbs and an undefined commercial yeast, and producing an equally sour and insipid beer. Besides the fact that one of the herbs—cinnamon—was likely not available to the Egyptians until more recently, the author's re-created libation hardly does justice to the range of brews that were enjoyed by ancient Egyptians, royalty and commoners alike. Even modern *bouza* (unrelated to English "booze"), which carries on ancient tradition and is served up by local street vendors throughout the country, is more richly textured, flavored, and high-powered than Kean's brew. I share Kean's enthusiasm for re-creating our human past, having devoted half of my archaeological career to discovering and resuscitating ancient ales and spirits from around the world (*1*). But I believe that he diminishes the contributions of professionals who have pioneered the field of experimental archaeology over the past half century while unduly elevating "rogue" amateurs.

An ancient Egyptian figurine shows a brewer preparing bread for beer.

To be sure, Kean recounts his forays into bread- and beermaking with literary verve and flair. Beginning in Africa 75,000 years ago, he travels progressively through time and space from South America to Turkey, Egypt, Polynesia, Italy, California, Scandinavia, Alaska, and China, finally reaching Mexico in the 1500s CE. Throughout the book, Kean intersperses his own experiments and those of other so-called "rogue archaeologists" with imaginative and highly dramatic tales meant to put the re-creations in context and bring the past alive. Human struggle and survival are highlighted in these stories, which often include gruesome details of murder, suicide, sacrifice, and plague.

One of the book's subjects is Roger Larsen, a carpenter by day in a small Mississippi town and an "Egypt enthusiast" in his off time. Kean recounts how Larsen erected miniature replicas of the Great Pyramid in his workshop to test his theory that this Wonder of the World was built with a rope-and-lever machine. His view runs counter to the standard hypothesis endorsed by most professional archaeologists, which proposes that gradually higher earthen ramps were constructed against the faces of the pyramid so that multiton limestone and granite blocks could be gradually pulled up the slope by rope on sleds under human and animal power. It is true that both hypotheses lack any direct archaeological, textual, or artistic evidence, but the ramp theory is generally thought to be more in keeping with the limited engineering prowess of about 4600 years ago.

Larsen takes a very hard-line approach in advocating his position and rejecting the standard hypothesis, which probably accounts for why his theory is yet to be published. He chalks this up to the intransigence of archaeologists looking down on amateurs and protecting their own turf, an underlying motif stressed by Kean throughout the book. This may make for more compelling reading, but it serves no useful purpose to set amateurs in opposition to professionals. Most archaeologists would agree that hobbyists have been a mainstay of this underfunded endeavor since its inception. As volunteers on excavations and in the lab, their contributions continue up to the present.

Experimental archaeology would have been better served by Kean had he articulated a coherent view of the field. At various turns, he describes archaeology as "the most stirring field in science" and then laments that "traditional" archaeology is "dull" and unexciting in its slow-motion excavation and detailed pottery study. But these tedious tasks provide the essential contemporaneous "facts on the ground" that enable re-creations of the past. When these data are combined with literary, artistic, and ethnographic perspectives and numerous ever-more-sensitive scientific approaches, well-defined and controlled experiments that truly advance our understanding of the past can be carried out and published. This knowledge may well lead to new taste sensations, alternative medicines, a better understanding of our shared biological and cultural heritage, and much more. Let us hope that Kean's spirited recounting of less-controlled experiments fires the minds of future experimental archaeologists in that direction. □

REFERENCES AND NOTES

1. P. E. McGovern, *Ancient Brews: Rediscovered and Re-created* (Norton, 2017).

10.1126/science.ady4803

The reviewer is the scientific director of the Biomolecular Archaeology Project at the University of Pennsylvania Museum, Philadelphia, PA, USA, and is at the Department of Anthropology, University of Pennsylvania, Philadelphia, PA, USA. Email: mcgovern@upenn.edu

RETRACTION

The Research Article "A bacterium that can grow by using arsenic instead of phosphorus" by F. Wolfe-Simon *et al.* (*1*) has been the subject of discussion and critique since its online publication in 2010. In 2011, *Science* published the print version of the paper accompanied by eight Technical Comments (*2*–*9*) and a Technical Response from Wolfe-Simon *et al.* (*10*), along with an Editor's Note (*11*). In 2012, *Science* published two papers that failed to reproduce the finding that the GFAJ-1 bacterium can grow using arsenic instead of phosphorus (*12*, *13*).

Science did not retract the paper in 2012 because at that time, Retractions were reserved for the Editor-in-Chief to alert readers about data manipulation or for authors to provide information about postpublication issues. Our decision then was based on the editors' view that there was no deliberate fraud or misconduct on the part of the authors. We maintain this view, but *Science*'s standards for retracting papers have expanded. If the editors determine that a paper's reported experiments do not support its key conclusions, even if no fraud or manipulation occurred, a Retraction is considered appropriate.

Over the years, *Science* has continued to receive media inquiries about the Wolfe-Simon Research Article, highlighting the extent to which the paper is still part of scientific discussions. On the basis of the 2011 Technical Comments and 2012 papers, *Science* has decided that this Research Article meets the criteria for retraction by today's standards. Therefore, we are retracting the paper. Author Ronald S. Oremland is deceased. Author Peter K. Weber disagrees with the Retraction. Authors Felisa Wolfe-Simon, Jodi Switzer Blum, Thomas R. Kulp, Gwyneth W. Gordon, Shelley (Hoeft) McCann, Jennifer Pett-Ridge, John F. Stolz, Samuel M. Webb, Paul C. W. Davies, and Ariel D. Anbar also disagree with the decision to retract and have posted an online eLetter, available at https://www.science.org/doi/10.1126/science.adu5488#elettersSection, explaining their objections.

H. Holden Thorp, Editor-in-Chief, *Science*

REFERENCES AND NOTES

1. F. Wolfe-Simon *et al.*, *Science* **332**, 1163 (2011).
2. D. W. Borhani, *Science* **332**, e1201255 (2011).
3. S. A. Benner, *Science* **332**, e1201304 (2011).
4. S. Oehler, Science **332**, e1201381 (2011).
5. I. Csabai, E. Szathmáry, *Science* **332**, e1201399 (2011).
6. B. Schoepp-Cothenet *et al.*, *Science* **332**, e1201438 (2011).
7. R. J. Redfield, *Science* **332**, e1201482 (2011).
8. P. L. Foster, *Science* **332**, e1201551 (2011).
9. J. B. Cotner, E. K. Hall, *Science* **332**, e1201943 (2011).
10. F. Wolfe-Simon *et al.*, *Science* **332**, e1202098 (2011).
11. B. Alberts, *Science* **332**, 1149 (2011).
12. T. J. Erb, P. Kiefer, B. Hattendorf, D. Günther, J. A. Vorholt, *Science* **337**, 467 (2012).
13. M. L. Reaves, S. Sinha, J. D. Rabinowitz, L. Kruglyak, R. J. Redfield, *Science* **337**, 470 (2012).

10.1126/science.adu5488

Reinstate the National Climate Assessment

In April, the Trump administration dismissed all authors working on the sixth National Climate Assessment (NCA) and canceled funding for the US Global Change Research Program (*1*)—the federal agency charged with coordinating the report. For more than three decades, the NCA has drawn on the best available science to understand how climate change affects every region of the country and every sector of the US economy. Government agencies and communities across the US depend on the report for nonpartisan, reliable information about escalating risks. Defunding the NCA undermines the nation's ability to prepare for global changes that affect the economy, public health, infrastructure, and national security while exacerbating risks for future generations (*2*).

The legislation authorizing the NCA passed unanimously in the Senate and by voice vote without objection in the House of Representatives in 1990 (*3*). Since that time, climate change has intensified. Atmospheric carbon dioxide concentrations have risen nearly 20% and exceeded the "red line" of 400 parts per million, beyond which climate scientists warn of increasingly irreversible damage (*4*). Global temperatures have increased by about 0.9°C, accounting for more than two-thirds of all warming since preindustrial times (*5*). Sea levels have risen more than 10 cm, and the rate of rise has more than doubled (*6*).

Addressing climate change involves consequential economic decisions and substantial scientific resources (*7*, *8*). Environmental and resource economists, who have spent their careers evaluating the costs and benefits of mitigation and adaptation, have called for the reinstatement of the NCA authors and the continuation of this critical work (*9*). The American Geophysical Union and the American Meteorological Society have organized a special collection spanning 29 peer-reviewed journals in an effort to "maintain momentum" on US climate assessment research (*10*), which is admirable but will likely fall short of the NCA's ability to promote comprehensive ecological, economic, and social sustainability and adaption to climate change. Elected officials and federal leadership must safeguard US resilience by protecting the independence of the NCA, ensuring its timely release, and fully funding its coordination and communication efforts.

Craig E. Landry[1], Lynne Lewis[2], Shana McDermott[3], Brian Vander Naald[4], Martin D. Smith[5], John C. Whitehead[6]

[1]Department of Agricultural and Applied Economics, University of Georgia, Athens, GA, USA. [2]Department of Agricultural and Resource Economics, Colorado State University, Fort Collins, CO, USA. [3]Department of Economics, University of Missouri, Columbia, MO, USA. [4]Department of Accounting, Economics, Finance, and Business Law, Drake University, Des Moines, IA, USA. [5]Nicholas School of the Environment, Duke University, Durham, NC, USA. [6]Department of Economics, Appalachian State University, Boone, NC, USA. Email: clandry@uga.edu

REFERENCES AND NOTES

1. S. Waldman, E&E News, "Trump dismisses scientists writing key climate report," *Scientific American*, 29 April 2025.
2. Intergovernmental Panel on Climate Change (IPCC), "Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," Core Writing Team *et al.*, Eds. (IPCC, 2023).
3. Global Change Research Act of 1990, Pub. L. No. 101-606, 104 Stat. 3096 (1990).
4. P. Friedlingstein *et al.*, *Earth Syst. Sci. Data* **14**, 4811 (2022).
5. H. Lee *et al.*, "Climate Change 2023: Synthesis Report, Summary for Policymakers. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," Core Writing Team *et al.*, Eds. (IPCC, 2023).
6. J. A. Church, N. J. White, *Surv. Geophys.* **32**, 585 (2011).
7. L. Barrage, W. Nordhaus, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2312030121 (2024).
8. K. Rennert *et al.*, *Nature* **610**, 687 (2022).
9. "Defending the National Climate Assessment" (2025); support-nca.org.
10. American Geophysical Union (AGU), "AGU and AMS join forces on special collection to maintain momentum of research supporting the U.S. National Climate Assessment," 2 May 2025.

10.1126/science.adz3885

NASA Earth Science Division provides key data

In May, the US administration proposed budget cuts to NASA, including a more than 50% decrease in funding for the agency's Earth Science Division (ESD) (*1*), the mission of which is to gather knowledge about Earth through space-based observation and other tools (*2*). The budget cuts proposed for ESD would cancel crucial satellites that observe Earth and its atmosphere, gut US science and engineering expertise, and potentially lead to the closure of NASA

research centers. As former members of the recently dissolved (*3*) NASA Earth Science Advisory Committee, an all-volunteer, independent body chartered to advise ESD (*4*), we warn that these actions would come at a profound cost to US society and scientific leadership.

NASA ESD measurements provide critical information supporting environmental and resource management and community decision-making. In collaboration with the National Oceanic and Atmospheric Administration (*5*), ESD develops and launches satellites that track weather and pollution. Weather forecasts, including artificial intelligence–based apps on cell phones, rely on this information and will suffer if there are fewer measurements available. ESD measurements are also used in near real-time to diagnose fire risk, support US wildland fire management, and track active fires and their impacts (*6*). Furthermore, the data collected by ESD are used to monitor croplands and rangelands, make commodity forecasts, and provide early-warning data for potential crop failures, water availability, and irrigation needs (*7*). ESD satellite data underpin drought, flood, and hurricane forecasting and are used to monitor water quality, fisheries, and harmful algal blooms (*8*, *9*). This information supports public water management, coastal communities, and commercial and recreational fishing. ESD satellites also track sensitive ecosystems, delivering important data to support wildlife and habitat conservation and management (*10*).

The US has long been the global leader in Earth-observing satellite technology (*11*). If the proposed cuts are realized, the US will surrender this leadership to countries in Asia and Europe. Given that the cuts would prevent the US from training and preparing the next generation of the scientific and technical workforce, the consequences would be long-lasting.

NASA ESD accounted for just 0.03% of US spending in 2024 (*12*). This investment returns its value many times over by improving predictions, by spurring technological innovation and high-tech jobs, and by forging the knowledge of the planet that is needed for short- and long-term planning.

Dylan B. Millet[1], Belay B. Demoz[2], Jennifer D. Watts[3], Robert Wright[4]

[1]Department of Soil, Water, and Climate, University of Minnesota, St. Paul, MN, USA. [2]Department of Physics, University of Maryland Baltimore County, Baltimore, MD, USA. [3]Woodwell Climate Research Center, Falmouth, MA, USA. [4]Hawai'i Institute of Geophysics and Planetology, University of Hawai'i at Mānoa, Honolulu, HI, USA. Email: dbm@umn.edu

REFERENCES AND NOTES

1. "Trump's proposed budget would mean 'disastrous' cuts to science," *Science*, 2 May 2025.
2. NASA Earth Science; https://science.nasa.gov/earth-science/.
3. L. McKenzie, "Agencies closing science advisory committees," American Institute of Physics, 20 March 2025.
4. NASA, Earth Science Advisory Committee; https://science.nasa.gov/researchers/nac/science-advisory-committees/esac/.
5. eoPortal, European Space Agency, Joint Polar Satellite System (JPSS); https://www.eoportal.org/satellite-missions/jpss.
6. NASA, FireSense project; https://cce.nasa.gov/firesense/.
7. S. Avant, "USDA-NASA's global view of Earth's soil holds many benefits," US Department of Agriculture (7 April 2015).
8. US Environmental Protection Agency, Cyanobacteria Assessment Network (CyAN); https://www.epa.gov/water-research/cyanobacteria-assessment-network-cyan.
9. NOAA Fisheries, Satellite data; https://www.fisheries.noaa.gov/national/science-data/satellite-data.
10. B. Burhans, "NASA and WMI partner to support biodiversity," Wildlife Management Institute, February 2025; https://wildlifemanagement.institute/outdoor-news-bulletin/february-2025/nasa-and-wmi-partner-support-biodiversity.
11. US Department of the Interior, "Earth's resources to be studied from space," 21 September 1966; https://www.usgs.gov/media/files/1966-earths-resources-be-studied-space-news-release.
12. NASA, Fiscal Year 2026 Budget Request (2025); https://www.nasa.gov/fy-2026-budget-request/.

COMPETING INTERESTS

The authors have received NASA funding for research as well as panel and proposal reviews.

Published online 10 July 2025
10.1126/science.adz6100

PHOTO: DENYS VYSHNEVSKIY

A sign warns, "Danger: landmines" in Ukrainian in a mined area in the Red Forest, located in the Chernobyl Exclusion Zone.

Weaponizing Europe's borders imperils wildlife

Five European Union countries that share borders with Russia are withdrawing (*1*) from the 1997 Anti-Personnel Mine Ban Convention, and one of them has also recently left the 2008 Convention on Cluster Munitions (*1*). Signatories to these conventions pledge to refrain from using, manufacturing, stockpiling, and supplying landmines and cluster munitions—persistent, indiscriminate, and uncontrolled weapons that can remain operative for decades and can be activated by humans, animals, and plants. Latvia, Lithuania, and Estonia (Baltic states), along with Finland and Poland, should seek alternative, more effective, humane, and measured approaches to national defense (*2*, *3*).

Landmines and cluster munitions achieve little for security (*3*), and they injure and kill mostly civilians (*2*, *3*) as well as domestic and wild animals (*4*, *5*). The risks they pose impede effective environmental monitoring, management, and restoration. The weapons, which pollute soil and water, degrade habitats of resident and migratory wildlife (*4*, *6*, *7*). Their deployment affects land use and ecosystem services, as they reduce access to and the value of the land for people and decrease the benefits of biodiversity.

The Baltic states, Poland, and Finland have recently built about 2000 km of barriers at borders with Russia and Belarus, thereby disrupting genetic exchange between populations of large mammals, such as herbivores (*8*) and large carnivores (*9*). The deployment of landmines and cluster munitions would further limit ecological connectivity.

Clearance of landmines and unexploded ordnance is slow and expensive (*1*, *3*, *4*, *6*). Their location is sometimes unknown, they can be moved by floods, and they may have plastic rather than metal casing, making them less detectable during demining activities. Modern "smart mines," designed to self-destruct, may fail and remain in the environment (*3*). Mine removal costs 100 to 300 times more than mine deployment (*4*), and given the need for detonation after removal or in their place of deployment, there will always be risk to life and land, complicating postconflict recovery (*6*, *10*). Mine cleanup in Croatia has cost more than 1 billion euros over the past 30 years and is still not complete. In Ukraine, now the world's most mined country (*11*), more than 30 billion euros and three decades will be required to remove landmines and unexploded ordnance (*11*), a challenge now compounded by a presidential decree that puts Ukraine on track to leave the Anti-Personnel Mine Ban Convention (*2*, *12*).

Baltic countries, Poland, and Finland should suspend the withdrawal process (*2*) and continue to refrain from using these weapons, as it will undermine not only global goals for a mine-free world but also nature conservation and human well-being. If such weapons are ultimately deployed, measures should be taken to improve safety and minimize environmental contamination and degradation. Establishing armament-free areas, such as peace parks, before any active conflict could serve as an effective strategy in this region.

Katarzyna Nowak[1], Jacopo Cerri[1], Svitlana Kudrenko[2,3], Alexander Kopatz[4], Izabela Stachowicz[1], Nuria Selva[5,6,7], Asmeret Asefaw Berhe[8], Ilpo Kojola[9], Nikica Šprem[10]

[1]Mammal Research Institute Polish Academy of Sciences, Białowieża, Poland. [2]Chornobyl Radiation and Ecological Biosphere Reserve, Kyiv, Ukraine. [3]Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway. [4]Norwegian Institute for Nature Research, Trondheim, Norway. [5]Estación Biológica de Doñana, Consejo Superior de Investigaciones Científicas, Sevilla, Spain. [6]Institute of Nature Conservation, Polish Academy of Sciences, Kraków, Poland. [7]Departamento de Ciencias Integradas, Facultad de Ciencias Experimentales, Centro de Estudios Avanzados en Física, Matemáticas y Computación, Universidad de Huelva, Huelva, Spain. [8]Department of Life and Environmental Sciences, University of California, Merced, CA, USA. [9]Natural Resources Institute Finland, Tampere, Finland. [10]Department of Fisheries, Apiculture, Wildlife Management, and Special Zoology, Faculty of Agriculture, University of Zagreb, Zagreb, Croatia. Email: knowak02@gmail.com

REFERENCES AND NOTES

1. M. Wareham, L. Lodenius, "Leaving the Landmine Ban Treaty puts civilians at risk," European Leadership Network, 16 June 2025.
2. United Nations, "Adhering to bans on mines only in peace time will not work: UN rights chief," 2 July 2025; https://news.un.org/en/story/2025/07/1165261.
3. C. Droege, M. Brehm, "Anti-personnel mines: the false promise of security through exceptionalism in war," *Humanitarian Law & Policy*, 13 March 2025.
4. C. T. Nachón, in *Landmines and Human Security: International Politics and War's Hidden Legacy*, R. A. Matthew *et al.*, Eds. (SUNY Press, 2012), pp. 191–207.
5. Z. Tucak, V. Kosovel, *Šumar. List* **129**, 501 (2005) [in Croatian].
6. A. A. Berhe, *Land Degrad. Develop.* **18**, 1 (2007).
7. J. P. Dudley *et al.*, *Conserv. Biol.* **16**, 319 (2002).
8. H. Bluhm *et al.*, *Divers. Distrib.* **29**, 423 (2023).
9. S. Nowak *et al.*, *Biol. Conserv.* **2024**, e01210 (2024).
10. J. D. Unruh, N. C. Heynen, P. Hossler, *Polit. Geogr.* **22**, 841 (2003).
11. "In Ukraine, tackling mine action from all sides to make land safe again," United Nations Development Programme, 14 October 2024.
12. "Ukraine on track to withdraw from Ottawa anti-personnel mine treaty, Zelenskiy decree shows," *Reuters*, 29 June 2025.

10.1126/science.adz1318

REVIEW SUMMARY

AGRICULTURE

Steering plant-soil feedback for sustainable agriculture

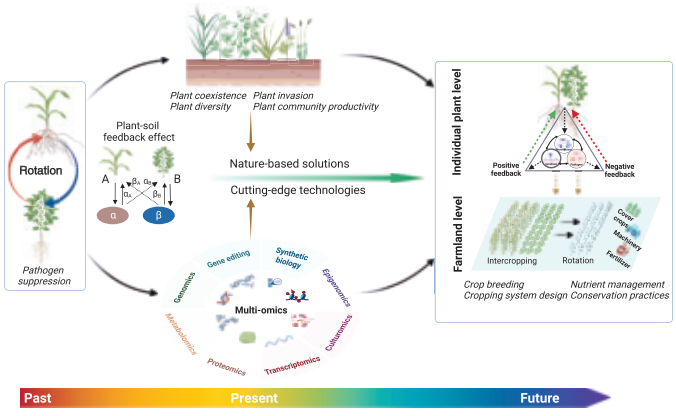
Guangzhou Wang, Wim H. van der Putten, John Klironomos, Fusuo Zhang*, Junling Zhang*

Full article and list of author affiliations: https://doi.org/10.1126/science.ads2506

BACKGROUND: Increased agricultural productivity is critical to meet the growing demand for food, feed, and fiber, but it comes with substantial environmental and resource trade-offs. In nature, feedback between plants and their soils can change the composition and productivity of vegetation. Efforts to make agriculture more sustainable by reducing chemical fertilizers and pesticides could be greatly aided by improving the natural capacity of soils to maximize crop yield. Ecological research has made considerable strides in understanding how plants interact with physicochemical and biological soil properties. Negative plant-soil feedback tends to prevail in early successional ecosystems, which are often the same type of habitats in which the wild ancestors of most modern crop species evolved. That negative feedback is a major crop productivity barrier in agriculture. The key question, therefore, asks how negative plant-soil feedbacks can be turned into positive ones, which exist in nature but are relatively rare in agriculture.

ADVANCES: Soil biota play a crucial role in plant-soil feedback interactions. Studies on plant-microbe interactions have revealed how chemical cues among microbes, between plants and microbes, and the signal transduction pathways within plants, underpin plant-soil feedback. Insights into the molecular basis underlying the relationships between plants, pathogens, and their antagonists may help the development of positive interactions at the individual plant level. Thus, identification and characterization of key genes and the underlying molecular mechanisms regulating plant-soil feedback are currently being studied with the ultimate objective of developing systems and technologies to attract, establish, and maintain beneficial microbiomes and/or suppress pests and pathogens. Root exudates or plant metabolites, such as flavonoids, quercetin, strigolactone, and coumarins act as signals or substrates for beneficial microbial growth, whereas defense compounds such as camalexin, indolic glucosinolates, and benzoxazinoids function as antimicrobials.

The current challenge is testing recent research findings under field conditions and integrating ecological principles and processes into agricultural management practices to optimize the delivery of multiple ecosystem services, rather than focusing on external inputs to maximize crop yield. For example, optimizing fertilizer applications (in terms of source, rate, timing, and placement) can enhance biological functions and minimize the need for external nutrient inputs. Plant-soil feedback-based strategies and their practical applications also require the use and further development of traditional approaches such as crop rotation, intercropping, cover cropping, and minimum tillage, all of which have been shown to support multiple ecosystem services without compromising yield. By these means, there is great potential to amplify positive plant-soil feedback effects by promoting symbiotic mutualists and other beneficial microbes that improve soil fertility and control pathogens.



Plant-soil feedback from agriculture to natural ecosystems and the pathways to sustainability. The concept of plant-soil feedback is widely applied in natural ecosystems in investigations of plant community dynamics. Crop rotation was developed to counteract negative plant-soil feedback caused by host-specific pathogens and nutrient shortage of crops grown in monoculture. Subsequently, a combination of technological developments and ecological principles have transformed the traditional, pathogen-focused approach of crop rotation into a broader synthesis for promoting soil health and ecosystem function. In the figure, "α_A" and "β_B" represent the plant-soil feedback effects of soil that has been conditioned by plant A and B on their own growth, respectively; "α_B" and "β_A" represent the reciprocal plant-soil feedback effects of soil conditioned by plants A and B, respectively. The sum of these feedback effects ($\alpha_A - \alpha_B - \beta_A + \beta_B$) indicates the magnitude and direction of the pairwise feedback. [Figure created using BioRender.com]

OUTLOOK: We review the potential of applying ecological principles to crop management to enhance positive plant-soil feedback effects in croplands while minimizing negative ones. Central to this approach is the integration of plant genomics, traits, and microbiomes with farming practices to foster mutually beneficial relationships between crops and their rhizosphere microbiomes. Exploring root- and rhizosphere-associated microbial communities offers promising opportunities for next-generation breeding strategies aimed at developing cultivars that more effectively leverage plant-microbe interactions to promote positive feedback. Ecologically based practices that optimize soil conditions and environmental factors support the development of sustainable agricultural systems that maintain high yields while preserving soil health and resilience. Realizing these goals in practice rests on decoding the complex and context-dependent interactions among management practices, soil variability, and environmental conditions under the pressures of climate change. □

*Corresponding author. Email: junlingz@cau.edu.cn (J.Z.); zhangfs@cau.edu.cn (F.Z.)
Cite this article as G. Wang *et al.*, *Science* **389**, eads2506 (2025). DOI: 10.1126/science.ads2506

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

ANIMAL COLORATION

The secret brilliance of birds

A wide array of biophysical mechanisms accounts for the astonishing diversity of coloration found in animals. Price-Waldman et al. uncovered a biophysical mechanism that explains variations in color brilliance in tanagers (genus *Tangara*) and other colorful birds. Analyses using hyper- and multispectral imaging as well as optical models revealed that "hidden" underlayments of achromatic feathers, either black or white, influence the brilliance of the colored upper layers. This mechanism appears to be widespread, pointing to its critical role in the evolution of birds, which are renowned for their extraordinary coloration. —Shahid Naeem

Sci. Adv. (2025) 10.1126/sciadv.adw5857

Colorful songbirds, such as this paradise tanager (*Tangara chilensis*), have underlying layers of black or white feathers that modulate their coloration.

PHOTO: MURRAY COOPERMINDEN PICTURES

MARINE HEATWAVES

The heat goes on

Ocean surface temperatures vary from year to year, experiencing heat waves like those felt on land, but 2023 saw an extraordinarily large increase in marine heat waves with no recent analog. Dong *et al*. report that 2023 set new records in the duration, extent, and intensity of these events by as much as three standard deviations above the historical average of the past four decades. The increasing trends in marine heat waves present intensifying dangers to ecological, social, and economic systems. —Jesse Smith

Science p. 369, 10.1126/science.adr0910

SIGNAL TRANSDUCTION

Diverging destinies of PI3Kα

Phosphoinositide 3-kinase (PI3K) has important roles in receptor signaling that control many processes, including cell growth and metabolism. Two papers in this issue target PI3K signaling in different ways and for different effects (see the Perspective by Solinas). Simanshu *et al*. developed a small molecule that inhibits activation of PI3Kα by members of the Ras family of small guanosine triphosphatases. The compound effectively inhibited tumor growth in mice, but because

insulin signaling does not require Ras proteins, it does so without hyperglycemic side effects. It was also effective against tumors promoted by excessive human epidermal growth factor receptor activity. Terayama *et al*. describe small-molecule PI3Kα activators that could regulate blood glucose concentrations in rat muscle cells in culture without the hormone insulin. These compounds functioned in an unusual manner by enhancing the binding of rat sarcoma (RAS) family proteins to PI3Kα by nearly three orders of magnitude. Their effectiveness in ameliorating hyperglycemia in a rat model of diabetes indicates that they may prove useful as therapeutic agents for the control of blood glucose without the use of insulin. —L. Bryan Ray

Science p. 409, 10.1126/science.adq2004; p. 402, 10.1126/science.adr9097
see also p. 348, 10.1126/science.adz4117

CARDIOVASCULAR BIOLOGY

Metabolic maladaptation in the heart

A high-fat diet increases the risk of cardiac steatosis and cardiomyopathy. Because obesity increases the activity and abundance of the protein phosphatase PTP1B, Sun *et al.* identified this enzyme as a potential target to mitigate the cardiovascular complications of excessive dietary fat. Mice lacking PTP1B in cardiomyocytes retained greater cardiac function when fed a high-fat diet. Moreover, cardiomyocytes in these mice accumulated less lipid and maintained fatty acid oxidation for energy production instead of switching to glucose utilization, which becomes dominant during pathological stress. —Wei Wong

Sci. Signal. (2024) 10.1126/scisignal.adp6006

PROTEIN DESIGN

Designing detection of diseased cells

Diseased and infected cells can be recognized through peptide antigens presented on the surface of the cell by major histocompatibility complex (MHC) proteins. Whereas the T cell receptor is nature's solution for recognizing peptide antigen–loaded MHCs, smaller proteins synthesized for this purpose could be transformative for diagnostic tools and immunotherapies. Liu *et al*., Johansen *et al*., and Householder *et al*. developed computational and experimental pipelines to facilitate the rapid design and validation of proteins that could recognize peptides from viral proteins, tumor-associated proteins, or neoantigens when presented by MHC molecules, with high specificity and low off-target recognition (see the Perspective by Hickok and Stromnes). When incorporated as the recognition domain for a chimeric antigen receptor and expressed in T cells, these de novo–designed peptide-MHC binders could induce signaling and on-target cell killing. —Sarah H. Ross

Science p. 386, 10.1126/science.adv0185; p. 380, 10.1126/science.adv0422; p. 375, 10.1126/science.adv3813; see also p. 349, 10.1126/science.adz6423

QUANTUM METASURFACES

Scaling multiport interferometers

The bunching and antibunching of interfering single photons is a fundamental quantum effect that underpins the development of optical-based quantum computing and communication. Extending this Hong-Ou-Mandel (HOM) effect to larger systems requires an increasing number of bulky optical components that would be practically infeasible. Yousef *et al*. report on the use of metasurfaces as a multiport HOM interferometer and related quantum correlation measurements (see the Perspective by Shcherbakov). They also introduce a graph-theoretic formalism that represents both metasurface-based quantum optics and the resulting nonclassical correlation landscape. Such graphs can be used for the design of scalable, low-decoherence quantum information infrastructures. —Ian S. Osborne

Science p. 416, 10.1126/science.adw8404; see also p. 354, 10.1126/science.adz5392

MARINE PROTECTED AREAS

Complexity of the real world

Marine protected areas (MPAs) are increasingly being established to protect marine ecosystems, in particular fish populations. However, these areas vary in the established level of protection, from complete prohibition of fishing and other uses to limiting these activities. It is also unclear to what degree these regions actually achieve the stated protection levels. In a pair of papers, authors using relatively similar data and approaches reveal seemingly contrary results. Seguin *et al.* found a continued unauthorized presence of fishing vessels in MPAs of all protection levels. By contrast, Raynor *et al.* found very little vessel activity in MPAs with the stricter protections. The authors discuss possible drivers underlying the differences (see the Perspective by Worm). —Sacha Vignieri

Science p. 396, 10.1126/science.adt9009; p. 392, 10.1126/science.ado9468; see also p. 352, 10.1126/science.adz8322

IN OTHER JOURNALS

Edited by **Corinne Simonti** and **Jesse Smith**

ANIMAL BEHAVIOR

Blowing bubbles

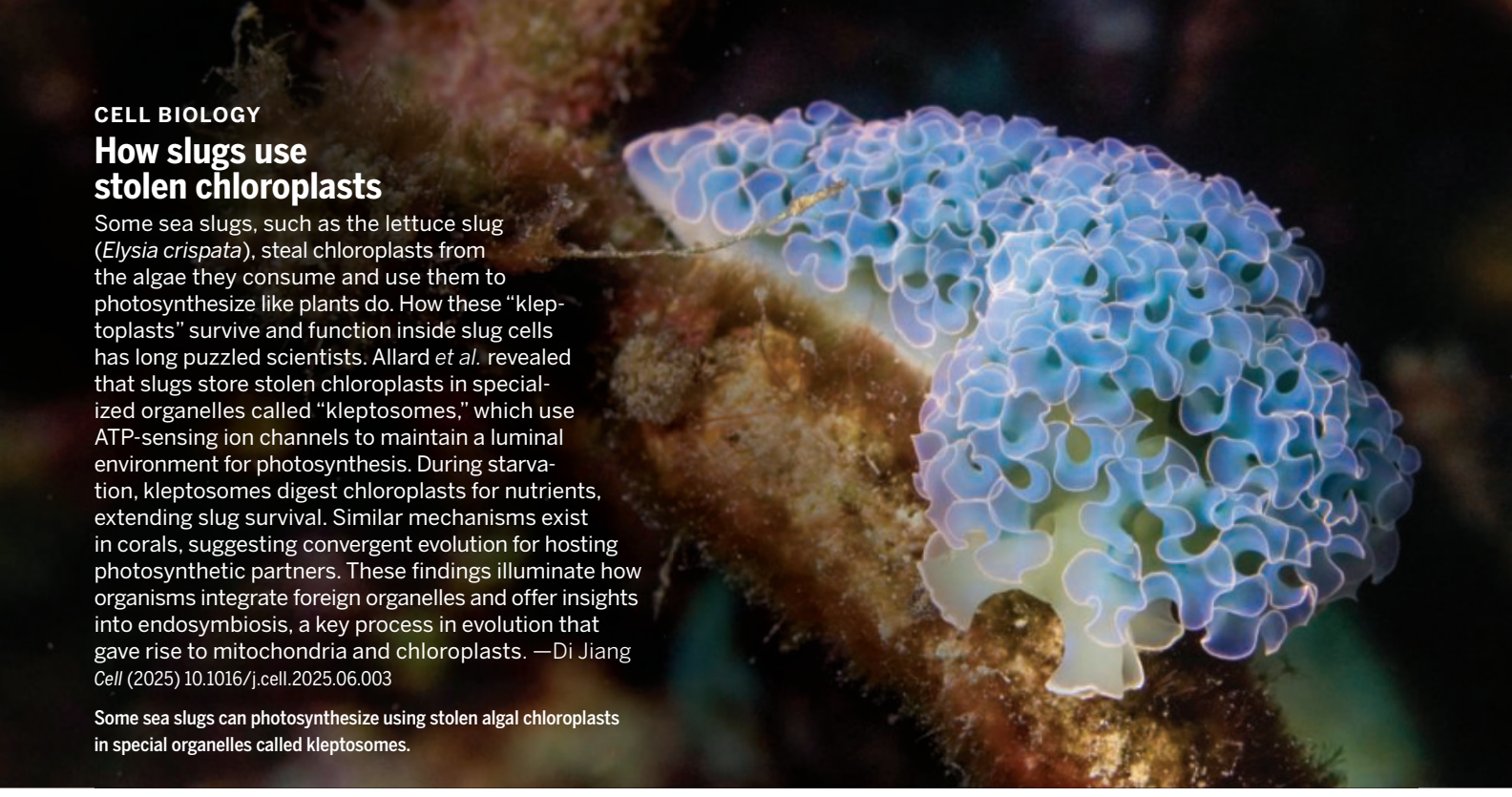
Whales have been observed blowing air-based shapes since people first started watching them. Much is known about the contexts in which odontocetes such as dolphins blow such shapes, but less is known about how baleen whales use these structures (although humpback whales have been seen using spiral shapes for foraging and air blasts on the breeding grounds). Sharpe *et al.* used drones, in conjunction with other observations and interviews, to further explore the specific production of spherical bubble rings in this species. Ten of the 12 episodes documented were made by whales in an inquisitive state in the presence of human observers, suggesting that there may be a relationship between this behavior and human receivers. —Sacha Vignieri

Mar. Mamm. Sci. (2025) 10.1111/mms.70026

IMMUNOLOGY

When not cutting it is beneficial

Toll-like receptors (TLRs) are cellular sensors that trigger immune defenses in response to molecules associated with pathogens, but not those that are intrinsic to cells. Bérouti *et al.* investigated why RNA modifications found in host cell RNA and those incorporated into mRNA therapeutics are not recognized by TLR7 and TLR8. Although modified RNA molecules and single nucleosides are not readily detected by the receptors, there is an additional safety mechanism in place in cells. The authors found that lysosomal endonuclease RNase T2, which is essential



CELL BIOLOGY

How slugs use stolen chloroplasts

Some sea slugs, such as the lettuce slug (*Elysia crispata*), steal chloroplasts from the algae they consume and use them to photosynthesize like plants do. How these "kleptoplasts" survive and function inside slug cells has long puzzled scientists. Allard *et al.* revealed that slugs store stolen chloroplasts in specialized organelles called "kleptosomes," which use ATP-sensing ion channels to maintain a luminal environment for photosynthesis. During starvation, kleptosomes digest chloroplasts for nutrients, extending slug survival. Similar mechanisms exist in corals, suggesting convergent evolution for hosting photosynthetic partners. These findings illuminate how organisms integrate foreign organelles and offer insights into endosymbiosis, a key process in evolution that gave rise to mitochondria and chloroplasts. —Di Jiang

Cell (2025) 10.1016/j.cell.2025.06.003

Some sea slugs can photosynthesize using stolen algal chloroplasts in special organelles called kleptosomes.

PHOTO: WILD HORIZONS/UNIVERSAL IMAGES GROUP/GETTY IMAGES

for generating RNA fragments that can trigger TLR7 and TLR8, could not cleave RNA molecules containing pseudouridine or N1-methylpseudouridine effectively, thus preventing potential ligands for the receptors from forming. —Sarah H. Ross

Cell (2025) 10.1016/j.cell.2025.05.032

BIOMATERIALS

Separating effects in ECM aging

Cardiac aging can lead to decreased organ functionality or disease, and this is often caused by changes to the extracellular matrix (ECM). Alterations to the ECM affect both the heart's biochemistry and its architecture and mechanics, making it difficult to determine which aspects most influence cell behavior. Sun *et al.* constructed a decellularized ECM-synthetic hydrogel hybrid scaffold that makes it possible to maintain native matrix composition while independently altering the mechanical properties. The authors seeded the scaffolds with murine primary cardiac fibroblasts and observed activation either by age (with stiffness determined by the hydrogel) or ECM ligand cues. When presented with conflicting cues, namely a young ECM attuned to be stiff like an aged one, the cues from the young matrix were able to dominate the cell response. —Marc S. Lavine

Nat. Mater. (2025) 10.1038/s41563-025-02234-6

POLARITON CHEMISTRY

Manipulating molecular assembly by light

Vibrational strong coupling (VSC) leverages light-matter strong coupling to alter intermolecular interactions, making it a promising tool for manipulating chemical reactivity across various fields, including chemical synthesis, material science, and even biological applications. However, many of the theoretically predicted capabilities of VSC have yet to be explored. Imai *et al.* demonstrated that intermolecular interactions between naphthalenediimide molecules could be altered under VSC to create distinctive supramolecular toroidal structures, some of which are inaccessible through conventional assembly methods. Their study demonstrates the ability of VSC to arrange supramolecules into distinct assembly structures. —Yury Suleymanov

J. Am. Chem. Soc. (2025) 10.1021/jacs.5c02960

CELL BIOLOGY

Getting a healthy start?

When development begins, embryos can face environmental challenges such as uterine infections. In these early stages before immune cells emerge, how early embryos respond to their first encounter with microbes remains unknown. Working with infection models in zebrafish and mouse embryos, Roncero-Carol *et al.* found that embryos can detect, ingest, and digest bacterial pathogens. Live imaging and sequencing revealed the activation of an immune program whereby embryonic surface epithelial cells generated protrusions that engulfed bacteria by phagocytosis. Similarly, donated human blastocysts could also eliminate bacteria associated with uterine diseases. Thus, immune defense is initiated in the preimplantation embryo, which may have implications for understanding human fertility. —Stella M. Hurtley

Cell Host Microbe (2025) 10.1016/j.chom.2025.05.025

GIANT IMPACTS

How to remove a planet's atmosphere

Collisions are common during the early stages of planet formation, and those between planet-sized bodies are referred to as giant impacts. These can remove some (or all) of a planet's atmosphere, but this process is poorly understood. Roche *et al.* simulated hundreds of giant impacts and tracked multiple mechanisms for atmospheric ejection. They found that large fractions of the atmosphere can be lost if the impact has sufficient mutual velocity, with the fraction lost also depending on the mass ratio and impact angle. The authors derived a scaling law that can estimate the atmospheric loss for any combination of these parameters. —Keith T. Smith

Planet. Sci. J. (2025) 10.3847/PSJ/add929

RESEARCH ARTICLE SUMMARY

IMMUNOLOGY

A human homolog of SIR2 antiphage proteins mediates immunity via the Toll-like receptor pathway

Delphine Bonhomme†, Hugo Vaysset†, Eirene Marie Q. Ednacot, Vasco Rodrigues, Yazan Salloum, Jean Cury, Anny Wang, Axel Benchetrit, Pierre Affaticati, Veronica Hernandez Trejo, Paul Vittot, Charlie Bories, Alexis Cornec, Jean-Pierre Levraud, Pedro P. Hernández, Philippe Benaroch, Benjamin R. Morehouse, Aude Bernheim†*, Enzo Z. Poirier†*

Full article and list of author affiliations: https://doi.org/10.1126/science.adr8536

INTRODUCTION: A common feature of living organisms is the evolutionary pressure exerted by pathogens, which has led to the diversification of immune mechanisms. Research in immunology largely focuses on species-specific defenses. The recent discovery of hundreds of bacterial immune systems revealed surprising homology links between certain bacterial and eukaryotic immune proteins, suggesting deep conservation across domains of life. These include cGAS-STING, viperins, and gasdermins, which are pivotal members of human innate immune pathways. However, the extent of immune conservation across the tree of life remains largely uncharted. In this work, we hypothesized that comparative immunology between bacteria and eukaryotes can be leveraged to discover previously uncharacterized immune proteins in humans and other eukaryotic species.

RATIONALE: Several bacterial proteins have been recently shown to use a silent information regulator 2 (SIR2) domain in antiphage defense. These defense systems, acting through the degradation of cellular metabolite NAD^+ (oxidized form of nicotinamide adenine dinucleotide), lead to cell death and confer immunity to the bacterial population. We hypothesized that this newly found immune function of the SIR2 domain may be conserved in eukaryotes.

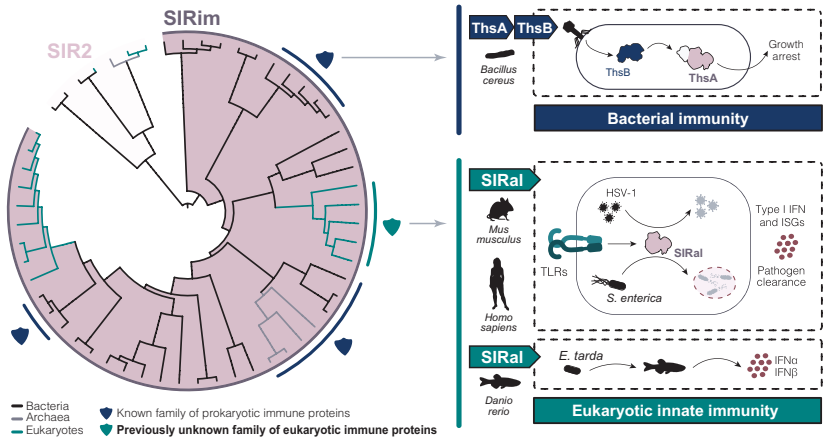
RESULTS: We used a systematic search for proteins with a SIR2 domain across bacterial, archaeal, and eukaryotic genomes to identify eukaryotic homologs of bacterial SIR2 antiphage proteins. Phylogenetic analyses revealed that SIR2 domain–containing proteins across the tree of life can be divided into two distinct subfamilies. One includes all homologs involved in housekeeping functions (sirtuins), whereas the other—termed SIRim—comprises all known bacterial immune proteins and uncharacterized eukaryotic proteins.

We hypothesized that the SIRim subfamily represents a lineage specialized in immune functions, with eukaryotic SIRims as potential previously unknown immune proteins. To test this hypothesis, we focused on FAM118B, a SIRim homolog found in vertebrates, including humans. FAM118B, renamed SIRal, has a poorly documented function and no known immune role. Innate immunity, the first line of defense against pathogens in humans, relies on the detection of infection by specific receptors, including Toll-like receptors (TLRs) that trigger an antipathogen inflammatory response. Using mouse and human cellular models, we demonstrate that SIRal is pivotal for innate immune activation within the TLR pathway. SIRal mediates the up-regulation of interferon-stimulated genes (ISGs) and proinflammatory cytokines and restricts the replication of herpes simplex virus 1 (HSV-1) and the bacterial pathogen *Salmonella enterica* in macrophages.

Given SIRal's role in mammalian innate immunity, we sought to determine whether other eukaryotic SIRims also participate in immunity. SIRim proteins are widely distributed, being present in 19% of available eukaryotic genomes and transcriptomes. Focusing on the zebrafish homolog, we show that zSIRal contributes to immunity in vivo in response to infection by bacteria *Edwardsiella tarda*. Finally, we assessed conservation of the enzymatic NADase activity. Akin to bacterial SIRims, we show that eukaryotic SIRims from different clades, including human SIRal, use NAD^+ in vitro.

CONCLUSION: Our study identifies SIRims, a previously uncharacterized family of immune proteins conserved across kingdoms of life. As part of the emerging framework of ancestral immunity—defined as conserved protein domains functioning in both bacterial and eukaryotic immunity—SIRims expand the known repertoire of ancestral immune domains. Within vertebrate SIRims, we uncovered SIRal as a pivotal protein of TLR-driven inflammation and antipathogen defense. We demonstrate that shifting the focus from a species-centric to a protein-centric view of immunity is fruitful to uncover previously unknown biology. This methodology empowers the exploration of immune organization in nonmodel eukaryotes, as well as the discovery of potential immunotherapeutic targets in humans. □

*Corresponding author. Email: enzo.poirier@curie.fr (E.Z.P.); aude.bernheim@pasteur.fr (A.Ber.) †These authors contributed equally to this work. **Cite this article as D. Bonhomme *et al.*, *Science* 389, eadr8536 (2025). DOI: 10.1126/science.adr8536**



The SIRim family of proteins is involved in immunity across the tree of life. The SIR2 family of proteins is conserved across the tree of life. Phylogenetic analyses revealed that all known SIR2 family proteins involved in bacterial immunity (e.g., ThsA from the Thoeris antiphage system) belong to the same subfamily that we named SIRim (left, brown). It is distinct from the subfamily that contains housekeeping SIR2 proteins (left, pink). We uncovered SIRal, a eukaryotic SIRim homolog found in vertebrates, notably humans, which contributes to innate immunity via the TLR pathway in macrophages. SIRal mediates the up-regulation of ISGs and proinflammatory cytokines and restricts the replication of HSV-1 and *S. enterica* (right). IFN, interferon.

RESEARCH ARTICLE SUMMARY

PROTEIN EVOLUTION

Genetics, energetics, and allostery in proteins with randomized cores and surfaces

Albert Escobedo, Gesa Voigt, Andre J. Faure, Ben Lehner*

Full article and list of author affiliations: https://doi.org/10.1126/science.adq3948

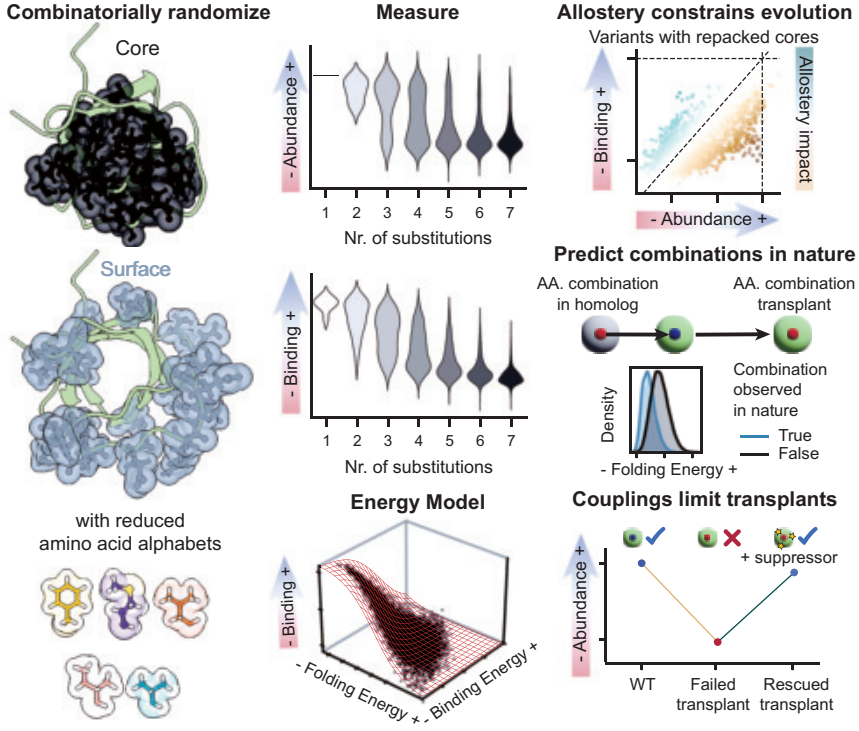
INTRODUCTION: Proteins typically contain hydrophobic amino acids buried in their cores and polar amino acids on their solvent-exposed surfaces. The rules governing which combinations of the 20 possible amino acids constitute stable and functional protein cores and surfaces are not well understood. This is partly because of the combinatorial explosion of possibilities when considering more than a few residues—experimental characterization of all combinations quickly becomes daunting.

RATIONALE: To better understand the genetic architecture of protein cores and surfaces, we designed experiments in which we quantified the stability of tens of thousands of proteins with randomized cores and surfaces, using reduced amino acid alphabets to bias toward stable combinations. For proteins with randomized cores, we also quantified their ability to bind to a ligand through a surface binding interface.

RESULTS: We found that very large numbers of proteins with randomized core or surface sequences are stable. However, we also observed that stable proteins with alternative core sequences quite frequently have impaired binding to a ligand; i.e., they are functionally impaired. We used our data to train energy models to accurately predict the stability and binding of proteins with randomized sequences. These models are simple and interpretable, with mutations having fixed additive energetic effects and a small contribution from energetic interactions between specific pairs of mutations. These energy models successfully identify the combinations of amino acids present in natural proteins that have evolved over more than a billion years, with only rare energetic interactions that we experimentally identify that prevent the transplantation of cores between highly diverged proteins.

CONCLUSION: Our results show that vast numbers of amino acid combinations can replace the core or surface of a small protein and that both the stability and binding of these proteins with randomized sequences can be predicted with simple energy models. These models also identify amino acid combinations present in natural proteins. However, changing the core of a small protein frequently disrupts its ability to bind a ligand, presumably through changes in surface conformation or altered dynamics. Indirect "allosteric" effects of mutations may thus be an important influence on the evolution of protein sequences. □

*Corresponding author. Email: bl11@sanger.ac.uk Cite this article as A. Escobedo *et al.*, *Science* **389**, eadq3948 (2025). DOI: 10.1126/science.adq3948



Randomizing protein cores and surfaces at scale. We measured intracellular abundance of thousands of proteins with randomized core and surface sequences. Very large numbers of the randomized proteins are stable, but proteins with randomized cores sometimes show reduced binding to a ligand because of allosteric effects. Simple and interpretable energy models trained on this data accurately predict protein stability and binding and amino acid combinations observed in nature, with only rare energetic couplings limiting the transplantation of cores between distantly related proteins. AA, amino acid.

RESEARCH ARTICLE SUMMARY

MOLECULAR BIOLOGY

Expansion in situ genome sequencing links nuclear abnormalities to aberrant chromatin regulation

Ajay S. Labade†, Zachary D. Chiang†, Caroline Comenho†, Paul L. Reginato, Andrew C. Payne, Andrew S. Earl, Rojesh Shrestha, Fabiana M. Duarte, Ehsan Habibi, Ruochi Zhang, George M. Church, Edward S. Boyden, Fei Chen, Jason D. Buenrostro*

Full article and list of author affiliations: https://doi.org/10.1126/science.adt2781

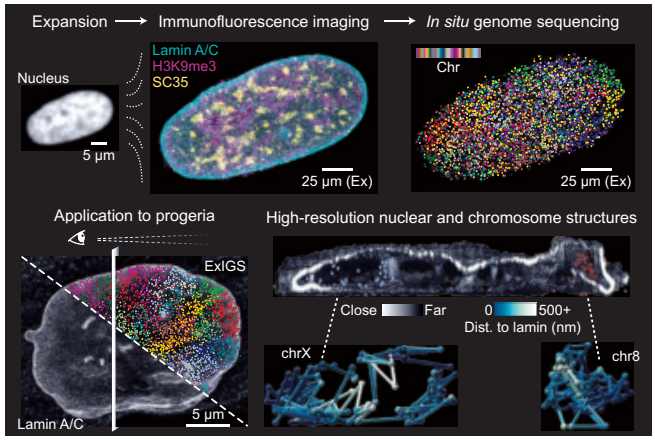
INTRODUCTION: Microscopy reveals the physical appearance of cells and enables clinical diagnoses, whereas genomics defines cell types using high-throughput molecular readouts. Although emerging spatial genomics approaches can link phenotypic disease markers to diverse-omics measurements, they often lack the resolution needed to precisely map associations within the cell nucleus, which is organized at nanometer scale. Thus, there remains a need for technologies capable of connecting high-resolution nuclear phenotypes with the spatial organization of the genome across cell types and diseases.

RATIONALE: Multiplexed imaging methods can simultaneously localize genomic regions and nuclear proteins, but their accuracy in measuring DNA-protein interactions is constrained by the diffraction limit of optical microscopy and probe-targeting restrictions. To overcome these limitations, we integrated expansion microscopy (ExM) and in situ genome sequencing (IGS) into a single approach, enabling simultaneous nanoscale imaging of nuclear proteins and spatially resolved genomic sequencing within physically expanded samples. This new technology, which we call expansion in situ genome sequencing (ExIGS), can link high-resolution nuclear phenotypes to spatial genome organization within individual cells

RESULTS: We first validated ExIGS in human fibroblast cells, demonstrating that our protocol uniformly expands nuclei while enhancing measurement of nanoscale DNA-protein interactions. We then applied ExIGS to fibroblasts derived from an individual with Hutchinson-Gilford progeria syndrome, allowing us to study how the characteristic nuclear abnormalities of this disease impact spatial genome organization. On the imaging side, expansion immunofluorescence of lamin A/C enabled us to quantitatively characterize lamin defects, including invaginations extending into the nuclear interior, which progressively accumulate in progeria fibroblasts across successive passages. On the sequencing side, three-dimensional (3D) localization analysis of genomic DNA fragments revealed that these defects primarily disrupt the functional compartmentalization of chromatin at local hotspots within individual nuclei rather than causing global disruption. Finally, paired expansion imaging of lamin A/C and RNA polymerase II revealed that lamin structures are generally associated with transcriptional repression and exhibit structural variation not only in progeria but also across normal tissue and aging contexts.

CONCLUSION: ExIGS unifies high-resolution imaging and sequencing within expanded nuclei, enabling direct characterization of how variations in nuclear morphology affect spatial genome organization. The local chromatin disruption hotspots that we observed in progeria fibroblasts provide new opportunities to investigate, such as whether lamin defects are also linked to chromatin dysregulation in normal aging. Moreover, our finding that lamin is broadly repressive regardless of structural configuration suggests that variation in nuclear shape may represent an underappreciated mechanism of gene regulation across biological contexts. We anticipate that ExIGS will serve as a generalizable imaging and sequencing platform for connecting high-resolution phenotypic variation to multiomic readouts across cells, tissues, and organisms. □

*Corresponding author. Email: jason_buenrostro@harvard.edu †These authors contributed equally to this work. Cite this article as A. S. Labade *et al.*, *Science* **389**, eadt2781 (2025). DOI: 10.1126/science.adt2781



ExIGS links nuclear morphology to spatial genome organization at high resolution. In ExIGS, nuclei are physically enlarged in an expandable hydrogel, nuclear proteins are imaged using immunofluorescence, and 3D genomic locations are identified using in situ sequencing. Applied to progeria, ExIGS enables high-resolution quantification of how characteristic lamin abnormalities affect 3D chromosome structures.

RESEARCH ARTICLE SUMMARY

LEPROSY

Pre-European contact leprosy in the Americas and its current persistence

Maria Lopopolo† and Charlotte Avanzi† *et al*.

Full article and list of author affiliations: https://doi.org/10.1126/science.adu7144

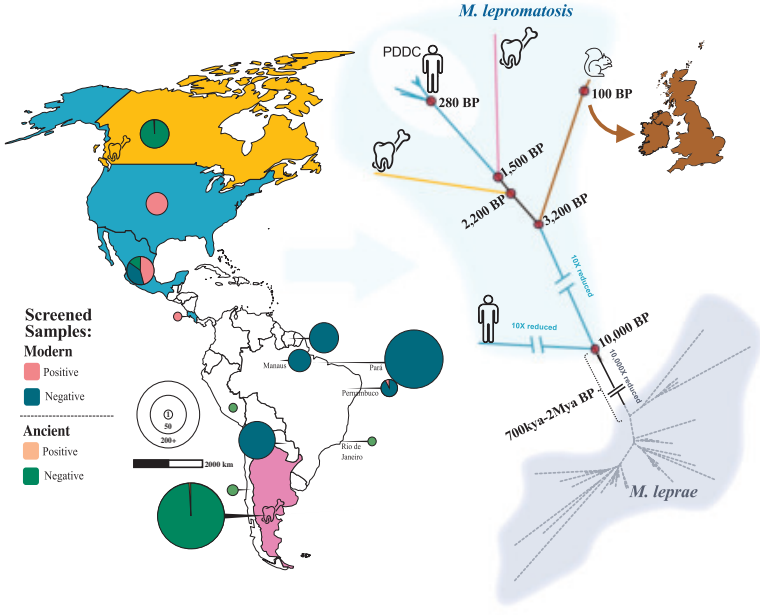
INTRODUCTION: Leprosy is a chronic infectious disease historically associated with *Mycobacterium leprae*, a pathogen whose emergence and global spread have been extensively studied through ancient and modern genomes. By contrast, the second causative agent of leprosy, *Mycobacterium lepromatosis*, was only identified in 2008 and is mostly reported in the Americas, with a few sporadic cases in Asia and in red squirrels from the British Isles. Although *M. leprae* was introduced to the Americas during European colonization, the evolutionary and epidemiological history of *M. lepromatosis*—including its origin and diversification—remains poorly understood, largely owing to the scarcity of genomic data.

RATIONALE: To investigate the history and diversity of *M. lepromatosis*–borne leprosy in the Americas, where most cases are reported today, we combined ancient and modern DNA analyses. We aimed to determine whether this pathogen existed in the Americas before European contact and to better understand its current distribution, phylogenetic structure, and temporal diversification. We performed a continent-wide screening of clinical samples from individuals with suspected leprosy and analyzed ancient human remains predating European arrival to provide a long-term historical perspective of the disease.

RESULTS: First, we generated a new high-quality reference genome of *M. lepromatosis* by assembling long-read data, which improved genotype calling accuracy in subsequent analyses. By screening 408 modern clinical samples from five American countries, we identified 34 positive cases—mainly from Mexico and the US—and reconstructed 24 genomes. Among 389 ancient samples, we detected *M. lepromatosis* DNA in three individuals—one from Canada and two from Argentina—radiocarbon dated 1310 to 860 years before the present, confirming its pre-contact presence in both North and South America. Phylogenetic analyses of ancient and modern genomes revealed five major clades, including three that were newly identified. Two clades are defined by the ancient strains from North and South America. A third, represented by two present-day US strains, forms a basal clade that diverged from all others around 10 millennia ago. Our results also indicate a likely American origin for the red squirrel clade in the British Isles, which appears to have expanded during the 19th century. Finally, a previously known clade—formerly represented by just three genomes—has now been expanded to 24 nearly identical strains. This clade dominates current diversity in Mexico and the US and likely expanded during the postcontact period. Using this newly expanded dataset, we re-estimated the divergence between *M. leprae* and *M. lepromatosis* to be at 2 million to 700,000 years ago—substantially more recent than previous estimates of ~14 million years.

CONCLUSION: This study provides definitive evidence that leprosy caused by *M. lepromatosis* existed in the Americas before European contact and was geographically widespread. We identified multiple distinct *M. lepromatosis* clades circulating across the continent, some of which persist today. Our findings suggest that this pathogen diversified within the Americas during the Late Holocene, underscoring the value of ancient genomics for understanding infectious disease history. Expanding the temporal and geographic sampling of both ancient and modern strains will be critical to better understand the evolutionary and epidemiological history of this neglected human pathogen. □

Corresponding author: Nicolás Rascovan (nicolas.rascovan@pasteur.fr) †These authors contributed equally to this work. Cite this article as M. Lopopolo *et al*., *Science* **389**, eadu7144 (2025). DOI: 10.1126/science.adu7144



Geographic distribution and diversity of *Mycobacterium lepromatosis*. Ancient and modern cases of *M. lepromatosis* were identified through continent-wide screening. Phylogenetic analysis reveals five distinct lineages, including ancient strains from both North and South America, demonstrating the pathogen's presence and diversification in the Americas before European contact. BP, years before the present; kya, thousands of years ago; Mya, millions of years ago; PDDC, present-day dominant phylogenetic clade.

RESEARCH ARTICLE SUMMARY

DRUG DEVELOPMENT

A neomorphic protein interface catalyzes covalent inhibition of RAS^{G12D} aspartic acid in tumors

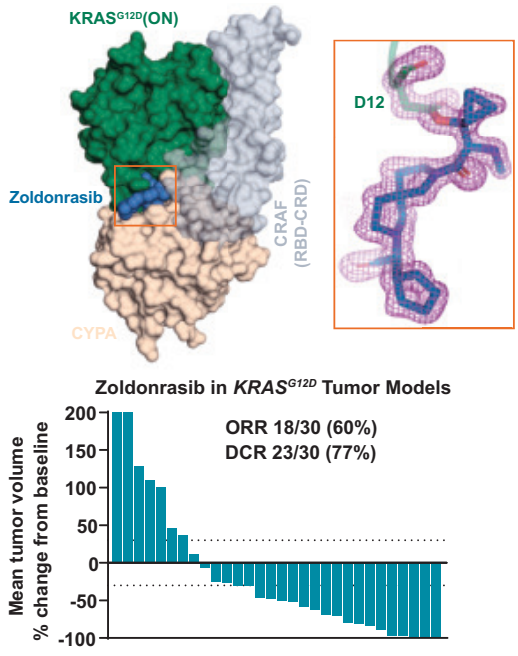
Caroline Weller†, G. Leslie Burnett†, Lingyan Jiang†, *et al*.

Full article and list of author affiliations: https://doi.org/10.1126/science.ads0239

INTRODUCTION: Mutant RAS is the primary oncogenic driver in RAS-addicted cancers. RAS proteins act as molecular switches that cycle between a guanosine triphosphate–bound "ON" state and a guanosine diphosphate–bound "OFF" state. Under normal circumstances, cell growth is tightly regulated by modulating the population of RAS in the ON versus the OFF state. In RAS-addicted cancers, mutations in RAS shift this balance towards the ON state, leading to increased RAS signaling and uncontrolled cell growth. Mutations in RAS are present in nearly a quarter of a million new cancer cases each year and are particularly frequent in pancreatic, colorectal, and lung cancers.

RATIONALE: Historically, RAS-addicted cancers have been challenging to treat with targeted drug therapies. Mutant-selective inhibitors that covalently and irreversibly inactivate mutant RAS are advantageous because they target the mutant protein present only in cancer cells and can achieve sustained target engagement even in the context of variable tumor exposures that often result from typical inhibitor dosing schedules. Such irreversible inhibitors of KRAS^{G12C}, in which glycine is replaced with a reactive thiol-containing cysteine residue, have been approved for both lung and colorectal cancers, but no similar therapies are available for the most common RAS mutation, KRAS^{G12D}. This mutation introduces a carboxylate-containing aspartic acid residue that has low reactivity and high abundance on protein surfaces compared with cysteine, thereby posing substantial hurdles to the design of irreversible inhibitors that have sufficient potency and selectivity yet maintain the properties necessary to enable once-daily oral dosing.

RESULTS: To create covalent inhibitors of KRAS^{G12D}, we used structure-based drug design to modify compounds that bind the abundant intracellular chaperone cyclophilin A (CYPA) and create a neomorphic protein-protein interface between CYPA and active RAS to covalently modify the D12 mutation located in the induced pocket at the interface. Precisely positioning reactive groups within this privileged environment enabled selective, enzyme-like rate enhancement of the covalent reaction between D12 and aziridine warheads that have low intrinsic reactivity. X-ray crystal structures and computational methods confirmed a role for the protein-protein interface in enabling selective reactivity. This approach yielded the investigational agent zoldonrasib (RMC-9805), currently undergoing clinical evaluation (NCT06040541), and the preclinical compound RMC-9945. These compounds efficiently covalently engaged RAS^{G12D} and potently suppressed oncogenic RAS signaling in RAS^{G12D}-mutant cancer cell lines in a CYPA-dependent manner. Consistent with the mechanism of action, the compounds exhibited low-potency noncovalent inhibition of RAS-driven proliferation in cell lines and patient-derived organoids with wild-type and non-G12D mutant RAS, and covalency conferred selectivity and durability of inhibition toward RAS^{G12D}. The chemical and metabolic stability was sufficient to enable once-daily oral dosing in mice, and zoldonrasib displayed marked antitumor activity in multiple preclinical models of *KRAS*^{G12D}-mutant pancreatic, lung, and colorectal cancers.



A selective covalent inhibitor of KRAS^{G12D}. An x-ray crystal structure shows zoldonrasib bound to CYPA and KRAS^{G12D}, with continuous electron density indicating a covalent bond formed with the D12 residue. This complex sterically blocks RAS interactions with effectors, such as CRAF, blocking oncogenic signaling and resulting in strong antitumor responses in many preclinical models of KRAS^{G12D}-mutant cancers. RBD, RAS binding domain; CRD, cysteine-rich domain; ORR, overall response rate; DCR, disease control rate.

CONCLUSION: Creation of a neomorphic protein-protein interface at RAS through chemical remodeling of the cellular chaperone CYPA selectively accelerated covalent bond formation between an inhibitor and an aspartic acid residue, enabling discovery of potent and irreversible inhibitors of the most common RAS mutation in human cancers. Zoldonrasib is an orally bioavailable, RAS(ON) G12D–selective covalent inhibitor that drives deep and durable tumor regressions in multiple preclinical models of KRAS^{G12D} cancers across indications. This strategy has the potential to greatly expand the repertoire of residues on cancer drivers or other proteins of therapeutic value that can be targeted by covalent warheads and may enable additional mechanisms of target modulation. □

Corresponding author. Email: Jingjing Jiang (jjiang@revmed.com); John E. Knox (jknox@revmed.com); David Wildes (pete@revmed.com) †These authors contributed equally to this work. **Cite this article as C. Weller *et al*., *Science* 389, eads0239 (2025). DOI: 10.1126/science.ads0239**

MARINE HEATWAVES

Record-breaking 2023 marine heatwaves

Tianyun Dong[1,2,3], Zhenzhong Zeng[1,2,3]*, Ming Pan[4], Dashan Wang[3], Yuntian Chen[1], Lili Liang[3], Shuai Yang[5], Yubin Jin[1,3,6], Shuxin Luo[1,2,3], Shijing Liang[3], Xiaowen Huang[3], Dongzhi Zhao[7], Alan D. Ziegler[8], Deliang Chen[9], Laurent Z. X. Li[10], Tianjun Zhou[11], Dongxiao Zhang[1]

The year 2023 witnessed an extraordinary surge in marine heatwaves (MHWs) across Earth's oceans, setting new records in duration, extent, and intensity, with MHW activity totaling 53.6 billion °C days square kilometer—more than three standard deviations above the historical norm since 1982. Notable events include the North Atlantic MHW (276-year return period) and the Southwest Pacific (141 years). Using ECCO2 (Estimating the Circulation and Climate of the Ocean-Phase II) high-resolution daily data, we conducted a mixed-layer heat budget analysis and identified region-specific drivers: enhanced shortwave flux and a shallower mixed layer in the North Atlantic and North Pacific, reduced cloud cover and increased advection in the Southwest Pacific, and oceanic advections in the Tropical Eastern Pacific. The 2023 MHWs highlight the intensifying impacts of a warm climate and the challenges in understanding extreme events.

Marine heatwaves (MHWs) are discrete and prolonged warm ocean temperature extremes that have the potential to substantially disrupt ecosystems (*1*, *2*), often leading to widespread coral bleaching and mortality events (*3*–*7*). These events also have profound economic and societal impacts (*8*–*11*), including declines in fishery productivity as well as disruptions to aquaculture and other marine-dependent industries (*12*, *13*). The sixth Intergovernmental Panel on Climate Change report outlines with high confidence the increasing frequency of MHWs in recent decades, alongside medium confidence in their widening occurrence, driven by ongoing anthropogenic climate change superimposed on natural climate variability (*14*). Since March 2023, many of the global oceans have experienced unusual warming of mean sea surface temperatures (SSTs), eclipsing all previously recorded observations (fig. S1) (*15*–*19*). In 2023, this unprecedented ocean warming had already triggered large-scale coral bleaching events (*20*), disrupted fisheries in multiple regions (*21*, *22*), and intensified stress on marine ecosystems (*23*), underscoring its immediate ecological and societal consequences. This sharp rise in global SST sparks concerns about the potential recurrence of extreme MHWs in a warming climate (*24*) and their possible indication of an approaching climatic tipping point (*25*, *26*).

The World Meteorological Organization declared 2023 as the "warmest year on record, by a huge margin" (*27*). That year also saw global ocean heat content surge to unprecedented levels, with particularly pronounced warming in the Southern Hemisphere and Atlantic Oceans (*28*). Regional reports highlighted extreme MHWs affecting varied areas, including the North Atlantic, the Tropical Pacific, the South Pacific, and the North Pacific (*29*–*32*). However, the origins, persistence, likelihood, intensification, and drivers of the 2023 global MHWs remain incompletely understood (*33*, *34*). Unraveling the mechanisms behind these processes—such as coupled air-sea interactions, dynamic ocean processes, and the modulation of climate variability modes (*35*–*39*)—is essential for developing effective adaptation and mitigation strategies.

To address this knowledge gap, we conducted a comprehensive investigation of the 2023 MHWs, analyzing their severity, magnitude, and driving mechanisms using gridded reanalysis products and satellite-derived variables. MHWs were identified as periods of at least five consecutive days during which SST exceeded the 90th percentile threshold of climatological values, following the widely accepted definition proposed by Hobday *et al.* (*40*), with the climatological baseline set from 1985 to 2014. Mean intensity was defined as the mean SST anomaly relative to the climatological mean during the MHW, whereas cumulative intensity referred to the sum of SST anomalies over the event's duration. To account for the spatial extent of individual MHWs, we applied the MHW activity index (*41*), which integrated the frequency, duration, intensity, and spatial extent of MHWs occurring within each extended year across the global ocean (see materials and methods). Additionally, the category intensity of a MHW event was classified according to scale descriptors: moderate (category I), strong (category II), severe (category III), and extreme (category IV) (*42*). To estimate the return level of a given event for a specific return period (T), we used the generalized extreme value distribution (GEV) to compute cumulative density functions (CDFs; $T = 1 - 1/\mathrm{CDF}$).

To better understand the mechanisms driving the unprecedented 2023 MHWs, we used the ECCO2 (Estimating the Circulation and Climate of the Ocean-Phase II) high-resolution (*43*) and Optimum Interpolation Sea Surface Temperature (OISST) satellite-derived dataset (*44*) to investigate the key physical processes at play. We also performed a mixed-layer heat budget analysis (*45*–*47*) to quantify the relative contributions of processes driving MHWs across the global ocean, including surface net heat flux, horizontal advection, vertical mixing, and ocean temperature tendency. This comprehensive approach leverages the strengths of ECCO2's capabilities and OISST's observational accuracy, providing critical insights into the variability and mechanisms sustaining MHWs across different regions (see materials and methods).

Unprecedented patterns of the MHWs in 2023

Our findings reveal that the global MHWs of 2023 exhibited the longest durations, widest extents, and highest intensities on record, representing a global event with a 50-year return period (<2% probability). The 2023 MHWs reached an unprecedented activity level of 53.6 billion °C days km^2, exceeding the 1982–2022 average by more than three standard deviations (Fig. 1A). The 2023 MHWs exhibited the longest duration (120 days; with an average of 35.72 days from 1982 to 2023) and the broadest spatial occurrence (0.32 billion km^2; 96% of the global marine area, compared with a historical average of 73.7%) and attained a maximum mean intensity of 1.3°C (exceeding the average of 0.98°C; Fig. 1B). Spatially, the global prevalence of MHWs increased substantially over 1982–2023, with a significant increasing trend of 9.03% per decade ($P < 0.001$, blue line in Fig. 1B), reaching its highest level in 2023. The increase in mean intensity paralleled the expansion MHW area, which increased at a rate of 0.12°C per decade ($P < 0.001$, yellow line in Fig. 1B). In 2023, MHW duration quadrupled

[1]Zhejiang Key Laboratory of Industrial Intelligence and Digital Twin, Eastern Institute of Technology, Ningbo, Zhejiang, China. [2]State Key Laboratory of Soil Pollution Control and Safety, Southern University of Science and Technology, China. [3]School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China. [4]Center for Western Weather and Water Extremes, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA. [5]Earth and Environmental Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [6]Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China. [7]School of Marine Sciences, Nanjing University of Information Science and Technology, Nanjing, China. [8]Andaman Coastal Research Station for Development, Kasetsart University, Ranong, Thailand. [9]Department of Earth System Science, Tsinghua University, Beijing, China. [10]Laboratoire de Météorologie Dynamique, CNRS, Sorbonne Université, Ecole Normale Supérieure, Ecole Polytechnique, Paris, France. [11]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China. ***Corresponding author. Email: zengzz@sustech.edu.cn**

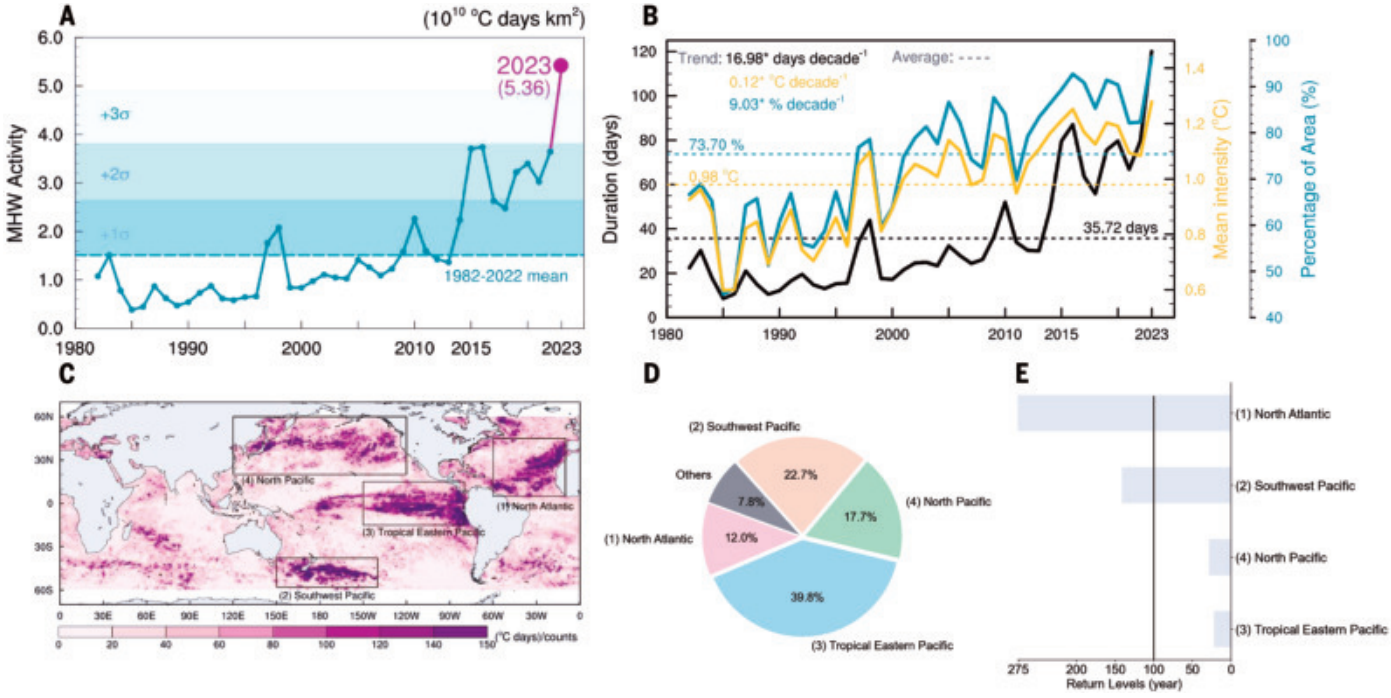


Fig. 1. Distribution of global MHWs from 1982 to 2023. (**A**) Global MHW activity from 1982 to 2023; darker, medium, and lighter shading indicate one, two, and three standard deviations, respectively, from the 1982 to 2022 mean. (**B**) Time series of duration (black line), mean intensity (yellow line), and the global percentage of ocean area experiencing MHWs (blue line) from 1982 to 2023. Horizontal dashed lines represent the long-term averages for each metric, with corresponding average values annotated above each dashed line. Trends for each variable are shown at the top of the panel, with asterisks (*) showing the 99.9% confidence level. (**C**) Spatial distribution of mean cumulative intensity in 2023 for each event. Notable regions are denoted by black rectangles. (**D**) Pie chart represents the spatial coverage proportion of cumulative intensity exceeding the first quartile across different regions. (**E**) Estimated return periods of MHWs that occurred in different regions in 2023. The return period is a statistical measurement typically used in climate studies and insurance to calculate the likelihood of a particular event occurring.

the average values registered from 1982 to 2022, after a significant increasing trend of 16.98 days per decade ($P < 0.001$, black line in Fig. 1B). This prolongation was a major contributor to the exceptionally high cumulative intensities of the 2023 global MHWs.

Beyond the global perspective, notable disparities in the mean cumulative intensity of individual MHW events were observed across four key oceanic regions: the North Atlantic Ocean, Tropical Eastern Pacific Ocean, North Pacific Ocean, and the Southwest Pacific Ocean in the vicinity of New Zealand (Fig. 1C). These regions collectively accounted for 90% of the global ocean warming above the 25th percentile (Fig. 1D). The North Atlantic Ocean experienced a record-breaking MHW event with a 276-year return period (i.e., <0.4% probability; Fig. 1E), consistent with findings of Kuhlbrodt *et al.* (*32*), who reported SSTs in this region reaching four standard deviations above the 1982–2011 mean (*48*). With the SST anomalies in the northeastern Atlantic Ocean peaking in June, the region has experienced multiple unusual MHWs over time, characterized by exceptionally high SSTs (*18*).

The MHW in the Southwest Pacific Ocean had an estimated return period of 141 years (0.7% probability). To assess the consistency between satellite-based and in situ observations, we examined data from a coastal station at Leigh in northern New Zealand. In 2022, this site recorded an extreme MHW event lasting 313 days—the most severe in its 57-year observational record (*49*). OISST satellite data corroborated these findings (fig. S2), demonstrating strong consistency between satellite and ground-based detection approaches. Expanding the analysis beyond the local Leigh site to the broader Southwest Pacific Ocean, we found that the 2023 MHW greatly exceeded the cumulative intensity of the 2022 event (fig. S3C). It also likely contributed to the intensification of Cyclone Gabrielle, which severely affected New Zealand in early 2023, highlighting the critical role of oceanic thermal anomalies in cyclone genesis and intensity (*50*, *51*).

The North Pacific Ocean experienced an MHW with a return period of 28 years (Fig. 1E and fig. S4). In the Tropical Eastern Pacific Ocean, milder MHW events in 2023 had return periods on the order of 20 to 30 years (Fig. 1E and figs. S3 and S4). Collectively, the prolonged duration of these events was a key factor contributing to the extremeness (cumulative intensity) of the global MHWs in 2023. Many of the 2023 MHWs exceeded previous yearly records by the largest margins, particularly in the North Atlantic Ocean and Southwest Pacific Ocean (figs. S3 to S5).

Tracking the evolution of MHWs in 2023

To understand the progression of MHWs in 2023, we analyzed their time series (Fig. 2), focusing on the onset, duration, and intensity variations. The spatial distribution and intensity categories of MHWs across key oceanic regions in 2023 are shown in Fig. 2A and range from moderate to extreme, reflecting the unprecedented scale and intensity of these events. In the North Atlantic, an MHW event began as early as June 2022 and persisted until the end of 2023, peaking in July 2023. This event lasted for 525 days with severe intensity (category III), which makes it the longest recorded MHW in the region (Fig. 2D and fig. S6A). The mean intensity exceeded 1.0°C throughout its life span, with localized temperatures surpassing 3.0°C (category IV, extreme MHW; Fig. 2, A and D). Notably, MHWs in this region are most frequent from July to October, with a historical occurrence probability exceeding 30%—meaning the temporal occurrence probability, i.e., MHWs occurred on the same calendar day in >30% of the years from 1982 to 2023 (Fig. 2D, colored dots at the bottom of the plot). The extraordinary duration and intensity of the 2023 MHWs underscore

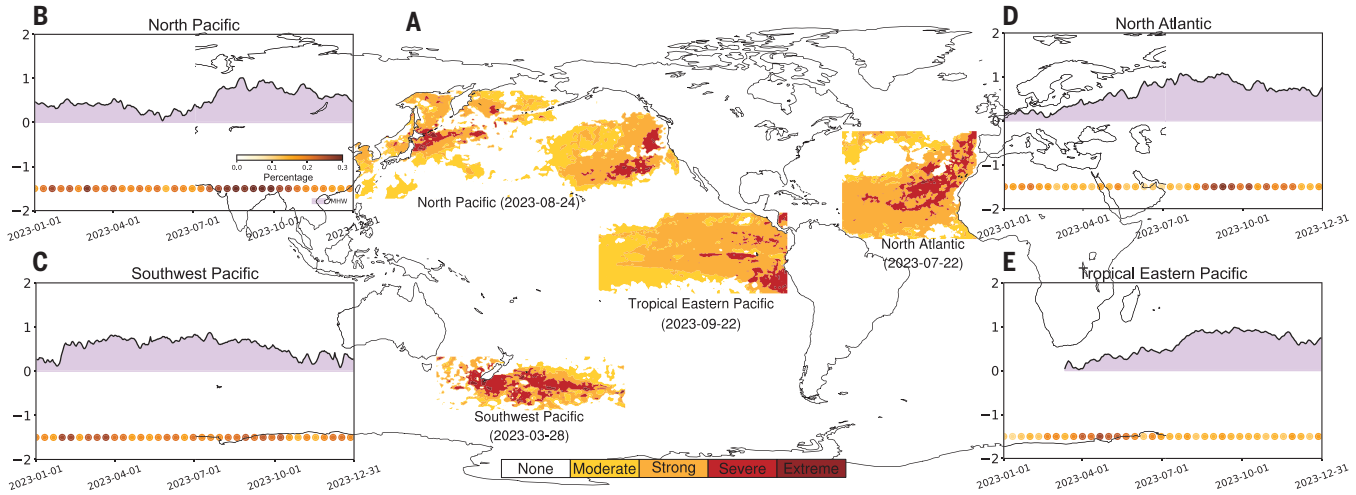


Fig. 2. Subset of representative MHW events in 2023. (**A**) Categories of MHW. (**B** to **E**) SST anomaly time series for each event (units, °C), spatially averaged over the area shown in (A). The intensity of the MHW is indicated with purple shading. Colored dots at the bottom of the plot represent the probability of MHW occurrence on a specific day over the period from 1982 to 2023. The probability is calculated for each day in the time series, and the dots are shown every 10 days for improved visualization and to avoid overplotting.

the major and persistent warming of the North Atlantic, which far exceeds the thermal tolerance of many marine organisms (*52*).

In the Southwest Pacific, an MHW event began on 17 November 2022 and reached peak intensity in March 2023 (fig. S6D). Historically, MHWs in this region are most frequent in January and March, with probabilities exceeding 30% (refer to the temporal occurrence probability, Fig. 2C). The 2023 MHW in the Southwest Pacific Ocean lasted for 410 days, with 26.7% of the event classified as severe (category III) and 44.7% as strong (category II), reaching a maximum mean intensity of 1.3°C (fig. S6D). Unlike the 2015–2016 and 2017–2018 Tasman Sea MHWs—which lasted 251 and 221 days, respectively (*53*, *54*)—the 2023 MHW not only surpassed these events in duration but also extended substantially eastward beyond the Tasman Sea, covering a much broader area, including regions east of New Zealand.

In the Tropical Eastern Pacific, an MHW event began forming on 12 March 2023 and intensified alongside the development of an Eastern Pacific–type El Niño event developing in June, ultimately peaking on 22 September 2023 (Fig. 2E). This event was classified as severe (category III) for 9.0% of its duration and strong (category II) for 58.0%, with a maximum mean intensity of 1.63°C (fig. S6B). May has historically been the month with the highest frequency of MHWs in this region, influenced and regulated by the El Niño–Southern Oscillation (ENSO) phenomenon. El Niño events typically weakened trade winds and reduced upwelling, amplifying sea surface warming and increasing MHW frequency (*38*). The 2023 event exemplified the pivotal role of El Niño–driven ocean-atmosphere interactions in shaping intensity and persistence in this climatically sensitive region.

Further, a warm anomaly in the North Pacific Ocean has sustained an uninterrupted MHW over the past 4 years (fig. S6C). This is a region of the ocean that has experienced higher warming than other basins since 2013 (*55*). The event was initiated on 20 May 2019 and peaked on 12 July 2022 with a maximum intensity of 1.51°C. The highest frequency of MHWs in this region occurs in August and September, aligning with a pronounced seasonal trend (Fig. 2B).

Mechanisms driving the 2023 MHWs in key regions

To investigate the mechanisms driving the unprecedented 2023 MHWs, we analyzed the high-resolution ECCO2 ocean reanalysis and the OISST satellite-derived dataset, complemented by a mixed-layer heat budget analysis to quantify the contributions of key physical processes (see "Ocean mixed layer heat budget" in the supplementary materials for details). The physical mechanisms sustaining these events into late 2023 are illustrated in Fig. 3, which explores both atmospheric and oceanic perspectives, elucidating the complex interactions contributing to the prolonged duration and cumulative intensity of the 2023 MHWs. The strong agreement between OISST and ECCO2 data [fig. S7; correlation coefficient (r) = 0.96, $P < 0.01$] supports the reliability of ECCO2 in replicating the MHW characteristics observed in OISST data, validating its utility for mechanistic analysis. The budget analysis decomposes the contributions of ocean temperature tendency, surface net heat flux, horizontal advection, and vertical mixing processes, providing region-specific insight into the physical drivers of the 2023 extremes.

The MHW in the North Atlantic Ocean preceded events in other regions, initiated by an anomalous atmospheric trough that weakened the prevailing surface easterlies and disrupted the North Atlantic High (Azores High), the region's dominant high-pressure system (fig. S8) (*56*, *57*). In 2023, the North Atlantic High intensity index (NAHI; see materials and methods) reached its lowest value in the past 42 years (fig. S8). The weakening of the Azores High reduced northerly winds, leading to a notable reduction in cloud cover (particularly low cloud cover, fig. S9) and a shallower mixed-layer depth (Fig. 3B). The substantially decreased cloud cover lowered regional albedo (*58*), increasing shortwave radiation that reaches the ocean and subsequently warming the upper ocean (fig. S10). Net surface heat flux emerged as the primary driver of warming (Fig. 3, A and C), with meridional advection (M.adv) playing an important role in the development and persistence of MHWs in the region (Figs. 3C and 4A). The combined effects of enhanced shortwave radiation and oceanic processes were critical in initiating and sustaining the North Atlantic MHW throughout 2023.

The MHW in the Southwest Pacific, which encompassed parts of the Tasman Sea and extended eastward, was primarily driven by a strong anticyclone that disrupted prevailing westerlies (fig. S11). Northeasterly winds on the eastern flank of the anticyclone weakened the westerlies, thereby enhancing meridional advection (Fig. 3F). A pronounced reduction in cloud cover and substantially increased shortwave radiation further amplified ocean warming, whereas other heat flux components (longwave radiation, sensible heat flux, and latent heat flux) showed minimal changes (fig. S10). This is consistent with the findings of Goessling *et al.* (*58*), which showed that a decrease in cloud cover and the associated reduction in albedo led to increased shortwave radiation absorption, contributing to the 2023 temperature anomalies. When

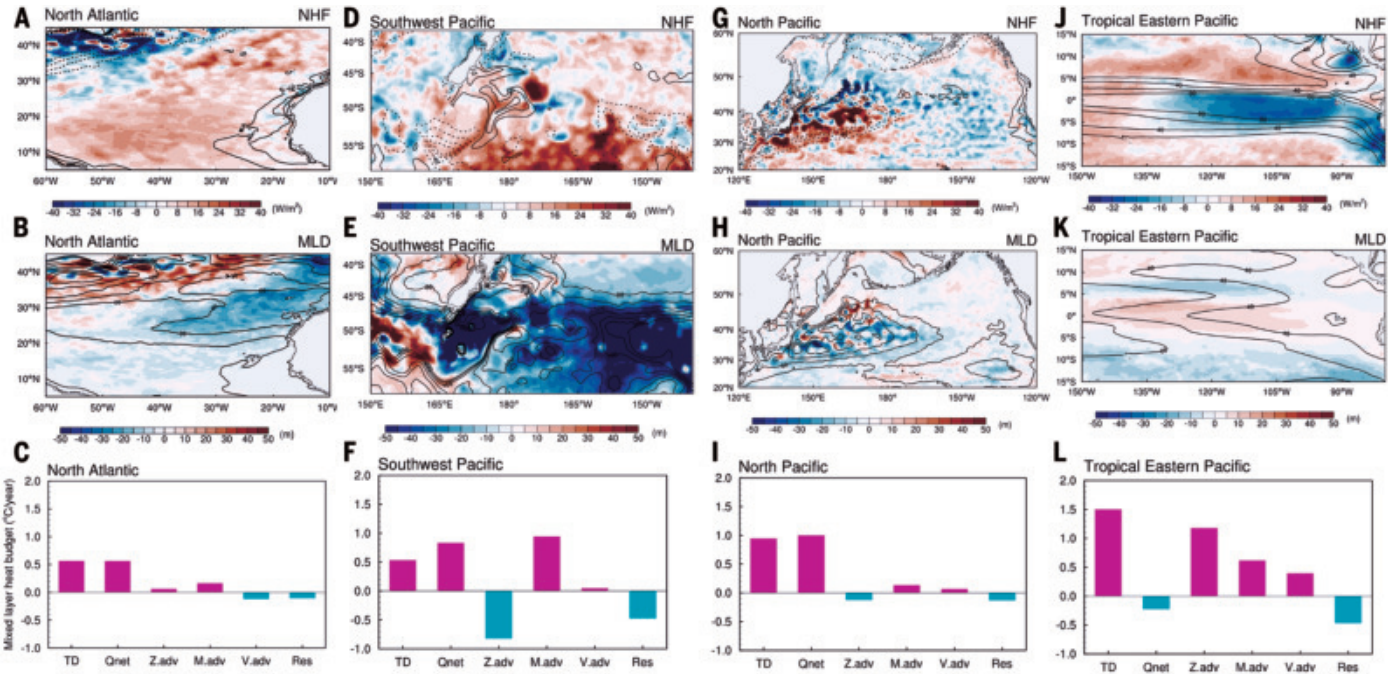


Fig. 3. The main contributing factors for MHWs in key regions. (**A** to **C**) Mean (contours) from January 1992 to December 2022 and 2023 anomalies (shading) are shown for net downward heat flux (NHF; units, W/m^2) (A), mixed-layer depth (MLD; units, m) (B), and ocean mixed-layer heat budget (units, °C/year) in North Atlantic (C). TD, mixed-layer ocean temperature tendency; Qnet, net surface heat flux term; Z.adv, zonal advection; M.adv, meridional advection; V.adv, vertical advection; Res, residual term. (**D** to **F**) Same as (A) to (C), but for the Southwest Pacific. (**G** to **I**) Same as (A) to (C), but for the North Pacific. (**J** to **L**) Same as (A) to (C), but for the Tropical Eastern Pacific.

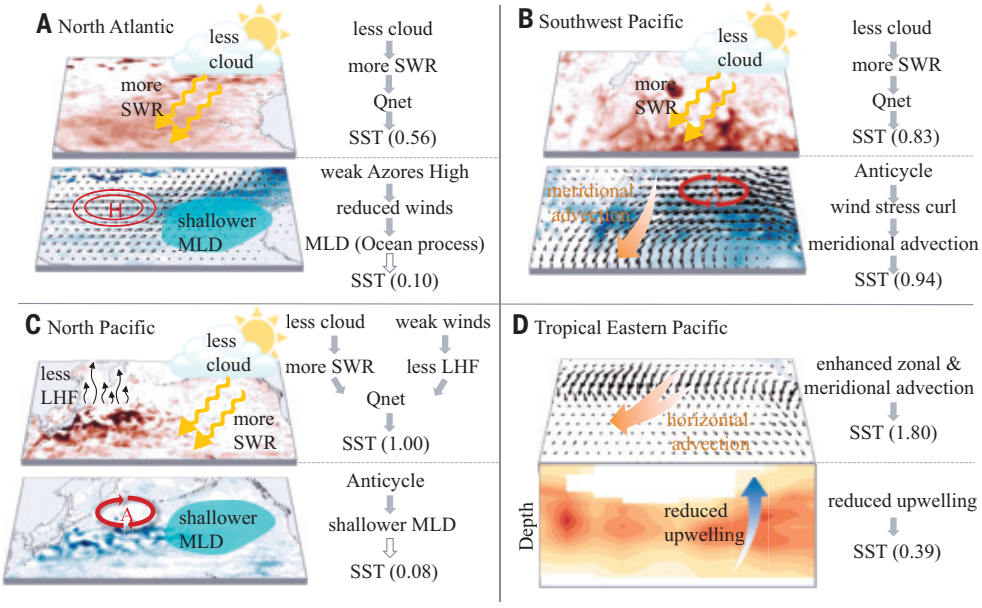


Fig. 4. Schematic representation of the mechanisms driving the 2023 MHWs. (**A** to **D**) Map shows key processes contributing to MHWs in the North Atlantic (A), Southwest Pacific (B), North Pacific (C), and Tropical Eastern Pacific (D). Variables SWR, MLD, and LHF represent shortwave radiation, mixed-layer depth, and latent heat flux, respectively. Qnet, horizontal advection, and vertical advection are defined by the mixed-layer heat budget equation (as detailed in the materials and methods; units, °C/year). The red circle with an "H" indicates a high-pressure system. Yellow arrows represent the increased shortwave radiation. Solid arrows indicate direct physical contributions to SST anomalies, while hollow arrows represent indirect or modulating influences. The red circle arrows labeled with "A" represent anticyclonic anomalies.

considering the ocean heat budget, warming was primarily driven by net downward heat flux, with meridional advection contributing positively and zonal advection (Z.adv) having a dampening effect (Figs. 3F and 4B). The MHW was not confined to the surface layers but extended its influence into the deeper ocean, reaching nearly 1000 m (fig. S12).

In the North Pacific, a notable increase in net downward heat flux (from the atmosphere to the ocean) over the shallower mixed layer across much of the region was the primary driver of the observed ocean warming in 2023. Reduced latent heat loss, and increased shortwave radiation due to changes in cloud cover and atmospheric conditions further amplified warming. Our analysis reveals that ~1.0°C of the average warming in 2023 was attributable to net surface heat flux, indicating its dominant role (Figs. 3I and 4C). Horizontal and vertical advection combined accounted for just 0.08°C of the total warming, suggesting a secondary role of oceanic advection. These results highlight the crucial role of atmospheric processes, particularly changes in shortwave radiation and latent heat flux, in shaping the North Pacific's thermal anomalies.

More than two-thirds of MHWs in the Tropical Eastern Pacific Ocean are historically associated with the ENSO (*38*). The positive correlation coefficient between NINO3.4 and MHWs in the Tropical Eastern Pacific is significant ($r = 0.64$, $P < 0.01$; fig. S13), underscoring the substantial influence of the El Niño phase on the occurrence of MHWs. As mentioned above, a notable transition from La Niña to El Niño conditions contributed to eastward-spreading warm SST anomalies and a greatly deepened thermocline in 2023 (*59*, *60*). The onset of an El Niño phase—typically characterized by weakening trade winds and suppressed upwelling of cold, nutrient-rich waters—is pivotal in driving the warming of surface waters in the Tropical Eastern Pacific (*61*). Our mixed-layer heat budget analysis reveals that these events were driven by a complex interplay of three-dimensional oceanic advections—encompassing both horizontal and vertical components—which played a crucial role in the overall warming of the region (Fig. 3L). These oceanic advections, in concert with the atmospheric changes associated with the El Niño phase, substantially contributed to the observed SST increase, further exacerbating the intensity and duration of MHWs in the Tropical Eastern Pacific. This interplay between atmospheric and oceanic processes underscored the intricate mechanisms behind MHWs in this region, with ENSO-driven changes in wind patterns and ocean circulation critically influencing the formation of these extreme temperature events.

The simultaneous occurrence of MHWs in 2023 may have been influenced by additional phenomena, including marine cloud brightening (*62*) and ocean current anomalies (*39*). Changes in atmospheric aerosol concentrations, driven by air quality measures, can affect cloud cover, which in turn affects radiative forcing and induces variations in ocean temperatures (*63*). Large-scale climate modes beyond ENSO, including Pacific Decadal Oscillation and Indian Ocean Dipole, are also important contributors to the modulation of oceanic and atmospheric teleconnections, which can alter surface winds, thermocline depth, and upwelling intensity (*38*). In addition, regional phenomena, such as boundary current shifts, mesoscale eddies, and vertical mixing suppression, further contribute to localized warming (*64*). The potential factors are varied and interconnected, underscoring the complexity of processes that drive MHW formation and their amplification under changing climate conditions.

Discussion and Summary

MHWs in 2023 exhibited record-breaking intensity, duration, and spatial extent across key oceanic regions. The North Atlantic MHW, lasting 525 days, revealed the scale of persistent ocean warming, whereas the Southwest Pacific MHW surpassed previous records with its extensive spatial coverage and prolonged persistence. In the Tropical Eastern Pacific, MHWs peaked at 1.63°C during El Niño development, and the North Pacific sustained an ongoing anomaly over 4 years. These variations point to the critical need for deeper investigation into the mechanisms initiating and sustaining the MHWs in each region to better comprehend their dynamics and impacts. Our mixed-layer heat budget analysis reveals region-specific drivers of the 2023 MHWs. In the North Atlantic, reduced cloud cover increased net downward shortwave flux, aligning with previous studies (*58*), and a weakened Azores High led to reduced wind anomalies, causing a shallower mixed-layer depth (Fig. 4A). The MHWs in this region were primarily driven by the combined effects of enhanced net downward shortwave flux (contributing nearly 0.56°C/year, Fig. 4A) and horizontal advection (adding ~0.10°C/year to SST warming). In the Southwest Pacific, notable reductions in low cloud cover increased net solar radiation (0.83°C/year), and wind stress curl associated with lower-level anticyclonic anomalies enhanced meridional advection, contributing nearly 0.94°C/year (Fig. 4B).

In the North Pacific, intensified shortwave flux and reduced latent heat loss drove persistent warming, with the net heat flux term accounting for 1.0°C/year of the SST increase, and vertical advection contributed an additional 0.10°C/year (Fig. 4C). Meanwhile, in the Tropical Eastern Pacific, MHWs were linked to ENSO-driven zonal and meridional (horizontal advection, 1.80°C/year), and vertical advection (0.39°C/year, Fig. 4D). These findings, summarized in Fig. 4, illustrate the diverse physical processes initiating and sustaining MHWs.

MHWs have profound implications for the Earth system, including ecological, social, and economic aspects. Ocean warming contributes to global sea level changes, with ~50% of this change attributed to the thermal expansion of seawater (*65*), thereby posing severe threats to low-lying islands and coastal zones through flooding (*8*). Crucially, the extraordinary MHWs of 2023 may represent a major shift in oceanic and atmospheric conditions, potentially indicating an early signal of a tipping point in Earth's climate system (*25*). These events can stress ecosystems beyond recovery thresholds, potentially triggering coral reef collapse, reducing species richness, increasing mortality rates (*9*, *66*), and causing redistribution of fish species (*67*–*70*), which has already led to the decline of key fisheries, such as the Pacific cod fishery (*68*, *71*).

For example, intensifying MHWs may influence marine phytoplankton biomass (*72*–*74*), particularly affecting ocean stratification and leading to variable responses in chlorophyll a concentrations across different latitudes during such events (*75*). MHWs can also affect fisheries by altering marine ecology, especially when anomalous warming is persistent (*8*, *76*, *77*). Persistent anomalous warming can disrupt marine ecosystems, as exemplified by the infamous "Blob" MHW in the Northeast Pacific (2013–2015). This event coincided with a substantial reduction in net primary productivity, decreased dissolved oxygen levels, increased acidification (*68*, *69*), and observable shifts in the distribution of marine organisms (*78*, *79*).

Other impacts include the destabilization of polar ice sheets (*80*) and the weakening of the Atlantic Meridional Overturning Circulation (*39*, *81*). Such changes amplify climate feedback, accelerating global warming and threatening foundational ecosystems such as kelp forests and seagrass meadows (*10*, *23*). The interaction between MHWs and tipping elements highlights the urgent need for enhanced Earth observation systems to monitor ocean physics and biogeochemistry, detect early warning signals, and improve predictive models (*25*, *26*, *82*).

Given that 89% of Earth's long-term energy imbalance is stored in the ocean, further research is essential to unravel the ocean's role in driving these unprecedented MHWs (*83*). Addressing this emerging type of climate extreme requires proactive adaptation strategies, enhanced predictive capabilities, and accelerated efforts to mitigate the growing impacts of climate change (*64*, *84*). Safeguarding ecological balance, economic stability, and human livelihoods in the face of intensifying MHWs must remain a global priority (*85*–*87*).

REFERENCES AND NOTES

1. N. Le Grix, J. Zscheischler, C. Laufkötter, C. S. Rousseaux, T. L. Frölicher, *Biogeosciences* **18**, 2119–2137 (2021).
2. E. C. J. Oliver *et al.*, *Nat. Commun.* **9**, 1324 (2018).
3. T. Wernberg *et al.*, *Science* **353**, 169–172 (2016).
4. S. G. Klein *et al.*, *Glob. Chang. Biol.* **28**, 1753–1765 (2022).
5. M. S. Thomsen *et al.*, *Front. Mar. Sci.* **6**, 84 (2019).
6. A. J. Pershing, K. E. Mills, A. Dayton, B. Franklin, B. R. C. Kennedy, *Oceanography (Wash. D.C.)* **31**, 152–161 (2018).
7. T. J. Goreau, R. L. Hayes, *Oxford Open Climate Change* **4**, kgae005 (2024).
8. W. Han *et al.*, *Nat. Commun.* **13**, 6410 (2022).
9. D. A. Smale *et al.*, *Nat. Clim. Chang.* **9**, 306–312 (2019).
10. N. J. Holbrook *et al.*, *Nat. Rev. Earth Environ.* **1**, 482–493 (2020).
11. J. C. Clements *et al.*, *Commun. Biol.* **8**, 735 (2025).
12. K. E. Smith *et al.*, *Science* **374**, eabj3593 (2021).
13. S. Berthou *et al.*, *Commun. Earth Environ.* **5**, 287 (2024).
14. Intergovernmental Panel on Climate Change*Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte *et al.*, Eds. (Cambridge Univ. Press, 2021).
15. L. Cheng *et al.*, *Adv. Atmos. Sci.* **41**, 1068–1082 (2024).
16. X. You, *Nature* **625**, 434–435 (2024).
17. Met Office, "Sea surface temperatures breaking records" (2023); https://blog.metoffice.gov.uk/2023/06/16/sea-surface-temperatures-breaking-records/.

18. ECMWF, "Record-breaking North Atlantic Ocean temperatures contribute to extreme marine heatwaves" (2023); https://climate.copernicus.eu/record-breaking-north-atlantic-ocean-temperatures-contribute-extreme-marine-heatwaves.
19. W. J. Ripple *et al.*, *Bioscience* **73**, 841–850 (2023).
20. O. Hoegh-Guldberg *et al.*, *Science* **382**, 1238–1240 (2023).
21. D. K. Kersting, C. García-Quintanilla, N. Quintano, I. Estensoro, M. D. M. Ortega-Villaizan, *Mediterr. Mar. Sci.* **25**, 578–585 (2024).
22. A. Nikolaou, E. Papadimitriou, E. Kiourani, S. Katsanevakis, *Mediterr. Mar. Sci.* **25**, 747–752 (2024).
23. K. E. Smith *et al.*, *Nat. Clim. Chang.* **15**, 231–235 (2025).
24. C. Laufkötter, J. Zscheischler, T. L. Frölicher, *Science* **369**, 1621–1625 (2020).
25. T. M. Lenton *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1786–1793 (2008).
26. D. I. Armstrong McKay *et al.*, *Science* **377**, eabn7950 (2022).
27. World Meteorological Organization, "WMO confirms that 2023 smashes global temperature record" (2024); https://wmo.int/news/media-centre/wmo-confirms-2023-smashes-global-temperature-record.
28. L. Cheng *et al.*, *Nat. Rev. Earth Environ.* **5**, 232–234 (2024).
29. Mercator Ocean, "Marine heatwave bullet in 4 September 2023" (2023); https://www.mercator-ocean.eu/bulletin/marine-heatwaves-europe-september-05-2023/.
30. Copernicus Marine Service, "The 2023 Northern Hemisphere summer marks record-breaking oceanic events" (2023); https://marine.copernicus.eu/news/2023-northern-hemisphere-summer-record-breaking-oceanic-events.
31. World Meteorological Organization (WMO), "State of the Global Climate 2023" (WMO-No. 1347, WMO, 2024); https://library.wmo.int/records/item/68835-state-of-the-global-climate-2023.
32. T. Kuhlbrodt, R. Swaminathan, P. Ceppi, T. Wilder, *Bull. Am. Meteorol. Soc.* **105**, E474–E485 (2024).
33. B. Huang *et al.*, *Geophys. Res. Lett.* **51**, e2024GL108369 (2024).
34. H. Sato *et al.*, *Sci. Rep.* **14**, 16100 (2024).
35. J. A. Benthuysen, E. C. J. Oliver, K. Chen, T. Wernberg, *Front. Mar. Sci.* **7**, 147 (2020).
36. H. A. Scannell, A. J. Pershing, M. A. Alexander, A. C. Thomas, K. E. Mills, *Geophys. Res. Lett.* **43**, 2069–2076 (2016).
37. E. Di Lorenzo, N. Mantua, *Nat. Clim. Chang.* **6**, 1042–1047 (2016).
38. N. J. Holbrook *et al.*, *Nat. Commun.* **10**, 2624 (2019).
39. X. Ren, W. Liu, *Geophys. Res. Lett.* **48**, e2021GL095941 (2021).
40. A. J. Hobday *et al.*, *Prog. Oceanogr.* **141**, 227–238 (2016).
41. A. Simon *et al.*, *Front. Mar. Sci.* **9**, 892201 (2022).
42. A. Hobday *et al.*, *Oceanography* **31**, 162–173 (2018).
43. D. Menemenlis, *et al.*, *Mercator Ocean Q. Newsl.* **31**, 13–21 (2008).
44. R. W. Reynolds *et al.*, *J. Clim.* **20**, 5473–5496 (2007).
45. T. Qu, *J. Geophys. Res.* **108**, 3242 (2003).
46. T. Kataoka, T. Tozuka, T. Yamagata, *J. Geophys. Res. Oceans* **122**, 8913–8932 (2017).
47. S. Lin, B. Dong, S. Yang, *Geophys. Res. Lett.* **51**, e2023GL106437 (2024).
48. K. von Schuckmann *et al.*, *State of the Planet* **4-osr8**, 2 (2024).
49. N. T. Shears, M. M. Bowen, F. Thoral, *N. Z. J. Mar. Freshw. Res.* 1–12 (2024).
50. J. Morton, "Cyclone Gabrielle: How 'abnormally warm' oceans are powering up incoming system," *NewstalkZB*, 9 February 2023; https://www.newstalkzb.co.nz/news/emergency/cyclone-gabrielle-how-abnormally-warm-oceans-are-powering-up-incoming-system/.
51. S. Radfar *et al.*, *Earths Futur.* **12**, e2024EF004935 (2024).
52. K. E. Smith *et al.*, *Nat. Commun.* **15**, 5052 (2024).
53. E. C. J. Oliver *et al.*, *Nat. Commun.* **8**, 16101 (2017).
54. S. E. Perkins-Kirkpatrick *et al.*, *Bull. Am. Meteorol. Soc.* **100**, S105–S110 (2019).
55. Z. Z. Hu, M. J. McPhaden, B. Huang, J. Zhu, Y. Liu, *Nat. Clim. Chang.* **14**, 929–931 (2024).
56. D. F. Schmidt, K. M. Grise, *J. Geophys. Res. Atmos.* **124**, 11188–11204 (2019).
57. R. W. Schlegel, E. C. J. Oliver, K. Chen, *Front. Mar. Sci.* **8**, 627970 (2021).
58. H. F. Goessling, T. Rackow, T. Jung, *Science* **387**, 68–73 (2025).
59. N. Jiang *et al.*, *Sci. Rep.* **14**, 2521 (2024).
60. Q. Peng, S. P. Xie, G. A. Passalacqua, A. Miyamoto, C. Deser, *Sci. Adv.* **10**, eadk8646 (2024).
61. K. Liu, K. Xu, C. Zhu, B. Liu, *J. Clim.* **35**, 877–893 (2022).
62. H. Hirasawa, D. Hingmire, H. Singh, P. J. Rasch, P. Mitra, *Geophys. Res. Lett.* **50**, e2023GL104314 (2023).
63. Y. Chen *et al.*, *Nat. Geosci.* **17**, 404–410 (2024).
64. A. Capotondi *et al.*, *Commun. Earth Environ.* **5**, 701 (2024).
65. C. Yue, S. Jevrejeva, Y. Qu, L. Zhao, J. C. Moore, *NPJ Clim. Atmos. Sci.* **6**, 135 (2023).
66. W. W. Cheung, R. Watson, D. Pauly, *Nature* **497**, 365–368 (2013).
67. K. Mills *et al.*, *Oceanography (Wash. D.C.)* **26**, 191–195 (2013).
68. M. J. Peterson Williams, B. Robbins Gisclair, E. Cerny-Chipman, M. LeVine, T. Peterson, *ICES J. Mar. Sci.* **79**, 573–583 (2022).
69. M. L. Pinsky, M. Fogarty, *Clim. Change* **115**, 883–891 (2012).
70. J. Pearce, N. Balcom, *J. Shellfish Res.* **24**, 691–697 (2005).
71. W. W. L. Cheung, T. L. Frölicher, *Sci. Rep.* **10**, 6678 (2020).
72. K. M. Noh, H.-G. Lim, J.-S. Kug, *Environ. Res. Lett.* **17**, 064034 (2022).
73. K. E. Smith *et al.*, *Annu. Rev. Mar. Sci.* **15**, 119–145 (2023).
74. M. S. Krishnapriya, H. Varikoden, P. Anjaneyan, J. Kuttippurath, *Mar. Pollut. Bull.* **197**, 115783 (2023).
75. A. Chauhan *et al.*, *Front. Mar. Sci.* **10**, 1177571 (2023).
76. M. M. Silveira, J. M. Donelson, M. I. McCormick, H. Araujo-Silva, A. C. Luchiari, *PeerJ* **11**, e15729 (2023).
77. F. Guzman, "Thousands of dead fish washed ashore beaches along the Texas Gulf Coast. Here's why," *USA Today*, 12 June 2023; https://www.usatoday.com/story/news/nation/2023/06/12/texas-gulf-coast-dead-fish/70312349007/.
78. T. Wernberg *et al.*, *Nat. Clim. Chang.* **3**, 78–82 (2012).
79. M. G. Jacox, M. A. Alexander, S. J. Bograd, J. D. Scott, *Nature* **584**, 82–86 (2020).
80. I. Joughin, R. Alley, *Nat. Geosci.* **4**, 506–513 (2011).
81. P. Ditlevsen, S. Ditlevsen, *Nat. Commun.* **14**, 4254 (2023).
82. K. von Schuckmann *et al.*, *Earth Syst. Sci. Data* **15**, 1675–1709 (2023).
83. T. Dong *et al.*, *Clim. Dyn.* **62**, 5593–5604 (2024).
84. A. J. Hobday *et al.*, *Nature* **621**, 38–41 (2023).
85. D. G. E. Gomes *et al.*, *Nat. Commun.* **15**, 1988 (2024).
86. N. Gruber, P. W. Boyd, T. L. Frölicher, M. Vogt, *Nature* **600**, 395–407 (2021).
87. S. C. Mogen *et al.*, *Geophys. Res. Lett.* **49**, e2021GL096938 (2022).
88. T. Dong, Z. Zeng, Code and data for "Record-breaking 2023 marine heatwaves," figshare (2025); https://doi.org/10.6084/m9.figshare.28778297.v1.

ACKNOWLEDGMENTS

This work was supported by the Center for Computational Science and Engineering at Southern University of Science and Technology and the High-Performance Computing Centers at Eastern Institute of Technology in Ningbo. We thank the principal researchers and teams for all the datasets and data-analysis methods used in the study. D.C. is supported by Tsinghua University (100008001). We appreciate the suggestions from S. Xie at Scripps Institution of Oceanography, UCSD. We acknowledge the anonymous reviewers for their detailed and helpful comments on the original manuscript. **Funding:** This study was supported by the startup fund provided by Shenzhen Science and Technology Project for Sustainable Development in Special Innovation (KCXFZ20230731093403008), Guangdong Basic and Applied Basic Research Foundation (no. 2022A1515240070), the National Natural Science Foundation of China (grants no. 42071022 and 42371026), China Meteorological Administration Climate Change Special Program (grant QBZ202316), and the startup and high-level special funds provided by the Southern University of Science and Technology (29/Y01296602, 29/Y01296122, 29/Y01296222, G030290001 and G03050K001), and the Shenzhen Key Laboratory of Precision Measurement and Early Warning Technology for Urban Environmental Health Risks (ZDSYS20220606100604008), and the funding agencies of Zhejiang Province and Ningbo Municipality through the program "Novel Technologies for Joint Pollution Reduction and Carbon Sequestration". **Author contributions:** T.D.: methodology, data curation, formal analysis, writing – original draft, and writing – review & editing. Z.Z.: conceptualization, supervision, methodology, formal analysis, writing – original draft, writing – review & editing, project administration, and funding acquisition. D.W. and Y.C.: formal analysis, project administration, and funding acquisition. M.P., L.L., S.Y., Y.J., S.Lu., S.Li., X.H., D. Zhao, A.D.Z., D.C., T.Z., L.Z.X.L., and D. Zhang: visualization and writing–review & editing. All authors discussed the concepts and reviewed the manuscript. The manuscript was drafted by T.D. and Z.Z. and edited by all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The daily SST data from the OISST are available from https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html. The Argo data were obtained from the China Argo Real-Time Data Centre (CARDC, https://www.argo.org.cn/index.php?m=content&c=index&a=lists&catid=27). The ERA5 dataset is available from https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels-monthly-means?tab=overview. The ECCO2 (Estimating the Circulation and Climate of the Ocean-Phase II) ocean reanalysis dataset is obtained from https://ecco.jpl.nasa.gov/drive/login?dest=L2RyaXZlL2ZpbGVzL0VDQ08yL2N1YmU5MI9sYXRsb25fcXVhcnRfOTBTOTBOL2RhaWx5. The programs used to generate all the results are NCL (6.6.2) and Python (3.9). Analysis scripts are available at https://doi.org/10.6084/m9.figshare.28778297.v1 (*88*). **License information:** Copyright © 2025 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adr0910
Materials and Methods; Figs. S1 to S14; References (*89*–*91*)

Submitted 14 June 2024; accepted 21 May 2025

10.1126/science.adr0910

PROTEIN DESIGN

De novo design and structure of a peptide-centric TCR mimic binding module

Karsten D. Householder[1,2], Xinyu Xiang[1], Kevin M. Jude[1], Arthur Deng[3], Matthias Obenaus[1], Yang Zhao[1], Steven C. Wilson[1]†, Xiaojing Chen[1], Nan Wang[1,4], K. Christopher Garcia[1,4,5]*

T cell receptor (TCR) mimics offer a promising platform for tumor-specific targeting of peptide–major histocompatibility complex (pMHC) in cancer immunotherapy. In this study, we designed a de novo α-helical TCR mimic (TCRm) specific for the NY-ESO-1 peptide presented by human leukocyte antigen (HLA)–A*02, achieving high on-target specificity with nanomolar affinity (dissociation constant K_d = 9.5 nM). The structure of the TCRm-pMHC complex at 2.05-Å resolution revealed a rigid TCR-like docking mode with an unusual degree of focus on the up-facing NY-ESO-1 side chains, suggesting the potential for reduced off-target reactivity. Indeed, a structure-informed in silico screen of 14,363 HLA-A*02 peptides correctly predicted two off-target peptides, yet our TCRm maintained peptide selectivity and cytotoxicity as a T cell engager. These results represent a path for precision targeting of tumor antigens with peptide-focused α-helical TCR mimics.

Peptide–major histocompatibility complex (pMHC) molecules on cancer cells are important targets for antigen-specific solid tumor immunotherapy (*1*, *2*). Class I MHC molecules present short peptide fragments (8 to 13 amino acids) from the intracellular proteome of a cell on its surface, providing a snapshot of the cell's internal state and whether it is healthy, cancerous, or infected (*3*, *4*). In humans, these molecules are encoded by the highly polymorphic human leukocyte antigen (HLA) genes, which determine which peptides are displayed to T cells. Cancer cells present mutant or self-peptides that are recognized by T cell receptors (TCRs) through a mechanism known as "MHC restriction" in which TCRs interact with both the MHC and the peptide composite surface (*5*). This recognition enables T cells to selectively detect and eliminate cancerous cells.

Therapeutic application of TCRs derived from the human immune system is hindered by their low affinity (*6*, *7*). To overcome these limitations, many groups have engineered TCR mimic (TCRm) antibodies: high-affinity molecules that replicate the MHC-restricted peptide recognition of TCRs (*8*–*14*). These mimics are valuable as cancer-specific probes and therapeutics. A recent example is tebentafusp-tebn, a bispecific T cell engager (TCE) approved in 2022 for uveal melanoma (*15*, *16*). Despite these advantages, the discovery of TCR mimics remains challenging. Unlike TCRs, antibodies have not been naturally selected for the optimal balance of peptide versus MHC binding selectivity. As a result, engineered TCRm antibodies can interact more prominently with the MHC helices than the peptide, leading to off-target toxicity (*17*–*21*). Although strategies such as "repurposing" TCRm antibodies have been described (*8*, *22*), fine-tuning these antibodies to have sufficient peptide specificity for use as therapies or probes is laborious and limiting the impact of this approach (*17*, *23*–*25*).

[1]Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, CA, USA. [2]Program in Immunology, Stanford University School of Medicine, Stanford, CA, USA. [3]Department of Computer Science, Stanford University, Stanford, CA, USA. [4]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA. [5]Department of Structural Biology, Stanford University School of Medicine, Stanford, CA, USA. ***Corresponding author. Email: kcgarcia@stanford.edu** †Present address: Xaira Therapeutics, Brisbane, CA, USA.

To address these obstacles, we sought to develop α-helical TCR mimics that could be rapidly programmed into cancer-specific probes or therapeutics. α-Helical bundles offer several advantages over antibodies: They are compact, modular, and present amino acid side chains on a rigid platform (*26*), reducing the design complexity observed with flexible complementarity-determining region (CDR) loops on antibodies. Because CDR loop flexibility contributes to TCR cross-reactivity (*27*–*29*), a more rigid scaffold could help minimize structural adaptation to alternative peptides. Additionally, state-of-the-art protein design tools, such as RFdiffusion and ProteinMPNN (*30*, *31*), favor α-helical folds and have demonstrated increased success rates for de novo protein design. We aimed to generate four-helix bundles, which are thermostable and known to mediate diverse protein-protein interactions across the natural proteome. Helical bundles, such as cytokines, are also ideal engineering scaffolds with proven robustness to amino acid substitutions along helical faces (*32*). By restricting our design space to this fold, we aimed to improve TCR mimic discovery rates by using RFdiffusion and ProteinMPNN.

Design of α-helical mini-TCR mimics with RFdiffusion and ProteinMPNN

As a first step, we generated 50 monomer folds (80 to 100 amino acids) with RFdiffusion and, by visual inspection, selected the most tightly packed helical bundles with a high likelihood of expression as recombinant proteins. We selected the fold best matched to this criterion (99 amino acids) to be used as a template scaffold for RFdiffusion fold-conditioning over the pMHC target. We selected the NY-ESO-1$_{157–165\ (C9V)}$ peptide presented on the HLA-A*02 allele as our model antigen, as it is a well-characterized tumor antigen (*2*, *23*, *33*) with structures in the Protein Data Bank (PDB) and robust experimental systems for downstream validation. Because crystal structures are not always available for a given pMHC antigen, we used AlphaFold (*34*) to predict the NY-ESO-1–HLA-A*02–β-2-microglobulin (β2M) structure, which together stabilize class I pMHC presentation. We specified the up-facing Met[4], Trp[5], Thr[7], and Gln[8] peptide residues as hotspots for design and masked scaffold loops to allow RFdiffusion to variably extend these regions during generation. We produced 100 distinct four-helix bundle folds docked over NY-ESO-1 HLA-A*02, each sampled for six amino acid sequences with ProteinMPNN, yielding 600 designs (data S1). AlphaFold2 then predicted each design docked with the target, and we ranked them by their interface predicted aligned error (iPAE) metric, as described in prior studies showing a correlation between iPAE values ≤10.0 and increased experimental success (*35*). iPAE measures AlphaFold's confidence in the relative positioning of amino acids across a protein-protein interface, with lower values indicating greater design confidence. Of the 100 scaffolds, 4 yielded at least one design under an iPAE threshold of 10.0 and, therefore, was considered a "hit" scaffold (Fig. 1A and data S1).

Our top scaffolds varied in helical bundle arrangement and size (99 to 129 amino acids) as well as in peptide versus MHC contact profiles. To expand sequence variant diversity, we generated 500 more sequences per hit with ProteinMPNN (2000 designs total) and scored each complex with AlphaFold2 (Fig. 1B and data S1). Scaffolds 1 and 2 yielded the highest rates of passing the iPAE threshold (132/500 and 318/500, respectively), and, upon visual inspection, we noticed desirable properties in scaffold 1 over the other scaffolds. Notably, it adopted a centered, TCR-like docking mode over the peptide-MHC target, favoring peptide-specific contacts. Moreover, all four peptide hotspots were predicted contacts in several of the top designs. Scaffold 1 appeared to best meet our design criteria, and we proceeded to screen its lowest iPAE-scoring binders.

Experimental validation of de novo TCR mimics specific to NY-ESO-1 HLA-A*02

The top five designs were displayed as single clones on yeast cells and stained with 100 nM NY-ESO-1 or MART-1 HLA-A*02 tetramer, with MART-1 serving as the negative control (Fig. 1C). We found that two

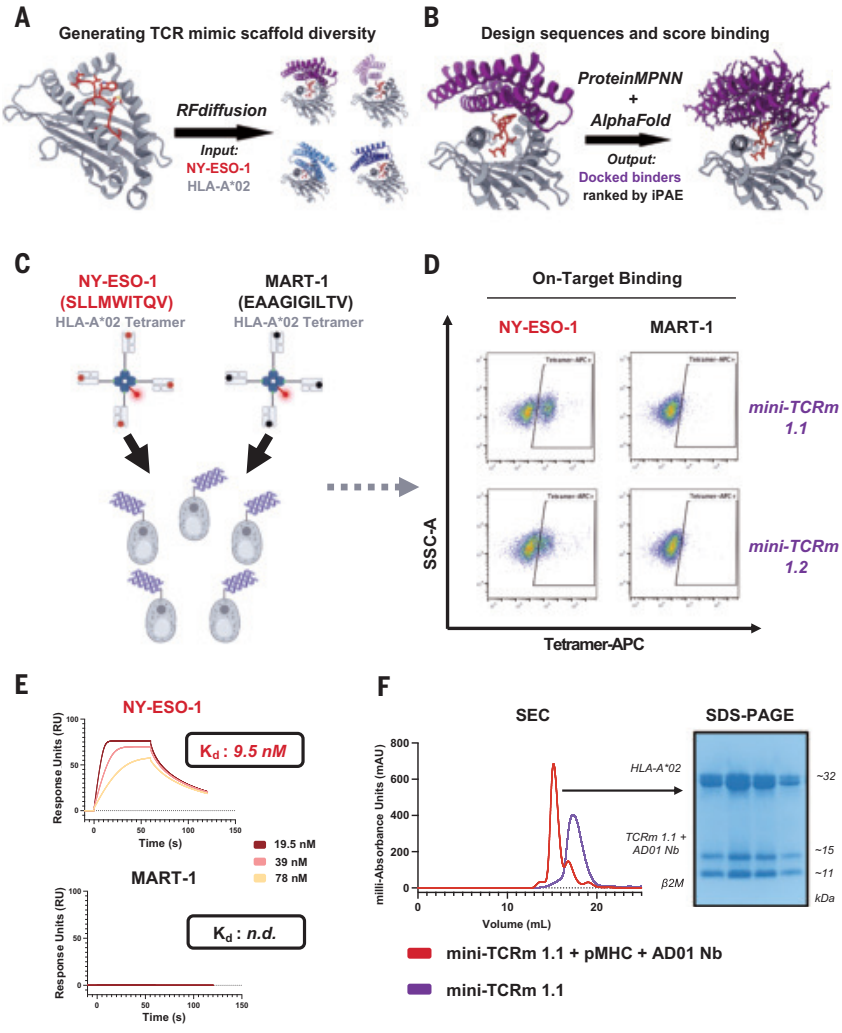


Fig. 1. Design and experimental validation of de novo mini-TCR mimics. (**A**) Schematic of RFdiffusion fold-conditioning pipeline generating four hit scaffolds with the NY-ESO-1 HLA-A*02 input (red and gray) out of 100 distinct folds sampled. (**B**) Schematic of pipeline using ProteinMPNN to design 500 sequences for scaffold 1 (purple) and predict likely binders with AlphaFold2. (**C**) Yeast display screening of top de novo designs with NY-ESO-1 (red) vs. MART-1 HLA-A*02 (black) tetramers. (**D**) Two of five evaluated binders staining NY-ESO-1 but not MART-1 HLA-A*02 tetramer by flow cytometry (see fig. S4A for gating strategy) (two independent experiments). (**E**) SPR analysis of mini–TCRm 1.1 analyte with immobilized NY-ESO-1 HLA-A*02 (red) versus MART-1 HLA-A*02 (black). Dissociation constants (K_d) indicated on sensograms. n.d., not determined. (**F**) SEC profiles of mini–TCRm 1.1 (purple) versus the coeluting complex (red) of mini–TCRm 1.1 with NY-ESO-1 HLA-A*02 and AD01 nanobody. SDS–polyacrylamide gel electrophoresis (PAGE) gel showing four fractions of the coeluting complex pooled for crystal screens.

out of five designs were stained by the NY-ESO-1 tetramer but not by MART-1 (Fig. 1D). We selected the strongest binder based on the highest mean fluorescence intensity (MFI) staining for NY-ESO-1 tetramer and designated it mini–TCRm 1.1. This binder was expressed as a recombinant protein in BL21 *Escherichia coli* and purified by Ni-NTA chromatography and size-exclusion chromatography (SEC). The recombinant protein was soluble with high yields and eluted as a single SEC peak (Fig. 1F, purple). Surface plasmon resonance (SPR) analysis revealed a K_d of 9.5 nM for NY-ESO-1 HLA-A*02 but no detectable affinity for MART-1 HLA-A*02 (Fig. 1E). This indicated that mini–TCRm 1.1 was peptide specific for NY-ESO-1, with undetectable trace affinity for the same MHC presenting a different peptide.

For structural studies, we complexed mini–TCRm 1.1 with refolded NY-ESO-1 HLA-A*02 and the AD01 anti-β2M nanobody, which served as a crystallization chaperone. The complex coeluted through SEC (Fig. 1F, red) and readily crystallized during screening.

High-resolution crystal structure reveals peptide-specific interactions

We determined the crystal structure of the complex at 2.05-Å resolution (Fig. 2, A and B, and table S1). The binder adopted a diagonal TCR-like docking footprint at a 40° angle relative to the peptide-MHC groove (Fig. 2C), burying 1216-Å^2 of surface area with 31% from peptide-specific contacts (table S2). Buried surface area (BSA) refers to the portion of the protein interface that becomes inaccessible to solvent upon binding, indicating the extent of molecular interaction between two proteins. The mini-TCR mimic used 14 residues within the groove of the A2 and A3 helices to form a concentrated network of hydrophobic and polar peptide contacts. The peptide-contacting side chains bent inward, creating a shape-complementary shell around the peptide side chains. In particular, the Met^4-Trp^5 peptide motif bulged prominently above the MHC groove and inserted into two hydrophobic pockets between the A2 and A3 helices (Fig. 2, B and D), stabilized by three hydrogen bonds (Fig. 2E). The binder also distributed hydrophobic contacts across both helices of HLA-A*02. These were supplemented by five hydrogen bonds and one salt bridge to the MHC, helping to orient the binder precisely over the peptide (Fig. 2F). Three of these hydrogen bonds (formed by Asn^{54}, Asn^{80}, and Arg^{57} of the binder) interact with residues inside the HLA-A*02 groove around the peptide's Trp^5 residue.

Comparison of the crystal structure and AlphaFold model revealed general agreement in docking orientation and hydrophobic pocket formation around the Met^4-Trp^5 bulge (root mean square deviation of 0.41 Å for 128 Cα atoms) (fig. S1A), albeit with substantial deviations in detail (table S3). Three differences affected peptide-specific contacts. First, AlphaFold predicted a hydrogen bond between Glu^{76} of the binder and Gln^8 of NY-ESO-1, which was absent in the crystal structure (fig. S1B). Second, AlphaFold failed to predict a hydrogen bond between the binder's Asn^{80} and the peptide's Thr^7, yet it was observed in the crystal structure (fig. S1C). Third, Arg^{57} of the binder made a hydrogen bond with the peptide's Gln^8 in the crystal structure, whereas AlphaFold predicted it to pair with the backbone of Ile^6 (fig. S1D). These pairwise bonding discrepancies in the predicted and actual structures underscore the value of structural validation in precisely refining de novo TCR mimic specificity.

Comparison with existing antibody TCR mimic and TCR

We also compared the structure of the mini-TCR mimic to an NY-ESO-1–specific Fab (3M4E5 Fab; PDB: 3HAE) and TCR (1G4 TCR; PDB: 2BNQ) (*23*, *33*). The mini-TCR mimic was compact, rigid, and less than half the height of the Fab or TCR (Fig. 3A). Unlike the Fab and TCR, which extended flexible loops into the MHC groove, our binder engaged the peptide's upward-facing side chains through a groove between its helices (Fig. 3B). Whereas the Fab and TCR used their CDR loops to create a large pocket simultaneously accommodating Met^4 and Trp^5 of the peptide, the designed binder used two shape-complementary pockets to capture them independently (Fig. 3C). Additionally, we compared binding footprints, identifying key contacts along the pMHC interface (Fig. 3B). The designed binder interacted with 19 MHC residues, compared with 14 for the Fab and 13 for the TCR. Nine residues on the MHC—Arg^{65}, Lys^{66}, Ala^{69}, Gln^{72}, Thr^{73}, and Val^{76} (α1 helix) as well as Lys^{146}, His^{151}, and Gln^{155} (α2 helix)—were shared across all three

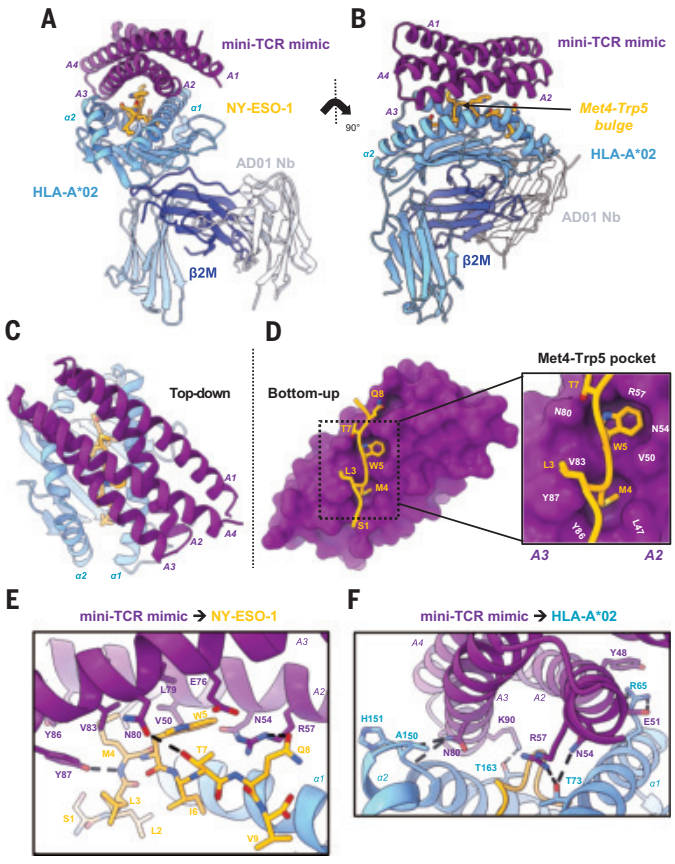


Fig. 2. High-resolution crystal structure reveals peptide-specific interactions. (**A**) Front view of mini-TCRm/pMHC complex (PDB ID 9MIN). A1 to A4 α helices of mini-TCRm (purple); α1 and α2 helices of HLA-A*02 and full chain (light blue); NY-ESO-1 peptide (gold); β2M (dark blue); AD01 nanobody (white). (**B**) Side view of complex with Met^4-Trp^5 bulge on the NY-ESO-1 peptide. (**C**) Top-down view of complex. (**D**) Bottom-up view of peptide-specific interactions with surface of mini-TCRm (purple). Box zooming in on NY-ESO-1's Met^4-Trp^5 motif (gold) fitting into mini-TCRm pockets formed between the A2 and A3 helices (key residues are labeled in white). (**E**) Hydrogen bonds (dashed lines) and key peptide-centric residues between mini-TCRm and NY-ESO-1 peptide. (**F**) Hydrogen bonds and salt bridges (dashed lines) between mini-TCRm and HLA-A*02. Single-letter abbreviations for the amino acid residues referenced throughout the figures are as follows: Q, Gln; T, Thr; W, Trp; L, Leu; M, Met; S, Ser; R, Arg; N, Asn; V, Val; Y, Tyr; E, Glu; I, Ile; H, His; A, Ala; K, Lys; F, Phe; D, Asp; F, Phe.

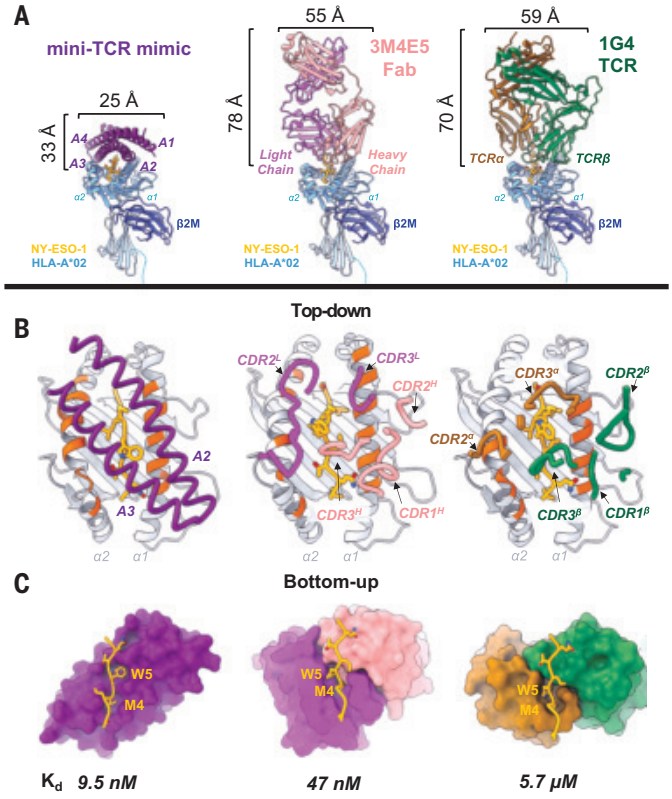


Fig. 3. Comparison with existing antibody TCR mimic and TCR. (**A**) Front-view comparison between mini-TCRm (purple), 3M4E5 Fab (magenta and salmon), and 1G4 TCR (brown and green) bound to NY-ESO-1 peptide (gold) and HLA-A*02 (light blue). Height and width of binders shown in angstroms (Å). (**B**) Top-down comparison of docking footprints of each binder with MHC contact residues (red orange) and NY-ESO-1 peptide side chains (gold). (**C**) Bottom-up comparison of interactions between the NY-ESO-1 peptide (gold) and binding pockets, with the Met^4-Trp^5 motif labeled. Dissociation constants (K_d) for each complex are shown at the bottom.

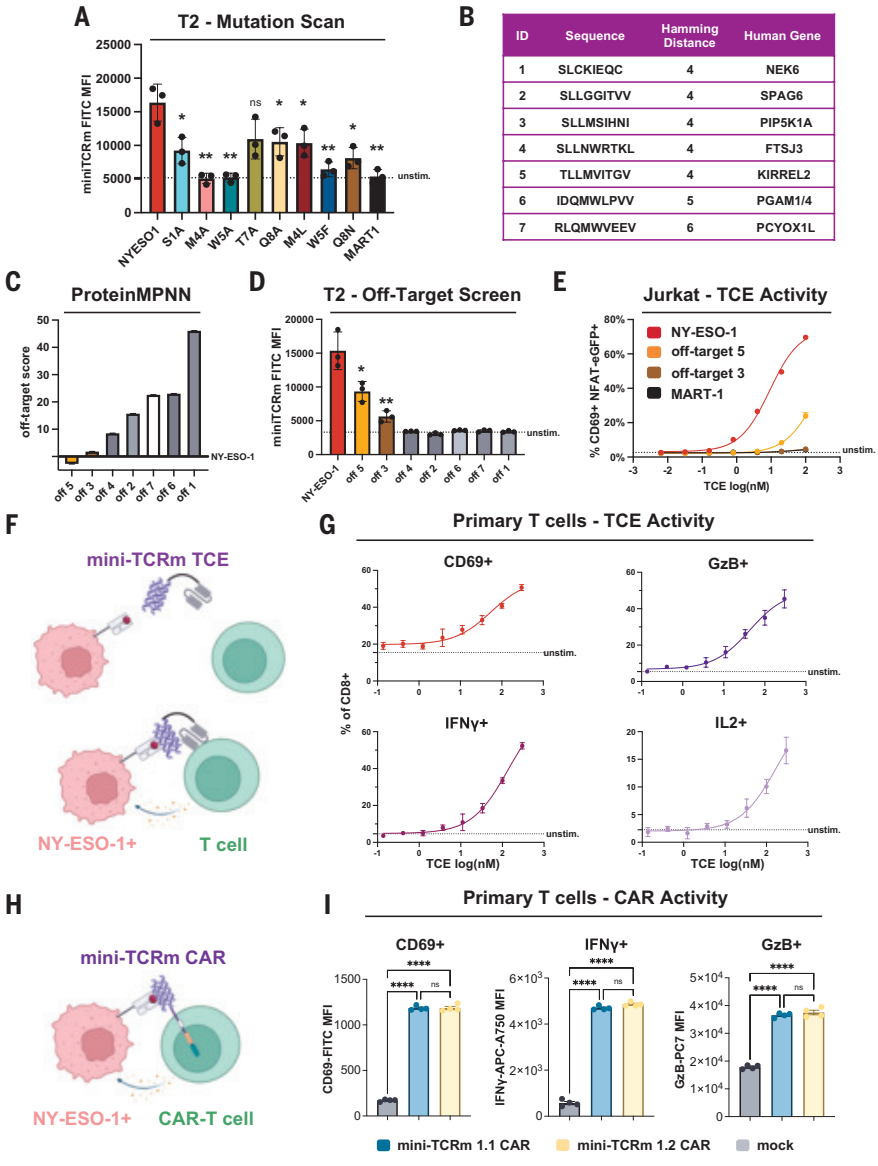
footprints and are classical MHC "anchor residues" used by the germline-encoded CDR loops of TCRs (*17*). Our mini-TCR mimic buried the largest total surface area but the smallest proportion of peptide surface relative to that of the Fab and TCR (table S2). Despite this, our mini-TCR mimic maintained peptide specificity and the highest binding affinity (Fig. 3C), suggesting that it formed focused peptide contacts while dispersing MHC interactions to stabilize binding. Furthermore, its rigid α-helical structure likely reduced the entropic penalty of binding by limiting conformational flexibility. Entropic penalty refers to the loss of freedom that occurs when a protein binds another protein, with smaller conformational rearrangements being more energetically favorable. By contrast, the Fab and TCR likely experienced greater entropic penalties with their flexible loops, which might contribute to less predictable cross-reactivity profiles.

Structure-guided identification of off-target peptides from the human proteome

Using our crystal structure as a template, we evaluated the specificity of our mini-TCR mimic, considering two potential applications. First, as a diagnostic tool used to probe cell surface antigen expression on tumors for patient stratification, where broad off-target assessment and rapid peptide evaluation would suffice. Second, for therapeutic development, where understanding cross-reactivity would inform the safety profile for clinical translation.

To confirm key peptide hotspots observed in the structure, we pulsed T2 cells with alanine-scanned NY-ESO-1 variants and stained with soluble mini-TCR mimic. Given its hydrophobic pockets, we hypothesized that other hydrophobic residues at peptide positions 4 and 5 could bind, so we included Leu^4 and Phe^5 mutants. Mutating Met^4 or Trp^5 to alanine abrogated binding, but Leu^4 or Phe^5 did not (Fig. 4A), confirming that the Met^4-Trp^5 bulge stabilized the complex, similar to the 3M4E5 single-chain variable fragment (scFv) (fig. S2). Thr^7 and Gln^8 alanine mutations had minimal effect when the Met^4-Trp^5 motif was present (Fig. 4A). Based on these findings, we anticipated that off-target peptides with the highest chemical similarity at NY-ESO-1$_{157–165\ (C9V)}$ peptide positions 1, 4, and 5 would be most likely to bind.

To test this, we conducted a Hamming distance search comparing the number of amino acid changes between the NY-ESO-1 peptide and 14,363 distinct HLA-A*02–presented nine–amino acid peptides detected by mass spectrometry from the MHC Motif Atlas (*36*). We restricted our search to this dataset, reasoning that MS detection provided a threshold for peptides found on the cell surface at sufficient quantities. After ranking the dataset (data S2), only 5 out of 14,363 peptides had the lowest Hamming distance of 4 mutations from NY-ESO-1 and were of human origin (Fig. 4B). All other peptides



ID	Sequence	Hamming Distance	Human Gene
1	SLCKIEQC	4	NEK6
2	SLLGGITVV	4	SPAG6
3	SLLMSIHNI	4	PIP5K1A
4	SLLNWRTKL	4	FTSJ3
5	TLLMVITGV	4	KIRREL2
6	IDQMWLPVV	5	PGAM1/4
7	RLQMWVEEV	6	PCYOX1L

Fig. 4. Mini-TCR mimic therapeutics exhibit peptide selectivity and cytotoxicity for NY-ESO-1 HLA-A*02. (**A**) Flow cytometry analysis of mini-TCRm–stained T2 cells pulsed with NY-ESO-1 peptide variants (see data S2 for sequences; see fig. S4B for gating strategy). MART-1 and no peptide were used as negative controls. Data points represent the geometric MFI averaged across technical replicates per experiment (bars represent mean ± SD; $n = 3$, $N = 3$). Statistical significance was calculated relative to NY-ESO-1 (*$P < 0.05$; **$P < 0.01$; ns, not significant; unpaired Student's t tests). (**B**) Table of candidate off-targets identified from the MHC Motif Atlas. (**C**) Ranking of candidate off-targets based on ProteinMPNN and the crystal structure (see data S2 for scores). (**D**) Flow cytometry of mini-TCRm–stained T2 cells pulsed with off-target candidates (see fig. S4B for gating strategy). Data points show the geometric MFI from technical replicates (mean ± SD; $n = 3$, $N = 2$). Statistical significance was calculated relative to NY-ESO-1 (*$P < 0.05$; **$P < 0.01$; unpaired Student's t tests). (**E**) Jurkat CD69$^+$ NFAT-eGFP$^+$ reporter activation following coculture with T2 pulsed with peptides and treated with mini-TCRm T cell engagers (see fig. S4C for gating strategy). Dose is shown in log(nM) (dots represent mean ± SD; $n = 3$, $N = 2$). (**F**) Schematic of a soluble mini-TCRm TCE (purple) activating T cells against an NY-ESO-1$^+$ cell (salmon). (**G**) Cytotoxicity dose response of mini-TCRm TCE in primary T cell cocultures with NY-ESO-1-pulsed T2 cells (see fig. S4D for gating strategy). Data represent the percentage of cells positive for the cytotoxicity marker relative to negative controls (dots represent mean ± SD; $n = 3$; an independent experiment with lower peptide dosage yielded similar results, though was not counted in $N = 2$). IFN, interferon; IL, interleukin. (**H**) Schematic of a mini-TCRm CAR T cell (CAR, purple) targeting an NY-ESO-1$^+$ cell (salmon). (**I**) Cytotoxic marker expression in mini-TCRm CAR T cells cocultured with NY-ESO-1$^+$ A375 cells for 48 hours (see fig. S4E for gating strategy). Mock T cells were activated, untransduced controls. Data represent geometric MFI in CAR$^+$ population (bars represent mean ± SD; $n = 4$, $N = 2$). Statistical significance was calculated for all pairs [****$P < 0.0001$; ns, not significant; one-way analysis of variance followed by Tukey's post hoc test].

with 5 or greater Hamming distance were considered less likely to adopt NY-ESO-1's target conformation and chemistry. A second Hamming distance search, which focused only on the critical Met^4-Trp^5 motif, found two additional peptides with an exact motif match but overall Hamming distances of 5 and 6. Together, these seven peptides were selected for off-target analysis (Fig. 4B).

We explored whether ProteinMPNN could rank these peptides by their compatibility with our crystal structure. By using the NY-ESO-1 backbone as input and conditioning on the molecular context of the mini-TCR mimic and HLA-A*02, ProteinMPNN assigned to each peptide an "off-target score" derived from its likelihood to fit with structural constraints. This score was calculated by subtracting each candidate off-target's likelihood to fit our crystal structure from that of NY-ESO-1. We hypothesized that more negative off-target scores (higher likelihoods to fit the structure) would identify real off-target peptides. Indeed, off-targets 3 and 5 exhibited the lowest scores (Fig. 4C), closely resembling that of NY-ESO-1, and were therefore predicted to bind the mini-TCR mimic. To confirm these predictions, we pulsed all peptides onto T2 cells and stained with our mini-TCR mimic. Consistent with the in silico analysis, cells pulsed with off-targets 3 and 5 bound the mini-TCR mimic, though more weakly than NY-ESO-1 (Fig. 4D).

Off-target 3 originated from PIP5K1A, a ubiquitously expressed kinase, whereas off-target 5 derived from KIRREL2, a transmembrane cell adhesion molecule predominantly found in pancreatic β cells. AlphaFold-predicted structures of these peptides bound to HLA-A*02 revealed that the Met^4 peptide residue contact was retained, along with a hydrogen bond between the mini-TCRm's Tyr^{87} and the Met^4 backbone, as well as a small polar residue at peptide position 1 (fig. S3). Interactions at peptide positions 5 to 8 varied slightly, subtly altering total BSA (table S2), which aligned with relative staining MFI in our T2 assay.

Mini-TCR mimic therapeutics exhibit peptide selectivity and cytotoxicity for NY-ESO-1 HLA-A*02

We then determined whether our binder could function as a soluble peptide-specific TCE by expressing and purifying it from mammalian cells. We designed our bispecific TCE as a fusion between the anti-CD3ε scFv from blinatumomab (*37*) and our mini-TCRm, connected by a short 5–amino acid glycine-serine linker. Blinatumomab, the first Food and Drug Administration–approved bispecific TCE, activates T cells by forming a synthetic immune synapse upon binding both CD3ε and its target antigen. In our construct, the anti-CD3ε scFv recruits T cells to engage target cells displaying the peptide-MHC antigen (Fig. 4F).

We pulsed T2 cells with NY-ESO-1, MART-1, off-target 3, or off-target 5 and cocultured in a 1:1 ratio with a Jurkat reporter cell line that expresses enhanced green fluorescent protein

(eGFP) upon TCR signaling. A dilution series of our TCE was added, and T cell activation was quantified the next day by measuring the percentage of $CD69^+$ $eGFP^+$ Jurkats through flow cytometry. In the presence of NY-ESO-1 but not MART-1 peptide, we observed T cell activation with a median effective concentration (EC_{50}) of 9.1 nM, and nearly 70% of Jurkats were activated with 100 nM engager (Fig. 4E). No activation was observed for off-target 3, whereas off-target 5 exhibited ~25% Jurkat activation with 100 nM engager but with a log-shifted EC_{50} of 174 nM. Between 1 and 10 nM, TCE activity was retained against the NY-ESO-1 peptide but not the off-target peptides, establishing a selectivity threshold. Our TCE also induced cytotoxic function in a dose-dependent manner when added to a coculture of primary human T cells with NY-ESO-1–pulsed T2 cells (Fig. 4G). Additionally, both mini-TCRm hits 1.1 and 1.2 functioned in chimeric antigen receptor T cell (CAR T cell) format (Fig. 4, H and I), driving activation and cytotoxicity against the A375 human melanoma cell line, which endogenously expresses NY-ESO-1 and presents the NY-ESO-$1_{157–165}$ peptide on HLA-A*02.

Discussion

In this work, we have shown that an alternative structural scaffold to antibodies offers a versatile approach for engineering peptide-specific TCR mimic probes and immunotherapies. By leveraging deep learning tools (*30*, *31*), we accelerated the discovery of functional TCR mimics, generating 2600 designs in 30 hours on a single graphics processing unit and identifying two candidate binders in 1 to 2 weeks. These mini-TCR mimics have potential as cancer-specific diagnostics; for example, when a new tumor antigen is identified, they could rapidly confirm its presentation on cancer cells, expediting antigen discovery.

For therapeutic applications, however, further development is required to eliminate off-target interactions. Although de novo binders offer a more advanced starting point than traditional antibody screening, comprehensive off-target profiling remains essential before translation. A scalable approach could combine computational off-target predictions with high-throughput screening, such as yeast display of peptide-MHC complexes, to systematically prioritize and validate all potential cross-reactive peptides. If predictive approaches prove reliable, then future mini-TCR mimics could be preoptimized for specificity during the design process. This could make them particularly valuable for personalized cancer therapies, where tumor antigens from exome sequencing (*1*, *2*) could guide real-time engineering of therapeutics in weeks to months.

Immunogenicity, or the ability to elicit antidrug antibodies (ADAs), is another key consideration for de novo proteins. Although ADAs can impact efficacy, many biologics remain effective owing to dose adjustments or low neutralizing incidence. As synthetic proteins, mini-TCR mimics require empirical testing to assess the extent of immunogenicity. A proactive approach would involve computational epitope prediction and minimization of immunogenic T and B cell motifs or engineering nonimmunogenic human proteins.

Beyond specificity and immunogenicity, the α-helical platform offers engineering and design advantages over traditional antibodies. Mini-TCR mimics could be engineered into multispecific or multivalent biologics with greater structural precision in geometry or orientation than the conventional variable heavy/variable light heterodimer of antibodies. This capability could enable therapeutic formats that fine-tune antigen sensitivity or T cell function. Additionally, their structural rigidity enhances computational design accuracy, as the energetic consequences of mutations are more predictable than in flexible antibody loops (*26*–*29*). Although crystallographic validation was crucial in this study, we anticipate that continued advancements in computational modeling will progressively replace the need for structural validation, such that it may not be necessary in future discovery efforts.

REFERENCES AND NOTES

1. N. Xie *et al.*, *Signal Transduct. Target. Ther.* **8**, 9 (2023).
2. C. A. Klebanoff, S. S. Chandran, B. M. Baker, S. A. Quezada, A. Ribas, *Nat. Rev. Drug Discov.* **22**, 996–1017 (2023).
3. J. Rossjohn *et al.*, *Annu. Rev. Immunol.* **33**, 169–200 (2015).
4. P. Cresswell, A. L. Ackerman, A. Giodini, D. R. Peaper, P. A. Wearsch, *Immunol. Rev.* **207**, 145–157 (2005).
5. K. C. Garcia, E. J. Adams, *Cell* **122**, 333–336 (2005).
6. Y. Li *et al.*, *Nat. Biotechnol.* **23**, 349–354 (2005).
7. P. D. Holler *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5387–5392 (2000).
8. X. Yang *et al.*, *Nat. Biotechnol.* **41**, 932–943 (2023).
9. Y. Sun *et al.*, *Sci. Immunol.* **8**, eadj5792 (2023).
10. E. H.-C. Hsiue *et al.*, *Science* **371**, eabc8697 (2021).
11. T. Dao *et al.*, *Sci. Transl. Med.* **5**, 176ra33 (2013).
12. P. Chames *et al.*, *J. Immunol.* **169**, 1110–1118 (2002).
13. A. Y. Chang *et al.*, *J. Clin. Invest.* **127**, 2705–2718 (2017).
14. L. Dubrovsky *et al.*, *OncoImmunology* **5**, e1049803 (2015).
15. S. Howlett, T. J. Carter, H. M. Shaw, P. D. Nathan, *Ther. Adv. Med. Oncol.* **15**, 17588359231160140 (2023).
16. J. C. Hassel *et al.*, *N. Engl. J. Med.* **389**, 2256–2266 (2023).
17. C. J. Holland *et al.*, *J. Clin. Invest.* **130**, 2673–2688 (2020).
18. G. P. Linette *et al.*, *Blood* **122**, 863–871 (2013).
19. B. J. Cameron *et al.*, *Sci. Transl. Med.* **5**, 197ra103 (2013).
20. M. R. Parkhurst *et al.*, *Mol. Ther.* **19**, 620–626 (2011).
21. E. Baulu, C. Gardet, N. Chuvin, S. Depil, *Sci. Adv.* **9**, eadf3700 (2023).
22. A. D. Skora *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9967–9972 (2015).
23. G. Stewart-Jones *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5784–5788 (2009).
24. K. M. Wright *et al.*, *Nat. Commun.* **14**, 5063 (2023).
25. J. Harper *et al.*, *PLOS ONE* **13**, e0205491 (2018).
26. D. N. Woolfson, *J. Biol. Chem.* **299**, 104579 (2023).
27. K. C. Garcia *et al.*, *Science* **279**, 1166–1172 (1998).
28. C. M. Ayres, D. R. Scott, S. A. Corcelli, B. M. Baker, *Sci. Rep.* **6**, 25070 (2016).
29. M. E. Birnbaum *et al.*, *Cell* **157**, 1073–1087 (2014).
30. J. L. Watson *et al.*, *Nature* **620**, 1089–1100 (2023).
31. J. Dauparas *et al.*, *Science* **378**, 49–56 (2022).
32. R. A. Saxton, C. R. Glassman, K. C. Garcia, *Nat. Rev. Drug Discov.* **22**, 21–37 (2023).
33. J.-L. Chen *et al.*, *J. Exp. Med.* **201**, 1243–1255 (2005).
34. J. Jumper *et al.*, *Nature* **596**, 583–589 (2021).
35. N. R. Bennett *et al.*, *Nat. Commun.* **14**, 2625 (2023).
36. D. M. Tadros, S. Eggenschwiler, J. Racle, D. Gfeller, *Nucleic Acids Res.* **51** (D1), D428–D437 (2023).
37. D. Nagorsen, P. Kufer, P. A. Baeuerle, R. Bargou, *Pharmacol. Ther.* **136**, 334–342 (2012).
38. K. D. Householder *et al.*, Raw diffraction images for mini-TCRm/pMHC structure, SBGrid Data Bank (2024); https://doi.org/10.15785/SBGRID/1151.

ACKNOWLEDGMENTS

We thank all members of the Garcia Lab for helpful discussions. We thank J. Nix at the Advanced Light Source (ALS) for assistance with x-ray data collection. The ALS is a Department of Energy Office of Science User Facility under contract no. DE-AC02-05CH11231. The Berkeley Center for Structural Biology is supported in part by the Howard Hughes Medical Institute (HHMI). Figures 1, A to C, and 4, F and H, were made with BioRender. We also thank V. Mallajosyula for valuable support with protein biochemistry. **Funding:** National Institute of Health grant 5R01AI103867-08 (K.C.G.); National Cancer Institute grant OT2CA297242 (K.C.G.); Yosemite Fund (K.C.G.); Parker Institute for Cancer Immunotherapy (K.C.G.); CRI Irvington Postdoctoral Fellowship (Y.Z.). This work was delivered as part of the MATCHMAKERS team, supported by the Cancer Grand Challenges partnership financed by CRUK (CGCATF-2023/100006), the National Cancer Institute (OT2CA297242), and the Mark Foundation for Cancer Research (K.D.H., K.M.J., K.C.G.). K.C.G. is an investigator of the HHMI. **Author contributions:** K.D.H. and K.C.G. conceived the project and wrote the manuscript; K.D.H., X.X., K.M.J., A.D., M.O., Y.Z., and S.C.W. designed and performed experiments; K.D.H., X.X., K.M.J., S.C.W. determined crystal structures; and X.C. and N.W. provided experimental reagents and protocols. All authors edited and approved the manuscript. **Competing interests:** K.C.G. is a cofounder of 3T Biosciences and consults for Xaira Therapeutics. All other authors declare that they have no competing interests. **Data and materials availability:** The crystallographic model and integrated intensities have been deposited in the RCSB Protein Data Bank with accession code 9MIN. Raw diffraction images have been deposited in the SBGrid Data Bank (*38*). All other data needed to evaluate the conclusions are available in the manuscript and supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adv3813
Materials and Methods; Figs. S1 to S6; Table S1 to S3; References (*39*–*50*); Data S1 and S2

Submitted 16 December 2024; accepted 9 May 2025

10.1126/science.adv3813

PROTEIN DESIGN

De novo-designed pMHC binders facilitate T cell–mediated cytotoxicity toward cancer cells

Kristoffer Haurum Johansen[1]*†, Darian Stephan Wolff[2]†, Beatrice Scapolo[1,2]†, Monica L. Fernández-Quintero[3]†, Charlotte Risager Christensen[1,2], Johannes R. Loeffler[3], Esperanza Rivera-de-Torre[2], Max D. Overath[2], Kamilla Kjærgaard Munk[1], Oliver Morell[2], Marie Christine Viuff[1], Iñigo Lacunza[1], Alberte T. Damm Englund[2], Mathilde Due[2], Anant Gharpure[3], Stefano Forli[3], Carlos Rodriguez Pardo[1], Tripti Tamhane[1], Emma Qingjie Andersen[2], Kasper Haldrup Björnsson[2], Jordan Sylvester Fernandes[2], Lasse Frank Voss[1], Suthimon Thumtecho[2], Andrew B. Ward[3], Maria Ormhøj[1], Sine Reker Hadrup[1]*†, Timothy P. Jenkins[2]*†

The recognition of intracellular antigens by CD8$^+$ T cells through T cell receptors (TCRs) is central for adaptive immunity against infections and cancer. However, the identification of TCRs from patient material remains complex. We present a rapid de novo minibinder (miBd) design platform leveraging state-of-the-art generative models to engineer miBds targeting the cancer-associated peptide-bound major histocompatibility complex (pMHC) SLLMWITQC/HLA-A*02:01 (NY-ESO-1). Incorporating in silico cross-panning enabled computational prescreening of specificity, and molecular dynamics simulations allowed for improved predictability of in vitro success. We identified a high-affinity NY-ESO-1 binder and confirmed its structure using cryo–electron microscopy, which, when incorporated in a chimeric antigen receptor, induced killing of NY-ESO-1$^+$ melanoma cells. We further designed and validated binders to a neoantigen pMHC complex, RVTDESILSY/HLA-A*01:01, with unknown structure, demonstrating the potential for precision immunotherapy.

T cells are immune cells capable of responding to intracellular antigens such as those derived from infections or cancerous mutations. CD8$^+$ T cells recognize antigens through their T cell receptor (TCR), which binds peptides presented by major histocompatibility complex class I (pMHC-I) on all nucleated cells. MHC-I molecules, encoded by human leukocyte antigen (HLA) genes, adopt a globular structure composed of three extracellular domains, α1, α2, and α3, and a β2-microglobulin (β2m) subunit. The α1 and α2 domains form a peptide-binding groove accommodating short peptides (typically nine to 11 amino acids) anchored at distinct positions, mostly second and C-terminal residues, and displays them for recognition by TCRs (*1*).

Cancer cells display antigens on pMHC-I that can be recognized by tumor-reactive T cells. TCRs isolated from such T cells have been tested in the clinical setting, with several clinical trials revealing durable responses in patients across several cancer indications (*1*). One such tumor-associated antigen, NY-ESO-1, contains a T cell epitope SLLMWITQC [NY-ESO-1$^{(157\text{-}165)}$] presented in the context of HLA-A*02:01 (*2*). NY-ESO-1, also known as cancer/testis antigen 1B (CTAG1B), is expressed in a wide range of tumors, including melanoma, lung, ovarian, and breast cancers, but not in normal tissues (except for the testis). It is structurally well characterized, and its broad, cancer-specific expression pattern makes it an attractive cancer immunotherapy target.

The identification of TCRs for therapeutic use is driven by the selection of relevant T cell populations from patients (*3*, *4*) or through expansion of T cell populations from the naïve repertoire toward a given antigen (*5*). These approaches are laborious, technically challenging, and potentially limited or biased by the TCR repertoire in the donors used for selection. TCRs are inherently cross-reactive to confer wider protection against diseases, so most of them cross-react to several peptides (*6*). However, predicting the behavior and potential cross-reactivity of such TCRs to pMHCs present on healthy tissue, when applied in hosts other than the donors they are selected from, remains a challenge (*7*, *8*).

To circumvent the challenges of TCR identification and to facilitate the clinical application of immunotherapies based on pMHC-I recognition, strong interest has emerged in developing alternative strategies to target infected or malignant cells presenting intracellular proteins on the pMHC-I complex. TCR-like molecules, which include TCR-like antibodies, TCR-bispecific engagers, and single-chain TCRs (*1*), are promising approaches. Such TCR-like molecules can be incorporated into chimeric antigen receptors (CARs) to induce T cell–mediated cytotoxicity toward a cell expressing a given pMHC-I target (*9*). TCR-like antibodies are currently found in either immune or synthetic libraries screened by display or other technologies. This process can take months to years and often yields very low success rates. Although some success has been reported, these approaches often struggle to generate antibodies that bind pMHC-I with high affinity. Hybridoma technology is limited by low throughput and labor-intensive protocols, and it skews toward the targeted HLA isoform rather than the peptide specificity (*10*). Phage display panning is faster and more cost-effective than hybridoma technology, but the libraries still rely on large, naïve antibody repertoires (*11*). Alternative methods have been explored in silico and in vitro (*12*, *13*), but these are limited by a lack of knowledge about the specificity of individual TCRs and the pMHC-I complex to which they bind (*8*). Furthermore, establishing the specificity of high-affinity binders by exploration of cross-reactive binding presents a substantial challenge for current in vitro approaches (*14*).

Structural models such as TCRen (*15*) and DeepAIR (*16*) have predicted TCR-pMHC interaction with improved accuracy. However, advances in generative protein design provide additional opportunities for the in silico development of alternative TCR-like pMHC-I binders. Denoising diffusion models, in particular, have optimized machine learning–guided protein design because of their ability to generate protein-binder backbones that are optimized toward their target through stepwise diffusion. This has led to substantial improvements in experimental success rates.

Among various de novo protein design approaches, RFdiffusion is considered one of the most promising tools due to its impressive average success rate of 10% of designed binders presenting ultrahigh target affinity (<10 nM) (*17*). The design pipeline consists of three independently trained models: (i) the diffusion model, which de novo generates a highly accurate scaffold, followed by (ii) ProteinMPNN (*18*), and finally (iii) AlphaFold2 (AF2), which can predict and evaluate a protein structure by its sequence (*19*). For the evaluation and down-selection of minibinder (miBd) designs, the AF2 interface predicted aligned error (iPAE) with a value <10 Å and a predicted local distance difference test (pLLDT) >90, as suggested by Bennett *et al.* (*20*), is widely applied (*17*, *21*, *22*). Here, we describe the power of these tools for rapidly developing de novo miBds to pMHCs that can facilitate T cell–mediated cytotoxicity toward cancer cells.

Results

Design and in silico cross-panning of NY-ESO-1 miBds

We sought to design miBds against SLLMWITQC/HLA-A*02:01 using a published crystal structure (Protein Data Bank ID 2BNR) (*23*) as input to RFdiffusion. To guide the binder interaction toward the

[1]Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark. [2]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark. [3]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA. ***Corresponding author. Email: tpaje@dtu.dk (T.P.J.); krisha@dtu.dk (K.H.J.); sirha@dtu.dk (S.R.H.)** †These authors contributed equally to this work.

surface-accessible residues of the peptide and not just the MHC-I backbone, we specified the residues L3, M4, W5, and T7 with upward-facing side chains as contact points for the RFdiffusion de novo design. Initially, we generated amino acid backbones for the miBds through RFdiffusion-denoising trajectories and interpreted sequences using ProteinMPNN (Fig. 1A). These sequence designs were evaluated by AF2 and filtered based on an iPAE <12 Å. To diversify the panel of designs, only miBds originating from different backbones were chosen. Next, the preselected miBds were sequence diversified by resubmission to ProteinMPNN (100 sequences per candidate), resulting in a final set of 44 miBds surpassing an iPAE of 7 Å and a pLDDT of 92. We found that most of the miBds consisted of tightly packed bundles of four α helices orientated almost exclusively parallel to the peptide (fig. S1, A to G).

To drive in silico design further toward specific recognition of the peptide, we leveraged AF2 to generate binding scores (iPAE and pLDDT) for all 44 miBds against a panel of single and double mutants of the wild-type peptide SLLMWITQC (Fig. 1B). The peptide variants presented on HLA-A*02:01 were modeled using AF3 and chosen to evaluate whether a minimal change of one or two key positions with exposed residues (mainly M4, W5, or T7) could disrupt the binding interaction between the miBds and the pMHC (*24*). Thus, we included peptide substitutions with a subtle change of the local physical-chemical properties, e.g., M4A and W5Y, as well as substantial changes of the local physical properties, by introducing negatively charged residues in position 4 (e.g., M4E and M4D). Furthermore, a variant peptide, identified in a human proteome search and predicted to have a strong affinity to the HLA-A*02:01, was added to the panel (fig. S2A) (*25*, *26*). Finally, applying the iPAE cutoff (<10 Å) to define a potential binding interaction between miBd and a variant pMHC, we ranked the miBds by the number of such undesired binding events and proceeded with the 44 top candidates for in vitro assessment.

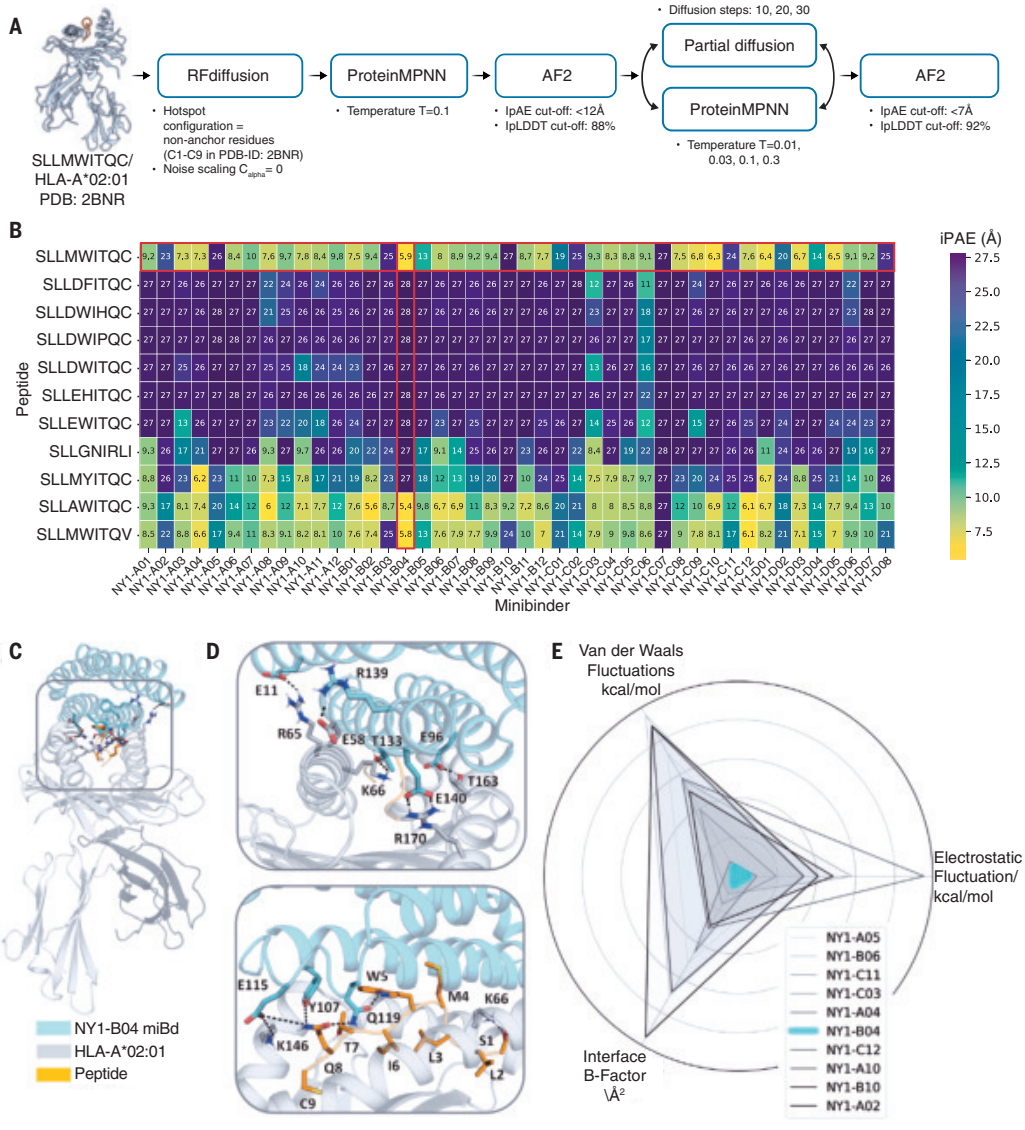


Fig. 1. Design of miBds against NY-ESO-1$^{(157\text{-}165)}$ presented on pMHC and down-selection with conformational diversity estimators from MD simulations. (**A**) Pipeline for the identification of miBds for the SLLMWITQC/HLA-A*02:01 pMHC complex using the crystal structure (PDB accession code 2BNR) as input, with specifications indicated. We generated 5500 backbone structures of the miBd (45 s each on an Nvidia A100 GPU with 40 GB). Next, the backbone structures were subjected to a combination of sequence interpretation using ProteinMPNN (four sequences per backbone) and AF2 structure prediction based on the sequence input for each miBd design (backbone with sequence) that took 20 s and 1.5 min, respectively. (**B**) Heatmap showing iPAE scores in Angstrom (Å) predicted using AF2 as the parameter for potential binding interaction of the 44 miBds selected for in vitro characterization toward the original SLLMWITQC peptide and 10 selected variant peptides. Numbers in each square represent the iPAE for the given complex, and red squares mark the target SLLMWITQC peptide (horizontal) and the NY-B04 miBd (vertical). (**C**) Characterization of dominant salt-bridge and hydrogen-bond interactions involved in stabilizing NY1-B04 with the target pMHC SLLMWITQC/HLA-A*02:01. (**D**) Magnification highlighting the salt-bridge and hydrogen-bond interactions of NY1-B04 (cyan) with HLA-A*02:01 (gray) (top) and the interactions of the peptide (orange) SLLMWITQC with NY1-B04 (bottom). (**E**) MD simulations based on down-selection using criteria such as fluctuations in electrostatic and van der Waals interactions combined with an interface–residue–B-factor analysis as estimators for conformational flexibility in the binding interface identify miBd NY1-B04 (cyan) as the most promising miBd-pMHC complex. For each miBd, 4 × 500 ns were calculated, with an average performance on a GPU of 200 ns/day.

Molecular dynamics identification of false positives and off-target binding likelihood

Protein design guided by machine learning can result in a high rate of false positives when predicted binders fail during experimental testing. Thus, identifying properties and features that can decrease the false-positive rate is crucial. Protein-protein binding is determined by an interplay of enthalpy and entropy, so considering both components is critical when characterizing and predicting protein-protein interfaces. To define the protein-protein interfaces for residue-wise flexibility (B-factor), electrostatic, and van der Waals interaction

energy fluctuations, we considered the salt-bridge and hydrogen-bond interactions for the miBd-pMHC complexes that occur at a rate of >10% throughout the simulations (Fig. 1C). The AF2-predicted complex highlighted salt-bridge and hydrogen-bond interactions of the NY1-B04 miBd with HLA-A*02:01 and the SLLMWITQC peptide (Fig. 1D). Residues E11, E96, T133, R139, and E140 of NY1-B04 formed long-lasting interactions with HLA-A*02:01, and residues Y107, E115, and Q119 interacted with the peptide. The radar plot based on flexibility and/or conformational entropy features showed that NY1-B04 had the lowest conformational diversity in the binding interface and the lowest interaction energy fluctuations, representing the most promising miBd-pMHC complex within the dataset (Fig. 1E). Therefore, we hypothesized that NY1-B04 was the most likely to bind the target pMHC, and we proceeded to test all 44 miBds to explore this prediction.

In vitro selection and specificity screening of SLLMWITQC/HLA-A*02:01 miBds

To validate binding and specificity toward pMHC-I experimentally, synthetic genes encoding all 44 miBd designs targeting SLLMWITQC/HLA-A*02:01 (data S1) were cloned into a lentiviral CD28-CD3ζ CAR vector for expression on the cell surface (Fig. 2A). Using a mammalian surface display system based on miBd-CAR–transduced Jurkat cells, we evaluated the capacity of the miBd library to bind to SLLMWITQC/HLA-A*02:01 tetramers (Fig. 2B and fig. S3A). From this library, we identified one candidate, NY1-B04, that was >2 $\log_2$FC enriched when comparing sorted cells with cells before sorting (Fig. 2, B and C, and data S2). NY1-B04 pMHC-binding characteristics aligned with the molecular dynamics (MD) predictions to deselect false-positive binders. NY1-B04 was individually cloned, and binding of SLLMWITQC/HLA-A*02:01 tetramers was confirmed in CD3 knockout (KO) Jurkat cells (Fig. 2D and fig. S3A). We further measured the binding affinity using biolayer interferometry and obtained a low nanomolar dissociation constant (K_d = 6.92 nM with an SD of 0.7 nM) (Fig. 2E).

To confirm the AF2-predicted binding pose of NY1-B04 to SLLMWITQC/HLA-A*02:01, we determined a cryo–electron microscopy (cryo-EM) structure of the complex (Fig. 2F and figs. S4 and S5). We obtained a 3.8-Å-resolution structure for the ~60-kDa complex (figs. S4 and S5, A to E), and found the same interacting residues in the binding interface as predicted by AF2 and MD simulations (Fig. 1, C to E, and fig. S5E). The overlay of the cryo-EM structure with the AF2 model (fig. S4) revealed an overall highly similar binding pose and angle of approach of NY1-B04 to the pMHC (overall C-alpha root mean square deviation (RMSD): 1.8 Å, peptide Ca-RMSD: 0.92 Å). Figure 2F shows a magnified image of the binding interface, highlighting the critical interacting residues of NY1-B04 with the peptide and emphasizing the importance of residues 1, 4, 5, 7, and 8 of the peptide in stabilizing the complex.

IMPAC-T cell–mediated cytotoxicity toward cancer cells

To evaluate the functional potential of NY1-B04 as a cancer-targeting molecule and not just as a target-binding molecule, we investigated whether the NY1-B04 miBd could facilitate the in vitro killing of cells presenting SLLMWITQC/HLA-A*02:01 complexes on their surface when incorporated into a CAR. We generated primary human T cells expressing an NY1-B04-CD8α hinge-CD28-CD3ζ CAR called

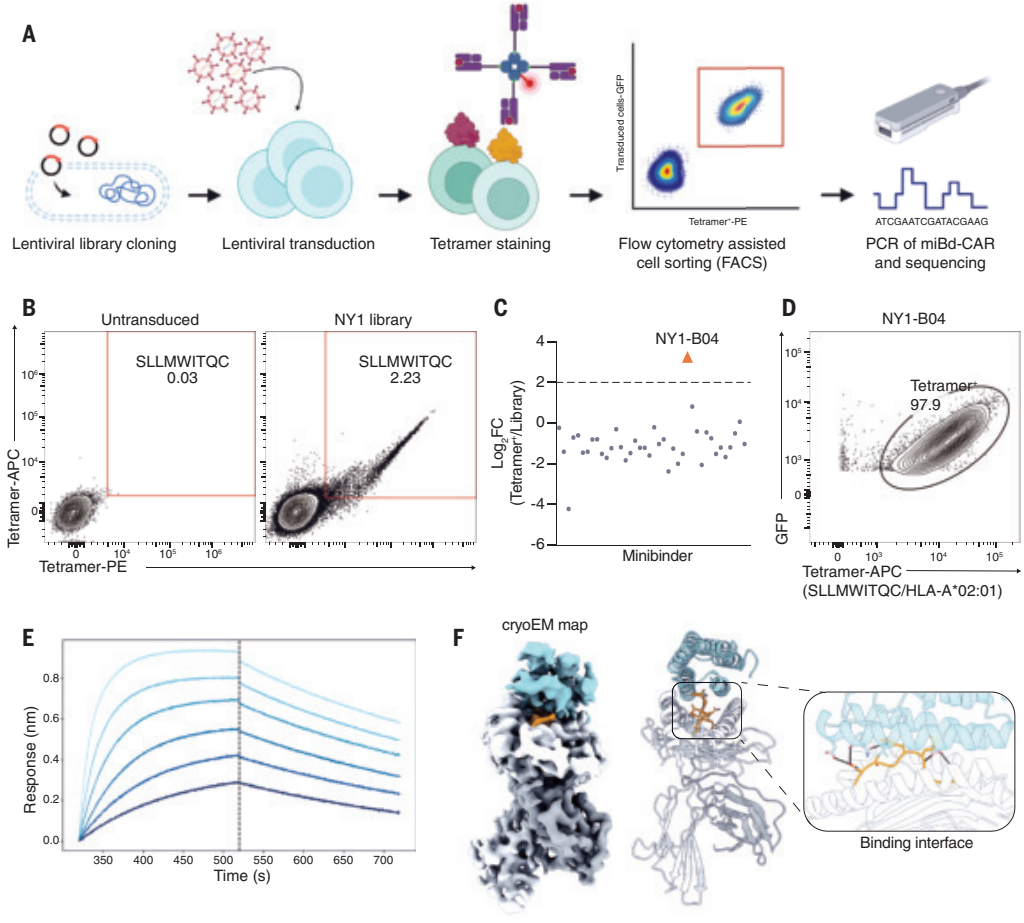


Fig. 2. Experimental selection and characterization of miBds. (**A**) Mammalian surface display workflow. (**B** and **C**) A library of 44 de novo–generated miBds (NY1) was lentivirally transduced into CD3 KO Jurkat cells. Cells were stained with SLLMWITQC/HLA-A*02:01 tetramers[using phycoerythrin (PE) and allophycocyanin (APC), double-positive cells were sorted, and the integrated miBds were sequenced. Data are representative of three screens (N = 3). (B) Flow plots showing the sorted cells gated on singlet/live/CD3$^-$/GFP$^+$ cells (fig. S3A) and (C) $\log_2$FC of miBds in the sorted tetramer$^+$ cells compared with the cells before sorting for the identification of NY1-B04 miBd targeting SLLMWITQC/HLA-A*02:01 (fig. S3A). Each dot represents a miBd in the library. miBds for which it was not possible to determine the $\log_2$FC were excluded from the plot. $\log_2$FC = 2 is marked with a dotted line. (**D**) NY1-B04 binding to SLLMWITQC/HLA-A*02:01 was confirmed by a single tetramer–APC stain gated on singlet/live/CD3$^-$/GFP$^+$ cells (as shown in fig. S3A). Data are representative of two stains (N = 2). (**E**) Affinity measurement of NY1-B04 with biolayer interferometry showing the association and dissociation of pMHC SLLMWITQC/HLA-A*02:01. Each shade of blue represents a concentration of the pMHC at a twofold serial dilution starting from 1 μM (lightest blue curve). 6xHis-tagged NY1-B04 was captured on HIS1K biosensor tips. Binding data were fitted globally using 1:1 model. The dissociation constant was determined to be 6.9 nM (SD of 0.7 nM). (**F**) Cryo-EM map (3.8-Å resolution) and structure model of NY1-B04 in complex with SLLMWITQC/HLA-A*02:01 (PDB accession code 9NNF) confirm the predicted binding mode from AF2 (fig. S3A). The NY1-B04 is shown in cyan, the peptide is shown in orange, and the MHC molecule is shown in gray. The magnification of the binding interface highlights the interactions of NY1-B04 with the peptide.

"immune-targeting minibinder for pMHC activated cytotoxic T" (IMPAC-T) cells by lentiviral transduction (Fig. 3, A and B). Subsequently, we evaluated the capacity of these cells to kill NY-ESO-1$^+$ HLA-A*02:01$^+$ A375 melanoma cells modified to express mCherry. After 24 hours of exposure to NY1-B04 IMPAC-T cells, A375 cell death was higher in IMPAC-T conditions compared with the nontransduced controls (Fig. 3, C and D, and movies S1 and S2). This provides proof of principle that de novo–designed miBds have the potential to be used in cell-based therapies.

Evaluation of cross-reactivity of NY1-B04

To investigate the cross-reactivity profile of NY1-B04, we stained the NY1-B04 IMPAC-T cells with a variety of pMHC tetramers. We assessed the ability of NY1-B04 to bind the peptide-HLA-A*02:01 variants tested during the in silico cross-panning (Fig. 3, E and F, and figs. S3A and S6A). NY1-B04 interacted with the single mutants (mutants marked in bold) SLL**A**WITQC (M4A), SLLM**Y**ITQC (W5Y), and SLLMWITQ**V** (C9V), of which M4A and C9V were ranked with a low iPAE (<10 Å). Furthermore, NY1-B04 also bound SLL**EH**ITQC to some extent but did not bind SLL**E**WITQC (Fig. 4, E and F). Structurally, the negatively charged M4E mutation disrupts binding, and the histidine in W5H could partially stabilize through a hydrogen bond (fig. S7A). Overall, we identified complete abrogation of NY1-B04 binding when introducing polar, negatively charged amino acids (aspartic acid and glutamic acid) in position 4. These data indicate that the in silico cross-panning was capable of correctly predicting binding for eight of 10 tested variants (accuracy $N = 10 = 8/10$) (fig. S7B).

To further investigate how subtle electrostatic changes, such as in alanine and glycine, and positively charged side chains, such as in arginine, affected the binding interactions, we performed single–amino acid substitution scans both in vitro and in silico. In vitro, we found that alanine and glycine substitutions on three of nine positions (positions L2, M4, and T7) decreased pMHC binding, whereas substitution with arginine disrupted or decreased binding in seven of nine positions (positions 1 to 7) (Fig. 3, G and H, and fig. S6B). The in silico binding predictions revealed that the binary categorization based on iPAE values led in 19 of 27 cases to the correct identification of binding and nonbinding (accuracy$_{N=27}$ = 19/27) (fig. S7C). Structurally, the NY1-B04 encloses position M4 and W5 in a combined hydrophobic pocket (M93, A97, A100, F123, L127, and A130) (fig. S7A). This entails the repulsion of polar residues, especially charged residues such as M4E, M4D, M4R, and W5R.

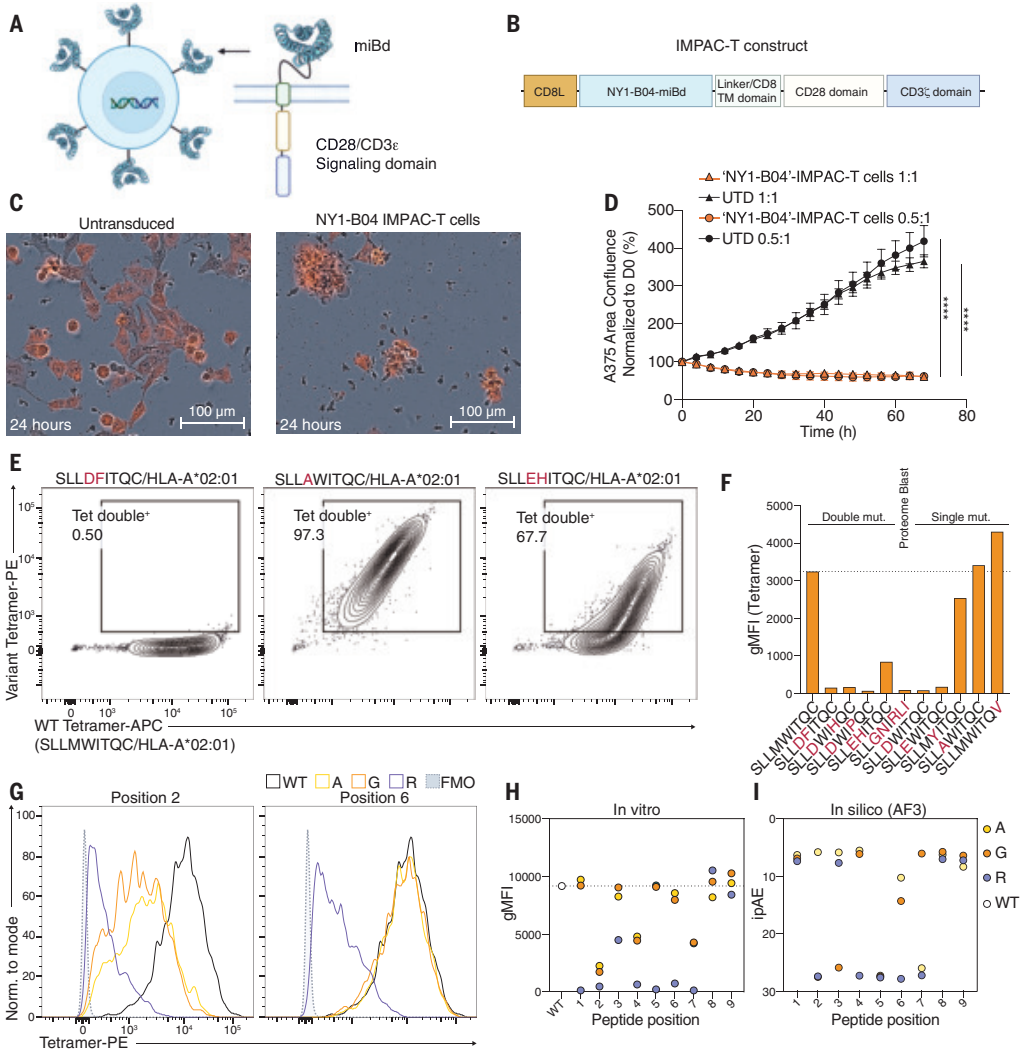


Fig. 3. IMPAC-T cell–mediated cytotoxicity against NY-ESO-1$^+$ A375 cancer cells. (**A**) Overview of the de novo IMPAC-T cell concept with miBds fused extracellularly to intracellular signaling domains. (**B**) Construct design of the NY1-B04-IMPAC construct with a CD8 leader, miBd, CD8a hinge region, and CD28/CD3ζ intracellular signaling domains. (**C**) Phase contrast images overlaid with mCherry fluorescence showing the mCherry$^+$ A375 cancer cells after 24 hours of incubation of untransduced (UTD) primary T cells (left) or primary NY1-B04-IMPAC-T cells (right). Scale bar, 100 µm. (**D**) Cell death of mCherry$^+$ A375 cancer cells was monitored and quantified over a period of ~2 days using Incucyte live-cell imaging, with the red fluorescence area normalized to the D0 area as a measure of confluency. The assay was set up with 0.5:1 and 1:1 effector-to-target ratios. (**E**) Cross-reactivity screening of NY1-B04 expressed in CD3 KO Jurkat cells against variant peptides showing representative plots of three examples of nonbinding and binding peptides after double staining with SLLMWITQC/HLA-A*02:01 tetramers conjugated to APC and variant tetramers conjugated to PE. Cells were gated on singlet/live/CD3$^-$/GFP$^+$ cells, as shown in fig. S3. (**F**) Bar plot showing representative geometric mean fluorescence intensity (gMFI) of single-stained variant tetramer stains. Mutations are shown in red, and the wild-type (WT) tetramer gMFI is indicated with a dotted line. (**G** and **H**) Representative histograms (G) and gMFI plot (H) of single substitution scanning data with alanine (A), glycine (G), and arginine (R) in primary T cells transduced with NY1-B04. In (H), gMFI is shown at each amino acid position of the peptide, with the dotted line indicating the gMFI of the WT peptide. (**I**) iPAE values for A, G, and R single–amino acid substitutions predicted with AF2 initial guess based on AF3 structure prediction of the pMHC. The *y* axis is inverted, given that a low iPAE value indicates predicted binding to match the data shown in (H). Data shown in (G) and (H) are gated on singlet/live/CD3$^+$/CD8$^+$/GFP$^+$ cells, as shown in fig. S6B. Data in (D) are means ± SD ($N = 3$ technical replicates) representative of two independent experiments ($N = 2$), and data in (E) and (F) are representative gMFIs of two experiments ($N = 2$). ****$P \leq 0.0001$ in (D) as evaluated by a Wilcoxon matched-pairs signed rank test comparing the area confluence of each time point between UTD and NY1-B04-IMPAC-T samples (0.5:1 and 1:1 are compared separately).

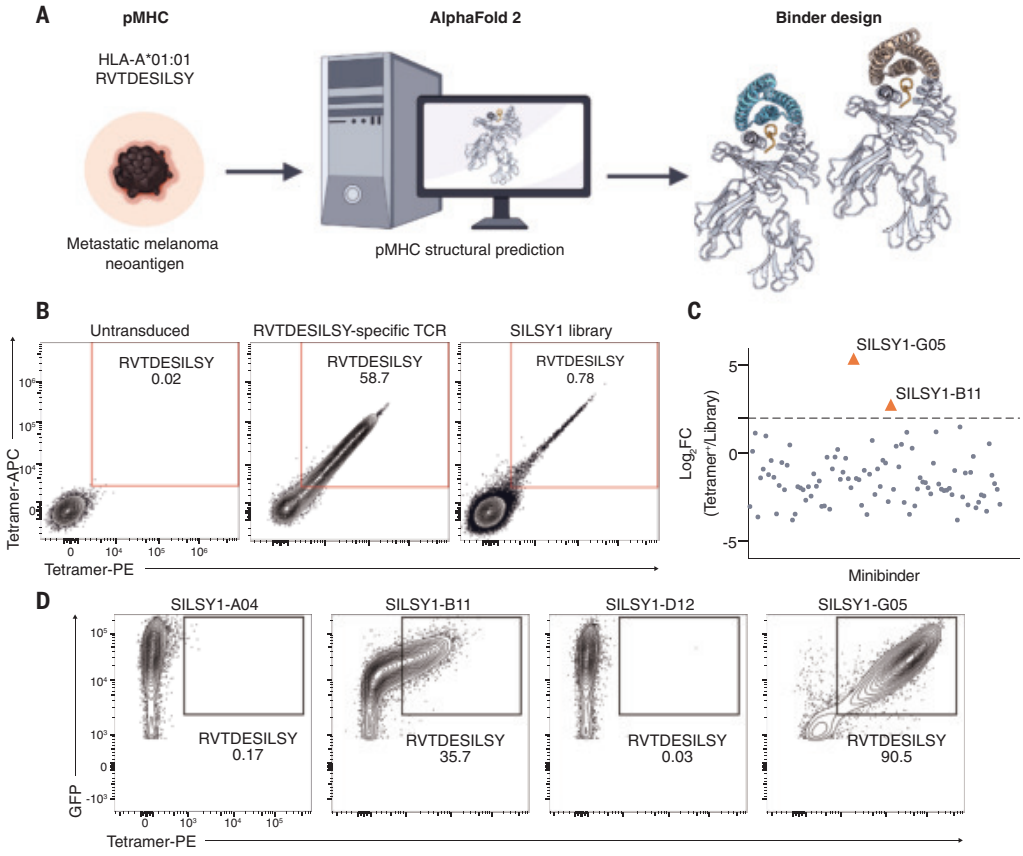


Fig. 4. Proof of concept of applying the design pipeline to neoantigens. (**A**) The structure of the neoantigen RVTDESILSY in complex with HLA-A*01:01 was predicted using AF2 and miBds designed following the previously described pipeline. (**B**) A library of 95 predicted miBds (SILSY1) was lentivirally transduced in CD3 KO Jurkat cells, and the cells were stained with RVTDESILSY/HLA-A*01:01 tetramers (PE and APC). Double-positive populations were sorted, and the integrated miBds were sequenced. (**C**) Log_2FC of miBds in the sorted population (tetramer$^+$) compared with the cells before sorting (library) for the identification of two miBds targeting RVTDESILSY/HLA-A*01:01. Each dot represents a miBd in the library. miBds for which it was not possible to determine the log_2FC were excluded from the plot. $Log_2FC = 2$ is marked with a dotted line. (**D**) Flow plots of RVTDESILSY/HLA-A*01:01 tetramer–PE-stained CD3 KO Jurkat cells expressing each of the four highest-scoring miBds from (B) and (C). Flow plots in (B) and (D) show cells gated on singlet/live/CD3$^-$/GFP$^+$ cells, as shown in fig. S3A. Data in (B) and (C) are representative data from two screens ($N = 2$), and data in (D) is from an experiment performed once.

Further, position L2 seems to determine the binding pose of the peptide in the MHC, which could be crucial for the interaction of E96 in NY1-B04 toward the backbone at position M4 (fig. S7A).

Extending the miBd discovery pipeline to neoantigens

A potential hurdle for broader application of machine learning–guided design of pMHC-I miBds is that crystal structures of pMHC-I rarely are available. Therefore, we sought to further test the limits of our approach by targeting the neoantigen RVTDESILSY derived from the mutation AKAP9^{P1796L} identified in metastatic melanoma, presented on HLA-A*01:01. The RVTDESILSY/HLA-A*01:01 complex was previously demonstrated as subject for T cell recognition in both blood and tumor-infiltrating lymphocyte infusion product in a patient with melanoma (*27*). On the basis of structure predictions, the neoantigen resulted in a targetable protruding leucine in the MHC-I groove (*27*). Because no experimental structure existed, we modeled the pMHC-I complex using AF2. Upon selecting the most confident structure based on AF2 (pLDDT of 95.7), we designed miBds using our pipeline (Fig. 4A).

We selected 96 miBds (data S1), introduced them for surface expression as a pool in CD3 KO Jurkat cells, and evaluated their binding to RVTDESILSY/HLA-A*01:01 tetramers in vitro. We found that the fluorochrome intensity of the tetramer$^+$ miBd library was comparable to that of the positive control primary T cells with CRISPR/Cas9 knock-in of a RVTDESILSY/HLA-A*01:01–specific TCR (Fig. 4B and fig. S3A).

Sequencing of the sorted population identified two distinct miBds (SILSY1-B11 and SILSY1-G05), constituting a success rate of ~2% (Fig. 4, C and D; fig. S3A; and data S2). This demonstrated the possibility of applying our design pipeline in the context of personalized cancer therapy for neoantigens for which experimental structures are not available.

Discussion

A challenge in cancer immunotherapy has been targeting pMHC with high specificity and efficacy (*28*). Traditionally, this has been achieved using TCRs identified by screening T cell responsiveness and specificity from cancer patients (*1*, *29*–*31*). However, this is extremely time and resource intensive, and the identified TCRs often lack the required specificity, which may either prevent treatment or provide a risk of fatal side effects (*28*). Thus, there has been a need to develop approaches to overcoming these challenges and unlocking the potential of targeted T cell therapy. We leveraged leading-edge protein design approaches to generatively design SLLMWITQC/HLA-A*02:01 pMHC-I miBds de novo. We modified current design pipelines (*17*, *21*, *22*) to design pMHC-I miBds while screening for specificity in silico with the aim of reducing the risk of off-target pMHC-I binding. We demonstrated that these rapid (<10 s per miBd-target complex) computational predictions largely correlated with in vitro specificity screening results, substantiating the potential that this approach has in de-risking T-cell therapy. Although further improvement in the current predictors and possibly a combination of several orthogonal metrics are needed, our results show the potential benefit of computational off-target binding screens. We propose that by using a mammalian surface display in a CAR construct when identifying binders, miBds that would not bind the target when presented in a CAR were deselected in the pooled in vitro screen. We further found that the miBds we designed facilitated target cell killing when using the miBd as the recognition element in a CAR format to generate IMPAC-T cells. This indicates that a key next step is the in vivo validation of CARs incorporating machine learning–guided pMHC miBds that can direct T cells toward the pMHC-I antigen of interest and activate engagement, signaling, and, consequently, cytotoxicity toward pMHC-I–expressing target cells.

Machine learning–guided design holds great promise for a rapid therapeutic discovery pipeline to target both shared cancer antigens, such as NY-ESO-1, and patient-specific pMHC targets, such as neoantigens (*32*). The timelines for identifying an miBd for both SLLMWITQC/

HLA-A*02:01, where the crystal structure is available, and for RVTDESILSY/HLA-A*01:01, where no prior pMHC-I structure was described, were both ~4 to 6 weeks. Given the success for both known crystallized and structurally predicted targets, we suggest that the pipeline we present here could be broadly applicable to the generation of miBds to other peptides and HLAs.

Another strength of this approach is that specific miBds were identified with a relatively high success rate (1:50), which we demonstrated could be increased by using MD approaches. We found that considering flexibility in the binding interface accelerated the identification of potential binders, and a direct comparison of the binding interfaces was facilitated by the high similarity of the interaction sites among these potential binders. Thus, considering interface flexibility and dynamics can substantially reduce the false-positive rate and further expedite the identification of the most promising miBd–pMHC-I complex. A limitation of applying de novo proteins in vivo is the risk that they could induce an immune response that could include endogenous TCR-driven immunogenicity against IMPAC-T cells and antidrug antibodies against de novo proteins. Because of their small size, miBds are unlikely to raise substantially larger immunogenicity than animal-derived antibodies, nonetheless, thorough in vitro specificity testing and well-structured in vivo studies evaluating and mitigating this risk will be needed.

Beyond cancer applications, the versatility of miBds makes them promising for targeting viral antigens presented on infected cells. This could be transformative for diseases such as HIV and hepatitis C, where broad-spectrum targeting remains challenging because of high mutation rates. Additionally, miBds could be used to block autoimmune TCR-pMHC interactions, as proposed for TCR-like antibodies (*33*), expanding their therapeutic scope. Further, the findings shown here suggest the use of pMHC-binding miBds as carriers of payloads or pMHC-targeting bispecific T cell engagers. Given the previous success of generatively designed miBds in vivo (*21*), these findings are likely to not only advance the field of cancer immunotherapy by introducing a rapid and precise method for de novo pMHC targeting but also pave the way for targeted patient-centric and personalized therapies.

REFERENCES AND NOTES

1. C. A. Klebanoff, S. S. Chandran, B. M. Baker, S. A. Quezada, A. Ribas, *Nat. Rev. Drug Discov.* **22**, 996–1017 (2023).
2. Y.-T. Chen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1914–1918 (1997).
3. R. Pétremand *et al.*, *Nat. Biotechnol.* **43**, 323–328 (2025).
4. M. Arnaud *et al.*, *Nat. Biotechnol.* **40**, 656–660 (2022).
5. E. Strønen *et al.*, *Science* **352**, 1337–1341 (2016).
6. A. K. Sewell, *Nat. Rev. Immunol.* **12**, 669–677 (2012).
7. F. Drost *et al.*, *Cell Genom.* **4**, 100634 (2024).
8. M. Nielsen *et al.*, *Immunoinformatics (Amst.)* **16**, 100045 (2024).
9. M. Yarmarkovich *et al.*, *Nature* **623**, 820–827 (2023).
10. A. Porgador, J. W. Yewdell, Y. Deng, J. R. Bennink, R. N. Germain, *Immunity* **6**, 715–726 (1997).
11. T. Dao *et al.*, *Sci. Transl. Med.* **5**, 176ra33 (2013).
12. H. Du *et al.*, *Nat. Biotechnol.* (2024).
13. G. Stewart-Jones *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5784–5788 (2009).
14. C. H. Lee *et al.*, *Front. Immunol.* **11**, 565096 (2020).
15. V. K. Karnaukhov *et al.*, *Nat. Comput. Sci.* **4**, 510–521 (2024).
16. Y. Zhao *et al.*, *Sci. Adv.* **9**, eabo5128 (2023).
17. J. L. Watson *et al.*, *Nature* **620**, 1089–1100 (2023).
18. J. Dauparas *et al.*, *Science* **378**, 49–56 (2022).
19. J. Jumper *et al.*, *Nature* **596**, 583–589 (2021).
20. N. R. Bennett *et al.*, *Nat. Commun.* **14**, 2625 (2023).
21. S. Vázquez Torres *et al.*, *Nature* **639**, 225–231 (2025).
22. S. Vázquez Torres *et al.*, *Nature* **626**, 435–442 (2024).
23. J.-L. Chen *et al.*, *J. Exp. Med.* **201**, 1243–1255 (2005).
24. J. Abramson *et al.*, *Nature* **630**, 493–500 (2024).
25. B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, *Nucleic Acids Res.* **48**, W449–W454 (2020).
26. M. Johnson *et al.*, *Nucleic Acids Res.* **36**, W5–W9 (2008).
27. N. P. Kristensen *et al.*, *J. Clin. Invest.* **132**, e150535 (2022).
28. G. P. Linette *et al.*, *Blood* **122**, 863–871 (2013).
29. R. A. Morgan *et al.*, *Science* **314**, 126–129 (2006).
30. W. Chour *et al.*, *Commun. Biol.* **6**, 528 (2023).
31. E. Giannakopoulou *et al.*, *Nat. Cancer* **4**, 1474–1490 (2023).
32. F. Huber *et al.*, *Nat. Biotechnol.* (2024).
33. L. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2656–2661 (2014).
34. M. L. Fernández-Quintero *et al.*, *MAbs* **15**, 2175319 (2023).

ACKNOWLEDGMENTS

We thank B. Liu, N. F. Greenwood, and J. Bonzanini from the Baker lab for discussions and knowledge exchange, as well as for timing submissions, and S. Sebbaha, research manager from the Flow Cytometry and Imaging Core (FLIC) at DTU Health Tech, for supporting flow cytometry experimental design and data analysis; and the EuroHPC Joint Undertaking for awarding us access to MeluXina, Luxemburg (M.L.F.Q.). BioRender was used for some of the figures. **Funding:** This work was supported by the Alliance program under a EuroTech Universities agreement (T.P.J); Innovation Fund Denmark [industrial PhD stipend 2052–00010B (2022) to D.S.W]; the Lundbeck Foundation (postdoctoral grant R347-2020-2174 to K.H.J); the Novo Nordisk Foundation (Bioscience and Biomedicine Project Grant 0087077 to S.R.H); and the European Research Council (ERC Consolidator Grant MIMIC 101045517 to S.R.H). **Author contributions:** T.P.J., K.H.J., and S.R.H. conceived the project, provided funding, and supervised the work. T.P.J., K.H.J., D.S.W., B.S., M.L.F.-Q., and S.R.H. wrote and reviewed the manuscript. D.S.W., C.R.C., J.S.F., and T.P.J. designed and in silico screened the pMHC miBds. D.S.W., M.D., and A.T.D.E. initiated the design of neoantigen pMHC miBds with input from C.R.C. E.Q.A. and D.S.W. analyzed the secondary structure composition and orientation of the miBds. K.H.J., B.S., and C.R.C. prepared binder libraries, performed mammalian surface display screening, and sequenced hits. M.L.F.-Q., A.G., and A.B.W. performed cryo-EM and data analysis. M.L.F.-Q., K.H.B., S.T., and J.R.L. performed affinity measurements and data analysis. M.L.F.-Q., J.R.L., S.F., and A.B.W. performed MD simulations and data analysis. K.H.J., B.S., C.R.C., M.C.V., I.L., C.R.P., and L.F.V. conducted cell assays and in vitro flow cytometric analysis of miBds. K.K.M. analyzed predicted miBd structures and assisted with sequencing and data analysis. D.S.W., M.D.O., and O.M. developed and improved in silico specificity screening protocols. M.O. designed CAR vectors. E.R.d.T. and T.T. expressed the miBds and MHC monomers. All authors reviewed and approved the final manuscript. **Competing interests:** K.H.J., D.S.W., B.S., S.R.H., and T.P.J. are inventors on a patent related to this work under filing number US 240809-P7151EP00. The remaining authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the study are available in the main text or the supplementary materials. The code for binder design using RFdiffusion has been published previously (*17*). The Pymol script for assessment and depiction of potential clashes of co-complexes was also published previously (*34*). The high-resolution structure of miBd B04 in complex with SLLMWITQC/HLA-A*02:01 was deposited in the Protein Data Bank with accession code 9NNF.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adv0422
Materials and Methods; Figs. S1 to S7; Tables S1 to S3; References (*35–50*); Data S1 and S2; Movies S1 and S2

Submitted 3 December 2024; accepted 16 May 2025

10.1126/science.adv0422

PROTEIN DESIGN

Design of high-specificity binders for peptide–MHC-I complexes

Bingxu Liu[1,2]*†, Nathan F. Greenwood[1,2]†, Julia E. Bonzanini[1,2,3]†, Amir Motmaen[1,2,3], Jeremy Meyerberg[4], Tao Dao[4], Xinyu Xiang[5], Russell Ault[6,7], Jazmin Sharp[1,2], Chunyu Wang[5], Gian Marco Visani[8], Dionne K. Vafeados[1,2], Nicole Roullier[1,2], Armita Nourmohammad[8,9,10,11], David A. Scheinberg[4,12], K. Christopher Garcia[5,13], David Baker[1,2,13]*

Class I major histocompatibility complex (MHC-I) molecules present peptides derived from intracellular antigens on the cell surface for immune surveillance. Proteins that recognize peptide–MHC-I (pMHCI) complexes with specificity for diseased cells could have considerable therapeutic utility. Specificity requires recognition of outward-facing amino acid residues within the disease-associated peptide as well as avoidance of extensive contacts with ubiquitously expressed MHC. We used RFdiffusion to design pMHCI-binding proteins that make extensive contacts with the peptide and identified specific binders for 11 target pMHCs starting from either experimental or predicted pMHCI structures. Upon incorporation into chimeric antigen receptors, designs for eight targets conferred peptide-specific T cell activation. Our approach should have broad utility for both protein- and cell-based pMHCI targeting.

Class I major histocompatibility complex (MHC-I) molecules are cell surface proteins that present peptides derived from intracellular proteins. The recognition of peptides displayed on MHC-I (pMHCI complexes) by the T cell receptor (TCR) enables the immune system to detect and destroy cells expressing proteins associated with disease, as a consequence of infection or mutations. From a therapeutic perspective, targeting pMHCs with engineered cells (*1*, *2*) or proteins—for example, bispecific T cell engagers (BiTEs) (*3*–*5*)—is attractive, as it provides an opportunity to distinguish cells on the basis of their intracellular proteins. TCRs can be used for such targeting, but, in many cases, TCRs with specificity to recognize cells expressing disease-associated proteins have not been identified. The diversity of MHC alleles (also referred to as human leukocyte antigen, or HLA, alleles) and antigenic peptides across patient populations necessitates the development of hundreds to thousands of effective and specific binders to achieve broad patient coverage. Current methods rely on empirical TCR screening from patient samples (*6*) or antibody-like constructs, such as single-chain variable fragment (scFv) libraries (*1*, *3*), which are costly, labor-intensive, and time-consuming. Efficient methods to generate small, stable proteins that recognize specific pMHCIs of interest could have considerable therapeutic utility for generating recognition domains in chimeric antigen receptors (CARs) for both cell-based and protein-based therapies. To be effective and safe in cell therapy settings, CARs that incorporate designed proteins must mediate specific activation by the target peptide loaded on the target HLA, but not by the thousands of other peptides from the proteome loaded on the same HLA, or by different HLAs—which poses a considerable challenge for binder discovery.

Recent advances in deep learning–driven binder design have enabled the structure-based design of small, stable binders with high affinity and specificity against a wide range of folded (*7*–*9*) and disordered (*10*, *11*) protein targets. We reasoned that if such binders could be generated against pMHCIs, there could be multiple advantages in stability, engineerability, and manufacturability, and we set out to explore the computational design of high-affinity and high-specificity pMHCI binders.

Computational design of pMHCI binders for high target specificity

To functionally substitute for TCRs, pMHC binders must be capable of distinguishing the target peptide from other proteome peptides when presented in the MHC groove (Fig. 1A). To generate such binders, we used the generative AI method RFdiffusion targeted to the outward-facing residues of the bound peptide to generate protein backbones that arc above the central peptide binding groove of MHC-I (Fig. 1B). From many independent trajectories (examples in fig. S1A), we selected backbones capable of hosting side chains that interact extensively with the peptide (*12*) (fig. S1B), optimized their sequences for folding and binding using ProteinMPNN, and selected those that AlphaFold2 (AF2) (*13*) predicted to fold and bind as designed. We then used ProteinMPNN to predict the sequence of the target peptide in the context of the designed binder and selected those designs for which the confidence of the on-target peptide sequence was higher than off-target sequences (fig. S1C). Finally, we used a fine-tuned AF2 model (*14*) to predict the structure of each design with the target peptides and closely related peptides in the proteome in complex with the MHC and selected designs predicted to bind the on-target peptide considerably more confidently than the off-target peptides (fig. S1D).

Designed binders bind to their target pMHCI with high affinity and specificity

To evaluate the ability of our design pipeline to generate binders to a wide range of pMHCI complexes, we selected a structurally diverse set of 11 complexes comprising HLA alleles A*01:01, A*02:01, A*03:01, and C*07:02 presenting 9- and 10-mer peptides from viral proteins, tumor-associated proteins, and neoantigens (data S1). For each pMHCI target, we obtained oligonucleotide pools encoding between 200 and 12,000 designs, displayed them on yeast, and selected those that specifically recognized the targeted pMHCI but not one to four closely related off-target peptides (table S1) on the same MHC using FACS (fluorescence-activated cell sorting) to collect cells displaying designs that bound on-target (labeled with one fluorescent marker) but not off-target (labeled with a different color marker) pMHCI tetramers (fig. S2A, lower-right quadrant) followed by next-generation sequencing enrichment analysis or clonal selection to determine their identities (Fig. 2A).

For eight of the pMHCI complexes, we used the de novo diffusion approach to generate designs with a range of topologies (Fig. 2A, left) and peptide binding interfaces (Fig. 2A, right) that bound the target peptide specifically, with reduced binding to off-target peptides on the same HLA (Fig. 2B). Predictions of the structures of the designed proteins in complex with their cognate pMHCI using AlphaFold3 (AF3) and Chai-1 (*15*, *16*) were very similar to the design models (table S2 and fig. S2, B and C, right).

To evaluate the accuracy of our design approach, we obtained a 2.2-Å crystal structure of binder mart1-3 with the MART-1 pMHCI complex. The crystal structure closely matched the design model at both the

[1]Department of Biochemistry, University of Washington, Seattle, WA, USA. [2]Institute for Protein Design, University of Washington, Seattle, WA, USA. [3]Bioengineering Graduate Program, University of Washington, Seattle, WA, USA. [4]Molecular Pharmacology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [5]Departments of Molecular and Cellular Physiology and Structural Biology, Stanford University School of Medicine, Stanford, CA, USA. [6]Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [7]Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [8]Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA. [9]Department of Physics, University of Washington, Seattle, WA, USA. [10]Department of Applied Mathematics, University of Washington, Seattle, WA, USA. [11]Fred Hutchinson Cancer Center, Seattle, WA, USA. [12]Weill Cornell Medicine, New York, NY, USA. [13]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA.
***Corresponding author. Email: bxliu@uw.edu (B.L.); dabaker@uw.edu (D.B.)**
†These authors contributed equally to this work.

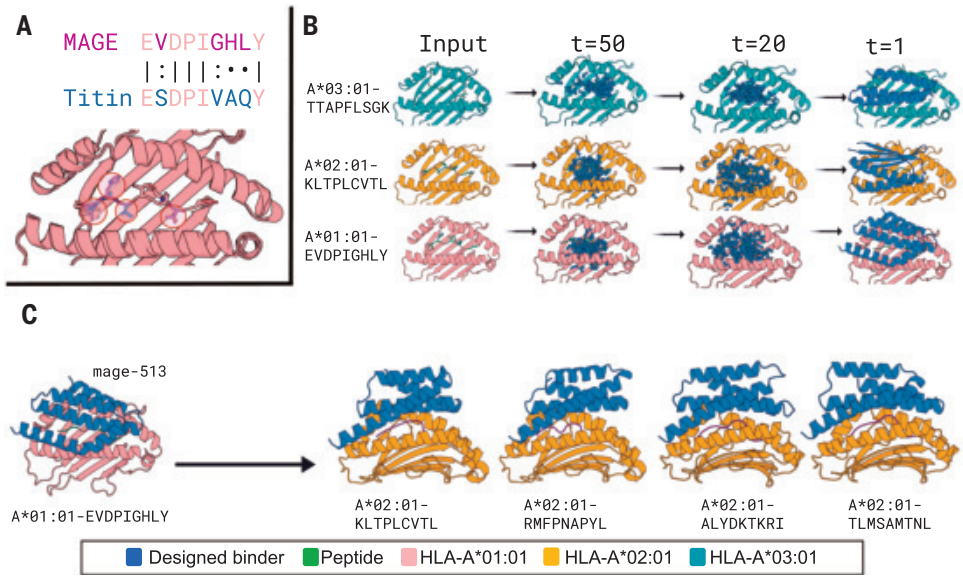


Fig. 1. Diffusion of pMHC binders. (**A**) pMHC structure and design challenge. The goal is to distinguish a target peptide (in this example, MAGE) from a closely related off-target (Titin). The positions that differ between the two peptides are circled in the pMHC structure model at the bottom. (**B**) Representative diffusion trajectories and design models for three different pMHC targets. Column 1, target pMHC; column 2, initial Gaussian noise initialization; columns 3 and 4, intermediate steps in diffusion denoising trajectories starting from completely random residue distributions; right column, fully denoised design model backbones. (**C**) Partial diffusion of starting scaffolds against related targets. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

backbone [C-alpha root mean square deviation (CαRMSD) = 0.4 Å] and side-chain level (interface all-atom RMSD = 0.4 Å across 15 interface residues), with the designed binder forming an extensive interface with the peptide (Fig. 2D, table S3, and fig. S3). Designed interactions between mart1-3 (design naming scheme is "lowercase target name–design number") and the MART-1 peptide recapitulated in the crystal structure include hydrophobic interactions with the L5 side chain and hydrogen bonds with the main chains of G6, I7, and T9 and the side chain of T9 (Fig. 2D).

Three of the pMHCI targets are viral peptides presented on A*02:01: SARS-CoV1/GLMWLSYFV (*17*), YFV/LLWNGPIAV (*18*), and HIV/KLTPLCVTL (*19*). To assess whether the designs could function when incorporated into CARs, we generated between 10 and 30 CAR constructs per design (*20*) and expressed them on the surface of Jurkat cells. For each of the three targets, we identified designs that distinguished the target peptides from closely related off-target peptides when displayed on yeast (Fig. 2B, top row) or when incorporated into CARs (Fig. 2B, rows 2 and 3). In the design models (Fig. 2A), these binders make extensive contacts with the target peptide (table S2). Design sars-6 buried the target peptide W4 and L5 in a hydrophobic pocket and made a network of flanking hydrogen bonds with the peptide backbone. Design yfv-2 made hydrogen bonds with N4 and hydrophobic interactions with I7 of the peptide. Design hiv-10 made hydrogen bonds with K1, T8, and the peptide backbone.

Four of the pMHCI design targets were tumor-associated antigens: WT1 (RMFPNAPYL/A*02:01) (*21*–*23*), PAP (TLMSAMTNL/A*02:01) (*24*), PHOX2B (QYNPIRTTF/C*07:02) (*1*), and the neoantigen CTNNb1 (TTAPFLSGK/A*03:01 with S5F mutation) (*25*). We identified specific designs for all four peptides (Fig. 2B), which interact extensively with the peptide targets (Fig. 2A). Design ctnnb1-15 makes extensive contact with the mutated F5 through cation-pi, pi-pi, and other hydrophobic interactions (Fig. 2A), likely contributing to specific binding for the S5F peptide versus the unmutated CTNNb1 peptide (Fig. 2B). Design phox2b-5, generated against a predicted pMHCI structure, makes hydrogen bonds to the peptide backbone and side chains, including bidentate interactions to R6 of the peptide backbone, rigidifying the peptide and contributing to the observed specificity against the R6A mutant (Fig. 2A).

Privileged scaffolds can be redesigned to target multiple pMHCs

While the de novo RFdiffusion approach can generate binders to pMHCI targets, it is computationally expensive, as only a small fraction of trajectories yield backbones that interact extensively with the peptide while largely avoiding the MHC. Given the overall similarity of pMHCI structures, we reasoned that designed backbones that effectively target one pMHCI complex could likely be repurposed to recognize other pMHCI targets. We used partial diffusion (*26*) to adapt the most promising scaffolds to additional pMHCI targets [Fig. 1C and data S1; partial diffusion carries out partial noising-denoising to sample around an input structure (*7*)]. We used partial diffusion starting from a designed binder against ALHGGWTTK/A*03:01 to generate a design (mage-513) that bound the melanoma-associated antigen 3 (MAGE-A3) peptide (EVDPIGHLY/A*01:01) (*27*, *28*) but not the off-target peptide (ESDPIVAQY/A*01:01) (*29*) derived from the cardiac Titin protein (Fig. 2C, top). From this scaffold, we used partial diffusion to design binders for other 9- and 10-residue tumor-associated antigens presented by A*02:01: the gp100 peptide YLEPGPVTA (*30*), the MART-1 peptide (A2L) ELAGIGILTV (*31*), and PRAME (ALYVDSLFFL) (*32*). For gp100, MART-1, and PRAME, we identified specific binders from partial diffusion that made extensive and diverse hydrophobic and hydrogen bonding interactions with the target peptide (Fig. 2C).

To evaluate the behavior of our designs as soluble proteins, we expressed them in *Escherichia coli* and purified the hiv-10 and mage-513 binder designs. The purified designs eluted as single peaks in size exclusion chromatography around the expected size (fig. S2, B and C, left). Surface plasmon resonance (SPR) showed that they bound their cognate pMHCIs with binding affinities in the single- to double-digit nanomolar range (fig. S2, B and C, middle).

Designed binders have high specificity

We evaluated the ability of the designs to confer specific T cell activation by the target pMHCI by incorporating them into CARs and expressing

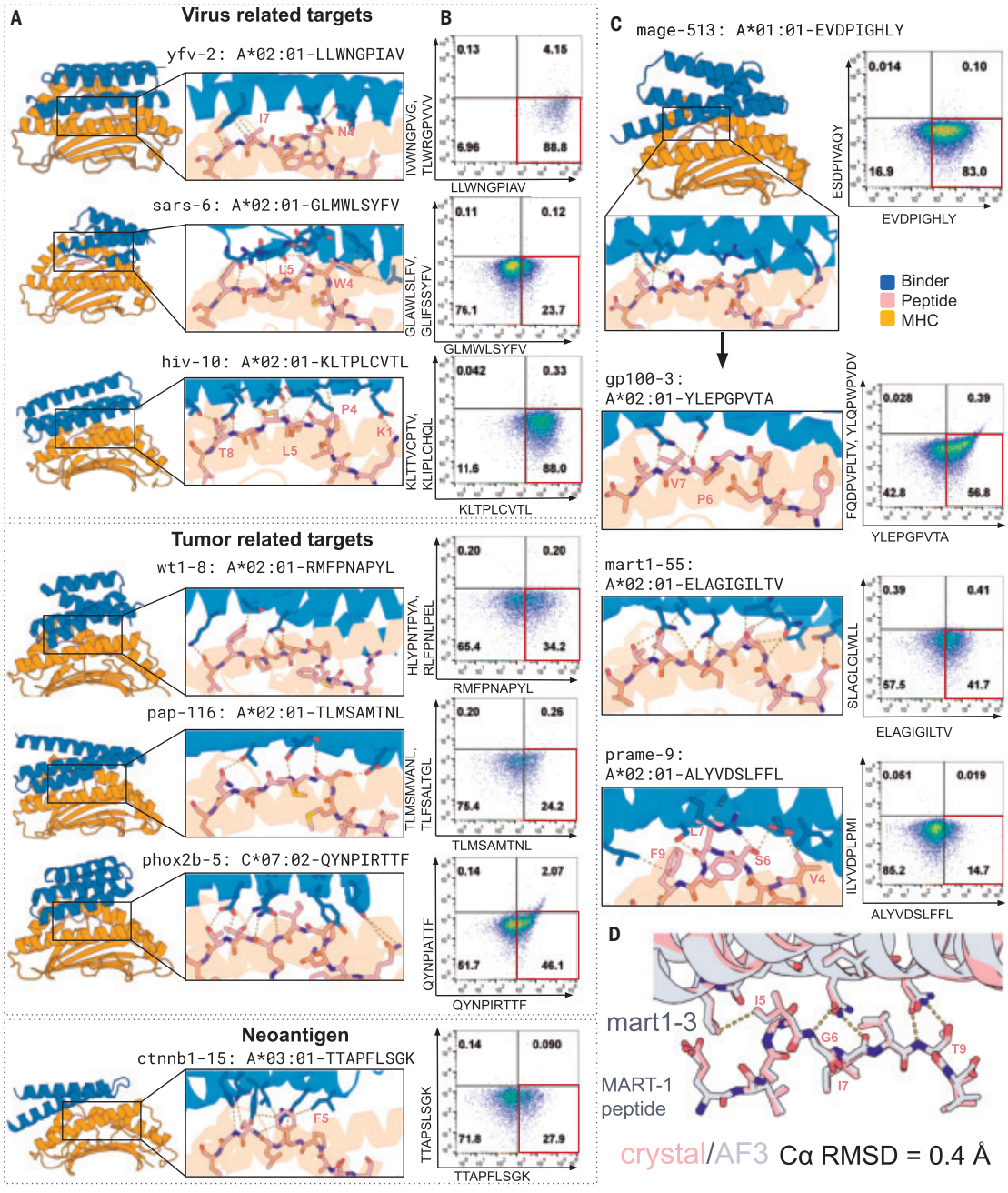


Fig. 2. Design models and binding specificity. (**A**) Design models. Left, overall structure; right, zoom in on peptide binding region. pMHCI in orange, peptide in sticks and salmon, and binder in blue. HLA allele and peptide sequence are specified above the zoom-in view. (**B**) Flow cytometry of cells displaying the design incubated with on-target pMHCI tetramer (*x* axis) and one or two off-target tetramers (*y* axis) at 10 nM concentration. Staining in the lower-right quadrant indicates specific on-target binding. Row 1, individual designs displayed on yeast; rows 2 to 7, CARs incorporating designs on Jurkat cells. (**C**) Partial diffusion of design at top to targets below (left panels) and corresponding Jurkat straining (right). (**D**) Crystal structure of mart1-3 binder overlaid with AF3 structure prediction.

them on Jurkat cells. These were incubated with 293T cells treated with on- or off-target peptides. 293T cells naturally express HLA-A*02:01 and C*07:02 and have an intact antigen presentation system, and hence present a wide range of peptides derived from self-proteins on their surfaces (represented as the DMSO condition in all Jurkat activation assays); requiring selective activation (measured by CD69 expression level) by pulsing peptides in this system is thus more stringent than using cells that cannot present self-peptides owing to defects in presentation.

Jurkat cells expressing CARs incorporating the mage-513 design were strongly and specifically activated by MAGE-A3 peptide stimulation but not by the off-target Titin peptide or peptides derived from the intracellular proteome on HLA-A*01:01 (Fig. 3A). Although they bound to the target tetramer specifically (fig. S4A), other designs showed weak (fig. S4B, mage-282) or background activation (fig. S4B, mage-4), likely due to responses to other peptides from the proteome loaded on the same HLA [only HLA-A*01:01–expressing 293T cells induced background activation (fig. S4B)].

In the design model (Fig. 3B), mage-513 engages the MAGE peptide through hydrogen bonds with the side chain of H7 and backbone of L8 (Fig. 3B) and through hydrophobic interactions (mediated by design L37 and L86) with the side chain of peptide L8 (there were also

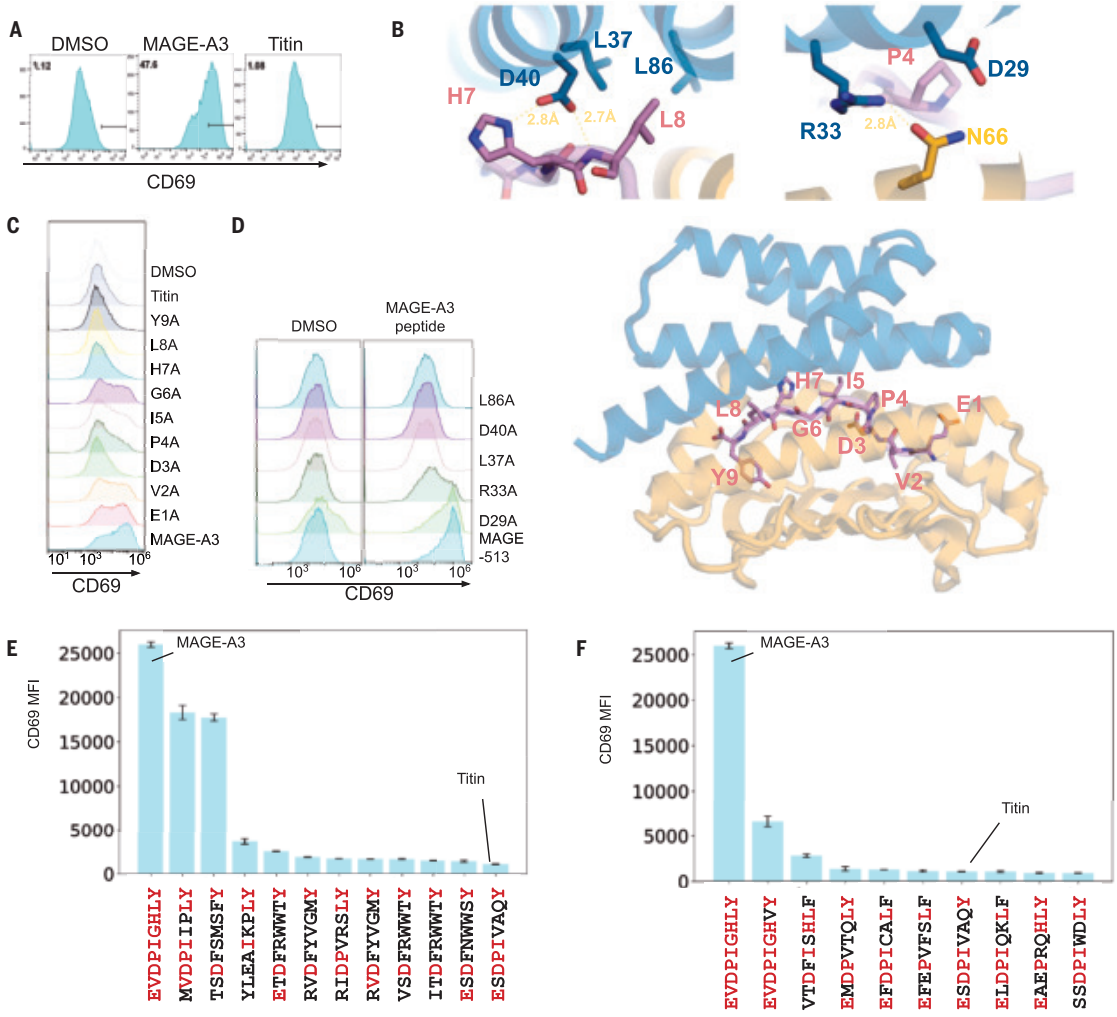


Fig. 3. Selective activation of T cells expressing designed CARs targeting MAGE-A3. Activation of Jurkat cells expressing the mage-513 CAR by 293T cells expressing HLA-A*01:01 pulsed with 5 μM of different peptides measured through CD69 expression level. (**A**) Histograms of CD69 expression levels after treatment with 5 μM MAGE-A3, the closely related Titin peptide, or DMSO; the MAGE peptide leads to considerable activation, whereas the Titin peptide is similar to the DMSO control. Horizontal black bars represent CD69 positive population; the fraction of cells within this range is indicated at top left of each panel. (**B**) Design models of the mage-513/MAGE-A3 peptide/HLA-A*01:01 complex (bottom panel) and zoom-in view of the key residues mediating the interaction (top panels). (**C**) Histograms of CD69 surface expression levels by mage-513 CARs after pulsing 293T cells with MAGE-A3 single alanine mutants (D3A indicates mutation of the Asp at position 3 in the peptide to alanine) or DMSO. (**D**) Histograms of CD69 expression levels after pulsing mage-513 variant CARs with MAGE-A3 peptide or DMSO. (**E** and **F**) CD69 levels of Jurkat cells expressing mage-513 CAR upon pulsing with top-ranked peptides from yeast binding screen (E) or from sequence similarity search (F). Residues that are identical to MAGE-A3 peptide at given positions are indicated in red. Experiments were done in duplicate. Error bars indicate standard deviations. MFI, mean fluorescence intensity.

hydrogen bonds between R33 of the design and N66 of the HLA) (Fig. 3B). D29 of the design is close to P4 of the peptide without making direct contact (Fig. 3B); we concluded that this residue likely played a gatekeeper role by clashing with off-target peptides with bulky residues at site 4. Consistent with the design model and predicted structure, mutating residues in the peptide that make extensive interactions to the binder (I5, H7, L8) to alanine disrupted TCR activation (Fig. 3C). CARs with alanine mutations in designed binder residues (L37, D40, L86), which interact most closely with the peptide, showed reduced activation compared with WT mage-513 upon stimulation (Fig. 3D). Mutation of D29 did not influence Jurkat activation upon pulsing with MAGE peptide but did have higher background activation when coincubated with HLA-A*01:01–expressing 293T cells, likely because of increased cross-reactivity, consistent with a gatekeeper role (Fig. 3D).

To further characterize the specificity of mage-513, we used it to probe a comprehensive library of covalently linked peptide-HLA-A*01:01 complexes displayed on yeast (*33*). The top-hit peptides (data S2) were further tested by target cell pulsing and coincubation with mage-513 CAR–expressing Jurkat cells, and three activating peptides were identified (peptides that bind on yeast but do not activate could be false positives arising from the covalent linkage in the yeast construct). Two of the three had outward-facing side chains similar to those of the MAGE-A3 peptide (Fig. 3E, shown in red). Therefore, we carried out an in silico scan of cross-activating peptides in the human proteome based on sequence similarity (*34*) and tested these for mage-513 CAR activation. The most activating peptides (EVDPIGHVY and VTDFISHLF) were among the most similar, sharing outward-facing residues I5, H7, and L8 (Fig. 3F), consistent with the alanine scanning results.

CARs incorporating binders against the gp100, MART-1, WT1, SARS, and HIV antigens likewise conferred target peptide–selective activation of CD69 signaling in Jurkat cells (Fig. 4, A to F, and fig. S5A). We further probed recognition specificity with peptide alanine scanning experiments. For most targets, pulsing 293T cells with cognate peptide activated at the same level as or at higher levels than pulsing with

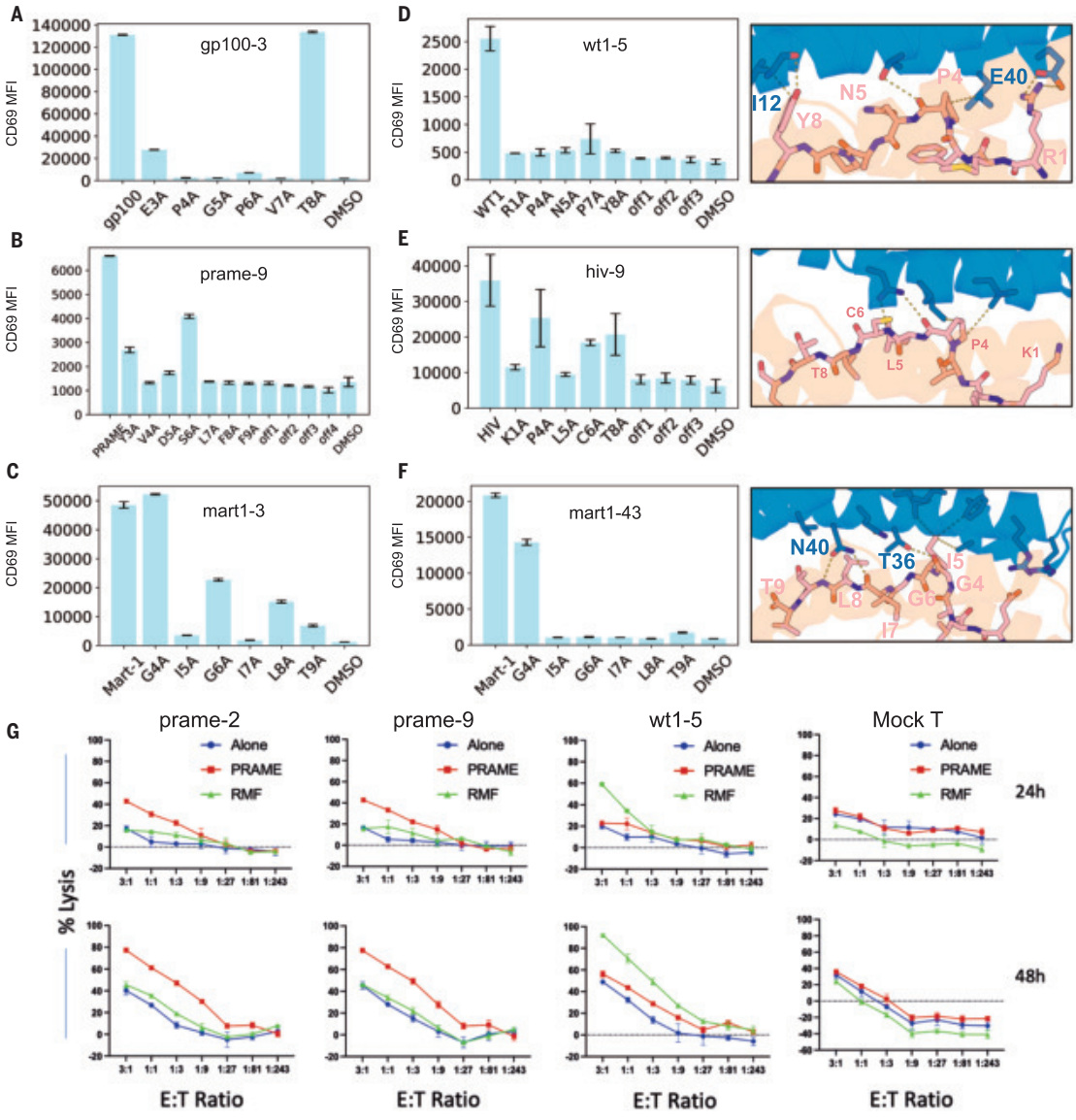


Fig. 4. Specific activation and cell killing by designed CARs in T cells through cognate pMHC complexes. Activation of Jurkat cells expressing the CARs by 293T cells with 5 μM of different peptides measured through CD69 expression level by staining (indicated by histogram or MFI). (**A** to **C**) CD69 MFI of (A) gp100-3, (B) prame-9, and (C) mart1-3 CARs with cognate peptide, alanine mutant peptides, or DMSO. (**D** to **F**) (Left) CD69 MFI of (D) wt1-5, (E) hiv-9, and (F) mart1-43. (Right) Zoom-in view of design model with pMHC target. Experiments done in duplicate. Error bars indicate standard deviations. (**G**) CD3 T cells transduced with indicated binder constructs were incubated with HLA-A*02:01+ T2 cells expressing firefly luciferase and green fluorescent protein and pre-pulsed with either the WT1 peptide (RMF) or PRAME at 10 μg/ml overnight at the indicated E:T ratios. Cytotoxicity was measured for each transduced cell (labeled above graph) by bioluminescence emission after adding substrate luciferin, after 24 or 48 hours of incubation. Green, RMF peptide (RMFPNAPYL) pulsed cells; red, PRAME peptide (ALYVDSLFFL) pulsed cells; blue, unpulsed cells. Each experiment was conducted twice with different human donors. One experiment is shown. Each data point and error bar are the mean ± SEM of triplicate measurements.

alanine peptide variants, and mutational effects were consistent with the design models. For gp100-3, substitution of residues P4, G5, P6, and V7 substantially reduced activation; in the design model of gp100-3 (Fig. 2C), there are multiple hydrogen bonds with the backbone of P6 and V7, which likely require precise positioning, making activation sensitive to mutations at the adjacent P4 and G5 (Fig. 4A). For wt1-5, E40 of the binder makes bidentate interactions with peptide residue R1, and the main chain of binder residue I12 makes a hydrogen bond with the side chain of peptide residue Y8, consistent with the alanine scanning results (Fig. 4D). For the MART-1 target, we performed alanine scanning experiments for two different binders, mart1-3 (Figs. 2D and 4C) and mart1-43 (Fig. 4F), which have binding docks centered over different regions of the peptide; this structural shift is reflected in the alanine scanning results observed. For hiv-9, the design makes hydrogen bonds with the peptide backbone and hydrophobic interactions with P4 and L5, likely holding the peptide in place and leading to higher specificity (Fig. 4E).

Binders against pMHCI targets without experimental structures have high specificity and mediate cell killing

As only a small fraction of potential pMHCI complexes have experimentally determined structures, we also explored the possibility of using predicted pMHCI structures as targets for binder design. The PRAME protein is highly expressed in multiple types of tumors, and a peptide derived from PRAME (ALYVDSLFFL) (*32*) is displayed on HLA-A*02:01. Despite the strong therapeutic relevance and wide patient population coverage, no high-resolution structure has been determined.

For the design, we experimented with using AF3-predicted pMHCI structures as starting points and identified two binders (prame-9 and prame-2) that specifically induced Jurkat activation upon incubation with 293T antigen-presenting cells and PRAME peptide pulsing (Fig. 4B and fig. S5A). In SPR experiments, the purified designs had dissociation constants for the target pMHCI of 96 and 35 nM, respectively (fig. S5B). Pulsing 293T cells with the original PRAME peptide and single alanine mutant version showed that binder prame-9 was quite specific (Fig. 4B). In the design model (Fig. 2C), the binder interacted with the side chains of V4, S6, L7, and F9 and the main chain of L7, consistent with the observed specificity. To test the generalizability of designing against predicted pMHCIs, we also tested CARs against PHOX2B and identified multiple binders, including phox2b-11 (fig. S5A), that enable specific activation upon incubation with 293T incubating with PHOX2B peptide.

To determine whether specific activation can be converted to specific killing, we transduced primary human T cells with CARs incorporating the designed binders against PRAME and WT1 described above. The designed binder CARs induced killing of HLA-A*02:01+ target T2 cells pulsed with cognate target peptide with higher killing efficiencies than noncognate peptide or nonpulsing controls (Fig. 4G), suggesting that designed binders can mediate specific killing.

Discussion

A central challenge in targeting pMHCI complexes is achieving high specificity. Screening TCRs from human T cells takes advantage of the natural repertoire but is limited by availability of donors with relevant HLA alleles, central tolerance that can eliminate TCRs most activated by endogenous targets, and the overall rarity of high-affinity TCRs (*1*, *6*). Engineering existing low-affinity TCRs can increase affinity but requires extra effort to ensure specificity to avoid toxicity (*29*). Screening binders from scFv libraries has primarily yielded nonspecific binders, with specific scFvs adopting similar docking geometry as TCRs (*1*, *3*), suggesting the importance of the docking interface for sufficient peptide contact and binding specificity. These methods have identified specific binders against multiple pMHCIs, but, given the high cost and low hit rate, they are difficult to employ against the large number of potential pMHCI therapeutic targets. Our de novo design approach builds on lessons learned from the binding modes of TCRs and successful scFvs by generating binders with pMHCI interaction interfaces focused on the presented peptide with limited contact with the MHC. We demonstrate that this enables robust design of specific binders for 11 diverse pMHCI targets; because of the simplicity of the structures (compared with scFvs and TCRs), further optimization can be carried out on the basis of the design models and predicted structures without experimental structure determination. Our design pipeline is readily applicable to a wide range of pMHCI targets, allowing for the generation of specific experimentally validated binders within weeks.

We expect the power of our approach to rapidly design specific binders to pMHCIs to continue to increase. First, it should be possible to learn from design campaigns the properties most correlated with specific cell activation and what scaffold geometries give the most effective readouts of sequence over the full peptide length. Second, as deep learning–based structure prediction, design, and model ranking methods improve, it should become possible to find designs with suitable affinities and specificities in testing dozens of candidates. With such advances, it should become possible to readily generate clinical-grade pMHCI binders to provide therapeutic benefits to broad patient populations.

REFERENCES AND NOTES

1. M. Yarmarkovich *et al.*, *Nature* **623**, 820–827 (2023).
2. R. Leidner *et al.*, *N. Engl. J. Med.* **386**, 2112–2119 (2022).
3. E. H.-C. Hsiue *et al.*, *Science* **371**, eabc8697 (2021).
4. P. Nathan *et al.*, *N. Engl. J. Med.* **385**, 1196–1206 (2021).
5. T. Dao *et al.*, *Nat. Biotechnol.* **33**, 1079–1086 (2015).
6. S. S. Chandran *et al.*, *Nat. Med.* **28**, 946–957 (2022).
7. M. Glögl *et al.*, *Science* **386**, 1154–1161 (2024).
8. R. J. Ragotte *et al.*, *Cell Rep.* **44**, 115760 (2025).
9. I. Sappington *et al.*, bioRxiv 2024.10.11.617496 [Preprint] (2024); https://doi.org/10.1101/2024.10.11.617496.
10. C. Liu *et al.*, bioRxiv 2024.07.16.603789 [Preprint] (2024); https://doi.org/10.1101/2024.07.16.603789.
11. K. Wu *et al.*, bioRxiv 2024.07.15.603480 [Preprint] (2024); https://doi.org/10.1101/2024.07.15.603480.
12. L. Cao *et al.*, *Nature* **605**, 551–560 (2022).
13. J. Jumper *et al.*, *Nature* **596**, 583–589 (2021).
14. A. Motmaen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2216697120 (2023).
15. Chai Discovery, bioRxiv 2024.10.10.615955 [Preprint] (2024); https://doi.org/10.1101/2024.10.10.615955.
16. J. Abramson *et al.*, *Nature* **630**, 493–500 (2024).
17. J. Liu *et al.*, *J. Infect. Dis.* **202**, 1171–1180 (2010).
18. A. Bovay *et al.*, *Mol. Immunol.* **125**, 43–50 (2020).
19. P. H. N. Celie *et al.*, *J. Am. Chem. Soc.* **131**, 12298–12304 (2009).
20. D. Bloemberg *et al.*, *Mol. Ther. Methods Clin. Dev.* **16**, 238–254 (2020).
21. C. J. Holland *et al.*, *J. Clin. Invest.* **130**, 2673–2688 (2020).
22. L. Gao *et al.*, *Blood* **95**, 2198–2203 (2000).
23. T. Dao *et al.*, *Sci. Transl. Med.* **5**, 176ra33 (2013).
24. B. M. Olson *et al.*, *Cancer Immunol. Immunother.* **59**, 943–953 (2010).
25. M. S. Miller *et al.*, *J. Biol. Chem.* **294**, 19322–19334 (2019).
26. S. Vázquez Torres *et al.*, *Nature* **626**, 435–442 (2024).
27. B. Gaugler *et al.*, *J. Exp. Med.* **179**, 921–930 (1994).
28. M. C. C. Raman *et al.*, *Sci. Rep.* **6**, 18851 (2016).
29. B. J. Cameron *et al.*, *Sci. Transl. Med.* **5**, 197ra103 (2013).
30. A. B. Bakker *et al.*, *Int. J. Cancer* **62**, 97–102 (1995).
31. Y. Kawakami *et al.*, *J. Exp. Med.* **180**, 347–352 (1994).
32. K. Rezvani *et al.*, *Blood* **113**, 2245–2255 (2009).
33. M. E. Birnbaum *et al.*, *Cell* **157**, 1073–1087 (2014).
34. M. H. Gee *et al.*, *Cell* **172**, 549–563.e16 (2018).
35. 77forest, 77forest/pmhci_binder_design: Pmhci_binder_design, v1.0.0, Zenodo (2025); https://doi.org/10.5281/zenodo.15169815.

ACKNOWLEDGMENTS

We acknowledge A. Minervina and P. Thomas for discussion and sharing plasmids of different HLA alleles; K. M. Jude for structural study; C. Correnti, P. Bradley, L. Tran, A. Swartz, K. Thompson, L. An, B. Coventry, S. Sharma, W. Chen, J. Sims, M. Bauer, and T. Schlichthaerle for helping with computation and experiments; and T. Spear, K. O'Reilly, and J. Maris for discussion. We also acknowledge the Jenkins lab and Hadrup group for their coordination of co-submission. Thank you to K. VanWormer and L. Goldschmidt for their continuous lab and computational support. **Funding:** The Audacious Project at the Institute for Protein Design (J.E.B., N.F.G.); Cancer Grand Challenges team MATCHMAKERS funded by Cancer Research UK [CGCATF-2023/100006 (C.W., K.C.G.), CGCATF-2023/100008 (B.L., D.B.)]; the National Cancer Institute [1OT2CA297242-01 (K.C.G.), 1OT2CA297288-01 (J.E.B., N.F.G., D.B.)]; and The Mark Foundation for Cancer Research (K.C.G., D.B.); Department of Defense, Defense Threat Reduction Agency grant HDTRA (B.L.); Microsoft Protein Prediction (B.L.); Washington Research Foundation Fellowship (N.F.G.); National Institutes of Health grant 2R01AI103867-11 (C.W., K.C.G.); Parker Institute for Cancer Immunotherapy (C.W., K.C.G.); Grants for Resident Innovation and Projects award from the Children's Hospital of Philadelphia (R.A.); National Institutes of Health MIRA award R35 GM142795 (G.M.V., A.N.); The CAREER award from the National Science Foundation no. 2045054 (G.M.V., A.N.); Howard Hughes Medical Institute (K.C.G., D.B.); National Institutes of Health P30 CA008748 and R35 CA241894 (D.A.S.); R50 CA265328 (T.D.). **Author contributions:** Conceptualization: D.B., A.M., N.F.G., B.L.; Computational screening strategies: D.B., N.F.G., B.L., A.M., G.M.V., A.N., J.E.B.; Experimental strategies: D.B., T.D., N.F.G., B.L., A.M., J.M., J.E.B., C.W., D.A.S., K.C.G.; Binder design: J.E.B., N.F.G., B.L., J.S., R.A., A.M.; Experimental testing: J.E.B., N.F.G., B.L., A.M., J.S., J.M., X.X., N.R., D.K.V.; Writing – original draft: D.B., N.F.G., B.L., J.E.B.; Writing – review & editing: D.B., N.F.G., B.L., J.E.B., A.N. **Competing interests:** A provisional patent (application number 63/779,176) covering minibinder sequences presented in this paper has been filed by the University of Washington. K.C.G. is a consultant for Xaira Therapeutics. D.A.S. is on a board of or has equity in or income from: Lantheus, Sellas Life Sciences, Iovance, Pfizer, Actinium Pharmaceuticals, OncoPep, Arvinas, Repertoire, Sapience, Atengen, and Eureka Therapeutics. T.D. is a consultant for Eureka Therapeutics. J.E.B. is a shareholder of Adimab LLC. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials. The mart-1 binder-antigen complex coordinates and structure factors are available in the Protein Data Bank under PDB ID 9O5S/pdb_00009o5s. Code and examples for the computational design of pMHC-I binders are available in Zenodo (*35*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adv0185
Materials and Methods; Figs. S1 to S5; Tables S1 to S3; References (*36*–*39*); MDAR Reproducibility Checklist; Data S1 and S2

Submitted 3 December 2024; accepted 16 May 2025

10.1126/science.adv0185

MARINE PROTECTED AREAS

Little-to-no industrial fishing occurs in fully and highly protected marine areas

Jennifer Raynor[1]*, Sara Orofino[2,3,4,5], Christopher Costello[2,3,4], Gavin McDonald[2,3,4], Juan Mayorga[4,6], Enric Sala[6]

There is a widespread perception that illegal fishing is common in marine protected areas (MPAs) due to strong incentives for poaching and the high cost of monitoring and enforcement. Using artificial intelligence and satellite-based Earth observations, we provide estimates of industrial fishing activity in fully and highly protected MPAs worldwide, in which such fishing is banned. We find little to no activity in most cases. On average, these MPAs had just one fishing vessel present per 20,000 square kilometers during the satellite overpass, a density nine times lower than that of the unprotected waters of exclusive economic zones.

Marine protected areas (MPAs) are a fundamental tool for ocean conservation and are expanding worldwide to meet the Global Biodiversity Framework's target of protecting at least 30% of the ocean by 2030 (30 × 30) (*1*, *2*). MPAs are established with various levels of protection; although some strictly prohibit extractive and damaging activities, many allow fishing, including destructive practices such as bottom trawling (*3*). By contrast, "no-take" or "fully protected" MPAs prohibit all fishing and "highly protected" MPAs ban industrial fishing but allow limited small-scale fishing (*3*). These stricter MPAs are considered the most effective for conservation (*3*, *4*) and are the focus of this study.

Paradoxically, the more successful a protected area is at fostering biological recovery, the greater the incentive to poach. This challenge is particularly pronounced in fully and highly protected MPAs where fish population recovery is highest (*5*); this recovery can benefit nearby fisheries (*5*–*8*) but also increases poaching incentives (*9*, *10*). Enforcing marine protection is costly, and many agencies lack the financial and technological resources to prevent illegal fishing (*11*, *12*).

These enforcement challenges fuel the widespread perception that many MPAs are "paper parks," designated in law but lacking true protection (*13*–*16*). Weak protection stems from two main drivers: First, some MPAs are minimally protected by design as they either permit industrial fishing or exist only as legal designations without implementation (*3*). A recent analysis found that one-third of the world's 100 largest MPAs provide minimal protection, making them "incompatible with the conservation of nature" (*16*). Similarly, 86% of MPAs in the European Union confer little protection (*17*). The second driver is insufficient monitoring and enforcement, which leads to low levels of compliance. Even fully protected MPAs in which all fishing is prohibited face widespread concerns about noncompliance (*13*–*16*). Regarding this issue, (*18*) examined fishing activity across all MPAs globally regardless of their protection level and found substantial fishing within their boundaries. Although their findings suggest that substantial fishing occurs inside MPAs, it remains unclear whether their results are driven by weak regulatory protections (driver 1) or poor compliance with otherwise strict regulations (driver 2).

Our study resolves this puzzle by estimating fishing activity inside the world's most strictly regulated MPAs. We overcome two key limitations of previous assessments: (i) We limit our study to areas with the highest levels of legal protection from industrial fishing, using newly available regulatory data from ProtectedSeas® Navigator (*19*); and (ii) we include MPA implementation dates from (*20*) to identify when fishing restrictions come into force, allowing us to limit our analysis to the period following an MPA's full implementation whenever possible. For example, the Palau National Marine Sanctuary was designated on 1 January 2015, but was not implemented (i.e., fishing was not prohibited) until 1 January 2020; fishing during this 5-year phase-out period was allowed.

To track industrial fishing activity, we leverage artificial intelligence and satellite-based Earth observations. We first measure fishing effort using the Automatic Identification System (AIS), which transmits real-time vessel positions, courses, and speeds to shore-based receivers and satellites (*21*). Machine learning algorithms analyze these signals to predict when vessels are actively fishing (referred to as "apparent" fishing effort), making AIS the primary data source for global industrial fishing activity (*21*). However, AIS has limitations, such as strategic tampering by captains (*22*) and the absence of many vessels from AIS records because they are not required to use the system, have turned off their transponders, and/or are operating in areas with poor signal reception (*23*).

We therefore complement AIS data with synthetic aperture radar (SAR) imagery from the Sentinel-1 satellite constellation. Unlike AIS, which relies on voluntary transmission, SAR actively detects vessels by sending radar pulses to the ocean surface and measuring the reflected signals. Machine learning algorithms then identify vessels from these reflectance patterns and classify them as fishing or nonfishing vessels (*23*). These algorithms detect most industrial vessels—those longer than 15 m—regardless of whether their AIS is active. The only way for vessels to evade detection would be to predict Sentinel-1 flyover times and exit MPAs at those times. However, our study covers 2017 to 2021, before Sentinel-1 was widely used to detect industrial fishing (*23*), which limits the scope for such strategic evasion.

We focus on the post-implementation period for all fully and highly protected MPAs within exclusive economic zones (EEZs) (fig. S1) in which fishing restrictions are in place year-round; 91% of these MPAs (by area) have the regulatory information required for our analysis and were implemented before 1 January 2021, the last year of our study period (fig. S2). Overall, our analysis includes 1380 MPAs spanning 7.9 million km^2, an area covering 2.2% of the global ocean (fig. S3A). Sentinel-1 imaged 638 of these MPAs at least once during the post-implementation period (fig. S3B); however, to limit small sample bias and potential false negatives, we focus on the 455 MPAs with high-frequency temporal and spatial imaging, defined as having at least 20 SAR images and at least 1 km^2 in the SAR footprint (*18*). These 455 MPAs span 3.2 million km^2, of which SAR imaged 0.7 million km^2; most (84%) of the unimaged area falls within three large and remote MPAs (fig. S4, A to C). SAR coverage tends to be relatively close to shore, where Sentinel-1 overpasses are concentrated (figs. S1 and S3); see the supplementary materials (SM) for further details on SAR coverage.

Results

AIS fishing hours

In total, we detected an average of 10,759 apparent fishing hours per year across all MPAs, dominated by a handful of large MPAs (fig. S5A). The most notable outlier is the Chagos Archipelago Marine Protected Area, which has about 2700 apparent fishing hours annually, consistent with previous reports of alleged illegal fishing (*24*, *25*) (figs. S5B and S6). Other high-effort MPAs include the South Georgia and the South Sandwich Islands and the Great Barrier Reef Marine Park (combined

[1]Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI, USA. [2]Bren School of Environmental Science and Management, University of California, Santa Barbara, Santa Barbara, CA, USA. [3]Marine Science Institute, University of California, Santa Barbara, Santa Barbara, CA, USA. [4]Environmental Markets Lab, University of California, Santa Barbara, Santa Barbara, CA, USA. [5]TNC: The Nature Conservancy, California Oceans Program, San Diego, CA, USA. [6]Pristine Seas, National Geographic Society, Washington, DC, USA. *Corresponding author. Email: jennifer.raynor@wisc.edu

with the Great Barrier Reef Coast Marine Park), each averaging around 900 hours per year. Several others—Camargue, Orange Shelf Edge Marine Protected Area, Pulau Seri Buat and Pulau Sembilang Marine Parks, Estrecho Island, Marado Island, Cape Canyon Marine Protected Area, and Northeast Canyons and Seamounts Marine National Monument—recorded between ~275 and 450 hours annually.

A closer examination of MPA-specific regulations reveals that some apparent fishing activity may be consistent with permitted uses while other cases suggest potential noncompliance. For instance, in the Northeast Canyons and Seamounts, the red crab and American lobster fleets were permitted to operate under a sunset provision during this period (*26*) and most detected fishing occurred in the canyons unit (fig. S6J), which overlaps with known crab and lobster habitat (*27*); however, none of the apparent fishing effort was classified as pots and traps, which are commonly used to target these species. Similarly, although small-scale professional gillnetting is allowed in Camargue, bottom-trawling is prohibited (*28*). Yet, 99.5% of the apparent fishing effort there was classified as trawling and only 0.3% as gillnets, indicating possible noncompliance (fig. S6D).

In contrast to these high-activity sites, most MPAs showed minimal or no fishing. Specifically, 1325 out of 1380 MPAs recorded less than 1 day (24 hours) of apparent fishing effort per year whereas 1084 MPAs registered no fishing activity (fig. S7A). However, because MPA size is highly skewed—just 20 MPAs account for 91% of total protected area—and most large MPAs show at least some fishing, the 96% of MPAs with 0 to 24 hours per year represent only 30% of total MPA area. This highlights the need for a standardized measure of "fishing density" within MPAs, described below.

To compare activity across MPAs that differ vastly in size, we use apparent fishing effort density, defined as average daily hours per 100 km^2 (Fig. 1A). By this density metric, Marado Island (Republic of Korea) had the highest apparent fishing effort (18 hours per day per 100 km^2), followed by Cha-gwido Island (Republic of Korea; 5 hours per day per 100 km^2), and Cat Harbor State Marine Conservation Area (USA; 4 hours per day per 100 km^2) (Fig. 2A and fig. S8). The remaining MPAs in the top 10 all had about 1 hour per day per 100 km^2. Yet, across all MPAs, the area-weighted average density was only 0.0004 hours per day per 100 km^2; this low number demonstrates that even in MPAs that have some apparent fishing effort, most of the protected area has minimal activity.

Overall, our AIS results show little to no apparent fishing effort in MPAs that prohibit industrial fishing, with the notable exceptions of a few large MPAs. However, the ability to tamper with AIS signals casts some doubt on relying solely on AIS, which motivates our SAR analysis.

SAR vessel detections

Our SAR analysis corroborates and extends the AIS results, revealing minimal industrial fishing vessel presence across most MPAs. One-quarter of MPAs (many of which are small and represent only 5% by area) had no fishing vessel detections (figs. S7B and S9A). But even among MPAs with some activity, detections were rare—the median MPA in this group recorded a fishing vessel on just 3% of imaged days (fig. S9B). Across all MPAs, we observed an average of 0.005 fishing vessel detections per 100 km^2 [95% confidence interval (CI): 0.003, 0.007], or approximately one detection per 20,000 km^2 (fig. S9C). By contrast, the fishing vessel density in unprotected areas of EEZs (fig. S1) was nearly nine times higher, at 0.042 detections per 100 km^2. This unprotected EEZ average lies well outside the 95% CI for MPAs, reinforcing the contrast in fishing vessel presence between protected and unprotected areas. These results are qualitatively robust to (i) using alternative methods to classify vessel detections as fishing or nonfishing vessels and (ii) removing the SAR sampling frequency restrictions. See the SM for further details.

Although a number of MPAs had high fishing vessel densities (primarily in East and South Asia; Figs. 1B and 2B), these were driven by a small number of vessel detections in relatively small MPAs. The highest densities occurred in three MPAs in the Republic of Korea—Munseom and Beomseom Islands, Gyeongnam Goseong County Finless Porpoise Water, and Sohwang Sand Dune Protected Area—each with about 2 to 3 fishing vessel detections per 100 km^2 (Fig. 2B). However, in each case, we detected only 6 to 18 fishing vessels in total (fig. S10, A to C), and fishing vessels were absent on 94 to 97% of imaged days. Because these MPAs have a small area imaged (~2 km^2 each), even a handful of detections results in a high average density. Only seven MPAs (1% of area imaged) had a fishing vessel detected on a majority of imaged days, including sites in Indonesia,

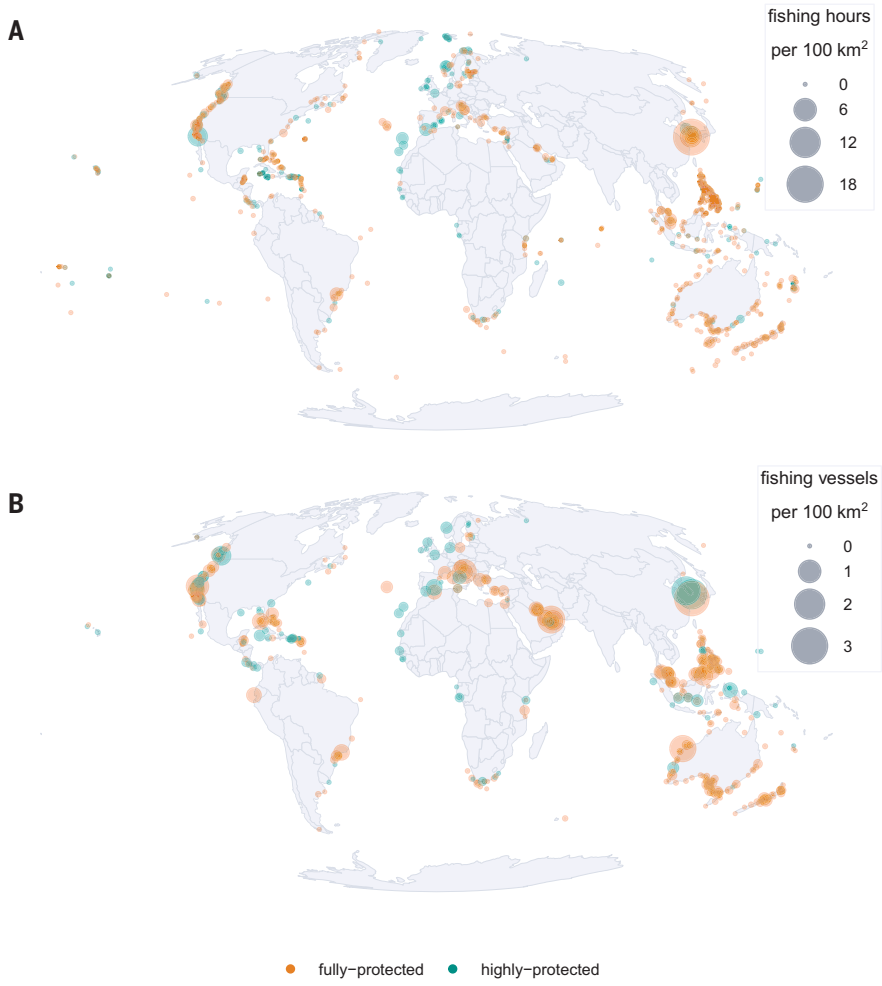


Fig. 1. Industrial fishing activity in fully and highly protected MPAs. (**A**) AIS apparent fishing effort density (average daily fishing hours per 100 km^2). (**B**) SAR fishing vessel density (fishing vessel detections per 100 km^2 at the average SAR flyover time). Point locations represent MPA centroids, constrained to occur within the MPA.

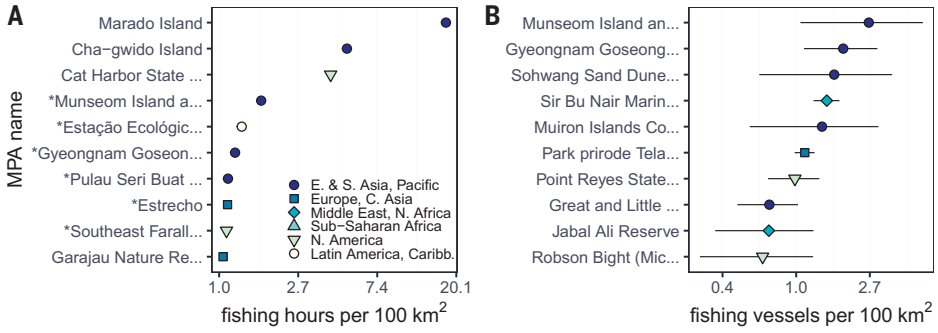


Fig. 2. Industrial fishing activity in fully and highly protected MPAs, by MPA rank. (**A**) AIS apparent fishing effort density (average daily fishing hours per 100 km^2). (**B**) SAR fishing vessel density (fishing vessel detections per 100 km^2 at the average SAR flyover time). Asterisks in (A) indicate MPAs that have high frequency SAR sampling (i.e., at least 20 images and 1 km^2 within the SAR footprint). Lines in (B) represent the 95% CI. For both panels, axes in are log-scaled but values are not log-transformed. Long MPA names are truncated for readability, indicated by ellipsis; full names are available in the captions of figs. S8 and S10.

the United Arab Emirates, the Philippines, Malaysia, Costa Rica, Ecuador, and Brazil.

Unlike AIS, SAR data is unlikely to miss substantial industrial fishing activity where it is occurring. Globally, the underlying algorithms detect and classify AIS-broadcasting fishing vessels with high accuracy (F1 scores of 0.97 and 0.91), and we find no evidence that performance is worse within MPAs (figs. S11 and S12). Among the few MPAs (<1% by area) where AIS indicated fishing activity but SAR did not, apparent effort was low overall (<6 hours per year for 23 out of 25 MPAs), suggesting that little meaningful fishing activity was missed by SAR. By contrast, SAR revealed vessel presence in 163 MPAs with no AIS activity, particularly in Southeast Asia, where AIS data are often incomplete as a result of poor reception or low adoption (*29*), indicating that SAR may be a valuable tool for filling known AIS data gaps. Finally, 91 MPAs had no detections from either method, providing high confidence that no industrial fishing occurred in those areas. More detailed comparisons of AIS and SAR fishing metrics are available in the SM.

Discussion

Our study focuses on MPAs that explicitly prohibit industrial fishing. SAR imagery detects an average of just one industrial fishing vessel per 20,000 km^2 in these areas, a density about nine times lower than in the unprotected waters of EEZs. This finding challenges the commonly held view that industrial fishing is widespread in MPAs (*13*–*16*) and undermines arguments to degazette MPAs or limit MPA expansion due to perceived ineffectiveness.

Our results refine and substantially expand on two recent studies that use SAR-based vessel detections in MPAs: The first study (*23*) develops the SAR vessel detection and classification algorithms and demonstrates their capability to track industrial fishing activity in two MPAs, the Galápagos Marine Reserve (which permits commercial fishing) and the Great Barrier Reef Marine Park. Although aligned in approach, our study is global in scope and distinguishes between levels of fishing protection, enabling us to assess fishing in places that specifically prohibit it. The second study (*18*) estimates fishing activity across MPAs of all protection levels, including those that allow fishing. As a result, they report an average fishing density within global MPAs (0.038 vessels per 100 km^2) that is almost eight times larger than our estimate for fully and highly protected MPAs (0.005; 95% CI: 0.003 to 0.007) and comparable to our estimate for unprotected EEZs (0.042) (fig. S13A). A key distinction between these studies lies in how protection levels are defined. We use actual regulatory data whereas (*18*) relies on self-reported International Union for Conservation of Nature (IUCN) management categories that often poorly reflect real protections. Despite this, (*18*) does find lower fishing activity in stricter IUCN categories, which is qualitatively consistent with our findings (fig. S13B). However, even their lowest estimate (0.009 fishing vessels per 100 km^2 in category III) is roughly double our estimate for fully and highly protected MPAs, underscoring the limitations of using IUCN categories as regulatory proxies. Overall, although (*18*) is complementary to our study, its inclusion of MPAs that allow fishing means that it cannot isolate activities in strictly protected areas as our study does.

Our analysis underscores the importance of using accurate data for MPA boundaries, levels of protection, and implementation dates (*30*, *31*). Without these critical data, estimates of fishing activity inside MPAs may be overestimated. Even small spatial delineation errors are problematic because vessels often fish near MPA boundaries to take advantage of spillover (*32*). Similarly, using legal designation dates instead of implementation dates could misclassify apparent fishing effort during phaseout periods as allegedly illegal. This issue is particularly concerning if fishing intensifies during the period following MPA announcement and before protections take effect (*33*). We address these data challenges and provide reliable estimates of fishing activity in MPAs, laying the groundwork for more informed and effective MPA management.

Even our finding of relatively little fishing inside MPAs may overstate alleged illegal activity for two main reasons: First, although the underlying machine learning models have high accuracy (*21*, *23*), they are subject to some classification error. Therefore, these datasets can aid monitoring and enforcement but generally do not constitute evidence of a crime in legal proceedings. Second, some MPAs include sunset provisions that allow certain fleets to continue fishing legally for a period of time after implementation. Despite these caveats, these datasets provide strong evidence of whether industrial fishing activity is likely occurring inside MPAs with the highest levels of regulatory fishing protection.

However, further exploration is needed to understand the underlying causal mechanisms for our findings. One possibility is that there is high demand to poach in these areas, but strict monitoring and enforcement effectively excludes fishing. This hypothesis is underpinned by the finding that fish biomass recovers inside some MPAs (*5*–*8*), particularly in those that are actively managed and strongly regulated (*3*, *4*). However, a competing hypothesis is that MPAs are located in places with little demand for fishing. This is underpinned by evidence that the strongest MPA regulations tend to be located in places where fishing pressure is already low (*34*, *35*). These competing hypotheses have different implications for the external validity of our results and therefore for future MPA designation and management. If past MPAs were sited in places with low fishing pressure, then future MPA expansions in more crowded locations may not have similarly low fishing activity.

One way to estimate these effects is to examine changes in fishing activity following MPA implementation using appropriate causal inference methods. However, most fully or highly protected MPAs were established before SAR data became available, limiting our ability to perform such analyses. Only 72 MPAs (5% by count, 19% by area) have at least one full year of pre-implementation data and only 47 of these were imaged by SAR; this is not a representative sample of global MPAs. Still, we can conduct some suggestive analysis with this limited subset. Both AIS and SAR data indicate that most of these MPAs had minimal fishing activity both before and after implementation:

61 MPAs (36% by area) had 0 to 1 day of total AIS fishing effort and 45 MPAs (99% by area imaged) had 0 to 35 total SAR fishing vessel detections in each period (fig. S14). A notable exception is the Palau National Marine Sanctuary (31% by area), which saw average annual apparent fishing effort drop from 51,000 hours to just 215 after implementation. Although these results are somewhat mixed, they suggest that many MPAs may have been placed in areas with little prior fishing. Distinguishing whether MPAs reduce fishing effort or are simply located in low-demand areas is essential for assessing their conservation impact.

Moreover, excluding industrial fishing may itself be insufficient to recover marine ecosystems in MPAs (*3*). Importantly, many MPAs allow artisanal fishing, which can deplete fish and invertebrate populations (*36*). SAR would usually fail to detect such vessels if they are smaller than 15 m. Although this does not undermine our ability to track industrial fishing, it prevents us from concluding that the ecosystems inside these MPAs are necessarily healthy or free from all fishing pressure.

As MPAs expand globally, efficient monitoring tools will be needed to support enforcement efforts and evaluate MPA effectiveness. Our approach—translating near-real-time vessel detections from multiple complementary data sources into actionable metrics of apparent fishing pressure—provides enforcement agencies with a cost-effective, scalable monitoring solution. These tools enhance monitoring not only within MPAs but also in exclusive economic zones, artisanal fishing zones, and other space-time closures. Moreover, by increasing the likelihood of detecting illegal fishing, they create a deterrent effect. As the likelihood of detecting illegal fishing increases, the incentive to fish illegally decreases, and as satellite technologies continue to improve, the capability to monitor even smaller vessels will further enhance effectiveness of these tools.

REFERENCES AND NOTES

1. E. Dinerstein *et al.*, *Sci. Adv.* **5**, eaaw2869 (2019).
2. Convention on Biological Diversity, "Kunming-Montreal Global Biodiversity Framework" (CBD/COP/15/L.25, UN environment programme, 2022); https://www.cbd.int/doc/c/e6d3/cd1d/daf663719a03902a9b116c34/cop-15-l-25-en.pdf.
3. K. Grorud-Colvert *et al.*, *Science* **373**, eabf0861 (2021).
4. B. Horta e Costa *et al.*, *Cell Rep. Sustain.* **2**, 100345 (2025).
5. S. Lester *et al.*, *Mar. Ecol. Prog. Ser.* **384**, 33–46 (2009).
6. S. Medoff, J. Lynham, J. Raynor, *Science* **378**, 313–316 (2022).
7. J. Lynham, J. C. Villaseñor-Derbez, *Science* **386**, 1276–1281 (2024).
8. M. J. Costello, *Sci. Mar.* **88**, e080 (2024).
9. K. D. Millage *et al.*, *Environ. Res. Lett.* **16**, 125001 (2021).
10. C. González-Andrés, J. L. Sánchez-Lizaso, J. Cortés, M. G. Pennino, *Mar. Policy* **119**, 104023 (2020).
11. B. Bergseth, J. C. Day, *Mar. Policy* **155**, 105728 (2023).
12. D. A. Gill *et al.*, *Nature* **543**, 665–669 (2017).
13. G. J. Edgar *et al.*, *Nature* **506**, 216–220 (2014).
14. M. Dureuil, K. Boerder, K. A. Burnett, R. Froese, B. Worm, *Science* **362**, 1403–1407 (2018).
15. V. Relano, D. Pauly, *Mar. Policy* **151**, 105571 (2023).
16. E. P. Pike *et al.*, *Conserv. Lett.* **17**, e13020 (2024).
17. J. Aminian-Biquet *et al.*, *One Earth* **7**, 1614–1629 (2024).
18. R. Seguin, F. Le Manach, R. Devillers, L. Velez, D. Mouillot, *Science* **388**, eado9468 (2025).
19. Navigator Data Download, ProtectedSeas® (2025); https://navigatormap.org/data-request.
20. G. McDonald, J. Bone, C. Costello, G. Englander, J. Raynor, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2400592121 (2024).
21. D. A. Kroodsma *et al.*, *Science* **359**, 904–908 (2018).
22. H. Welch *et al.*, *Sci. Adv.* **8**, eabq2109 (2022).
23. F. S. Paolo *et al.*, *Nature* **625**, 85–91 (2024).
24. C. Collins *et al.*, *Front. Mar. Sci.* **8**, 650276 (2021).
25. J. L. Harris, G. M. W. Stevens, *Mar. Policy* **163**, 106110 (2024).
26. U.S. President, Proclamation, "Northeast Canyons and Seamounts Marine National Monument, 2016, Proclamation 9496 of September 21, 2016," Federal Register 81, no. 183 (September 21, 2016): 65161–65167, https://www.govinfo.gov/content/pkg/FR-2016-09-21/html/2016-22921.htm
27. P. J. Auster, B. C. Hodge, M. P. McKee, S. D. Kraus, *Front. Mar. Sci.* **7**, 566 (2020).
28. L. Le Direach *et al.*, "Création d'une réserve marine dans le golfe de Beauduc. Etat initial. Années 2014-2015," *Partenariat Parc naturel régional de Camargue & GIS Posidonie publ.* (France, 2015), p. 79.
29. M. Taconet, D. Kroodsma, J. A. Fernandes, "Global Atlas of AIS-based fishing activity - Challenges and opportunities" (Food and Agriculture Organization of The United Nations, 2019); .
30. S. Lippi, C. Piroddi, M. Graziano, A. Di Franco, *Front. Mar. Sci.* **11**, 1369447 (2024).
31. J. Aminian-Biquet *et al.*, *Npj Ocean Sustain.* **4**, 3 (2025).
32. J. B. Kellner, I. Tetreault, S. D. Gaines, R. M. Nisbet, *Ecol. Appl.* **17**, 1039–1054 (2007).
33. G. R. McDermott, K. C. Meng, G. McDonald, C. J. Costello, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5319–5325 (2019).
34. R. Devillers *et al.*, *Aquat. Conserv.* **25**, 480–504 (2015).
35. R. Devillers *et al.*, *Aquat. Conserv.* **30**, 1758–1764 (2020).
36. L. Schiller, J. J. Alava, J. Grove, G. Reck, D. Pauly, *Aquat. Conserv.* **25**, 431–446 (2015).
37. J. Raynor, S. Orofino, C. Costello, J. Mayorga, E. Sala, Little-to-no industrial fishing occurs in fully and highly protected areas, Zenodo (2025); https://doi.org/10.5281/zenodo.15866359.
38. Marine Conservation Institute [and Partners], MPAtlas (Marine Conservation Institute, 2025); https://marine-conservation.org/mpatlas/.
39. Office of National Marine Sanctuaries, Marine Protected Areas (MPA) Inventory 2023-2024, Inport (2025); https://www.fisheries.noaa.gov/inport/item/69506.
40. P. Wessel, W. H. F. Smith, GSHHG, Version 2.3.7 (2017); https://www.soest.hawaii.edu/pwessel/gshhg/index.html.
41. UNEP-WCMC, IUCN, Protected Planet: The World Database on Protected Areas (WDPA, 2024); https://www.protectedplanet.net/en.
42. Flanders Marine Institute, Maritime Boundaries Geodatabase, version 12 (2023); https://www.marineregions.org/.
43. R Core Team, R: The R Project for Statistical Computing (R Foundation for Statistical Computing, 2024); www.R-project.org/
44. W. Landau, *J. Open Source Softw.* **6**, 2959 (2021).
45. K. Ushey, H. Wickham, renv: Project Environments,R package version 1.1.4 (2023); https://cran.r-project.org/web/packages/renv/index.html
46. H. Wickham *et al.*, *J. Open Source Softw.* **4**, 1686 (2019).
47. E. Pebesma, *R J.* **10**, 439 (2018).
48. E. Pebesma, R. Bivand, *Spatial Data Science: With Applications in R* (Chapman & Hall/CRC, 2023), ed. 1

ACKNOWLEDGMENTS

ProtectedSeas provided data on MPA boundaries and classifications. Global Fishing Watch provided data for AIS fishing effort and SAR vessel detections. J. Sletten, T. Vincent, V. Zetterlind, and D. Brannigan at ProtectedSeas and D. Kroodsma, F. Paolo, T. Clavelle, and Z. Wei at Global Fishing Watch provided invaluable technical expertise and analytical support for using these key datasets. K. Wang, K. Kim, and J. Park provided assistance with evaluating MPA classifications in the Republic of Korea. **Funding:** This work was funded by the following: National Geographic Pristine Seas supported J.R., S.O., C.C., J.M., and E.S. J. Arnhold supported C.C. and G.M. **Author contributions:** Conceptualization: C.C. and E.S. Data curation: J.R., S.O., and G.M. Formal analysis: J.R., S.O., C.C., and G.M. Funding acquisition: E.S. and C.C. Investigation: J.R., S.O., and C.C. Methodology: J.R., S.O., and C.C. Project administration: J.R. and C.C. Software: J.R., S.O., and G.M. Supervision: J.R. and C.C. Validation: J.R., S.O., and G.M. Visualization: J.R., S.O., and G.M. Writing – original draft: J.R. Writing – review & editing: J.R., S.O., C.C., G.M., J.M., and E.S. **Competing interests:** J.R. and E.S. are current trustees and C.C. is a former trustee of Global Fishing Watch, one of the primary data providers for this study. C.C. is currently employed at Environmental Defense Fund. All other authors declare that they have no competing interests. **Data and materials availability:** All code necessary for fully reproducing the analysis is available at Zenodo (*37*). Most data necessary for reproducing the analysis are also contained directly within this repository. These include: ProtectedSeas MPA area boundaries, levels of fishing protection, and designation dates (*19*); The Marine Protection Atlas MPA boundaries and levels of fishing protection (*38*); National Oceanic and Atmospheric Administration Marine Protected Areas Inventory boundaries (*39*); Global Self-consistent Hierarchical High-resolution Geography (GSHHG) land mass data (*40*); and all AIS apparent fishing effort (*21*) and SAR fishing vessel detection (*23*) point locations within fully and highly protected MPAs from Global Fishing Watch. The only datasets not available directly in the repository but which can be publicly downloaded are the World Database on Protected Areas boundaries [available from (*41*)], Marine Regions Version 12 EEZ boundaries [available from (*42*), and European MPA boundaries and protection levels [available from (*17*)]. For our analysis we used the R programming language version 4.4.1 (*43*). We managed pipeline reproducibility using the targets package (*44*) and package reproducibility using the renv package (*45*). We extensively used the tidyverse suite of packages for data wrangling operations (*46*) and the sf (*47*) and stars (*48*) packages for spatial operations. All spatial area and distance calculations were performed using the Mollweide equal-area projection.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adt9009
Materials and Methods; Supplementary Analyses; Figs. S1 to S15; References (*49*–*53*); MDAR Reproducibility Checklist

Submitted 15 October 2024; accepted 25 May 2025

10.1126/science.adt9009

MARINE PROTECTED AREAS

Global patterns and drivers of untracked industrial fishing in coastal marine protected areas

Raphael Seguin[1,2]*, Frédéric Le Manach[2], Rodolphe Devillers[3], Laure Velez[1], David Mouillot[1]

Marine protected areas (MPAs) are expanding worldwide, but industrial fishing within their boundaries remains poorly understood. By combining vessel Global Positioning System data with satellite imagery, we detected industrial fishing in 47% of coastal MPAs worldwide between 2022 and 2024, with two-thirds of vessel detections (67%) untracked by public monitoring. These untracked vessels were most prevalent (80%) in MPAs with the most restrictive management category, which aims to prohibit all extractive activities. The presence and density of fishing vessels were mainly driven by the size and remoteness of MPAs rather than their management category itself. Including untracked vessels increased estimated fishing effort within MPAs by 28%, to an average of 4 hours per square kilometer per year. These results highlight the continued presence of industrial fishing vessels in MPAs and an important gap in our understanding of fishing pressure on protected areas worldwide.

Marine protected areas (MPAs) are an essential tool to counteract the pervasive degradation of marine ecosystems and sustain their contributions to humankind (*1*–*3*). In 2022, guided by scientific recommendations, 196 countries endorsed the objective to effectively protect 30% of the ocean by 2030 (*4*, *5*). Although >9% of the ocean is currently covered by MPAs, <3% is fully or highly protected as of July 2025, despite these levels being recognized as the most effective for biodiversity conservation or restoration (*6*–*13*). More precisely, the International Union for Conservation of Nature (IUCN) classifies MPAs into six management categories ranging from strictly protected areas that exclude all extractive activities (categories I and II) to those allowing some nonindustrial activities (categories V and VI) along a gradient of restrictions. Overall, industrial fishing activities are excluded from all MPA categories according to the IUCN guidelines (*14*, *15*). However, although MPAs are effective when these guidelines are well implemented, many MPAs still legally allow industrial fishing within their boundaries (*16*–*18*), which can result in higher fishing pressure on threatened and protected species inside compared with those outside (*19*, *20*). Such extractive activities limit the ability of MPAs to counteract threats to biodiversity, livelihoods, and climate change (*14*, *17*, *18*, *21*–*23*).

At the regional level, many governments use proprietary vessel monitoring systems (VMS)—closed-data tracking systems—to monitor specific fishing fleets within their territorial waters (*24*). Since 2015, a new publicly available data stream based on signals emitted by automatic identification systems (AIS) has enabled more transparent, global-scale tracking of industrial fishing activities (*24*, *25*), including within MPAs (*20*). However, not all vessels are required to use AIS, many conceal their activities by switching off their AIS device, and poor satellite reception in some areas creates "blind spots" across the ocean (*26*–*28*). As a result, most of the global fishing fleet remains untracked (*27*, *29*).

Recent advances in the detection of industrial fishing activities—achieved by coupling synthetic aperture radar (SAR) satellite imagery, AIS data, and artificial intelligence—reveal that AIS-untracked fishing fleets are more widespread than previously thought across the world's coastal waters (*29*). As a result, MPAs that appear free of AIS-tracked industrial vessels may still face substantial fishing pressure from untracked vessels (*29*). Given the detrimental impacts of industrial fishing—tracked or untracked—on MPA effectiveness and associated benefits, there is an urgent need to assess and better understand the extent of industrial fishing activities within MPAs.

In this study, we used a recently published dataset combining satellite imagery, AIS data, and deep learning models (*29*) to quantify tracked and untracked fishing vessel detections within the world's coastal MPAs. We identified socioeconomic and environmental variables associated with the presence and density of both AIS-tracked and -untracked vessels. Because the presence of vessels alone does not necessarily indicate fishing activity, we also predicted fishing effort within coastal MPAs worldwide using a two-stage hurdle model that incorporated AIS-derived fishing effort, vessel detections, and a comprehensive set of socioeconomic and environmental variables.

The industrial fishing fleet in the world's MPAs

Using Global Fishing Watch data (materials and methods), a total of 2,002,273 industrial fishing vessel detections (potentially representing multiple instances of the same vessel) were obtained from coastal SAR images worldwide between 2022 and 2024, of which 185,962 (9.3%) were within MPAs (Fig. 1). Owing to the incomplete spatial coverage of SAR imagery for most large and remote MPAs (*29*), this study focuses exclusively on coastal MPAs larger than 1 km^2, which make up 41.2% of designated MPAs and 17.4% of global MPA coverage, where most industrial fishing is concentrated (*29*–*31*) (figs. S1 to S5). Of the 6021 coastal MPAs considered in this study, 2836 (47.1%) were visited by industrial fishing vessels detected through SAR satellite imagery, 53.1% lacked a reported management plan, and 52% did not declare an IUCN category (materials and methods, figs. S5 to S9, and tables S1 and S2).

Because coastal areas are not all imaged with the same effort, we divided the number of vessel detections by the number of SAR images taken in 2022 to 2024 over each location to obtain standardized vessel detections. The density of fishing vessel detections inside each MPA was then calculated by dividing this standardized detection count by the MPA's surface area. The density of fishing vessel detections was significantly higher within IUCN-declared MPAs compared with "unprotected" areas of exclusive economic zones (EEZs) (*P* value of Dunn test < 0.001; Fig. 2A, fig. S10, and table S3). However, given that many MPAs are small, a low number of vessel detections can result in a high vessel density per area, skewing the average density within MPAs upward. The number and coverage of MPAs varied widely among countries (fig. S11). Notably, countries with a high gross domestic product (GDP) tend to have a greater coastal MPA coverage, coinciding with an increased likelihood of vessel tracking and detection (*8*, *25*). Thus, whereas countries such as China, India, Japan, or North Korea showed the highest number of fishing vessel detections within their EEZs (figs. S12 and S13), Japan, France, the United Kingdom, and Spain had the largest number of detections within MPAs (figs. S14 and S15). EEZs with the highest average density of detections inside MPAs were those of Belgium, the Netherlands, and China (figs. S16 to S20). The proportion of detections within EEZs found inside MPA boundaries ranged from 0% (mostly countries with low fishing or little to no MPAs) to 64% (Germany)—i.e., 64% of all fishing vessel detections in the German EEZ were inside MPAs—with a global mean of 7% (SD = 18%; Fig. 2B).

We then focused on untracked fishing vessel detections—i.e., vessels appearing on satellite imagery but whose locations could not be matched with AIS records (Fig. 1 and Fig. 2C). Between 2022 and 2024, a total of 1,205,675 untracked vessel detections were identified within and outside MPAs. Untracked vessels were present in 90% of the MPAs in which vessels were detected (2539 of 2836 MPAs; fig. S21 and table S4). There was a strong positive correlation between the number of untracked and tracked vessel detections within MPAs [correlation

[1]MARBEC, Université de Montpellier, CNRS, IFREMER, IRD, Montpellier, France. [2]BLOOM, Paris, France. [3]Espace-Dev (IRD, Université de Montpellier, Université de Guyane, Université de La Réunion, Université de Nouvelle Calédonie), Montpellier, France. *Corresponding author. Email: raphaelseguin@protonmail.com

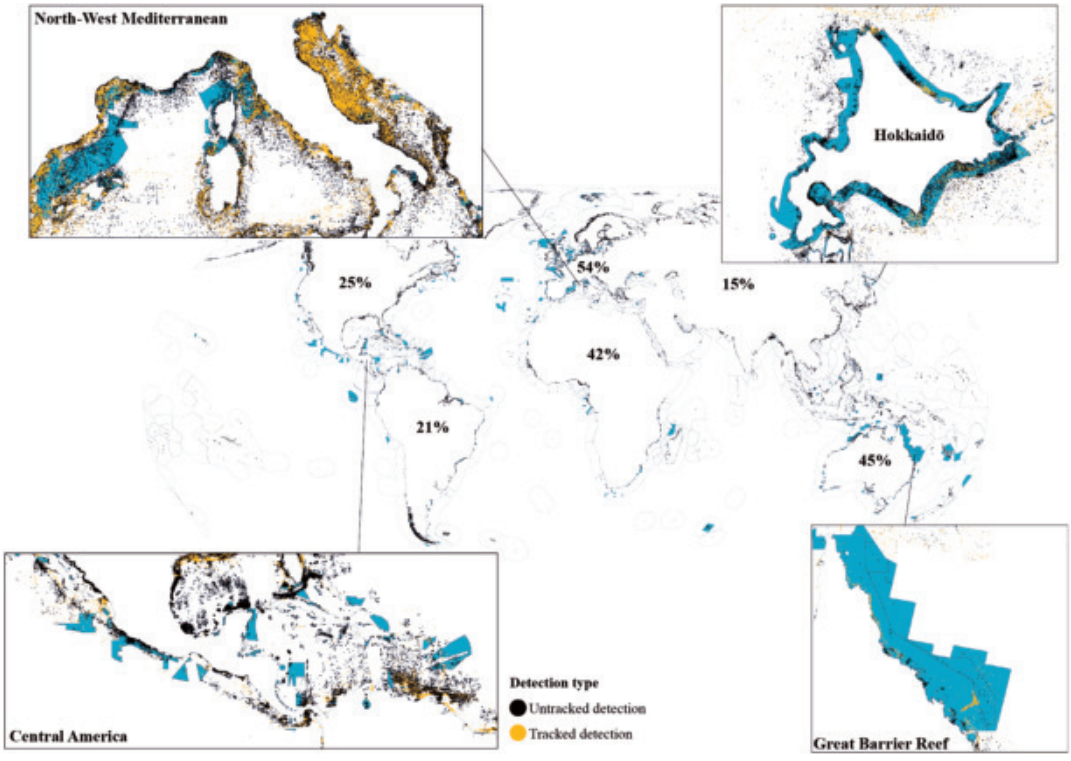


Fig. 1. High-resolution satellite imagery reveals tracked and untracked industrial fishing vessel detections inside coastal MPAs. Global map shows MPAs (blue areas) overlapping with available SAR imagery. For each continent, we highlight the percentage of AIS-tracked fishing vessel detections (North America includes Central American countries). For the inset maps, orange dots represent tracked fishing vessel detections, and black dots represent untracked detections.

coefficient (r) = 0.76, $P < 0.001$; Fig. 2C]. The percentage of untracked detections within MPAs ranged from 0% (i.e., all detections were tracked) to 100% (i.e., all detections were untracked), with an average of 67% (SD = 35%). In comparison, 68% of detections in unprotected EEZs were untracked on average (SD = 28%). The highest percentage of untracked detections was observed in MPAs of IUCN category I, where 80% of all detections were untracked (SD = 33%), whereas the lowest was for MPAs of IUCN category VI, where 64% of all detections were untracked (SD = 38%). The percentage of untracked detections found inside MPAs within EEZs ranged from 0 to 79% for Germany, which means that 79% of all untracked detections were inside MPAs, with a mean of 10% (SD = 18%; figs. S22 to S25).

Because MPAs vary widely in terms of size, location, and context, we modeled the presence or absence of industrial fishing vessels in MPAs with a declared IUCN category (n = 2858 MPAs) using a binomial model with a spatial autocorrelation term, 11 socioenvironmental variables, and MPA characteristics (materials and methods and figs. S26 to S28). The model was highly accurate [area under the curve (AUC) = 0.79; tables S5 to S7] and revealed that, as expected, MPA size was by far the strongest predictor of fishing vessel presence (t = 16.2, $P < 0.001$), with larger MPAs more likely to contain detections, whereas MPAs situated further away from cities were less likely to be visited by vessels (t = −3.9, $P < 0.001$). Therefore, although the percentage of MPAs with no vessel detections was higher for the most restrictive IUCN categories I and II (74% and 66%, respectively) compared with that for MPAs of IUCN categories V and VI (52%), we found no significant effect of the IUCN management category on the presence of industrial fishing within MPAs when accounting for size, location, and the socioenvironmental context (Fig. 3A), except for a weak effect of IUCN category II (t = 2.24, P = 0.03). Hence, MPAs less visited by industrial fisheries primarily tend to be smaller and/or remote (table S8 and figs. S29 and S30).

We then modeled the standardized number of fishing vessel detections in MPAs with a declared IUCN category using a gamma regression model and the same variables, including the spatial term. The model was accurate [coefficient of determination (R^2) = 0.79, n = 1076] and revealed that the number of vessel detections was positively associated with MPA size and sea surface temperature (t = 29 and 3.3, respectively; $P < 0.001$; Fig. 3A). Furthermore, the number of fishing vessel detections significantly decreased with a longer travel time from the nearest city (t = −6.9, $P < 0.001$; Fig. 3B). Although the average density of vessel detections was lower for the most restrictive IUCN categories I and II (2.7×10^{-4} fishing vessel detections per square kilometer) compared with that for MPAs of IUCN categories V and VI (5.9×10^{-4} and 7.4×10^{-4} fishing vessel detections per square kilometer, respectively), we again found no significant effect of the IUCN management category on the number of detections within MPAs (Fig. 3B, fig. S31, and table S9). Similar results were obtained when modeling the number of untracked fishing vessel detections only (Fig. 3A and table S10).

Estimating the real fishing effort in the world's MPAs

Fishing vessel detections from SAR cannot provide a direct measure of active fishing effort because vessels may transit through MPAs without fishing. Nevertheless, 89% of the MPAs with vessel detections between 2022 and 2024 (2544 MPAs) exhibited fishing activity on the basis of AIS-based fishing effort provided by Global Fishing Watch (*32*). The number of tracked detections displayed a strong positive correlation with the number of fishing hours inside MPAs (r = 0.82, $P < 0.001$, on a log-log scale), which was similar across large marine ecosystems (LMEs) and IUCN categories (figs. S32 and S33). Furthermore, fishing effort can be accurately predicted using satellite-based vessel monitoring data and a set of environmental and socioeconomic variables (*33*). We thus predicted the presence and the level of fishing effort performed by both tracked and untracked industrial vessels in the world's MPAs by including untracked detections (materials and methods).

According to AIS data, 53% of the 5875 coastal MPAs used for prediction experienced industrial fishing effort between 2022 and 2024, with more

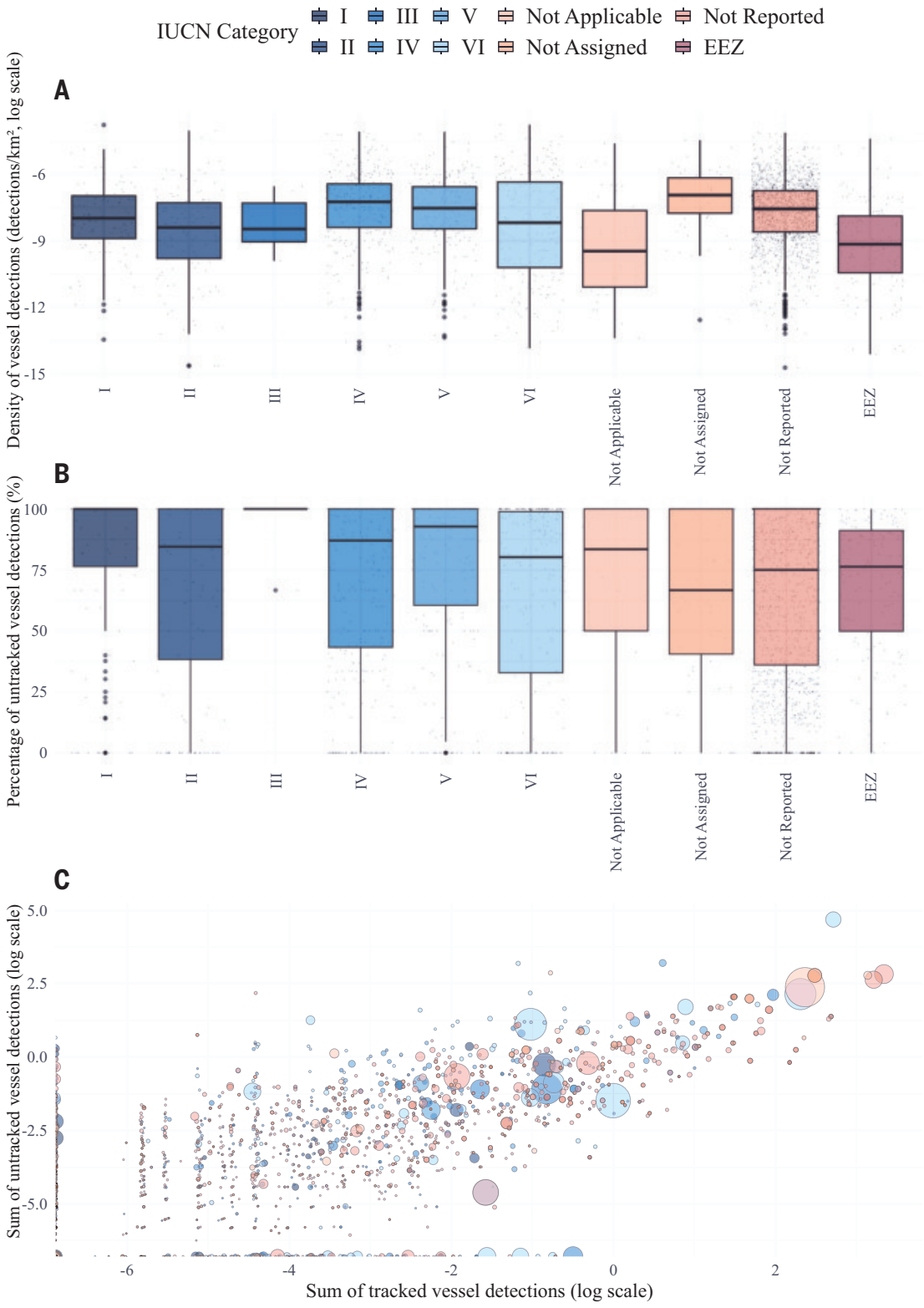


Fig. 2. Density and size distribution of fishing vessel detections inside and outside coastal MPAs with at least one fishing vessel detection worldwide. (**A**) Density of vessel detections (log scale, excluding MPAs with zero detections) in each IUCN category and EEZs without MPAs. Density is the number of detections divided by the number of satellite images in an MPA or EEZ and the corresponding surface, providing a standardized number of detections per SAR acquisition per square kilometer. Each dot represents one MPA, and only MPAs with at least one detection are plotted. (**B**) Percentage of untracked fishing vessel detections compared with tracked vessel detections inside and outside MPAs. Each dot represents one MPA, excluding MPAs with zero detections. (**C**) Relationship between the sum of tracked and untracked fishing vessel detections inside MPAs on a log-log scale. Each dot represents one MPA, dot size represents MPA size, and dot color represents the IUCN category. In each boxplot, each dot represents an MPA. The black horizontal line is the median, and the box edges (hinges) mark the 25th and 75th percentiles. The whiskers extend to the largest and smallest values within 1.5× the interquartile range from the hinges.

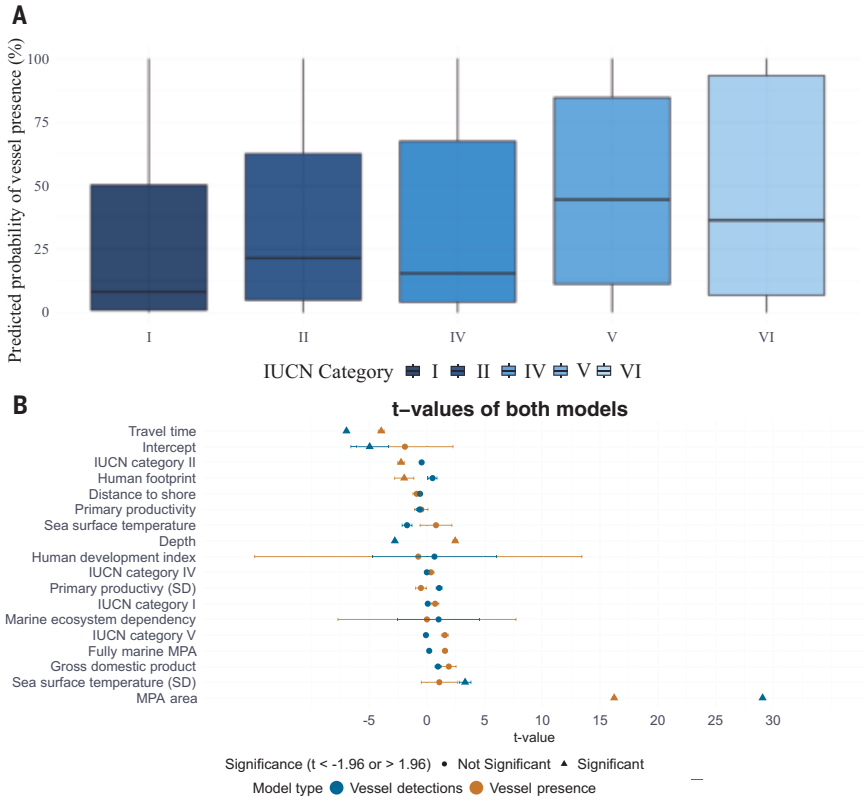


Fig. 3. Results of the models predicting the presence and number of tracked and untracked fishing vessel detections inside coastal MPAs of different IUCN categories. (**A**) Predicted probabilities of fishing vessel presence within MPAs across IUCN categories using the fitted presence-absence binomial model (n = 2858). (**B**) Effect sizes and associated confidence intervals for the models' variables predicting the number of untracked fishing vessel detections (blue) and tracked fishing vessel detections (orange), with shapes representing significance according to $t < -1.96$ or $t > 1.96$ ($P < 0.05$). Models were computed on all MPAs with a declared IUCN category (n = 1076 for all vessels, n = 935 for untracked vessels). For all models, we removed MPAs in category IUCN III owing to the small number of MPAs in this category in our dataset.

than 6 million hours of fishing every year. The first binomial model (n = 5875), based on AIS-based fishing effort 1 year before (e.g., 2023 for 2024), accurately predicted the presence of industrial fishing effort in MPAs for a given year (F1 = 0.84, accuracy = 0.84, averaged across 10-fold spatial cross-validation; fig. S34 and tables S11 and S12) using a set of environmental and socioeconomic variables. The presence of tracked vessel detections had a major contribution to the model, with MPAs containing such detections having a higher probability of being fished (t = 10.3, 8.9, and 7.4 in 2022, 2023, and 2024, respectively; $P < 0.001$; fig. S35).

The regression model predicting the level of industrial fishing effort inside MPAs, in terms of hours, was highly accurate, with an average R^2 of 0.84 across 10-fold spatial cross-validations between 2022 and 2024 (fig. S36 and tables S13 and S14). To test the model's ability to predict fishing effort outside of its training range, we performed an out-of-sample procedure, predicting fishing effort on a subset of MPAs not used for training but having values higher than the 90th quantile (n = 131, 124, and 122 for 2022, 2023, and 2024, respectively). Again, the model had a high cross-validation performance (R^2 = 0.94 between 2022 and 2024), which shows its capacity to predict fishing effort outside of the range of values that it has been trained on. The number of tracked vessel detections also significantly correlated with fishing effort (t = 14.7, 14.6, and 13.9 in 2022, 2023, and 2024, respectively; $P < 0.001$; figs. S37 and S38).

We then included untracked fishing vessels in our fitted model by updating our variables representing the presence and number of tracked vessel detections with the total number of detections (tracked and untracked). We could thus predict the presence of industrial fishing effort for all MPAs with at least one untracked vessel detection and for which no fishing effort was detected according to AIS and the level of fishing effort for all MPAs with at least one untracked vessel detection. In 2022, the percentage of MPAs with fishing effort increased from 54 to 59% when including untracked fishing vessels (289 more MPAs with industrial fishing effort), from 52 to 56% in 2023 (259 more MPAs), and from 53 to 57% in 2024 (233 more MPAs). Including untracked fishing vessels increased the level of fishing effort by 28% inside MPAs worldwide (5,406,822 hours), reaching 9,495,789 hours in 2022, 7,753,741 hours in 2023, and 7,426,704 in 2024, for a cumulative total of 24,676,234 hours (Fig. 4 and fig. S39). The countries with the highest additional predicted untracked fishing effort inside MPAs were France, Germany, and the United Kingdom, with 1,092,804, 846,052, and 812,684 hours, respectively. The highest percentage increases were in Japan (79%), the United States (70%), Germany (49%), and Belgium (47%) (Fig. 4A and fig. S40). For MPAs with a declared IUCN category, the highest predicted levels of untracked fishing effort were observed in MPAs of IUCN categories V and VI, with an additional 701,689 and 529,257 hours, respectively, of fishing predicted in 2022 to 2024, and the lowest predicted levels were observed in MPAs of categories III and I, with 410 and 11,337 hours, respectively (Fig. 4 and figs. S41 and S42).

Because MPAs span a large range of surface areas, we also calculated the density of predicted fishing effort as the average fishing effort performed by both tracked and untracked vessels per year divided by MPA size. Fishing effort density inside MPAs ranged from 0 to 210 hours km^{-2} $year^{-1}$ (mean = 3.9; SD = 11 hours km^{-2} $year^{-1}$). The highest average densities were observed in MPAs categorized as V and VI (mean = 2.34 and 2.30; SD = 6.86 and 10.32 hours km^{-2} $year^{-1}$), and the lowest in IUCN categories I and II (mean = 0.87 and 1.37; SD = 5.19 and 8.00 hours km^{-2} $year^{-1}$; Fig. 4C).

Discussion

Using a global dataset combining AIS and SAR imagery, we assessed the presence and activity of industrial fishing fleets across the world's coastal MPAs encompassing both publicly and nonpublicly tracked vessels, the latter making up 72 to 76% of the world's industrial fishing vessels (*29*). Our findings suggest that the density of industrial fishing vessel detections was significantly higher inside MPAs compared with that in unprotected national jurisdictions. This result is not new, at least for European MPAs (*18*–*20*), and is not so counterintuitive because many industrial fisheries legally operate with MPAs, irrespectively of IUCN guidelines (*17*, *18*); coastal MPAs cover some of the most favorable and profitable fishing grounds, such as nurseries or reproduction spots (*34*); and, when illegal, the potential gain of fishing often outweighs the potential cost of being caught (*35*). Industrial fishing vessels were detected in 47% of the world's coastal MPAs. IUCN categories I and II, the most restrictive management categories, had a lower proportion of MPAs with fishing vessel detections and lower average densities of detections compared with categories V and VI, which suggests an apparent barrier to industrial fisheries (*36*). However, after controlling for MPA remoteness, size, and other socioenvironmental factors, our model found little to no effect of IUCN management category on the presence or density of industrial fishing vessels. This highlights that, globally, the spatial patterns of industrial fishing fleets within MPAs are mainly driven by their remoteness or size, which limit access to fishing

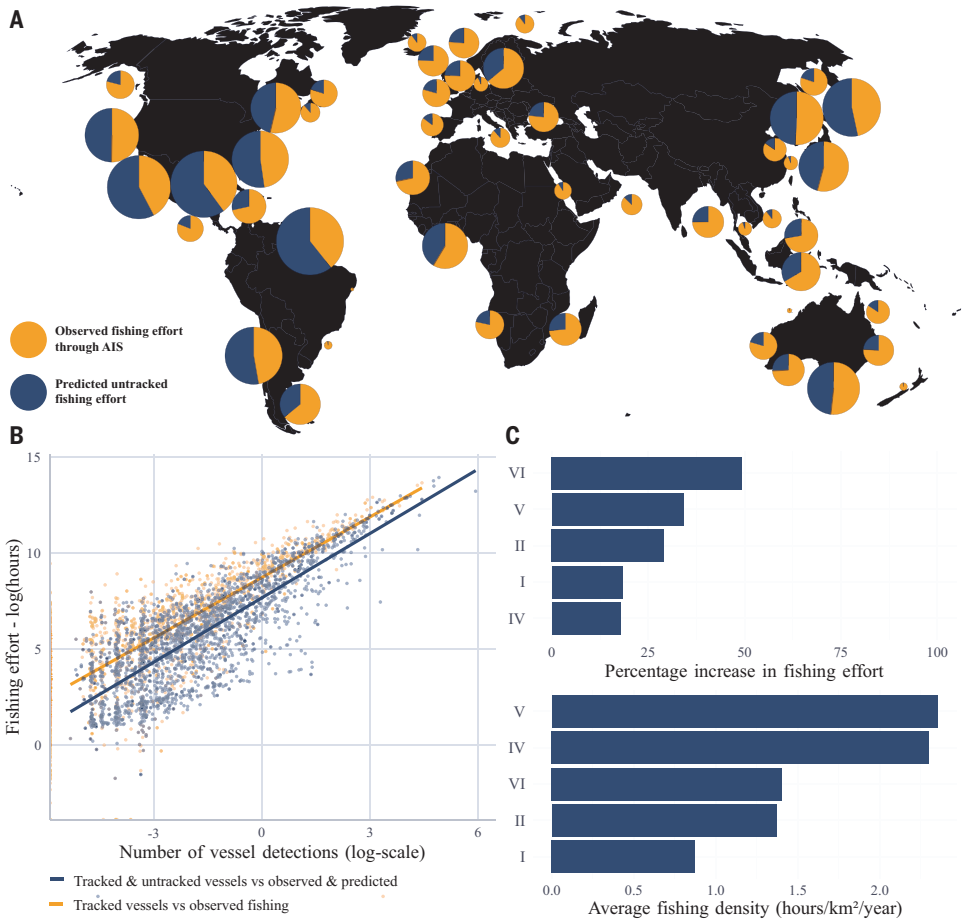


Fig. 4. Predictions of industrial fishing effort across the world's coastal MPAs using both tracked and untracked vessel detections. (**A**) Tracked and untracked predicted fishing effort across LMEs. Each circle corresponds to one LME, and circle size is proportional to the percentage of increase in fishing effort when including untracked fishing vessels detected with satellite imagery. Within each circle, the pie chart shows the proportion of tracked (orange) versus untracked (blue) fishing effort. (**B**) Correlation between the number of tracked vessel detections and the observed fishing effort through AIS (in orange) and the cumulated tracked and untracked number of detections and the predicted fishing effort (in blue), with each dot representing an MPA. (**C**) Percentage increase in fishing effort when including the untracked fishing fleet and average annual density of observed and predicted fishing effort (hours per square kilometer per year) in 2022 to 2024 for all IUCN categories with at least 30 MPAs.

grounds, rather than by self-declared IUCN categories. In other words, although these categories do exhibit lower levels of industrial fishing (*36*), the real drivers of this apparent protection seem related to the MPA's placement rather than the IUCN management category. This provides empirical support to the suggestion that many fully and highly protected areas are "residual" (*37*)—i.e., strategically placed in areas that are either remote or of little interest to industrial activities to minimize conflict, which allows for easy political victories and conservation gains in terms of protection coverage but no change to "business as usual" for the marine industry (*38*). These placement strategies minimize conservation effectiveness, as they tend to provide little protection to the species and ecosystems most exposed to anthropogenic threats (*37*, *39*).

Furthermore, the true number of fishing vessels is likely underestimated within MPAs for three main reasons. First, we only detect vessels larger than 15 m, and most vessels between 15 and 20 m are missed by Sentinel-1 SAR imagery (*29*), which means that small-scale fisheries—responsible for >40% of global fish catches (*31*)—are almost entirely excluded, even though they almost entirely operate in the geographical area studied in this work, i.e., coastal waters (*31*). Second, SAR images are typically acquired around 6 a.m. or 6 p.m. local time, which provides a narrow temporal window rather than continuous daily coverage. Moreover, owing to the limited spatial coverage of SAR imagery, only coastal MPAs were considered, whereas many fully and highly protected MPAs, typically large and remote (*17*), could not be captured in this study but are potentially under fishing pressure as well. Third, we use a highly conservative threshold to classify vessel detections as fishing vessels, which means that our results underestimate the true number of fishing vessels and thus represent a lower bound of industrial fishing within MPAs (supplementary text). The inclusion of higher-resolution satellite imagery with more-frequent revisit periods and larger geographic footprints, such as Sentinel-2 or Planet Labs, could improve the detection of smaller vessels and increase the spatial and temporal coverage of fisheries in the near future (*27*). Finally, although IUCN categories are known to be imperfect proxies for MPA effectiveness (*14*), they remain the main authoritative and comparative classification system for MPAs documented at the global level on the basis of intended management objectives (*26*) and can provide important explanatory power in the modeling of conservation benefits (*40*). However, given the discrepancy between self-declared IUCN management categories and actual observed protection from industrial fishing at a global scale, and given the fact that more than half of the world's coastal MPAs had no declared IUCN category, our results show that relying solely on IUCN categories to assess MPA effectiveness is misleading and highlight the need for more direct and standardized measures of protection, such as the Protected Seas classification, to accurately assess MPA effectiveness (*6*, *11*). A study using the Protected Seas framework, AIS-based fishing effort, and vessel detections from SAR imagery found little to no industrial fishing activity inside fully and highly protected MPAs, which suggests that, at least regarding industrial fishing, implemented regulations are largely being complied with (*36*). Our assessment of vessel density and effort within MPAs may also contribute to revisiting IUCN categories and contextualizing compliance with their regulations.

Our findings also indicate that untracked fishing vessels are present in most MPAs globally, with 90% of MPAs with tracked vessel detections also containing untracked detections. On average, 67% of all vessel detections within MPAs are untracked, a proportion similar to unprotected EEZs (68%). However, the percentage of untracked vessels peaks in IUCN category I MPAs (80%), which are intended to ban all extractive activities (*15*). This result could suggest illegal activities or a discrepancy between the self-declared IUCN management category and the actual protection status implemented at the national level (*17*, *18*). Furthermore, even if the presence of fishing vessels does not mean that they are actively fishing, their mere presence has been shown to hinder efficient protection of marine predators, with no-entry MPAs being the most effective kind of protection for shark populations (*41*–*43*). Beyond direct exploitation, industrial fishing vessels cause noise pollution, which can affect the behavior, physiology, and reproduction of most marine animals (*44*).

Although SAR imagery does not directly measure fishing effort, given that some vessels may simply be in transit and not actively fishing, our analysis shows a strong positive correlation between vessel detections and AIS-based fishing effort. Our models estimate a 28% increase in fishing effort in the world's MPAs when accounting for untracked vessels, if we assume that untracked vessels contribute equally to fishing effort as tracked vessels, meaning that the same proportion of vessels is in transit or in fishing activity. Because the model is trained on previously tracked AIS fishing effort, it will inherently learn patterns from fleets and regions with high AIS tracking, whereas regions with low AIS tracking may have systematically lower reported effort. These regions are likely to have high levels of untracked fishing (e.g., Southeast Asia and West Africa), which means that the model may underestimate untracked effort in these areas (Fig. 4). However, this approach allowed us to predict more realistic fishing effort values at a global scale, even if some local situations may contradict the observed trend. Therefore, although our predictions are cautious estimates, the strong correlation between vessel detections and fishing effort highlights a substantial gap in the current understanding of fishing activity. Our findings suggest that industrial fishing effort inside MPAs has been underestimated worldwide, even in regions with high AIS coverage, such as Europe (*26*).

The establishment of MPAs often prioritizes the achievement of international agreements, conservation targets, and political objectives, without substantive progress toward genuine ocean protection (*39*, *45*–*48*). Simply designating an area as protected without clear legal definitions, a standardized protection framework, and enforcement capacity may result in little benefit toward the intended goals of MPAs—namely, protecting biodiversity (*7*, *8*, *49*), sustaining artisanal fisheries (*2*, *50*–*54*), alleviating poverty (*55*), and preserving natural carbon sinks (*2*, *56*). To avoid confusion and counterproductive claims around conservation target achievements, the IUCN has developed guidelines delineating what constitutes an MPA (*57*). According to IUCN guidelines, all MPAs, regardless of their IUCN category, should prohibit all industrial activities, including industrial fishing (*58*). Given the lack of barriers against industrial fisheries provided by even MPAs with strict biodiversity conservation targets (IUCN categories I and II), we suggest that adherence to these MPA standards is imperative to realize optimal benefits for marine life. Protected Seas complements these IUCN guidelines and proposes an efficient and science-based framework to categorize and evaluate new or existing MPAs, where all categories ban industrial activities (*8*, *36*). We argue that MPAs failing to meet these standards should not be counted toward international targets because they cannot effectively safeguard biodiversity (*59*) or standing populations of reproductive individuals (*60*). Adhering to and complying with established MPA standards should result in more tangible benefits for biodiversity conservation, artisanal fisheries, food security, livelihoods, and climate change mitigation.

REFERENCES AND NOTES

1. C. Marcos *et al.*, *Front. Mar. Sci.* **8**, 613819 (2021).
2. E. Sala *et al.*, *Nature* **592**, 397–402 (2021).
3. A. J. Nowakowski *et al.*, *Nat. Sustain.* **6**, 1210–1218 (2023).
4. B. C. O'Leary *et al.*, *Conserv. Lett.* **9**, 398–404 (2016).
5. A. C. Hughes, R. E. Grumbine, *Front. Environ. Sci.* **11**, 1281536 (2023).
6. O. Aburto-Oropeza *et al.*, *PLOS ONE* **6**, e23601 (2011).
7. E. Sala, S. Giakoumi, *ICES J. Mar. Sci.* **75**, 1166–1168 (2018).
8. S. Lester *et al.*, *Mar. Ecol. Prog. Ser.* **384**, 33–46 (2009).
9. S. Giakoumi *et al.*, *Sci. Rep.* **7**, 8940 (2017).
10. P. P. Molloy, I. B. McLean, I. M. Côté, *J. Appl. Ecol.* **46**, 743–751 (2009).
11. M. J. Costello, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2412543121 (2024).
12. D. A. Gill *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2313205121 (2024).
13. L. Sanchez *et al.*, *Ecol. Lett.* **27**, e14418 (2024).
14. K. Grorud-Colvert *et al.*, *Science* **373**, eabf0861 (2021).
15. J. Day *et al.*, *Guidelines for Applying the IUCN Protected Area Management Categories to Marine Protected Areas* (IUCN, ed. 2, 2019).
16. M. J. Costello, B. Ballantine, *Trends Ecol. Evol.* **30**, 507–509 (2015).
17. E. P. Pike *et al.*, *Conserv. Lett.* **17**, e13020 (2024).
18. J. Aminian-Biquet *et al.*, *One Earth* **7**, 1614–1629 (2024).
19. A. L. Perry, J. Blanco, S. García, N. Fournier, *Front. Mar. Sci.* **9**, 811926 (2022).
20. M. Dureuil, K. Boerder, K. A. Burnett, R. Froese, B. Worm, *Science* **362**, 1403–1407 (2018).
21. C. Costello *et al.*, *Nature* **588**, 95–100 (2020).
22. C. C. Hicks *et al.*, *Nature* **574**, 95–98 (2019).
23. G. Mariani *et al.*, *Sci. Adv.* **6**, eabb4848 (2020).
24. D. A. Kroodsma *et al.*, *Science* **359**, 904–908 (2018).
25. D. J. McCauley *et al.*, *Science* **351**, 1148–1150 (2016).
26. M. Taconet, D. Kroodsma, J. A. Fernandes, *Global Atlas of AIS-Based Fishing Activity: Challenges and Opportunities* (Food and Agriculture Organization of the United Nations, 2019).
27. J. Park *et al.*, *Sci. Adv.* **6**, eabb1197 (2020).
28. H. Welch *et al.*, *Sci. Adv.* **8**, eabq2109 (2022).
29. F. S. Paolo *et al.*, *Nature* **625**, 85–91 (2024).
30. C. Warren, D. J. Steenbergen, *Ocean Coast. Manage.* **202**, 105498 (2021).
31. X. Basurto *et al.*, *Nature* **637**, 875–884 (2025).
32. Global Fishing Watch, Global AIS-based Apparent Fishing Effort Dataset, version 3.0.0, data set, Zenodo (2025). https://doi.org/10.5281/zenodo.14982712
33. G. McDonald, J. Bone, C. Costello, G. Englander, J. Raynor, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2400592121 (2024).
34. D. Belhabib *et al.*, *Fish Fish.* **21**, 379–392 (2020).
35. C. González-Andrés, J. L. Sánchez-Lizaso, J. Cortés, M. G. Pennino, *Mar. Policy* **119**, 104023 (2020).
36. J. Raynor *et al.*, *Science* **389**, 392–395 (2025).
37. R. Devillers *et al.*, *Aquat. Conserv.* **25**, 480–504 (2015).
38. D. Mouillot *et al.*, *Nat. Commun.* **15**, 9007 (2024).
39. R. Devillers *et al.*, *Aquat. Conserv.* **30**, 1758–1764 (2020).
40. L. Benedetti-Cecchi *et al.*, *Nat. Commun.* **15**, 1822 (2024).
41. A. C. Nisi *et al.*, *Science* **386**, 870–875 (2024).
42. J.-B. Juhel *et al.*, *Sci. Rep.* **9**, 2897 (2019).
43. A. J. Frisch, J. R. Rizzari, *Front. Ecol. Environ.* **17**, 145–150 (2019).
44. C. M. Duarte *et al.*, *Science* **371**, eaba4658 (2021).
45. G. J. Edgar *et al.*, *Nature* **506**, 216–220 (2014).
46. B. S. Halpern, *Nature* **506**, 167–168 (2014).
47. V. Relano, D. Pauly, *Mar. Policy* **151**, 105571 (2023).
48. A. N. Rife, B. Erisman, A. Sanchez, O. Aburto-Oropeza, *Conserv. Lett.* **6**, 200–212 (2013).
49. IPBES, Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, version 1, Zenodo (2019). https://doi.org/10.5281/zenodo.6417333
50. S. D. Gaines, S. E. Lester, K. Grorud-Colvert, C. Costello, R. Pollnac, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18251–18255 (2010).
51. B. S. Halpern, S. E. Lester, J. B. Kellner, *Environ. Conserv.* **36**, 268–276 (2009).
52. A. Hastings, L. W. Botsford, *Science* **284**, 1537–1538 (1999).
53. J. Lynham, A. Nikolaev, J. Raynor, T. Vilela, J. C. Villaseñor-Derbez, *Nat. Commun.* **11**, 979 (2020).
54. M. J. Costello, *Sci. Mar.* **88**, e080 (2024).
55. S. Desbureaux *et al.*, *Conserv. Lett.* **17**, e13048 (2024).
56. C. M. Roberts *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6167–6175 (2017).
57. J. A. Fitzsimons, *Conserv. Lett.* **4**, 340–345 (2011).
58. IUCN, *IUCN Resolutions, Recommendations and other Decisions* (IUCN, 2016).
59. J. E. Cinner *et al.*, *Science* **368**, 307–311 (2020).
60. J. Hadj-Hammou *et al.*, *Nat. Commun.* **15**, 6105 (2024).

ACKNOWLEDGMENTS

The World Database on Protected Areas (WDPA) provided data on MPA boundaries and classifications. Global Fishing Watch provided data for AIS fishing effort and SAR vessel detections. We thank Global Fishing Watch and D. Kroodsma for their technical expertise and counsel to use these datasets. We thank the four reviewers for their thoughtful and constructive feedback, which greatly improved the quality of this manuscript. **Funding:** This study benefited from the IA-Biodiv ANR project FISH-PREDICT (ANR-21-AAFI-0001-01) and the BioBoost+ project funded by Biodiversa+, a European Partnership cofunded by the European Union under Horizon Europe, notably through financial support from the ANR and the OFB. **Author contributions:** Compilation of environmental and socioeconomic data: L.V., R.S.; Data analysis: R.S., D.M.; Study design: R.S., R.D., D.M.; Writing and editing: R.S., R.D., F.L.M., L.V., D.M. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All vessel and fishing effort data are freely available through the Global Fishing Watch data portal at https://globalfishingwatch.org/datasets-and-code. All data used in the models are publicly available following the links in the supplementary materials. Codes are available at https://github.com/RaphSeguin/MPA_untracked_fishing.

This research was funded in whole or in part by the French National Research Agency (ANR) (ANR-21-AAFI-0001-01), a cOAlition S organization. The author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ado9468
Materials and Methods; Supplementary Text; Figs. S1 to S42; Tables S1 to S14; References (*61*–*71*); MDAR Reproducibility Checklist

Submitted 7 May 2024; resubmitted 26 September 2024; accepted 25 May 2025

10.1126/science.ado9468

SIGNAL TRANSDUCTION

Molecular glues that facilitate RAS binding to PI3Kα promote glucose uptake without insulin

Koji Terayama[1]†, Shinji Furuzono[1]†, Nicole Fer[2], Wupeng Yan[2], Lucy C. Young[3], Daniel J. Czyzyk[2], Ruby Goldstein de Salazar[3], Masato Sasaki[4], Akihiro Uozumi[1], Masahiro Konishi[1], Shoichi Kanda[1], Yoshitaka Sogawa[1], Mitsuhiro Yamaguchi[5], Takashi Tsuji[5], Junichi Kuroyanagi[5], Mayumi Hayashi[6], Yuji Ogura[6], Dhirendra K. Simanshu[2]*, Kazuishi Kubota[6]*, Jun Tanaka[1]*, Frank McCormick[2,3]*

While exploring strategies to control blood glucose concentrations in diabetes, we identified so-called molecular glues D223 and D927 that promote glucose uptake in the absence of insulin. They act by increasing the binding affinity of phosphoinositide 3-kinase α (PI3Kα) catalytic subunit p110α to canonical small guanosine triphosphatase RAS proteins and to RRAS, RRAS2, and MRAS by three orders of magnitude. The compounds bind to the RAS-binding domain of p110α, stabilizing the secondary structures of the PI3Kα in a RAS-binding conformation and forming direct interactions with RAS residues tyrosine-40 and arginine-41. In vivo, D927 mimicked the effects of insulin: It rapidly lowered blood glucose concentrations, enhanced glucose metabolism in normal and Zucker fatty rats, and improved hyperglycemia in models of type 1 and type 2 diabetes, even in insulin-deficient diabetic animals.

Phosphoinositide 3-kinase α (PI3Kα) has a prominent role in various aspects of signal transduction and insulin homeostasis. Activating mutations in the *PIK3CA* gene—which encodes the catalytic subunit (p110α) of PI3Kα—occur at high frequency in malignancies (*1*), mosaic tissue overgrowth syndromes (*2*), venous malformations (*3*), and brain abnormalities associated with epilepsy (*4*). Consequently, several small-molecule inhibitors of PI3Kα are being developed (*5*), with one, alpelisib, approved by the Food and Drug Administration (FDA) in combination with the estrogen receptor antagonist fulvestrant for the treatment of hormone receptor–positive, human epidermal growth factor receptor 2–negative (HER2−) metastatic breast cancers harboring *PIK3CA* mutations and for *PIK3CA*-related overgrowth spectrum. Additionally, a small-molecule PI3Kα activator has been described for potential use in protecting cardiac tissue from ischemia and reperfusion injury (*6*), situations in which acute activation of the PI3Kα pathway promotes cell survival and tissue regeneration.

In skeletal muscle, heart, and adipose tissues, PI3Kα transduces signals from the insulin receptor to 3-phosphoinositide–dependent kinase 1 (PDK1) and AKT, promoting the translocation of glucose transporter type 4 (GLUT4) to the plasma membrane to facilitate glucose uptake. Defects in this process lead to insulin resistance, type 2 diabetes mellitus, and metabolic syndrome (*7*). Furthermore, the loss of a single allele of the *PIK3CA* gene in mice causes glucose intolerance and hyperinsulinemia, which underscores the central role of PI3Kα in glucose homeostasis and insulin signaling (*8*).

[1]Cardiovascular Metabolic Research Laboratories, Daiichi Sankyo Co., Ltd., Tokyo, Japan. [2]National Cancer Institute RAS Initiative, Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Frederick, MD, USA. [3]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA, USA. [4]Organic and Biomolecular Chemistry Department, Daiichi Sankyo RD Novare Co., Ltd., Tokyo, Japan. [5]Medicinal Chemistry Research Laboratory, Daiichi Sankyo Co., Ltd., Tokyo, Japan. [6]Discovery Science and Technology Department, Daiichi Sankyo RD Novare Co., Ltd., Tokyo, Japan. ***Corresponding author. Email: dhirendra.simanshu@nih.gov (D.K.S.); kazuishi.kubota@daiichisankyo.com (K.K.); jun.tanaka@daiichisankyo.com (J.T.); frank.mccormick@ucsf.edu (F.M.)** †These authors contributed equally to this work.

Discovery of GLUT4 activators and their selective PI3Kα-AKT activation

We aimed to identify small molecules that activate glucose uptake as an approach to moderating blood glucose concentrations in patients with various types of diabetes. We screened a library for compounds that caused GLUT4 translocation in the absence of insulin in differentiated rat L6 myotube cells that had been engineered to express myc-tagged GLUT4 (L6-GLUT4 myc). A compound identified in this screen was derivatized to increase its ability to promote GLUT4 translocation (*9*). Representative active compounds D223 and D927, along with an inactive compound D484, are shown in Fig. 1A. The two active compounds, D223 and D927, have identical scaffolds, with D223 containing an *N-tert*-butoxycarbonyl group in place of the methoxy group in D927. Both active compounds stimulated GLUT4 translocation (Fig. 1B) in differentiated L6-GLUT4 myc myotubes in a concentration-dependent manner like insulin (Fig. 1C).

Insulin stimulates the PI3Kα-AKT pathway by recruiting insulin receptor substrate 1 (IRS-1) and the RAF–extracellular signal–regulated kinase (ERK1/2) pathway through the activation of RAS proteins (*10*, *11*). The metabolic activities of insulin, such as the stimulation of GLUT4 translocation and the induction of protein synthesis, are mediated by the PI3Kα-AKT pathway, whereas activation of the protein kinase RAF to the ERK1/2 pathway is related to cell proliferation (*12*, *13*). In rat L6 myotubes, D927 activated the PI3Kα-AKT pathway, measured by increased phosphorylation of AKT (Ser^{473} and Thr^{308}) and 70-kDa ribosomal protein S6 (p70S6) kinase (Thr^{389}), but had no substantial effect on the RAF-ERK1/2 pathway, measured by amounts of phosphorylated ERK1 and ERK2 (P-ERK1/2) and phosphorylated mitogen-activated protein kinase (P-MEK) (Fig. 1D and fig. S1A). Furthermore, D927 did not affect levels of the catalytic subunit (p110α) of PI3Kα.

Treatment of L6 myotubes with D927 or insulin increased phosphorylation of AKT (Ser^{473} and Thr^{308}), p70S6 kinase (Thr^{389}), and eukaryotic initiation factor 4E–binding protein 1 (4EBP1) (Thr^{37} or Thr^{46}). However, only insulin treatment led to increased phosphorylation of MEK and ERK (fig. S2). Treatment with D927 and insulin resulted in a synergistic effect on AKT activity (Fig. 1E and fig. S1B). Additionally, D927 increased activity of the PI3Kα pathway, but not the RAF-MAPK pathway, in a panel of human cancer cells, including cells with PTEN mutations, activating mutations in the helical domain or the kinase domain of PIK3CA, or activating mutations in KRAS (Fig. 1F and fig. S1C).

Selective enhancement of PI3Kα-RAS interactions by D223 and D927

We used chemical proteomics to explore the molecular mechanism by which D223 and D927 activate the PI3Kα pathway without a signal from the insulin receptor. D223 bound to immobilized beads was incubated with L6-GLUT4 myc lysates, and the associated proteins were analyzed by mass spectrometry (data S1). PI3Kα was identified as a major associated molecule. Neither compound activated recombinant PI3Kα in vitro (fig. S3A), which indicates that another factor is necessary to mediate the activation of PI3Kα in cells (*14*, *15*). To identify this factor, we analyzed proteins associated with PI3K immunoprecipitated from L6-GLUT4 myc treated with insulin, D223, or D927 by mass spectrometry (data S2). The active compounds (D223 or D927), unlike the inactive compound (D484), increased binding of the RAS-related protein RRAS2. This interaction was confirmed by immunoprecipitation of endogenous proteins (Fig. 2A). KRAS was also detected in our mass spectrometry analysis, and IRS-1 recruitment to PI3Kα was observed in insulin-treated cells, as expected (fig. S3B).

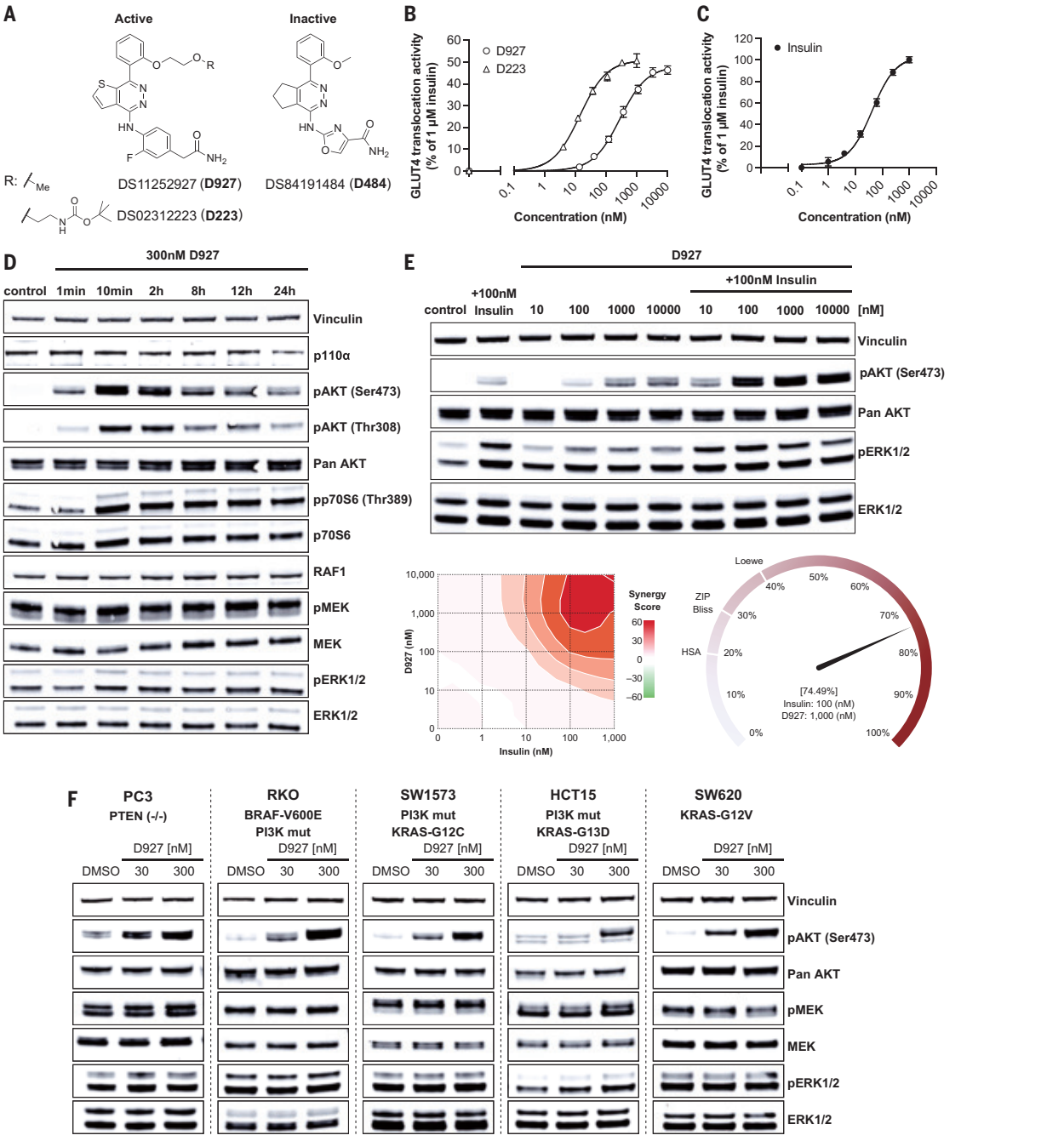


Fig. 1. Effects of D223 and D927 compounds on GLUT4 translocation and PI3Kα signaling pathway. (**A**) Chemical structures of D927, D223, and D484. (**B** and **C**) GLUT4 translocation activities on L6-GLUT4 myc myotubes after treatment with compounds D223 and D927 (B) or insulin (C). Data are shown as mean values ± SEMs of relative GLUT4 translocation activities compared with those of 1 μM insulin ($n = 3$ for D223 and insulin, $n = 4$ for D927). Results are representative of two independent experiments with similar results. (**D**) Time course of PI3Kα pathway activation after D927 addition in differentiated L6 myotubes. L6 cells were differentiated for 5 days in 2% fetal bovine serum–Dulbecco's minimum essential medium (FBS DMEM). Myotubes were serum-starved overnight and treated with 300 nM D927. Results are representative of three independent experiments with similar results. Quantification of these data is shown in fig. S1A. (**E**) Combination of D927 and insulin in L6 myotubes resulted in a synergistic effect on the activation of AKT. L6 cells were differentiated for 5 days in 2% FBS DMEM. Myotubes were serum-starved overnight and treated with insulin and D927 alone or in combination for 10 min. Western blot dose response of 100 nM insulin with or without 10, 100, 1000, or 10,000 nM of D927 is shown (top). Quantification of these data is shown in fig. S1B. Homogeneous time-resolved fluorescence (HTRF) AKT assays were performed with 1, 10, 100, and 1000 nM insulin and 10, 100, 1000, and 10,000 nM D927 alone or in combination for 10 min. Synergy scores derived from the zero interaction potency (ZIP) model using synergyfinderplus.org are positive across consecutive dosage areas (red), confirming synergistic effects from insulin and D927. An example barometer plot shows the observed response (percent activation) of 100 nM insulin and 1000 nM D927, which are beyond the reference noninteraction additive effects calculated by four different models [ZIP, Bliss, Loewe, and highest single agent (HSA)] that are marked on the plot (bottom). All results are representative of two independent experiments with similar results. (**F**) D927 increased the activity of the PI3Kα pathway in a panel of cancer cells. Cell lines were treated with 30 or 300 nM D927 for 5 min. Results are representative of two independent experiments with similar results. Quantification of these results is shown in fig. S1C. Relevant mutations in these cell lines are PTEN −/− (PC3 cells), PIK3CA H1047R and BRAF V600E (RKO cells), PIK3CA K111E and KRAS G12C (SW1573), PIK3CA E544K/D549N and KRAS G13D (HCT15), and KRAS G12V (SW620). DMSO, dimethyl sulfoxide.

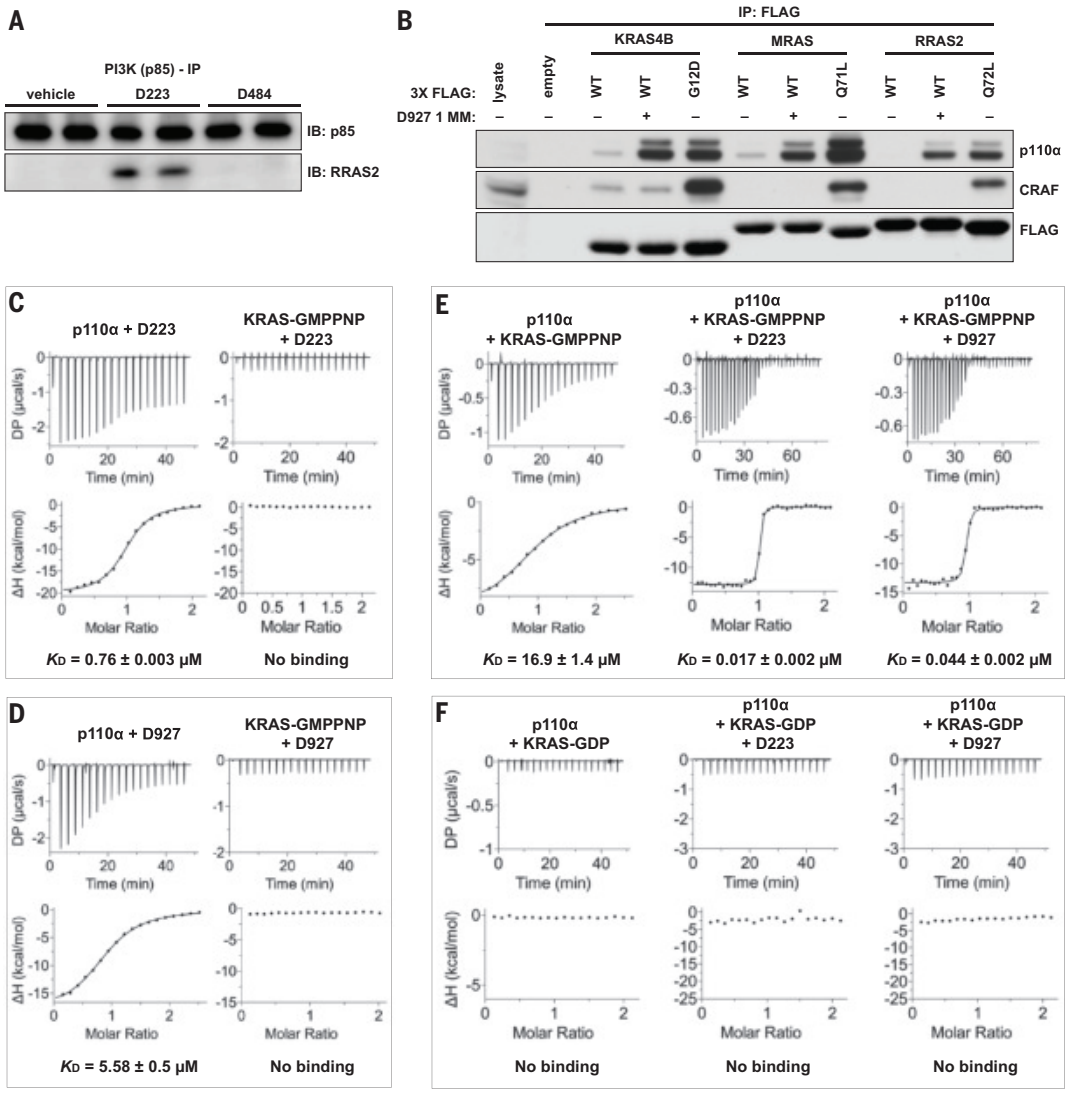


Fig. 2. D927 increases the binding affinity of multiple RAS proteins to p110α. (**A**) Association of PI3Kα and RRAS2 after treatment of L6-GLUT4 myc myotubes with compounds. L6-GLUT4 myc myotubes starved for 3 hours were treated with vehicle, 1 μM D223, and 10 μM D484 for 5 min. Cell lysate were incubated with an antibody to p85, and immunoprecipitated protein G (immunoglobulin G) beads were analyzed by immunoblotting using an antibody to RRAS2. Results are representative of two independent experiments with similar results. (**B**) D927 acts as a molecular glue for multiple RAS proteins. FLAG-tagged wild-type (WT) KRAS4B, KRAS4B G12D, wild-type MRAS, MRAS Q61L, wild-type RRAS2, and RRAS2 Q72L proteins were transfected into HEK293 cells. After 24 hours, cells were treated with D927 for 1 hour. FLAG-tagged proteins were immunoprecipitated and probed for association with endogenous p110α, endogenous CRAF, or FLAG to measure relative amounts of expression. (**C**) Isothermal titration calorimetry (ITC) measurements showing the binding affinity of D223 with p110α and KRAS(GMPPNP). DP, differential power; ΔH, enthalpy change. (**D**) ITC measurements showing the binding affinity of D927 with p110α and KRAS(GMPPNP). (**E**) ITC measurements showing the binding affinity of KRAS(GMPPNP) with PI3Kα in the absence and presence of D223 or D927. (**F**) ITC measurements showing the binding affinity of KRAS(GDP) with PI3Kα in the absence and presence of D223 or D927. [(C) to (F)] The K_D values shown are presented as the means ± standard deviations derived from two replicates.

D927 promoted binding of wild-type KRAS and an oncogenic KRAS mutant G12D (Gly^{12}→Asp) to endogenous p110α when expressed in human embryonic kidney (HEK) 293 cells (Fig. 2B). Binding to KRAS G12D was significantly higher than that to wild-type KRAS, presumably because this mutant is predominantly in a guanosine 5′-triphosphate (GTP)–bound form, as discussed below. D927 had no effect on the binding of these RAS proteins to CRAF, which indicates that the compounds do not act by increasing GTP binding to RAS and that binding of KRAS to p110α in the presence of D927 does not deplete cells of KRAS(GTP) and thus reduce binding to CRAF. D927 also promoted the binding of MRAS and RRAS2 to p110α in a manner consistent with GTP dependence (Fig. 2B). Furthermore, D927 enhanced the binding of all members of the RAS superfamily previously shown to bind p110α (KRAS4A, KRAS4B, RRAS, RRAS2, and MRAS) but not to RIT1, RAP1, RAP2, RHEB, or RAC2 (fig. S4).

D223 and D927 showed no measurable binding affinity to active KRAS bound to the nonhydrolysable GTP analog, guanosine-5′-(β,γ-imido)triphosphate (GMPPNP) (Fig. 2, C and D). However, both compounds bound directly to p110α with a dissociation constant (K_D) of 0.76 and 5.6 μM, respectively. The compounds exhibited no binding to PI3Kβ, PI3Kγ, PI3Kδ, or CRAF (fig. S5). Active KRAS bound to p110α with a relatively low affinity (K_D = 17 μM), but the addition of D223 and D927 increased this affinity by nearly three orders of magnitude, yielding K_D values of 17 and 44 nM, respectively (Fig. 2E). Binding of KRAS to p110α was strictly dependent on GMPPNP. The guanosine diphosphate (GDP)–bound form of KRAS failed to bind to p110α, even

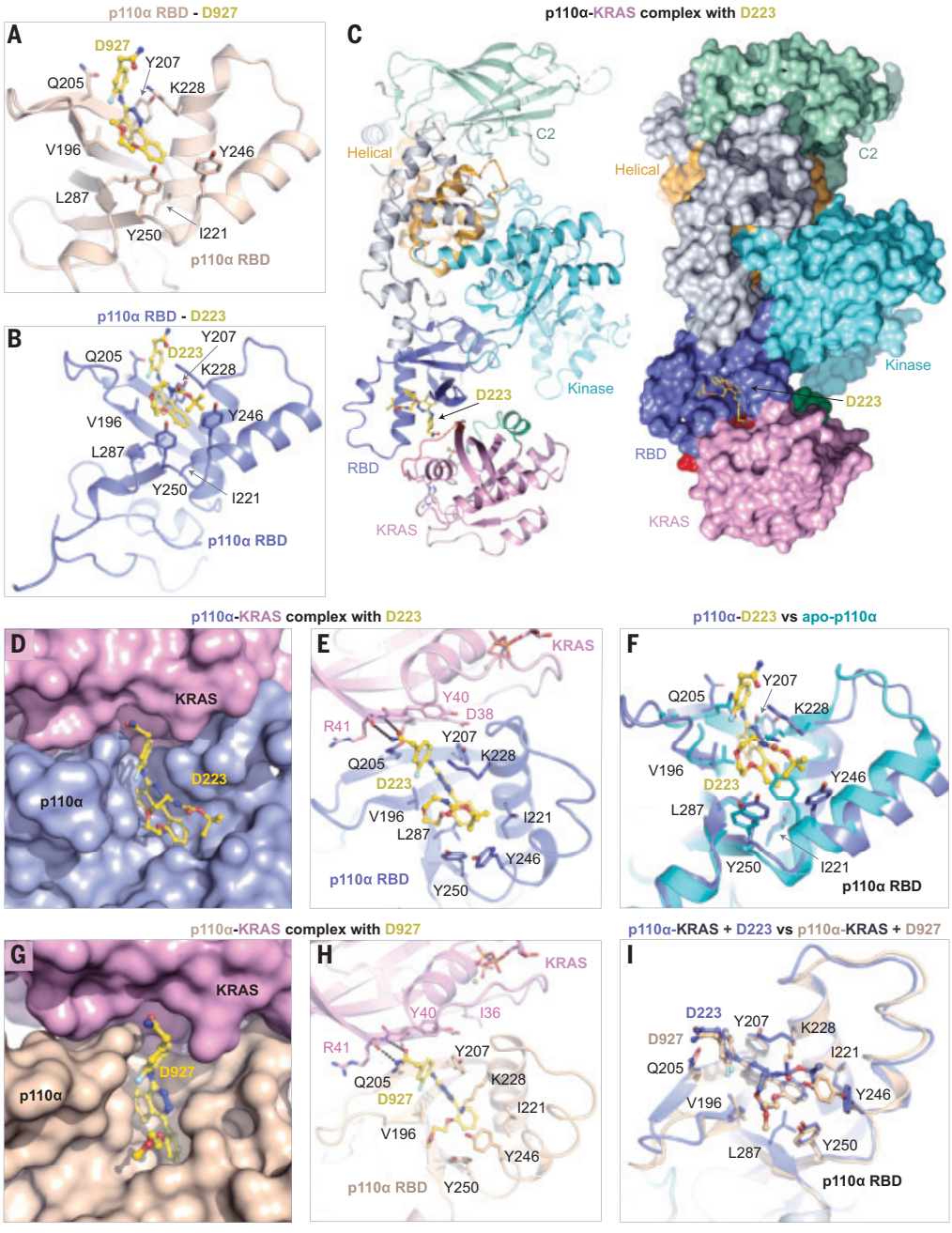


Fig. 3. Binding pocket in p110α and the structural basis of D223 and D927 enhancing RAS binding to p110α. (**A**) Crystal structure of the p110α-RBD in complex with D927, highlighting interactions between D927 and p110α residues. (**B**) Crystal structure of the p110α-RBD in complex with D223, highlighting interactions between D223 and p110α residues. (**C**) The overall structure of the KRAS-p110α complex with D223 is shown in cartoon and surface representations. The p110α-RBD, C2, helical, and kinase domains are colored blue, pale green, orange, and cyan, respectively. KRAS is colored pink, with switch-I and switch-II regions highlighted in red and green, respectively. GMPPNP and glue D223 are shown as sticks, and Mg^{2+} ions as spheres. (**D**) Surface representation of the KRAS-p110α interface showing the D223 binding pocket. (**E**) Enlarged view of the glue binding pocket showing p110α and KRAS residues interacting with D223. (**F**) Structural superposition of the p110α-KRAS complex with D223 and the apo-structure of p110α-RBD (PDB: 6VO7), showing conformational changes in p110α residues forming the induced pocket for glue binding. (**G**) Surface representation of the KRAS-p110α interface showing the D927 binding pocket (PDB: 9C15). (**H**) Enlarged view of the D927 binding pocket, showing p110α and KRAS residues interacting with D927. (**I**) Structural superposition of the p110α-KRAS complexes with D223 and D927, showing similarities and differences in how these two glue compounds interact with p110α residues. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

in the presence of D223 or D927 (Fig. 2F). Therefore, D223 and D927 appear to act as molecular glues that bind directly to PI3Kα and, in cells, recruit active RAS proteins to bind PI3Kα and promote activation.

Structural mechanism of molecular glue–mediated PI3Kα-RAS interaction

Crystal structures of D927 and D223 bound to the RAS-binding domain (RBD) of p110α revealed the glue compound binding to a pocket between the β sheet and α1 helix of the RBD, formed by residues V196, Q205, Y207, I221, K228, Y246, Y250, and L287 (Fig. 3, A and B; fig. S6; and table S1). These p110α residues interacted with D223 and D927 primarily through hydrophobic and van der Waals interactions (fig. S6). To validate these interactions, we performed alanine-scanning mutagenesis of key residues (V196, Q205, Y207, I225, and K228) within the glue-binding pocket of p110α-RBD and assessed their binding affinity to D927. These experiments demonstrated that the Y207A mutation completely abolished glue binding to p110α-RBD, whereas the I225A and K228A mutations reduced binding affinity to one-half or one-fourth that of the wild-type protein. By contrast, V196A and Q205A mutations exhibited minimal effect on D927 binding (fig. S7). Structural comparison with apo-p110α-RBD (without ligand bound) showed that the D927 or D223 binding pocket does not exist in apo-p110α, and the binding of D927 and D223 induced formation of a pocket in the RBD near the interaction interface for binding to RAS proteins (fig. S8).

To understand how D223 and D927 achieve nearly three orders of magnitude of increased affinity between RAS and p110α, we solved the crystal structure of p110α (residues 105 to 1068) in complex with KRAS in the presence of D223 (Fig. 3C and table S1) or D927 [Protein Data Bank (PDB) ID: 9C15] (*16*). At the RAS-p110α interface, the amide groups of D223 and D927 formed hydrogen bonds with the main chain atoms of R41 (KRAS) and interacted through van der Waals forces with the side chain of Y40 (KRAS) (Fig. 3, D to H, and fig. S9). Upon binding to p110α, the glue compounds stabilized the first two α helices and the intervening loop in the RBD, arranging

interface residues into an optimal conformation for RAS protein binding (Fig. 3F). Consequently, these interactions and the stabilization of secondary structures by D223 and D927 in p110α enhanced the affinity between RAS and p110α. The additional increase in affinity observed with D223 compared with D927 apparently arises from the interactions formed by the *N-tert*-butoxycarbonyl group in D223 with p110α (Fig. 3, D and E). The corresponding methoxy moiety in D927 pointed toward the solvent and made minimal interaction with p110α-RBD (Fig. 3, G to I). KRAS residues that interacted with p110α at the interface are largely conserved among RAS family members HRAS, NRAS, KRAS, RRAS, RRAS2, and MRAS, which aligns with the observed effects of D927 and D223 on these interactions (Fig. 2B and fig. S9).

Evaluating RAS family guanosine triphosphatases' contributions to D927-induced PI3Kα activation

To investigate the effects of D927 on PI3Kα activation by wild-type KRAS, we used mouse embryo fibroblasts (MEFs), in which *Hras* and *Nras* had been deleted and in which the *Kras* locus could be excised after the

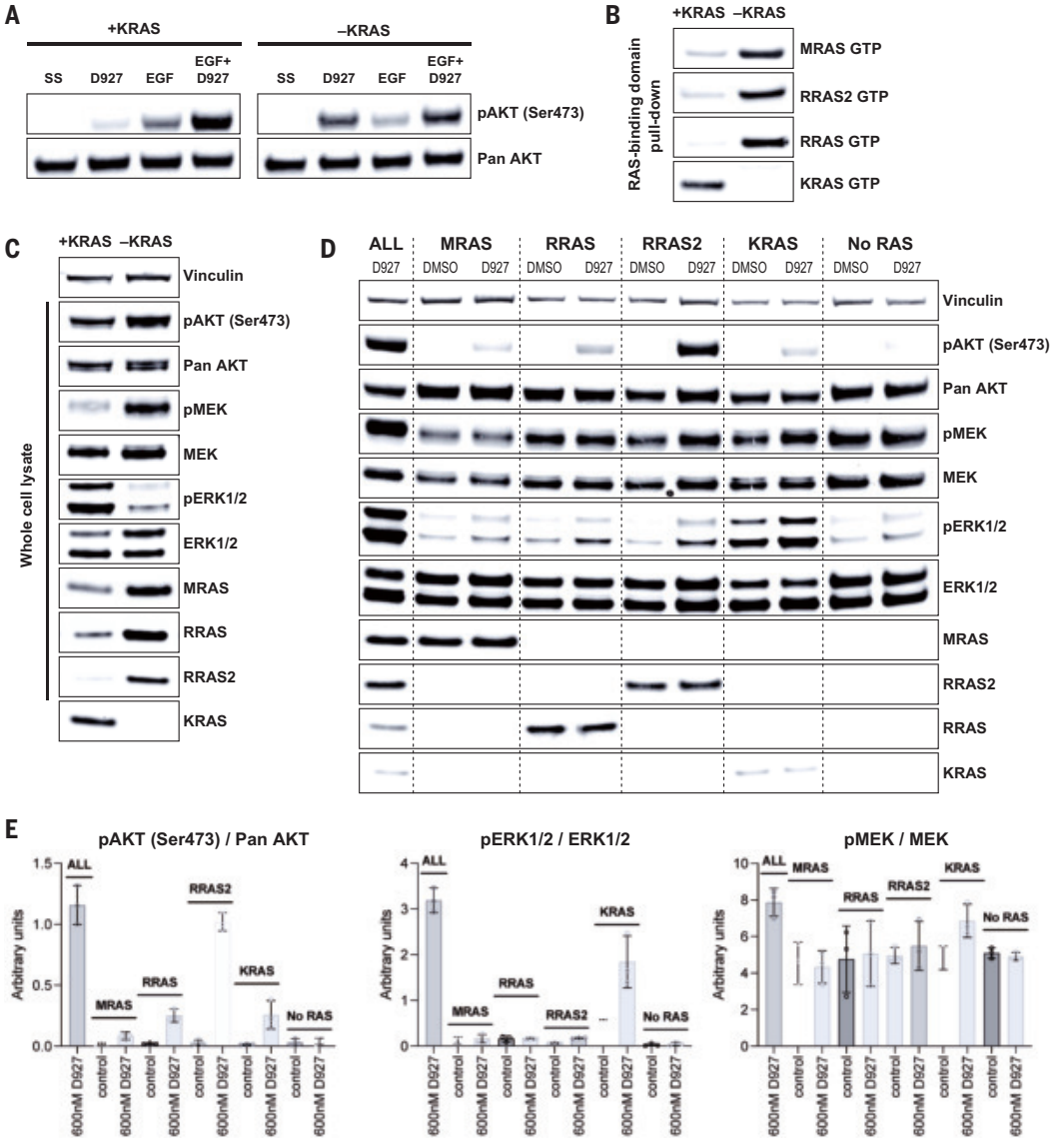


Fig. 4. Dependence of D927 on GTP binding of individual RAS family proteins for PI3Kα activation in MEFs. (**A**) Removal of *Kras* in a diploid parental *Hras/Nras* null DU1473 MEF line results in a marked increase of P-AKT induced by D927. DU1473 MEF cells were 4-hydroxy tamoxifen–treated for 11 days to excise *Kras*. MEFs with KRAS (+KRAS) or without KRAS (−KRAS) protein expression were then serum-starved (SS) overnight and treated with 1 μM D927 for 30 min or EGF or both for 5 min. Results are representative of three independent experiments. (**B**) MRAS, RRAS, and RRAS2 accumulate in their GTP states after KRAS depletion in MEFs. DU1473 MEF cells were 4-hydroxy tamoxifen–treated for 11 days to excise *Kras*. MEFs with KRAS (+KRAS) or without KRAS (−KRAS) protein expression were harvested, and GTP-loaded KRAS, MRAS, RRAS, or RRAS2 were isolated by association with RBD. Results are representative of three independent experiments. (**C**) MRAS, RRAS, and RRAS2 accumulate at the protein level in MEFs after KRAS protein depletion. DU1473 MEF cells were 4-hydroxy tamoxifen–treated for 11 days to excise *Kras*. MEFs with KRAS (+KRAS) or without KRAS (−KRAS) protein expression were harvested, and whole-cell lysates were analyzed. Results are representative of three independent experiments. (**D**) RRAS2 accounts for virtually all D927-induced P-AKT in MEFs. CRISPR knockout was used to generate a panel of MEFs from the *Kras*-excisable MEFs described previously, expressing only a single member of the RRAS family (MRAS, RRAS, or RRAS2) after treatment with 4-hydroxy tamoxifen for 11 days. A DU1473 MEF line expressing KRAS, MRAS, RRAS, and RRAS2 protein (labeled as "ALL") and a MRAS, RRAS, and RRAS2 knockout line treated with 4-hydroxy tamoxifen to remove *Kras* (labeled as "No RAS") were used as controls. All lines were treated with 600 nM D927 for 5 min. (**E**) Quantification of three independent experiments for (D).

addition of 4-hydroxy tamoxifen (*17*). These cells were deprived of serum to render the KRAS protein GDP bound. D927 induced a minimal increase in P-AKT in these cells (Fig. 4A). Addition of epidermal growth factor (EGF) to promote the accumulation of KRAS in its GTP-bound state increased amounts of P-AKT. This was further increased after the addition of D927, consistent with the effects of D927 depending on active, GTP-bound RAS (Fig. 4A). Unexpectedly, the level of P-AKT induced by D927 in serum-starved cells increased after the removal of *Kras* by the addition of 4-hydroxy tamoxifen. Because these cells lack HRAS and NRAS, this P-AKT may derive from active RRAS, RRAS2, or MRAS that accumulated after KRAS depletion. Indeed, these proteins accumulated in their GTP-bound states after KRAS depletion (Fig. 4B). This accumulation was due, at least in part, to increased protein abundance (Fig. 4C). MRAS accumulates in cells after KRAS inhibition (*18*), but the accumulation of RRAS and RRAS2 has not been described. Little is known about the regulation of RRAS and RRAS2 proteins. They are not regulated by SOS proteins [RAS guanine exchange factors (GEFs)] and exist to a relatively high degree in their GTP-bound states without any input from

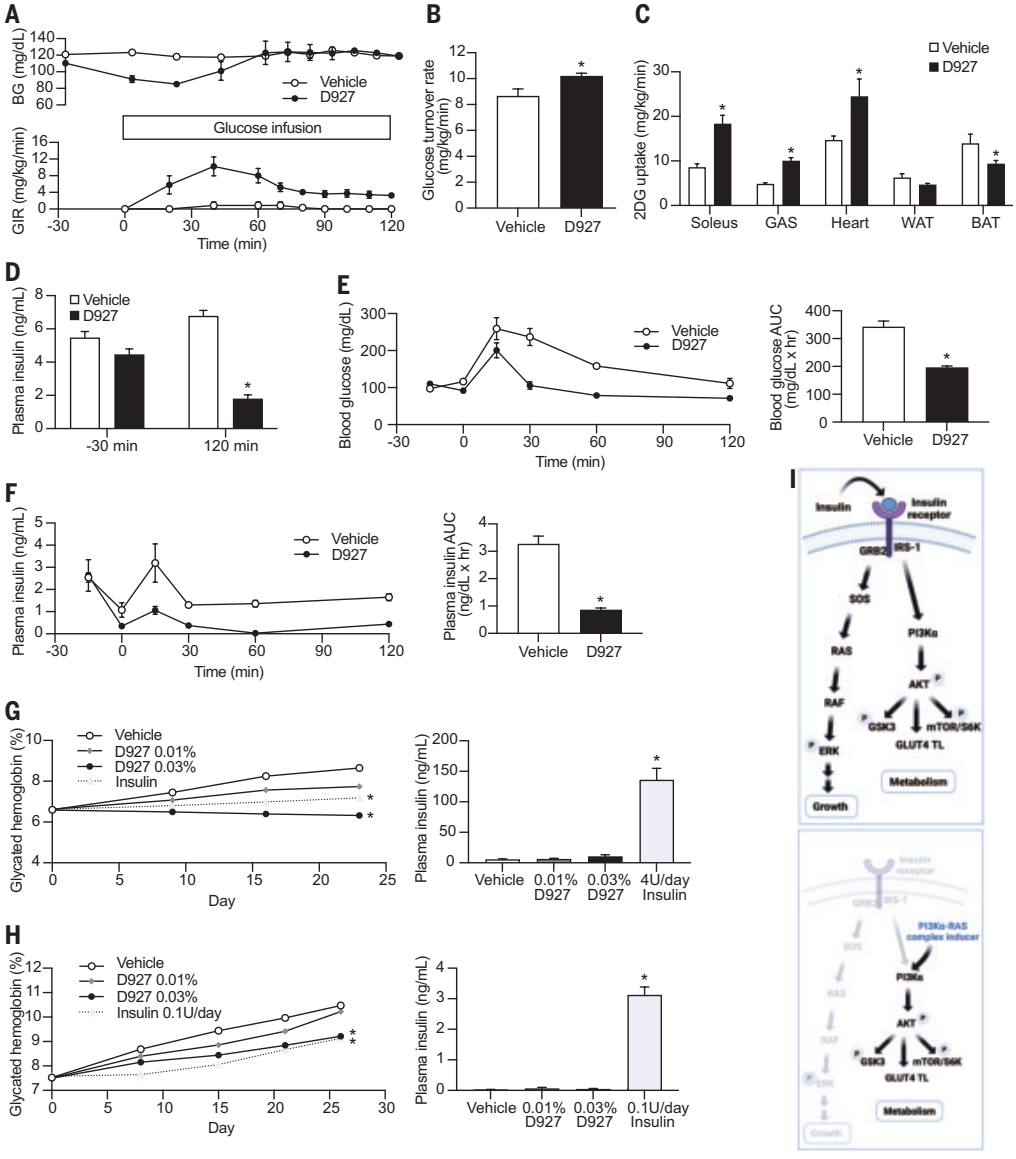


Fig. 5. Oral administration of D927 improved glucose intolerance and ameliorated hyperglycemia in diabetic models. (**A** to **D**) Euglycemic clamp study without insulin infusion in Zucker fatty rats. Vehicle or 30 mg/kg of D927 was orally administered to overnight-fasted rats at 30 min before the start of glucose infusion. Results are from one representative experiment designed based on several independent preliminary experiments. (A) Blood glucose (BG) and glucose infusion rates (GIR) during the clamp. (B) Glucose turnover rates in the steady state (90 to 120 min). (C) 2-Deoxy-glucose (2DG) uptake during the clamp. GAS, gastrocnemius; WAT, white adipose tissue; BAT, brown adipose tissue. (D) Plasma insulin at 30 and 120 min. [(B) to (D)] Data are shown as mean values ± SEMs. **P* < 0.05 versus control by Wilcoxon rank sum test (*n* = 5). (**E** and **F**) Blood glucose (E) and plasma insulin (F) levels during oGTT in *db/db* mice. Vehicle or 10 mg/kg D927 was orally administered 15 min before glucose (1 g/kg) loading. Area under the curve (AUC) from 0 to 120 min was calculated. Data are shown as mean values ± SEMs (*n* = 5). AUC values were compared with the vehicle group using the Wilcoxon rank sum test (**P* < 0.05). Results are from one representative experiment designed based on several independent preliminary experiments. (**G** and **H**) Long-term administration of D927 on *db/db* mice (G) and STZ-B6 mice (H). Glycated hemoglobin during the administration (left) and plasma insulin on the last day of administration (right) are shown as mean values ± SEMs (*n* = 7 for *db/db* mice and *n* = 8 for STZ-B6 mice). Glycated hemoglobin and plasma insulin on the last day of the administration were compared with those in the vehicle group by the Steel test (**P* < 0.05). Results are from one representative experiment designed based on several independent preliminary experiments. (**I**) Schematic illustration of intracellular signal transduction of insulin- and molecular glue–enhanced PI3Kα-RAS complex formation, where RAS refers to all isoforms, including RRAS2.

GEFs, possibly because they have high intrinsic GDP off rates (*19*). MRAS has been studied more extensively (*20*), but the mechanism regulating MRAS abundance is poorly understood.

To analyze the contribution that each RAS family member makes toward D927-induced P-AKT, we used CRISPR to generate a panel of MEFs from the *Kras*-excisable MEFs described above, each expressing only a single member of the RRAS family. D927 increased P-AKT accumulation through RRAS, RRAS2, MRAS, and KRAS, but RRAS2 accounted for nearly all of the P-AKT activity induced by D927 in parental cells (Fig. 4, D and E). Cells in which all six RAS proteins were deleted showed little, if any, response to D927, which indicates that other, more distantly related proteins or proteins unrelated to the RAS family are not involved in D927-mediated activation of PI3Kα. Furthermore, these data show that wild-type KRAS can activate PI3Kα, at least if D927 is present.

In vivo insulin-mimicking effects on glucose homeostasis in animal models

To investigate the in vivo pharmacological effects of D223 and D927 on glucose homeostasis, we performed euglycemic clamps on insulin-resistant Zucker fatty rats. A single oral administration of D927 rapidly decreased blood glucose concentration (Fig. 5A). At steady state (90 to 120 min), glucose infusion rates (Fig. 5A) and glucose turnover rates (Fig. 5B) were higher in D927-treated rats. During this period, glucose uptake significantly increased in the soleus and gastrocnemius skeletal muscles as well as the heart (Fig. 5C), despite decreased plasma insulin levels in D927-treated rats (Fig. 5D).

We further evaluated the acute and chronic effects of D927 in obese and insulin-resistant type 2 diabetic *db/db* mice. In an oral glucose tolerance test (oGTT), D927 acutely improved hyperglycemia independently of insulin secretion (Fig. 5, E and F). Chronic administration of D927 for 24 days resulted in a significant reduction in plasma glycated hemoglobin, a marker of chronic hyperglycemia, comparable to the effect of insulin (Fig. 5G). To determine whether the antidiabetic effects of D927 apply to a type 1 diabetes model, D927 was evaluated in streptozotocin (STZ)–treated mice, which have minimal insulin. Administration of D927 for 27 days resulted in a significant reduction in plasma glycated hemoglobin, comparable to the effect of insulin, even in insulin-deficient type 1 diabetic mice (Fig. 5H). Food intake did not change, and the increases in body and tissue weight correlated with the improvement in glycemic control, like the outcomes in animals treated with insulin in these studies (fig. S10). We also examined the effects of D927 on normal animals. D927, like insulin, lowered blood glucose levels in a time- and dose-dependent manner in normal Sprague-Dawley (SD) rats (fig. S11A). Furthermore, D927 promoted phosphorylation of AKT in the soleus, gastrocnemius, liver, heart, and white adipose tissue in normal C57BL/6j mice, similar to insulin (fig. S11B). Further evaluation of safety would be necessary before clinical testing, but D927 could potentially ameliorate dysregulated glucose homeostasis independently of insulin in diabetic patients.

Summary

We present the discovery of two compounds, D223 and D927, that increase the affinity of RAS binding to PI3Kα by ~500-fold. This effect applies to canonical RAS proteins as well as to RRAS, RRAS2, and MRAS, of which RRAS2 is the major isoform affected by D927 in cells, reflecting its high affinity for PI3Kα as well as its abundance in cells in its GTP-bound form.

Our glue compounds were discovered in a screen for compounds that mimic insulin and promote glucose uptake in diabetic individuals where insulin homeostasis is compromised. These compounds exerted robust antidiabetic effects in vivo. They have also enabled us to solve the structure of KRAS bound to PI3Kα (*16*) and to develop molecular breaker compounds that disrupt this interaction as potential cancer therapeutics. One such compound, BBO-10203, is undergoing clinical trials (*21*). As expected from the data presented therein, this compound prevents the activation of PI3Kα by RAS proteins but does not affect glucose uptake and insulin homeostasis.

REFERENCES AND NOTES

1. Y. Samuels *et al.*, *Science* **304**, 554 (2004).
2. M. J. Lindhurst *et al.*, *Nat. Genet.* **44**, 928–933 (2012).
3. P. Castel *et al.*, *Sci. Transl. Med.* **8**, 332ra42 (2016).
4. A. Roy *et al.*, *eLife* **4**, e12703 (2015).
5. B. Vanhaesebroeck, M. W. D. Perry, J. R. Brown, F. André, K. Okkenhaug, *Nat. Rev. Drug Discov.* **20**, 741–769 (2021).
6. G. Q. Gong *et al.*, *Nature* **618**, 159–168 (2023).
7. P. D. Brewer, E. N. Habtemichael, I. Romenskaia, C. C. Mastick, A. C. Coster, *J. Biol. Chem.* **289**, 17280–17298 (2014).
8. L. C. Foukas *et al.*, *Nature* **441**, 366–370 (2006).
9. T. Tsuji *et al.*, *Bioorg. Med. Chem. Lett.* **29**, 1785–1790 (2019).
10. C. M. Taniguchi, B. Emanuelli, C. R. Kahn, *Nat. Rev. Mol. Cell Biol.* **7**, 85–96 (2006).
11. M. L. Wei, P. Duan, Z. M. Wang, M. Ding, P. Tu, *Mol. Med. Rep.* **16**, 6690–6696 (2017).
12. W. Zhang, H. T. Liu, *Cell Res.* **12**, 9–18 (2002).
13. F. Chang *et al.*, *Int. J. Oncol.* **22**, 469–480 (2003).
14. E. H. Walker, O. Perisic, C. Ried, L. Stephens, R. L. Williams, *Nature* **402**, 313–320 (1999).
15. J. Yu *et al.*, *Mol. Cell. Biol.* **18**, 1379–1387 (1998).
16. D. Czyzyk *et al.*, *Nat. Commun.* **16**, 525 (2025).
17. M. Drosten *et al.*, *EMBO J.* **29**, 1091–1104 (2010).
18. Y. Adachi *et al.*, *Nat. Cancer* **4**, 829–843 (2023).
19. Y. Ohba *et al.*, *J. Biol. Chem.* **275**, 20020–20026 (2000).
20. L. C. Young, P. Rodriguez-Viciana, *Cold Spring Harb. Perspect. Med.* **8**, a033621 (2018).
21. D. K. Simanshu *et al.*, *Science* **389**, 409–415 (2025).

ACKNOWLEDGMENTS

We thank O. Ubukata and I. Iino at Daiichi Sankyo RD Novare; D. Esposito, B. Gillette, S. Messing, M. Drew, C. Grose, and their colleagues from the RAS Initiative at the Frederick National Laboratory (FNL) for assistance with recombinant protein production; W. Burgan, K. Powell, and B. Xu (FNL) for help with cell line development and genomic analysis; D. Bonsor (FNL) for figure preparation; and M. Yi (FNL) for bioinformatic analysis of the synergy assay. We also acknowledge H. Matsunaga, S. Yamamoto, and K. Araki (Daiichi Sankyo) for providing various RAS expression plasmids and J. Harada (Daiichi Sankyo) for consultation on the statistical analysis. **Funding:** The x-ray crystallographic work is based on research conducted at the NE-CAT beamline (24-ID-E) at Advanced Photon Source funded by NIH/NIGMS (P30 GM124165) and the CBMS beamline (17-ID-2) at NSLS-II funded by NIH/NIGMS (P30 GM133893) and Department of Energy (DOE) Office of Biological and Environmental Research (KP1607011). The beamtime at NSLS-II was facilitated through the NECAT BAG proposal no. 311950. The Eiger 16M detector on 24-ID-E is funded by an NIH-ORIP HEI grant (S10OD021527). This research used resources of the Advanced Photon Source, a US DOE Office of Science User Facility operated by Argonne National Laboratory under contract no. DE-AC02-06CH11357. NSLS2 is a US DOE Office of Science User Facility operated under contract no. DE-SC0012704. This project has been funded in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract no. 75N91019D00024. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services nor does the mention of trade names, commercial products, or organizations imply endorsement by the US government. **Author contributions:** K.T. and S.F. conceived of the project, performed in vitro studies, and analyzed the data. M.S., M.H., and Y.O. performed the chemical proteomics experiments. A.U., S.K., M.K., and Y.S. performed the in vivo studies. M.Y., T.T., and J.K. designed and synthesized the series of compounds. W.Y., D.J.C., and D.K.S. carried out ITC and crystallographic studies. L.C.Y. and R.G.d.S. performed pull-down experiments. N.F. performed Western blot experiments and developed the diploid parental MEF line and MEF CRISPR knockout panel. K.T., Y.S., K.K., D.K.S., F.M., and J.T. wrote the manuscript. D.K.S., K.K., J.T., and F.M. supervised the project. All authors discussed the results and commented on the manuscript. **Competing interests:** F.M. is a consultant for the following companies: Amgen, Daiichi Sankyo Ltd., Exuma Biotech, Ideaya Biosciences, Kura Oncology, and Quanta Therapeutics. F.M. has an ownership interest, including stock options in BridgeBio and Quartz. F.M. is the scientific director of the National Cancer Institute RAS Initiative at the FNL for Cancer Research and Leidos Biomedical Research, Inc. F.M. has research agreements with Boehringer-Ingelheim and Roche. N.F. is a shareholder in BridgeBio. S.F., M.Y., T.T., and J.K. are inventors on the patent and patent application (WO/2018/221433) held and submitted by Daiichi Sankyo Co., Ltd., that covers D223 and D927. M.H., Y.O., and K.K. are currently employees of Daiichi Sankyo Co., Ltd. M.S. is currently an employee of PeptiDream Inc. J.K. is currently an employee of Mitsubishi UFJ Research and Consulting Co., Ltd. The authors declare no other competing interests. **Data and materials availability:** The atomic coordinates and structure factors have been deposited in the Protein Data Bank and are available under the accession numbers 9CML (p110α-RBD bound to D223), 9CMK (p110α-RBD bound to D927), and 9CMV (p110α-KRAS complex with D223). Mass spectrometry raw data have been deposited in a public repository (jPOST; https://repository.jpostdb.org) under the accession numbers JPOST000627 and JPOST000626. All data are available in the main text or the supplementary materials. D223 and D927 are available from J.T. under a material transfer agreement with Daiichi Sankyo Co., Ltd.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adr9097
Materials and Methods; Figs. S1 to S11; Table S1; References (*22*–*32*); MDAR Reproducibility Checklist; Data S1 and S2

Submitted 21 July 2024; resubmitted 11 December 2024; accepted 20 May 2025
10.1126/science.adr9097

SIGNAL TRANSDUCTION

BBO-10203 inhibits tumor growth without inducing hyperglycemia by blocking RAS-PI3Kα interaction

Dhirendra K. Simanshu[1]*†, Rui Xu[2]†, James P. Stice[2], Daniel J. Czyzyk[1], Siyu Feng[2], John-Paul Denson[1], Erin Riegler[2], Yue Yang[3], Cathy Zhang[2], Sofia Donovan[2], Brian P. Smith[1], Maria Abreu-Blanco[1], Ming Chen[2], Cindy Feng[2], Lijuan Fu[2], Dana Rabara[1], Lucy C. Young[4], Marcin Dyba[1], Wupeng Yan[1], Ken Lin[2], Samar Ghorbanpoorvalukolaie[5], Erik K. Larsen[1], Wafa Malik[5], Allison Champagne[1], Katie Parker[5], Jin Hyun Ju[6], Stevan Jeknic[6], Dominic Esposito[1], David M. Turner[1], Felice C. Lightstone[3], Bin Wang[2], Paul M. Wehn[2], Keshi Wang[2], Andrew G. Stephen[1], Anna E. Maciag[1], Aaron N. Hata[5], Kerstin W. Sinkevicius[2], Dwight V. Nissley[1], Eli M. Wallace[2], Frank McCormick[1,4]*, Pedro J. Beltran[2]*

BBO-10203 is an orally available drug that covalently and specifically binds to the rat sarcoma (RAS)–binding domain of phosphoinositide 3-kinase α (PI3Kα), preventing its activation by HRAS, NRAS, and KRAS. It inhibited PI3Kα activation in tumors with oncogenic mutations in *KRAS* or *PIK3CA* and in tumors with human epidermal growth factor receptor 2 (HER2) amplification or overexpression. In preclinical models, BBO-10203 caused significant tumor growth inhibition across multiple tumor types and showed enhanced efficacy in combination with inhibitors of cyclin-dependent kinase 4/6 (CDK4/6), estrogen receptor (ER), HER2, and KRAS-G12C mutant, including in tumors harboring mutations in Kelch-like ECH-associated protein 1 (KEAP1) and serine/threonine kinase 11 (STK11). Notably, these antitumor effects occurred without inducing hyperglycemia, because insulin signaling does not depend on RAS-mediated PI3Kα activation to promote glucose uptake.

Phosphoinositide 3-kinases (PI3Ks) function in many aspects of cell physiology, including growth, differentiation, survival, and migration. PI3Kα is regulated by receptor tyrosine kinases and contributes to insulin homeostasis. The catalytic subunit of PI3Kα, p110α (encoded by the *PIK3CA* gene), is recruited to activated receptors through its direct interaction with the PI3Kα regulatory subunit p85, or in the case of insulin and insulin-like growth factor 1 (IGF-1), through p85 binding to insulin receptor substrate proteins (*1*). Deletion of p110α causes embryonic lethality (*2*), whereas mice heterozygous for p110α survive but are glucose intolerant and suffer from hyperinsulinemia and hyperphagia, among other phenotypes associated with defective insulin signaling (*1*).

Lipid kinase activity associated with v-Src and polyomavirus middle T oncoproteins (*3*–*6*) led to the identification of PI3K (*7*). Subsequently, three classes and multiple isoforms of PI3Ks were characterized. An oncogenic form of PI3Kα (*PIK3CA*) was discovered in an avian retrovirus (*8*), and activating mutations in *PIK3CA* were later identified in human tumors (*9*). These mutations occur in 24 to 46% of endometrial cancers, 20 to 32% of breast cancers, 20 to 27% of bladder cancers, and at notable frequencies in most other cancer types (*10*).

Canonical rat sarcoma (RAS) proteins bind directly to PI3Ks (*11*) but with lower affinity than to rapidly accelerated fibrosarcoma (RAF) kinases (*12*). Oncogenic RAS proteins activate PI3Kα when overexpressed (*13*). Furthermore, disrupting RAS-PI3Kα interaction through mutations T208D (T208→D) and K227A in the PI3Kα RAS-binding domain (RBD) severely impairs tumor formation driven by KRAS-G12D (*14*). Systemic disruption of RAS-PI3Kα interaction in mice is well tolerated and does not provoke hyperglycemia (*15*). Tumors driven by oncogenic mutants of epidermal growth factor receptor (EGFR) regress after genetic disruption of RAS-PI3Kα binding (*15*), and neoangiogenesis is impaired (*16*). These findings indicate that RAS binding to PI3Kα is not essential in normal cells, or for insulin homeostasis, but is vital for tumor formation and maintenance.

We describe an orally available drug, BBO-10203, that phenocopies the PI3Kα T208D-K227A mutant. It binds covalently to PI3Kα near the RAS binding site and blocks its interaction with the canonical RAS proteins HRAS, NRAS, and KRAS. It is effective in multiple mouse models and, notably, does not affect insulin homeostasis.

BBO-10203 binds covalently and specifically to p110α

The crystal structures of MRAS, RRAS2, and KRAS bound to p110α have been solved, with the KRAS-p110α complex stabilized by the molecular glue compound D927 (*17*). Figure 1A shows a structure of KRAS bound to nonhydrolyzable guanosine triphosphate (GTP) analog (GMPPNP) in complex with p110α, highlighting its interaction primarily through the RBD (PDB: 9C15). RAS binding to the p110α RBD does not alter the conformation of the kinase domain, supporting a proposed model in which RAS activates PI3Kα by stabilizing contacts between the p110α and the plasma membrane, where the substrate phosphatidylinositol 4,5-bisphosphate (PIP2) resides, without affecting kinase activity directly or allosterically (*17*–*19*).

Compound D927, used to crystallize the KRAS-p110α complex, was identified through phenotypic screening for antidiabetic compounds (*20*). It functions as a molecular glue by enhancing RAS-p110α interactions (*17*, *20*). Structural analysis revealed that D927 binds by inducing a pocket in the p110α RBD, with its aniline moiety interacting with KRAS residues Y40 and R41 (*17*). Given this structural insight, we tested whether substituting the flexible aniline group in D927 with a rigid moiety linked to the pyridazine ring would sterically clash with KRAS Y40, thereby disrupting the KRAS-p110α interaction (fig. S1). Structure-based drug design and molecular modeling led to the discovery of compound 1, which breaks instead of stabilizes the KRAS-p110α interaction. A considerable increase in cell potency was achieved by introducing an acrylamide electrophile that specifically targets the native cysteine C242 in the p110α RBD binding pocket, leading to compound 2. Subsequent optimization of potency and pharmacokinetic properties led to the development candidate BBO-10203.

BBO-10203 (Fig. 1B) covalently binds to p110α, as determined by matrix-assisted laser desorption/ionization–time-of-flight (MALDI-TOF) mass spectrometry (Fig. 1C and fig. S2). The crystal structure of PI3Kα RBD complexed with BBO-10203 shows covalent engagement with C242 via an acrylamide warhead (Fig. 1D and table S1). Key residues in PI3Kα RBD that interact with BBO-10203 include Q205, Y207, I225, K228, Y246, and Y250 (fig. S3), which are adjacent to mutations T208D and K227A, known to disrupt RAS-p110α binding (*14*). Figure 1E shows that BBO-10203, covalently bound to PI3Kα, would sterically clash with KRAS, preventing its binding.

BBO-10203 is specific for PI3Kα and exhibits no affinity for the other three PI3K isoforms: PI3Kβ, PI3Kδ, and PI3Kγ (fig. S4A). Comparisons

[1]National Cancer Institute (NCI) RAS Initiative, Cancer Research Technology Program, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Frederick, MD, USA. [2]BridgeBio Oncology Therapeutics, South San Francisco, CA, USA. [3]Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, USA. [4]Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA, USA. [5]Massachusetts General Hospital Cancer Center, Boston, MA, USA, and Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [6]BridgeBio Pharma, San Francisco, CA, USA. ***Corresponding author. Email: dhirendra.simanshu@nih.gov (D.K.S.); frank.mccormick@ucsf.edu (F.M.); pedro.beltran@bridgebiooncology.com (P.J.B.)** †These authors contributed equally to this work.

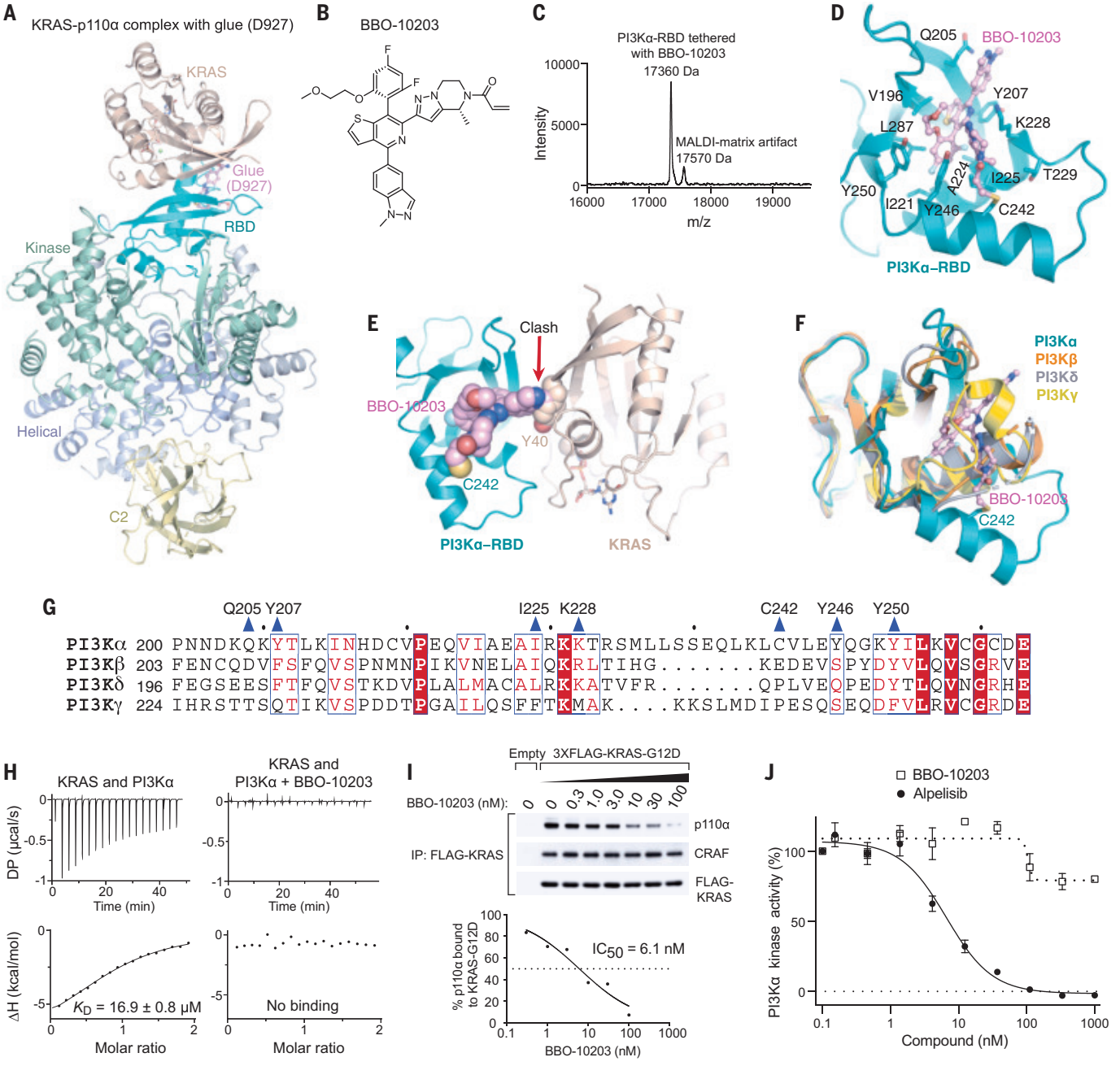


Fig. 1. BBO-10203 covalently binds PI3Kα-RBD, preventing interaction with HRAS, KRAS, and NRAS. (**A**) Structure of the KRAS-p110α complex (PDB: 9C15), stabilized by glue D927 and showing interaction between p110α-RBD (cyan) and KRAS (wheat). The C2, helical, and kinase domains are colored yellow, blue, and green, respectively. (**B**) Chemical structure of BBO-10203. (**C**) Mass chromatogram of PI3Kα-RBD covalently labeled with BBO-10203 after 30 min; a minor peak at 17569.5 mass/charge ratio (m/z) is a MALDI matrix artifact generated during ionization. (**D**) Crystal structure of PI3Kα-RBD (cyan, cartoon) covalently labeled with BBO-10203 (pink, ball and stick). Key PI3Kα residues interacting with BBO-10203 are illustrated in stick representation. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; and Y, Tyr. (**E**) Model of KRAS in complex with PI3Kα-RBD covalently bound to BBO-10203 shows steric clash between BBO-10203 and KRAS Y40. (**F**) Superposition of the RBD structures of PI3Kα, PI3Kβ, PI3Kδ, and PI3Kγ highlights differences at the BBO-10203 binding site. (**G**) Sequence alignment of the RBDs of all four PI3K isoforms shows deletions and lack of conservation of key residues (indicated by triangles above the aligned sequences), which form the BBO-10203 binding pocket in PI3Kα. (**H**) Isothermal titration calorimetry measurements showing the binding affinity of KRAS (GMPPNP) for both unlabeled and BBO-10203–labeled forms of PI3Kα. DP, differential power. (**I**) BBO-10203 disrupts the interaction between KRAS-G12D and p110α, with an IC_{50} of 6 nM. HEK 293T cells were transfected with 3XFLAG-KRAS-G12D. Twenty-four hours later, cells were treated for 2 hours with BBO-10203. After immunoprecipitation (IP) with anti-FLAG antibodies, Western blots were probed for endogenous p110α and CRAF. A representative blot from $n = 3$ replicates is shown. (Bottom) Dose-response curve of p110α bound to KRAS, using quantification of Western blot bands. (**J**) PI3Kα kinase activity assay comparing BBO-10203 and alpelisib. Dose titration of inhibitors was tested with recombinant PI3Kα, lipid substrate, and ATP, measuring ATP/adenosine diphosphate (ADP) conversion through luciferase activity, reported as percent activity.

of the sequence and structure of the RBD of PI3Kα with the other three PI3K isoforms elucidate this specificity. The other isoforms have deletions in this region and differ at critical residues within the binding site, and the cysteine (C242) to which BBO-10203 binds covalently is present only in the α isoform (Fig. 1, F and G).

KRAS binds to p110α with a dissociation constant of 17 μM under physiological salt and pH conditions (Fig. 1H), decreasing to 2.2 μM under lower salt and pH conditions (*17*, *21*), as measured with isothermal titration calorimetry. BBO-10203, covalently bound to the p110α RBD, completely blocked KRAS binding (Fig. 1H), with similar effects observed for HRAS and NRAS (fig. S4, B and C). BBO-10203 blocked the interaction of KRAS-G12D with PI3Kα in human embryonic kidney (HEK) 293T cells with a median inhibitory concentration (IC_{50}) of 6 nM but did not affect KRAS-G12D binding to the protein kinase CRAF (Fig. 1I). Covalent modification of p110α by BBO-10203 did not alter its kinase activity in vitro (Fig. 1J), unlike the PI3Kα inhibitor alpelisib, which binds in the adenosine triphosphate (ATP) pocket and potently inhibits kinase activity (*18*).

Responses to BBO-10203 in specific cell genotypes

BBO-10203 inhibited the PI3Kα pathway, as indicated by reduced phosphorylation of protein kinase AKT in multiple cell lines (Fig. 2A and table S2). Responses clustered into three groups: a nonresponse group characterized by phosphatase and tensin homolog (PTEN) loss, an intermediate group that includes KRAS-G12X mutations and mutations in *PIK3CA* (helical or kinase domains), and a highly responsive group with human epidermal growth factor receptor 2 (HER2) amplification. Figure 2B illustrates that most HER2-amplified cell lines showed near-complete inhibition of phosphorylated AKT (pAKT) by BBO-10203 with IC_{50} values <10 nM (table S2).

The inhibited growth of HER2-amplified tumors by BBO-10203 was unexpected. Inhibition of the interaction between RAS proteins and PI3Kα causes regressions in tumors driven by activated EGFR, suggesting that RAS activation of PI3Kα might extend beyond KRAS mutant cancers (*15*). However, multiple noncanonical RAS-related proteins, including RRAS, RRAS2, and MRAS, can activate PI3Kα (*13*, *22*), and it is not known which proteins are important in this context. The effects of BBO-10203 on pAKT and cell viability were less pronounced in cells driven by receptor tyrosine kinases other than HER2. In HER2-amplified tumors, canonical RAS proteins have been excluded (*23*), suggesting that other noncanonical RAS proteins are involved in this process. The extent to which multiple noncanonical RAS proteins may be involved in each cell line, along with the ability of BBO-10203 to inhibit them, might be responsible for the heterogeneity observed within the group. Regardless of the precise mechanism, the effects of BBO-10203 on these cells are quite strong and suggest a potential role in treating patients suffering from tumors of this genotype.

Hotspot mutations in *PIK3CA* occur in the helical and kinase domains. Kinase domain mutations, particularly H1047R, are considered less dependent on RAS than are helical domain mutations (E542K

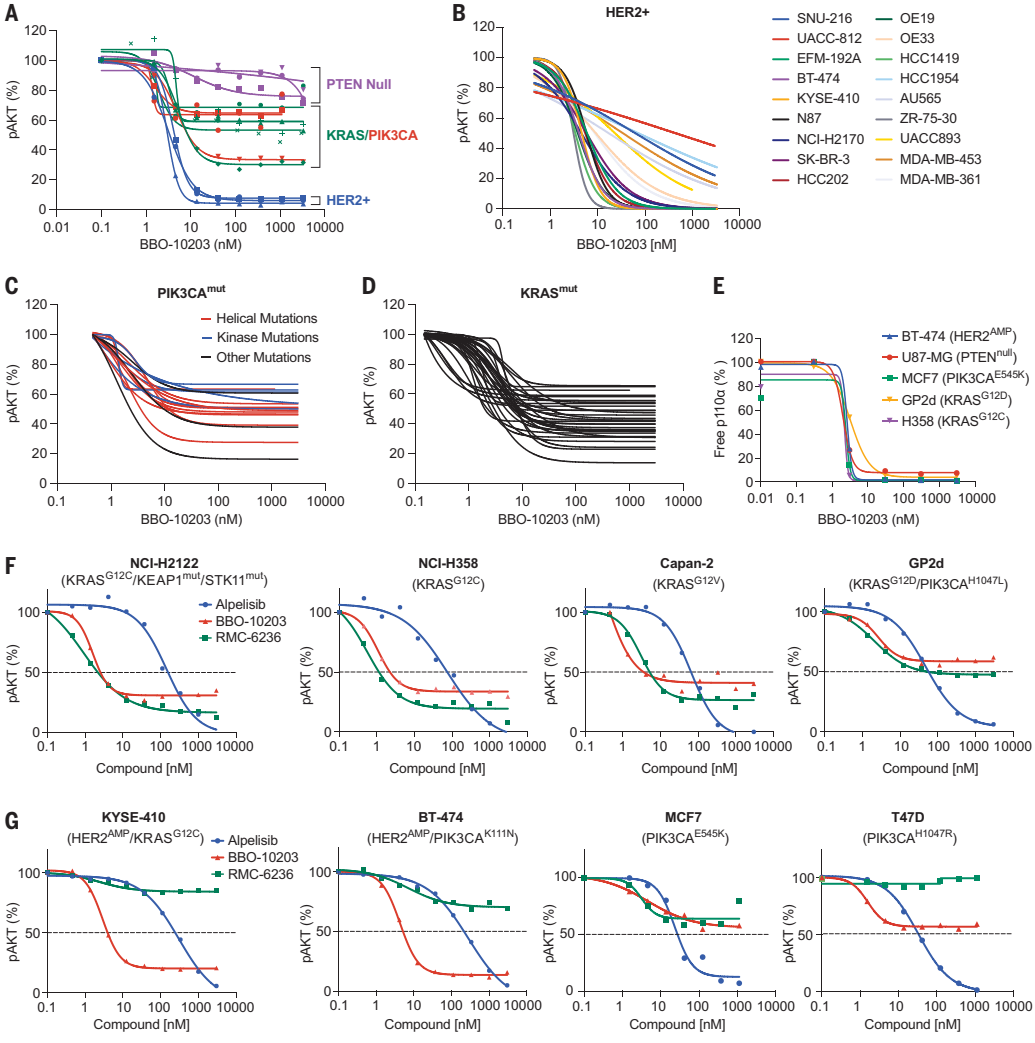


Fig. 2. BBO-10203 activity in various cell lines with diverse cancer genotypes. (**A**) Inhibition of phospho-AKT by BBO-10203 in cancer cell lines carrying different genetic variations, including PTEN loss-of-function mutations (purple), HER2 amplification (blue), *PIK3CA* mutations (red), and *KRAS* mutations (green). (**B**) Inhibition of phospho-AKT by BBO-10203 in HER2-amplified or overexpressing cancer cell lines. (**C**) Inhibition of phospho-AKT by BBO-10203 in *PIK3CA*-mutated cancer cell lines, encompassing helical domain mutations (red), kinase domain mutations (blue), and others (black). (**D**) Inhibition of phospho-AKT by BBO-10203 in KRAS-mutated cancer cell lines. (**E**) Target engagement of P110α by BBO-10203 in cancer cell lines with different genotypes, including HER2-amplified, KRAS-G12X, PTEN-null, and PIK3CA-mutated, using a customized Meso Scale Discovery assay to measure the unbound fraction of p110α. (**F**) Inhibition of phospho-AKT by BBO-10203 (red), the PI3Kα kinase inhibitor alpelisib (blue), and the RAS(ON) inhibitor RMC-6236 (green) in KRAS-mutated cancer cell lines, including NCI-H2122 (G12C), NCI-H358 (G12C), Capan-2 (G12V), and GP2d (G12D). (**G**) Inhibition of phospho-AKT by BBO-10203 (red), the PI3Kα kinase inhibitor alpelisib (blue), and the RAS(ON) inhibitor RMC-6236 (green) in HER2-amplified or *PIK3CA*-mutated cancer cell lines, including KYSE-410 ($HER2^{AMP}$), BT-474 ($HER2^{AMP}$), MCF7 ($Helical^{E545K}$), and T47D ($Kinase^{H1047R}$).

or E545K) (*8*). In our panel of cell lines, both classes of mutation responded to BBO-10203, although kinase domain mutants were indeed less sensitive (Fig. 2C and table S2).

We tested a panel of human cell lines with KRAS mutations and observed pAKT inhibition ranging from 40 to 70% (Fig. 2D and table S2). The incomplete response could result from partial target engagement or could be due to RAS-independent pathways leading to PI3Kα activation. To test these possibilities, we developed a target engagement assay measuring free p110α after exposure to BBO-10203, revealing complete or near-complete engagement across all genotypes (Fig. 2E). Specifically, target engagement was observed in the PTEN-null glioblastoma cell line U87-MG, where pAKT remained unaffected; in the GP2d pancreatic cancer (KRAS-G12D), NCI-H358 lung cancer (KRAS-G12C), and MCF7 breast cancer (PI3Kα-E545K) cell lines, where pAKT was partially reduced; and in the BT-474 breast cancer cell line with HER2 amplification, where pAKT was completely inhibited. We conclude that variations in response to BBO-10203 reflect varying dependencies on RAS proteins for PI3Kα activation.

We compared BBO-10203 with RMC-6236, which inhibits active canonical RAS proteins HRAS, NRAS, and KRAS (*24*). In four KRAS G12 mutant cell lines, both compounds produced highly similar responses, particularly in terms of maximum response achieved (E_{max}), representing the maximum inhibition achievable, regardless of dose. By contrast, RMC-6236 had little effect on cell lines with HER2 amplification, whereas BBO-10203 completely inhibited pAKT, suggesting that PI3Kα activation in these cells likely depends on noncanonical RAS proteins (Fig. 2F). In a breast cancer cell line, MCF7—which has a PI3Kα helical domain mutation (E545K)—both RMC-6236 and BBO-10203 caused partial reduced pAKT levels, indicating RAS-independent signaling. In another breast cancer cell line, T47D, harboring a PI3Kα kinase domain mutation (H1047R), BBO-10203 similarly produced a partial response, whereas RMC-6236 had no effect, further implicating contributions from noncanonical RAS proteins to PI3Kα signaling (Fig. 2G).

In all the cell lines tested, alpelisib inhibited AKT activity completely only at concentrations of 1 μM or higher. At these concentrations, alpelisib not only targets PI3Kα but also effectively inhibits PI3K δ and γ isoforms (*21*). These data indicate that inhibition of multiple PI3K isoforms is necessary for complete inhibition of AKT activity (*22*). In most cases, BBO-10203 inhibited PI3Kα signaling more potently than alpelisib but did not achieve complete inhibition of AKT activity (E_{max}) because its effect only reflects inhibition of RAS-driven PI3Kα activity. Despite differences in mechanism, potency, and E_{max}, when alpelisib was used at concentrations mostly selective for PI3Kα, RNA sequencing analysis and Western blot analysis revealed almost identical changes in global transcription and posttranslational downstream signaling. The similarities confirm that BBO-10203, like alpelisib at concentrations below 1 μM, does indeed act through inhibition of PI3Kα (figs. S5 and S6). Furthermore, BBO-10203 showed excellent selectivity in the Eurofins Discovery SafetyScreen44, KINOMEscan screening panel, and cellular cysteine profiling assays (figs. S7 and S8). The specificity of BBO-10203 for C242 on PI3Kα was further assessed with a nanoluciferase bioluminescence resonance energy transfer (NanoBRET) protein-protein interaction assay. BBO-10203 showed strong activity in disrupting the interaction between PI3Kα and KRAS, with an IC_{50} of 3 nM (fig. S9). When the C242 in PI3Kα was mutated to a serine, BBO-10203 no longer potently inhibited

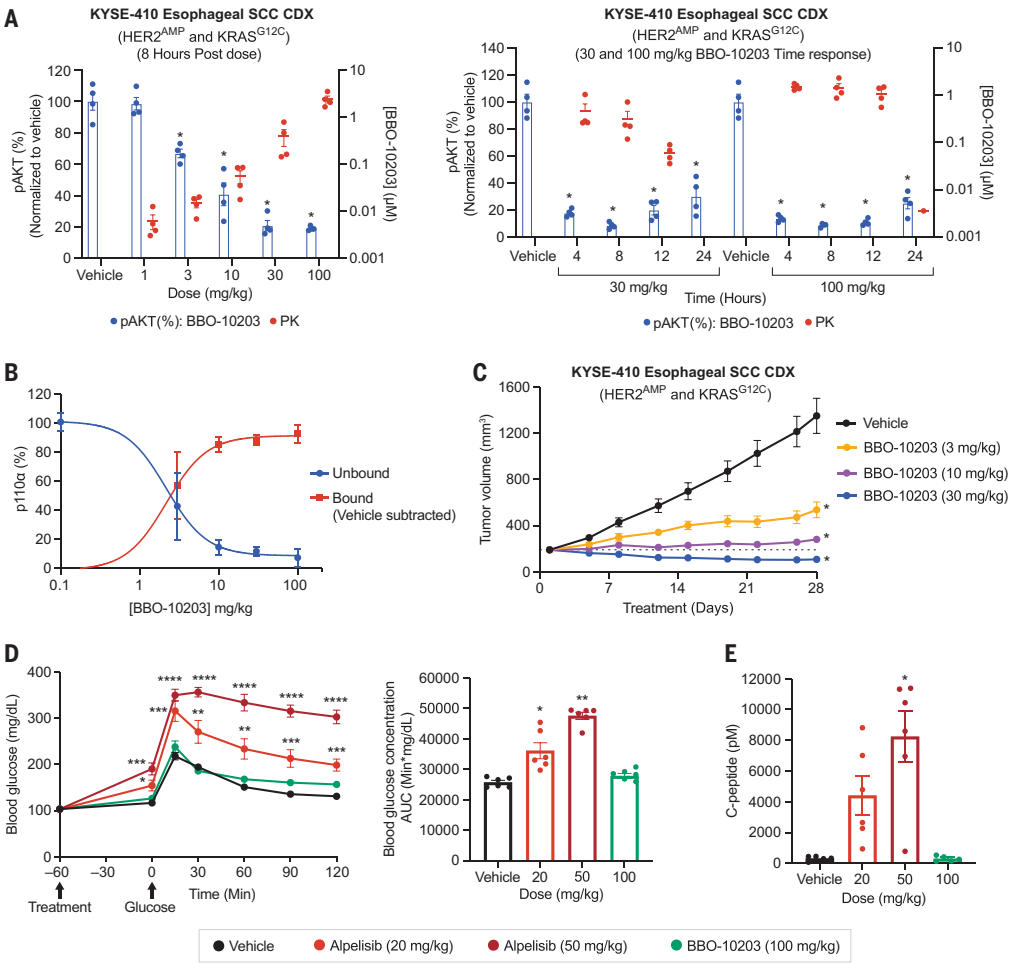


Fig. 3. Inhibition of PI3Kα activity without inducing hyperglycemia. (**A**) Tumor pAKT levels in KYSE-410 tumor-bearing mice at the indicated time points after a single dose of the indicated levels of vehicle or BBO-10203. The indicated groups showed significant inhibition of pAKT compared with the vehicle group (*$P < 0.0001$). (**B**) p110α target engagement of BBO-10203 in KYSE-410 tumors was measured with a Western blot–based band-shift SMaSh assay. A dose-dependent response in p110α target engagement was observed 8 hours after dosing. (**C**) KYSE-410 tumor-bearing mice were dosed with the indicated levels of vehicle or BBO-10203. All treatment groups exhibited significant antitumor activity compared with the vehicle group (*$P <$ 0.0001). (**D**) An oral glucose tolerance test was conducted with fasted male C57BL/6 mice after a single dose of the indicated levels of vehicle, alpelisib, or BBO-10203. The indicated groups demonstrated a significant increase in blood glucose concentrations compared with the vehicle group at each timepoint (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$) (left panel) and in blood glucose concentration area under the curve (AUC) compared with the vehicle group (*$P < 0.001$, **$P < 0.0001$) (right panel). (**E**) Serum C-peptide levels were measured from the mice in (D) at 150 min after the glucose dose. The indicated groups showed a significant increase in C-peptide levels compared with the vehicle group (*$P < 0.001$).

the PI3Kα (C242S)– KRAS interaction, with an IC_{50} of 2000 nM, a 650-fold decrease in potency.

Target engagement in vivo and effects on glucose uptake

BBO-10203 demonstrated good oral bioavailability in mice with a dose-proportional increase in plasma levels, enabling its evaluation in xenograft tumor models (fig. S10). BBO-10203 inhibited accumulation of pAKT in the KYSE-410 cell line–derived xenograft (CDX) model of esophageal squamous cell carcinoma, in a dose-dependent manner from 3 to 100 mg/kg, and showed sustained inhibition over 24 hours at 30 and 100 mg/kg (Fig. 3A). A single 30 mg/kg dose of BBO-10203 achieved complete target engagement in KYSE-410 tumors as measured in a Western blot-based streptavidin mass shift (SMaSh) assay (Fig. 3B). Dose-dependent and effective antitumor activity was observed after daily oral treatment with BBO-10203 in the KYSE-410 CDX model, with tumor regression observed at 30 mg/kg (Fig. 3C). Both the KRAS-G12C inhibitor sotorasib and the HER2 kinase inhibitor tucatinib showed minimum efficacy in this model, and the PI3Kα kinase inhibitor alpelisib led to tumor stasis at the clinically relevant dose of 20 mg/kg (fig. S11).

The PI3Kα kinase inhibitor alpelisib is Food and Drug Administration (FDA)–approved for treating hormone receptor–positive (HR+), HER2-negative (HER2−) breast cancer in which the *PIK3CA* is mutated (*25*). Despite its efficacy in tumors expressing mutated *PIK3CA*, alpelisib also inhibits wild-type PI3Kα, resulting in significant on-target side effects, including hyperglycemia. BBO-10203 does not provoke hyperglycemia, because insulin signaling to PI3Kα does not depend on RAS (*26*). Figure 3D shows that in an oral glucose tolerance test, glucose accumulates in a dose-dependent manner in mice treated with alpelisib, but mice treated with 100 mg/kg of BBO-10203 (three times the efficacious dose) show no accumulation of glucose and no increase in serum C-peptide (Fig. 3E). C-peptide measurements are a commonly used surrogate for insulin measurements owing to the short half-life of insulin (*27*).

BBO-10203 causes tumor regression in combination with targeted therapies

BBO-10203 caused tumor regressions in the KYSE-410 model, but in most models, combinations of treatments are required to achieve tumor regression. Increased pAKT signaling provides acute resistance to multiple targeted therapies such as trastuzumab in HER2-positive (HER2+) breast cancer patients, estrogen receptor (ER) antagonists and cyclin-dependent kinase 4/6 (CDK4/6) inhibitors in HR+ and HER2− breast cancer patients, and KRAS-G12C inhibitors in non–small cell lung cancer patients with KRAS-G12C mutant tumors.

Combinations of BBO-10203 with these standard-of-care therapies are predicted to prevent pAKT-driven resistance while being well-tolerated. A combination of BBO-10203 with the HER2 inhibitor trastuzumab caused inhibition of cell proliferation at concentrations of each agent that are ineffective on their own (Fig. 4A). Near complete and sustained inhibition of pAKT for 72 hours was achieved by the combination of these agents, without affecting p110α expression (Fig. 4B). Similar results were achieved in another HER2-amplified cell line, ZR-75-30, with the same combination. BBO-10203 was also effective in combination with the ER antagonist fulvestrant, the CDK4/6 inhibitor palbociclib, and a novel KRAS-G12C ON inhibitor, BBO-8520 (Fig. 4, C and D) (*28*).

Analysis of PI3Kα inhibitors in cell lines may underrepresent their potential as therapeutics in vivo, because growth in vitro often

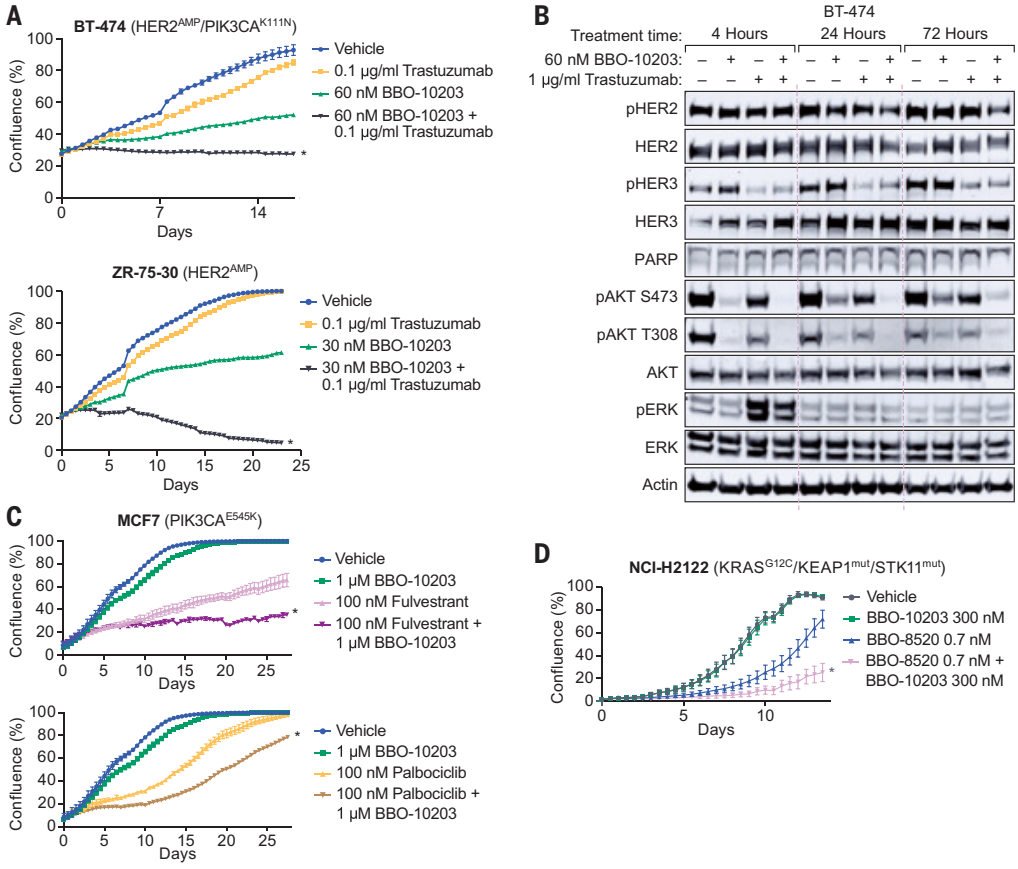


Fig. 4. In vitro activity of BBO-10203 in combination with multiple agents. (**A**) Growth inhibition from mono- and combination therapies of BBO-10203 and trastuzumab was assessed in clonogenic assays, using HER2+ cancer cell lines BT474 and ZR-75-30. Cell confluence was measured every 12 hours on an Incucyte S3 imaging system for 2 to 3 weeks continuously. For (A), (C), and (D), confluence (± SEM) with two-way repeated measures analysis of variance (ANOVA) followed by a Tukey multiple comparison test showed that the indicated treatment groups had a significant activity for the combination treatment compared with each respective monotherapy (*$P < 0.001$). (**B**) Western blot analysis of the upstream and downstream signaling changes of PI3K/AKT and mitogen-activated protein kinase (MAPK) pathways at various time points after treatment with BBO-10203, trastuzumab, or a combination of both. (**C**) Growth inhibition from mono- and combination therapies of BBO-10203 with fulvestrant or palbociclib was evaluated in clonogenic assays, using the ER+/HER2-/*PIK3CA*-E545K cancer cell line MCF7. Cell confluence was measured every 12 hours on an Incucyte S3 imaging system for 4 weeks continuously. (**D**) Growth inhibition from mono- and combination therapies of BBO-10203 with the KRAS-G12C (ON) inhibitor BBO-8520 was assessed in clonogenic assays, using the cancer cell line NCI-H2122. Cell confluence was measured every 12 hours for 2 weeks continuously on an Incucyte S3 imaging system.

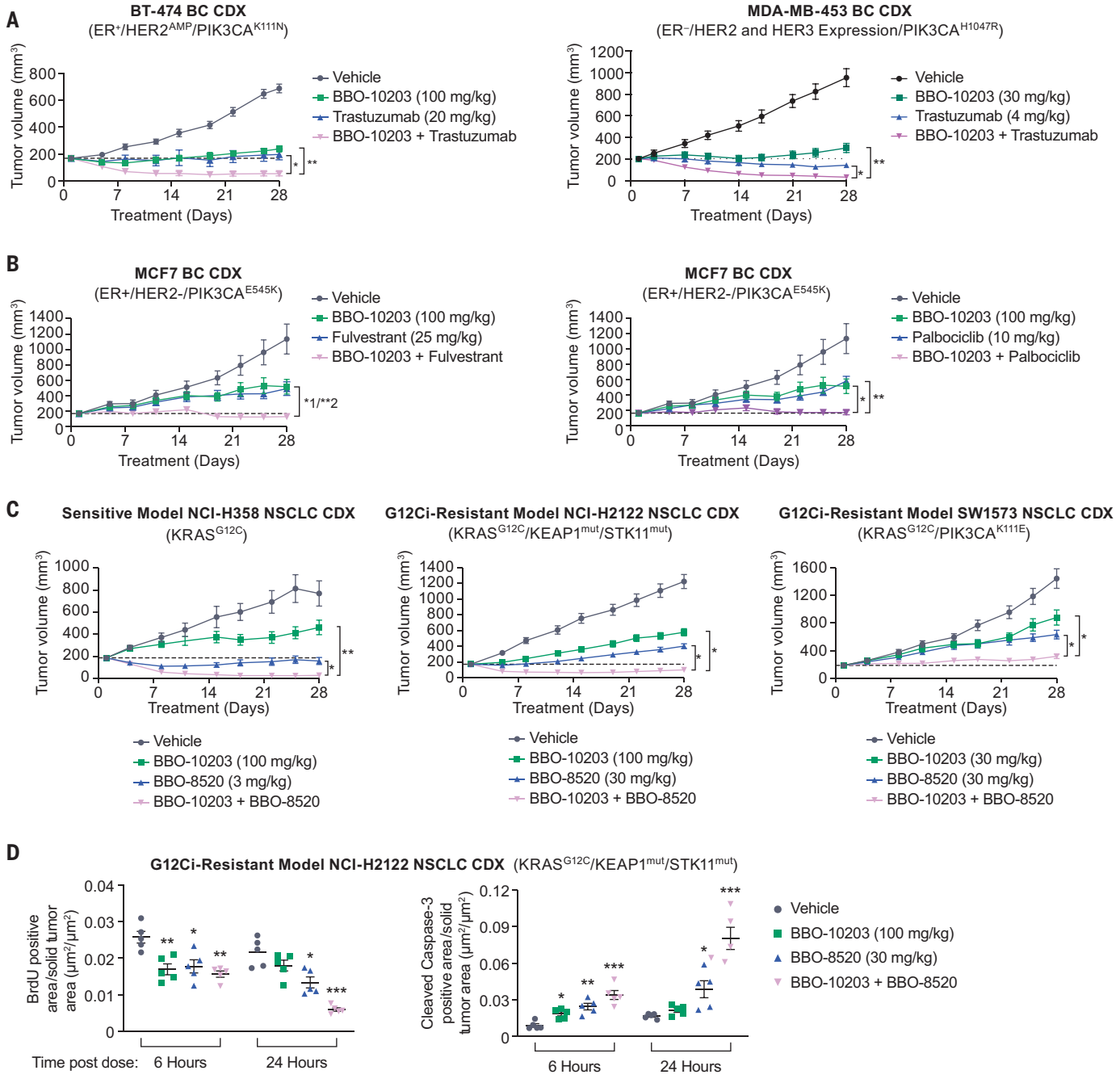


Fig. 5. In vivo activity of BBO-10203 in combination with multiple agents. (**A**) BT-474 (left panel) or MDA-MB-453 (right panel) tumor-bearing mice were dosed with the indicated levels of vehicle, BBO-10203, trastuzumab, or the combination of BBO-10203 and trastuzumab. The indicated monotherapy versus combination groups had a significant difference in antitumor activity from day 5 to 28 (*$P < 0.05$, **$P < 0.0001$) (left panel) and from day 3 to 28 (*$P < 0.01$, **$P < 0.001$) (right panel). (**B**) MCF7 tumor-bearing mice were dosed with the indicated levels of vehicle, BBO-10203, fulvestrant, or the combination of BBO-10203 and fulvestrant (left panel), or with vehicle, BBO-10203, palbociclib, or the combination of BBO-10203 and palbociclib (right panel). The indicated monotherapy versus combination groups had a significant difference in antitumor activity from day 5 to 28 [*$P < 0.01$ (1: fulvestrant) and **$P < 0.001$ (2: BBO-10203) (left panel); (*$P < 0.01$, **$P < 0.001$) (right panel)]. (**C**) NCI-H358 (left panel), NCI-H2122 (middle panel), and SW1573 (right panel) tumor-bearing mice were dosed with the indicated levels of vehicle, BBO-10203, BBO-8520, or the combination of BBO-10203 and BBO-8520. The indicated monotherapy versus combination groups had a significant difference in antitumor activity from day 4 to 28 (*$P < 0.01$, **$P < 0.0001$) (right panel), from day 5 to 28 (*$P < 0.0001$) (middle panel), and from day 4 to 28 (*$P < 0.001$) (left panel). (**D**) Tumor levels of BrdU (left panel) and cleaved caspase-3 (right panel) were measured from NCI-H2122 tumor-bearing mice at 6 or 24 hours after a single dose of the indicated levels of vehicle, BBO-10203, BBO-8520, or the combination of BBO-10203 and BBO-8520. The indicated groups had a significant change in levels compared with the vehicle group [(*$P < 0.01$, **$P < 0.001$, ***$P < 0.0001$) (left panel); (*$P < 0.05$, **$P < 0.001$, ***$P < 0.0001$) (right panel)].

provokes activation of PI3Kα through integrin engagement. Therefore, we tested the combinations described above in xenograft models. Tumor stasis or regressions were achieved in the selected representative models in combination with trastuzumab, fulvestrant, palbociclib, or BBO-8520 (Fig. 5, A to C).

To evaluate the mechanism of action driving the efficacy in the NCI-H2122 CDX model, BBO-10203 and BBO-8520 were dosed alone or in combination. Although amounts of bromodeoxyuridine (a measure of cell proliferation) significantly decreased and amounts of cleaved caspase-3 increased 6 hours after a single dose of BBO-10203, the combination of BBO-10203 and BBO-8520 exhibited much larger effects at both 6 hours and 24 hours. These results demonstrated that BBO-10203 has tumor-intrinsic effects on cell proliferation and survival and that the combination with a KRAS-G12C inhibitor enhances these effects (Fig. 5D).

Summary

Mutations in *PIK3CA* occur frequently in cancer along with driver oncogenes and, in model systems, potentiate driver activity. For example, the expression of an activated allele of *PIK3CA* accelerates the onset and penetrance of tumors driven by BRAF-V600E in mice (*29*). In another example, treatment of mice with KRAS-G12D–driven tumors using PI3K inhibitors demonstrated synergy with mitogen-activated protein kinase kinase (MEK) inhibitors in reducing established tumors (*30*). We show that BBO-10203, combined with KRAS inhibitors, causes the regression of tumors unresponsive to KRAS inhibitors alone, including those with Kelch-like ECH-associated protein 1 (KEAP1) and serine/threonine kinase 11 (STK11) mutations. We also observed tumor regression in response to BBO-10203 in combination with drugs targeting CDK4/6, ER, and HER2. BBO-10203 achieved these effects without detectable toxicity and, most importantly, without provoking hyperglycemia. In tumors with HER2 amplification, BBO-10203 completely inhibited PI3Kα, revealing a pathway essential for tumor cell survival, likely involving a noncanonical RAS protein. BBO-10203 is currently being tested in patients and is expected to provide substantial clinical benefits through a differentiated mechanism that alleviates toxicities that have previously restricted use of PI3Kα inhibitors.

REFERENCES AND NOTES

1. L. C. Foukas *et al.*, *Nature* **441**, 366–370 (2006).
2. L. Bi, I. Okabe, D. J. Bernard, A. Wynshaw-Boris, R. L. Nussbaum, *J. Biol. Chem.* **274**, 10963–10968 (1999).
3. Y. Sugimoto, M. Whitman, L. C. Cantley, R. L. Erikson, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2117–2121 (1984).
4. I. G. Macara, G. V. Marinetti, P. C. Balduzzi, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2728–2732 (1984).
5. D. R. Kaplan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3624–3628 (1986).
6. M. Whitman, D. R. Kaplan, B. Schaffhausen, L. Cantley, T. M. Roberts, *Nature* **315**, 239–242 (1985).
7. M. Whitman, C. P. Downes, M. Keeler, T. Keller, L. Cantley, *Nature* **332**, 644–646 (1988).
8. H. W. Chang *et al.*, *Science* **276**, 1848–1850 (1997).
9. Y. Samuels *et al.*, *Science* **304**, 554 (2004).
10. R. Arafeh, Y. Samuels, *Semin. Cancer Biol.* **59**, 36–49 (2019).
11. P. Rodriguez-Viciana *et al.*, *Nature* **370**, 527–532 (1994).
12. H. Nakhaeizadeh, E. Amin, S. Nakhaei-Rad, R. Dvorsky, M. R. Ahmadian, *PLOS ONE* **11**, e0167145 (2016).
13. P. Rodriguez-Viciana, C. Sabatier, F. McCormick, *Mol. Cell. Biol.* **24**, 4943–4954 (2004).
14. S. Gupta *et al.*, *Cell* **129**, 957–968 (2007).
15. M. M. Murillo *et al.*, *Cell Rep.* **25**, 3545–3553.e2 (2018).
16. M. M. Murillo *et al.*, *J. Clin. Invest.* **124**, 3601–3611 (2014).
17. D. Czyzyk *et al.*, *Nat. Commun.* **16**, 525 (2025).
18. M. Zhang, H. Jang, R. Nussinov, *Chem. Sci.* **11**, 5855–5865 (2020).
19. T. C. Buckles, B. P. Ziemba, G. R. Masson, R. L. Williams, J. J. Falke, *Biophys. J.* **113**, 2396–2405 (2017).
20. K. Terayama *et al.*, *Science* **389,** 402–408 (2025).
21. N. G. Martinez *et al.*, *J. Mol. Biol.* **433**, 166838 (2021).
22. C. Fritsch *et al.*, *Mol. Cancer Ther.* **13**, 1117–1129 (2014).
23. A. Ruiz-Saenz *et al.*, *Cancer Res.* **78**, 3645–3658 (2018).
24. J. Jiang *et al.*, *Cancer Discov.* **14**, 994–1017 (2024).
25. T. Wilhoit, J. M. Patrick, M. B. May, *J. Adv. Pract. Oncol.* **11**, 768–775 (2020).
26. C. Taha, A. Klip, *J. Membr. Biol.* **169**, 1–12 (1999).
27. E. Leighton, C. A. Sainsbury, G. C. Jones, *Diabetes Ther.* **8**, 475–487 (2017).
28. A. E. Maciag *et al.*, *Cancer Res.* **84** (suppl. 7), ND07 (2024).
29. R. P. Charles, J. Silva, G. Iezza, W. A. Phillips, M. McMahon, *Mol. Cancer Res.* **12**, 979–986 (2014).
30. J. A. Engelman *et al.*, *Nat. Med.* **14**, 1351–1356 (2008).

ACKNOWLEDGMENTS

We thank B. Gillette, S. Messing, M. Drew, C. Grose, and their colleagues from the RAS Initiative at the Frederick National Laboratory for their assistance in preparing recombinant proteins. We thank the support from the Lawrence Livermore National Laboratory's institutional Computing Grand Challenge Program. **Funding:** The x-ray crystallography work is based on research conducted at the NE-CAT beamlines, funded by NIH/NIGMS (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by an NIH-ORIP HEI grant (S10OD021527). This research utilizes resources from the Advanced Photon Source, a US Department of Energy Office of Science User Facility operated by Argonne National Laboratory under contract no. DE-AC02-06CH11357. This project has been funded in part with federal funds from the National Cancer Institute (NCI), NIH, under contract no. 75N91019D00024, and Lawrence Livermore National Laboratory under contract no. DE-AC52-07NA27344 (Release: LLNL-JRNL-863848). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US government. **Author contributions:** Conceptualization: D.K.S., F.M., D.V.N., R.X., P.J.B., E.M.W.; Formal analysis: R.X., D.K.S., K.S., J.J., J.P.S., S.F., S.G., C.Z., D.J.C., A.N.H., P.J.B.; Investigation: D.J.C., S.F., E.R., C.Z., A.C., B.P.S., L.F., C.F., D.R., E.K.L., J.-P.D., K.P., L.C.Y., M.C., M.D., S.G., S.D., S.J., W.M., W.Y., Y.Y.; Methodology: B.W., C.F., D.J.C., E.R., K.L., M.A.-B., M.D., P.M.W., P.J.B., S.F., S.G., S.D., Y.Y.; Project administration: D.K.S., R.X., D.V.N., P.J.B., E.M.W.; Resources: D.V.N., E.M.W., P.J.B., A.N.H., F.C.L., F.M.; Supervision: K.S., D.K.S., R.X., J.P.S., K.L., A.E.M., A.G.S., D.E., D.M.T., K.W., S.F., A.N.H., F.C.L., D.V.N., E.M.W., P.J.B., F.M.; Writing – original draft: F.M., P.J.B., D.K.S., R.X.; Writing – review & editing: F.M., P.J.B., D.K.S., R.X., E.M.W., D.V.N. **Competing interests:** All BridgeBio Oncology Therapeutics (BBOT) authors are employees and stockholders of BBOT, a private company. F.M. is a consultant for Amgen, Daiichi Sankyo, Exuma Biotech, Ideaya Biosciences, and Quanta Therapeutics. F.M. is a consultant and cofounder (with ownership interest, including stock options) of BridgeBio. F.M. serves as the scientific advisor for the NCI RAS Initiative at the Frederick National Laboratory for Cancer Research/Leidos Biomedical Research. F.M. has received research grants from Daiichi Sankyo and Gilead Sciences and serves as the principal investigator for research agreements with Boehringer Ingelheim and Roche. A.N.H. has received research grants from BridgeBio, Amgen, Bristol-Myers Squibb, C4 Therapeutics, Eli Lilly, Novartis, Nuvalent, Pfizer, and Scorpion Therapeutics and has served as a compensated consultant for Amgen, Engine Biosciences, Nuvalent, Oncovalent, Pfizer, TigaTx, and Tolremo Therapeutics. R.X., B.W., E.M.W., P.M.W., P.J.B., J.P.S., K.S., Y.Y., F.C.L., D.K.S., D.M.T., D.J.C., and F.M. are coinventors on a patent application titled "Compounds having a T-structure formed by at least four cycles for use in the treatment of cancer and other indications," application no. WO2023/154282. The other authors declare that they have no competing interests. **Data and materials availability:** The atomic coordinates and structure factors of the PI3Kα-RBD covalently bound to BBO-10203 have been deposited in the Protein Data Bank and can be accessed using the accession number 9B4U.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adq2004
Materials and Methods; Figs. S1 to S11; Tables S1 and S2; References (*31*–*46*); MDAR Reproducibility Checklist

Submitted 6 May 2024; resubmitted 25 November 2024; accepted 14 March 2025; published online 12 June 2025

10.1126/science.adq2004

QUANTUM METASURFACES

Metasurface quantum graphs for generalized Hong-Ou-Mandel interference

Kerolos M. A. Yousef, Marco D'Alessandro, Matthew Yeh, Neil Sinclair, Marko Loncar, Federico Capasso*

Multiphoton interference and entanglement are fundamental to quantum information science, yet extending these effects to higher-dimensional systems remains challenging given the imperfections and complexity of scaling conventional linear-optical setups. We present a generalized Hong-Ou-Mandel effect using metasurfaces and graph theory, achieving controlled multiphoton bunching, antibunching, and entanglement across parallel Jones matrix–encoded spatial modes—all within a single-layer metasurface. A graph-theoretic dual framework is introduced that simultaneously encodes the metasurface-based multiport interferometer designs and its resulting nonclassical correlations, enabling the direct translation of linear quantum optical networks into a single-layer metasurface. We also demonstrate the ability of metasurfaces to produce multipath-entangled states and perform transformations equivalent to higher-order Hadamard interferometers. Our results underscore metasurface quantum graphs for scalable, low-decoherence quantum information infrastructure.

In linear quantum optics (LQO), multiphoton multipath interference is fundamental to crafting complex quantum states, which are crucial for advancements in quantum computing and developing technologies (*1*–*4*). When a single photon passes through a multiport interferometer, its output is simply described by a state representing the coherent superposition of all possible exit paths. However, nontrivial quantum effects arise when two or more individual photons interfere, as evidenced by bosonic bunching and antibunching (*5*, *6*). A key manifestation of this phenomenon is known as the Hong-Ou-Mandel (HOM) effect (*7*). In particular, when two individual photons, represented by the creation operators $\hat{a}^{\dagger}_{\mathrm{H},k_1}$ and $\hat{a}^{\dagger}_{\mathrm{H},k_2}$ (Fig. 1A), indistinguishable in all degrees of freedom except the spatial modes (labeled as k_1 and k_2), enter a 50:50 nonpolarizing beam splitter, they bunch and exit together through the same port. This behavior also demonstrates quantum entanglement in the number state (Fock) basis. This phenomenon underpins a wide range of applications, such as the construction of logic gates and the verification of Bell-state entanglement (*8*), laying the foundational basis for universal linear optical quantum computing (*9*). However, extending the two-photon, two-output HOM effect to systems with $n > 2$ output ports using conventional free-space optics usually requires an impractically large number of beam splitters, scaling with $O(n^2)$ (*10*–*13*).

Metasurfaces—planar arrays of subwavelength elements—represent a compelling platform for controlling amplitude, phase, and polarization in an ultrathin form factor (*14*, *15*). Such precise manipulation of light in a compact footprint opens new opportunities for mitigating loss and decoherence inherent in conventional optical setups (*13*, *15*–*18*). Indeed, metasurfaces have already enabled quantum photonic devices, including compact beam splitters, wavefront shapers, and polarization-based state tomography (*19*–*25*). Yet, developing a comprehensive framework to map conventional LQO networks, described by higher-dimensional unitaries, onto a single-layer metasurface remains a nontrivial task (*10*, *13*). As the number of photons and output ports increases, the scattering problem becomes more complex, with an exponential growth in both the interfering amplitudes and possible output events (*26*–*30*).

To address this complexity, it is desirable to adopt an intuitive and visual framework for understanding the intricate dynamics of multiphoton quantum interference. Graph theory offers both a pictorial and mathematical toolset (*31*) to study and predict pairwise correlations in quantum experiments and their outcomes. Graph states, for example, are key resources for one-way quantum computing (*32*), and more recently, they have enabled fault-tolerant quantum computing schemes based on linear optics and beyond (*33*, *34*). These advances motivate us to establish a direct correspondence between metasurface designs for quantum optics—or equivalent parallel optics—and their outcomes, using graph theory to describe multiphoton interference.

Concept

We start by discussing a simple variation of the original HOM experiment, where photons are distinguished by their polarization rather than spatial modes (*35*). In particular, we consider two photons in mode k_1 described by creation operators $\hat{a}^{\dagger}_{\mathrm{H},k_1}$ and $\hat{a}^{\dagger}_{\mathrm{V},k_1}$ (Fig. 1B), with orthogonal polarizations—horizontal (H) and vertical (V)—passing through a half-wave plate (HWP) with a fast axis oriented at 22.5° with respect to the H-polarized photon. The HWP rotates the H and V photon polarizations to 45° and 135° [commonly referred to as diagonal (D) and antidiagonal (A) polarizations], respectively. When these two photons in the D and A bases collinearly enter a polarization beam splitter (PBS), which projects the input into the H and V bases, HOM interference is observed, resulting in a maximally entangled state in the Fock (photon number) basis.

To extend the setup configuration to include more than two output ports, we introduce a scheme that leverages metasurfaces capable of encoding parallel Jones matrix transformations across multiple diffraction orders (*18*). The metasurface consists of spatially varying nanostructures, each defined by an independent Jones matrix J. By carefully designing the matrix-valued transmission function $J(x)$ of the metasurface, acting as a periodic metagrating, its Fourier decomposition $J(x) = \sum_k J_k e^{ik_x x}$ enables each diffraction order to function as an independent polarization analyzer in the far field defined by the Fourier coefficient J_k, which is itself a Jones matrix (*18*, *36*). For detailed discussion on the metasurface design strategy using matrix Fourier optics, see supplementary note 1 (*37*).

Thus, it is possible to condense the cascaded HWP and PBS setup in Fig. 1B into a single-layer metasurface with two diffraction orders acting as D and A analyzers described by Jones matrices $J_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $J_2 = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, respectively. As such, the two photons are now projected into the D and A basis, $\hat{a}^{\dagger}_{\mathrm{H},k_1} \rightarrow \frac{1}{\sqrt{2}}\left(\hat{a}^{\dagger(1)}_{\mathrm{D}} + \hat{a}^{\dagger(2)}_{\mathrm{A}}\right)$ and $\hat{a}^{\dagger}_{\mathrm{V},k_1} \rightarrow \frac{1}{\sqrt{2}}\left(\hat{a}^{\dagger(1)}_{\mathrm{D}} + \hat{a}^{\dagger(2)}_{\mathrm{A}}\right)$, where the superscript refers to photons exiting from diffraction orders 1 and 2, respectively. The resulting two-photon state becomes.

$$|\psi_{\mathrm{out}}\rangle = \frac{1}{2}\left(\hat{a}^{\dagger(1)}_{\mathrm{D}} + \hat{a}^{\dagger(2)}_{\mathrm{A}}\right)\left(\hat{a}^{\dagger(1)}_{\mathrm{D}} - \hat{a}^{\dagger(2)}_{\mathrm{A}}\right)|0,0\rangle = \frac{1}{2}\left(\hat{a}^{\dagger(1)}_{\mathrm{D}}\hat{a}^{\dagger(1)}_{\mathrm{D}} - \hat{a}^{\dagger(2)}_{\mathrm{A}}\hat{a}^{\dagger(2)}_{\mathrm{A}}\right)|0,0\rangle \quad (1)$$

where the bunching of photons in diffraction orders 1 and 2 indicates maximal path entanglement. Additional quantum correlations can be

Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. *Corresponding author. Email: capasso@seas.harvard.edu

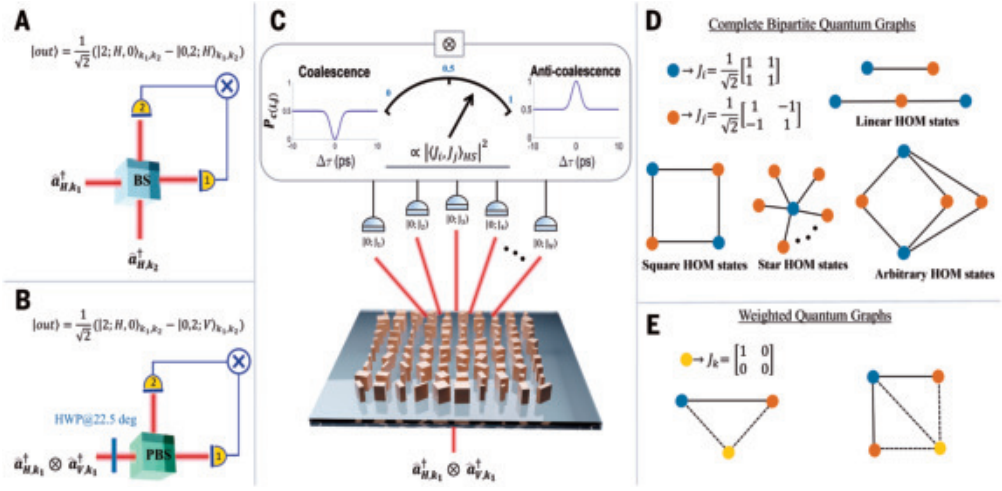


Fig. 1. Generalized quantum interference using metasurfaces and graph theory. (**A**) HOM effect: Two indistinguishable photons, represented by creation operators $\hat{a}^{\dagger}_{\mathrm{H},k_1}$ and $\hat{a}^{\dagger}_{\mathrm{H},k_2}$, enter a 50:50 nonpolarizing beam splitter (BS) and bunch into the same port, demonstrating quantum entanglement in the Fock basis. (**B**) Modified HOM setup using polarization: Two photons, initially orthogonally polarized in the same spatial mode k_1, pass through a half-wave plate (HWP) at 22.5°, rotating to diagonal (D) and antidiagonal (A), then enter a polarization beam splitter (PBS) collinearly, which projects the input into H and V, resulting in maximal path entanglement. (**C**) Metasurface-enabled multioutput HOM interference: Building on (B), a one-dimensional metagrating applies parallel Jones matrix transformations across multiple diffraction orders, each acting as an independent polarization analyzer J_q. Coincidence probabilities $P_{\mathrm{c}(i,j)}$ between orders i and j depend on the Hilbert-Schmidt inner product $\left|\langle J_i, J_j\rangle_{\mathrm{HS}}\right|^2$, indicating the orthogonality of their Jones matrices. If $\left|\langle J_i, J_j\rangle_{\mathrm{HS}}\right|^2 = 0$, a HOM dip (coalescence) appears; if $\left|\langle J_i, J_j\rangle_{\mathrm{HS}}\right|^2 = 1$, an anti-HOM peak (anti-coalescence) occurs. (**D** and **E**) Graph-theoretic representations of metasurface-based quantum correlations, where nodes represent matrix-valued diffraction orders and edges mark pairwise correlations. In (D), blue and orange vertices correspond to orthogonal D and A analyzers corresponding to J_i and J_j, respectively. In (E), introducing J_k nonorthogonal to J_i and J_j creates a weighted quantum graph with dashed edges signifying partial entanglement.

introduced simply by designing new diffraction orders, each characterized by its own Jones matrix. For instance, adding an extra diffraction order J_3 behaving as a D analyzer, such that $\hat{a}^{\dagger}_{\mathrm{H},k_1} \rightarrow \frac{1}{\sqrt{3}}\left(\hat{a}^{\dagger(1)}_{\mathrm{D}} + \hat{a}^{\dagger(2)}_{\mathrm{A}} + \hat{a}^{\dagger(3)}_{\mathrm{D}}\right)$ and $\hat{a}^{\dagger}_{\mathrm{V},k_1} \rightarrow \frac{1}{\sqrt{3}}\left(\hat{a}^{\dagger(1)}_{\mathrm{D}} - \hat{a}^{\dagger(2)}_{\mathrm{A}} + \hat{a}^{\dagger(3)}_{\mathrm{D}}\right)$, the output quantum state now reads

$$|\psi_{\mathrm{out}}\rangle = \frac{1}{3}\left(\hat{a}^{\dagger(1)}_{\mathrm{D}}\hat{a}^{\dagger(1)}_{\mathrm{D}} - \hat{a}^{\dagger(2)}_{\mathrm{A}}\hat{a}^{\dagger(2)}_{\mathrm{A}} + \hat{a}^{\dagger(3)}_{\mathrm{D}}\hat{a}^{\dagger(3)}_{\mathrm{D}} + 2\hat{a}^{\dagger(1)}_{\mathrm{D}}\hat{a}^{\dagger(3)}_{\mathrm{D}}\right)|0,0,0\rangle \quad (2)$$

Here, we observe that all cross-terms representing coincident detection of photons in separate paths cancel each other out except for $\hat{a}^{\dagger(1)}_{\mathrm{D}}\hat{a}^{\dagger(3)}_{\mathrm{D}}$. Consequently, the pairwise correlation of diffraction orders {1,2} and {2,3} exhibits path entanglement, resulting in a HOM dip, while {1,3} leads to a peak in coincidence measurements, often referred to as an anti-HOM effect (*5*).

Let J_q denote the Jones matrix transformation of the qth diffraction order. For two collinear photons in input mode k_1, with arbitrary polarizations α_1 and α_2, injected into a metasurface with N diffraction orders, the metasurface is characterized by the composite Jones matrix $J_{\mathrm{c}} = \frac{1}{\sqrt{N}}\oplus_{q=1}^{N} J_q$ (where $\oplus$ and $\otimes$ denote the direct sum and tensor product, respectively). This composite matrix accounts for the equal probability of photons being diffracted into any of the N orders. The transformed quantum state can then be expressed as

$$|\psi_{\mathrm{out}}\rangle = \left(J_{\mathrm{c}}\hat{a}^{\dagger}_{\alpha_1,k_1}\right) \otimes \left(J_{\mathrm{c}}\hat{a}^{\dagger}_{\alpha_2,k_1}\right)|0\rangle^{\otimes N} \quad (3)$$

To analyze the general behavior of the twofold correlation $P_{\mathrm{c}(i,j)}$ (coincidence probability) between diffraction orders i and j—defined as the probability of simultaneously detecting a single photon at both diffraction orders—we describe the action of the metasurface as simultaneously projecting multiphoton quantum states into a Hilbert-Schmidt space $\mathfrak{B}(\mathcal{H}_d)$ (*38*–*41*). This space comprises linear operators (Jones matrices $\{J_q\}$) acting on the d-dimensional Hilbert space $\mathcal{H}_d$, which describes each photon's polarization state. From Eqs. 1 and 2, we observe that if the Jones matrices $\{J_q\}$ implemented render both photons indistinguishable in every mode q, the HOM effect can be realized by photon bunching. This criterion is met for all Jones matrices $J_q \in \mathfrak{B}(\mathcal{H}_d)$, corresponding to orthogonal (Hermitian) and oblique (non-Hermitian) projectors, satisfying the idempotency condition $\left(J_q^2 = J_q\right)$, which ensures that applying the transformation multiple times does not alter the state. We also observe that the twofold correlation $P_{\mathrm{c}(i,j)}$ between i and j is proportional to the overlap of their corresponding Jones matrix operators, which is quantified using the Hilbert-Schmidt inner product $\langle J_i, J_j\rangle_{\mathrm{HS}} \in \mathfrak{B}(\mathcal{H}_d)$, given by

$$\langle J_i, J_j\rangle_{\mathrm{HS}} = \frac{\mathrm{Tr}\left(J_i^{\dagger}J_j\right)}{\sqrt{\mathrm{Tr}\left(J_i^{\dagger}J_i\right)\mathrm{Tr}\left(J_j^{\dagger}J_j\right)}} \quad (4)$$

where Tr denotes the trace operation, and $J_i^{\dagger}$ is the conjugate transpose of J_i. The inner product $\langle J_i, J_j\rangle_{\mathrm{HS}}$ measures the degree of orthogonality between Jones matrices J_i and J_j: The more orthogonal they are, the less overlap exists between their corresponding subspaces, manifesting as a HOM dip (or an anti-HOM peak for complete overlap), as in Eq. 2. As such, it is possible to decouple the effect of the metasurface-based interferometer from the influence of the input state. In the case of more than two correlated diffraction orders, owing to the conservation of probability, $P_{\mathrm{c}(i,j)}$ can thus be modeled in terms of HOM (bunching) and anti-HOM (antibunching) contributions, which depend on $\langle J_i, J_j\rangle_{\mathrm{HS}}$ and an input distinguishability parameter $\eta_{i,j}$

$$P_{\mathrm{c}(i,j)} = \frac{1}{2}\left[\left(1 - \left|\langle J_i, J_j\rangle_{\mathrm{HS}}\right|^2\right)(1-\eta_{i,j}) + \left|\langle J_i, J_j\rangle_{\mathrm{HS}}\right|^2(1+\eta_{i,j})\right] \quad (5)$$

Where $\eta_{i,j}$ is expressed as

$$\eta_{i,j} = \gamma_{1,2}(\Delta\tau)\left(1 - \left|\frac{\langle \alpha_1|\alpha_2\rangle}{\||\alpha_1\rangle\|\||\alpha_2\rangle\|}\right|^2\right)\left(1 - \frac{1}{2}\sum_{q\in\{i,j\}}^{2}\left|2\left|\frac{\langle \alpha_{12+}|J_q^{\mathrm{H}}\otimes J_q^{\mathrm{H}}|\alpha_{12+}\rangle}{\||\alpha_{12+}\rangle\|\|J_q^{\mathrm{H}}\otimes J_q^{\mathrm{H}}|\alpha_{12+}\rangle\|}\right|^2 - 1\right|^2\right) \quad (6)$$

Here, $\gamma_{1,2}(\Delta\tau)$ is related to the spectrotemporal overlap integral between the two photons as a function of delay $\Delta\tau$. For single photons with Gaussian spectral profiles, $\gamma_{1,2}(\Delta\tau) = \chi e^{-\frac{(\sigma\Delta\tau)^2}{2}}$, where $0 \le \chi \le 1$ is obtained from the mode overlap between the two photons (*42*, *43*). The second factor in $\eta_{i,j}$ (Eq. 6) quantifies the degree of modal orthogonality between the input photons, while the third factor relates to the change-of-basis transformation governing the interference strength between the photons. Here, $|\alpha_{12+}\rangle = \frac{1}{\sqrt{2}}\left(|\alpha_1, \alpha_2\rangle + |\alpha_2, \alpha_1\rangle\right)$ is the total symmetric wavefunction due to the bosonic nature of the two-photon input. In the case of linear polarizations, this last factor can be written as $1 - \langle\cos^2(2\theta)\rangle$, where θ is the difference in polarization angles between the input state $|\alpha_{12+}\rangle$ and the measured output state $J_q^{\mathrm{H}} \otimes J_q^{\mathrm{H}}|\alpha_{12+}\rangle$,

induced by the projective action of the Hermitian part $J_q^{\mathrm{H}} = \sqrt{J_q^{\dagger} J_q}$ of the two-photon Jones matrix $J_q^{\mathrm{H}} \otimes J_q^{\mathrm{H}}$. Since there are two possible projective measurements for calculating the twofold correlation, the resulting expression includes a sum of two terms, capturing the averaged effect.

Figure 1C is a schematic representation of a metasurface that performs parallel transformations on two input photons with orthogonal polarizations (H and V) across N matrix–valued diffraction orders. At $\Delta\tau = 0$, when the metasurface is designed such that the last factor in the distinguishability parameter $\eta_{i,j}$ (Eq. 6) equals one ($\theta = 45°$)—effectively placing each of the two orthogonal input polarizations into equally weighted superpositions for maximum interference—the twofold correlation between pairs of diffraction orders reveals that photon bunching and antibunching can be controlled by adjusting the degree of orthogonality in matrix space between the output modes, as quantified by $\langle J_i, J_j \rangle_{\mathrm{HS}}$. Specifically, if $\left| \langle J_i, J_j \rangle_{\mathrm{HS}} \right|^2 = 0$, the coincidence probability simplifies to

$$P_{\mathrm{c}(i,j)} = \frac{1}{2}\left(1 - \eta_{i,j}\right) \tag{7}$$

resulting in a HOM dip in $P_{\mathrm{c}(i,j)}$ (see Fig. 1C, upper left panel), signifying maximal path entanglement due to photon bunching, provided the photons are temporally indistinguishable ($\eta_{i,j} = 1$). Conversely, if the output modes are fully nonorthogonal in matrix space $\left(\left| \langle J_i, J_j \rangle_{\mathrm{HS}} \right|^2 = 1 \right)$, a peak appears in the coincidence probability according to

$$P_{\mathrm{c}(i,j)} = \frac{1}{2}\left(1 + \eta_{i,j}\right) \tag{8}$$

indicating anti-HOM behavior due to photon antibunching, as required by the conservation of probability (Eq. 5); also see Fig. 1C, upper right panel). This behavior can be described graph-theoretically—we construct a graph with N vertices, each representing an output mode, and the pairwise relationships between vertices are captured in the corresponding $N \times N$ adjacency matrix $\mathbf{A}$ (*31*) [see supplementary note 2 (*37*) for detailed derivations]. The element $A_{i,j}$ is related to the quantum correlation between modes i and j as follows

$$A_{i,j} = \left(1 - \delta_{i,j}\right)\left(1 - P_{\mathrm{c}(i,j)}\right) \text{ where } \delta_{i,j} = \begin{cases} 1 \text{ if } i = j; \\ 0 \text{ if } i \neq j \end{cases} \tag{9}$$

When the distinguishability parameter $\eta_{i,j} = 1$ (as per Eqs. 5 and 6), Fig. 1, D and E, presents graph-theoretic dual representations of both the metasurface design and the resulting quantum correlations, controlled solely by the value of $\left| \langle J_i, J_j \rangle_{\mathrm{HS}} \right|^2$. In Fig. 1D, we illustrate instances of complete bipartite graphs where the vertices are divided into two disjoint sets, representing two orthogonal Jones matrices $\{J_i, J_j\}$—such that every vertex in one set is connected to all vertices in the other, with no edges within the same set (*31*). These vertices correspond to matrix-valued diffraction orders (paths), and the edges E capture the pairwise correlations between them (Fig. 1D).

The first vertex set, V_1 (blue nodes), represents diffraction orders associated with D-analyzers J_i, while the second vertex set, V_2 (orange nodes), corresponds to A-analyzers J_j. The sizes of these subsets are n and m for V_1 and V_2, respectively. Edges E (representing maximal path entanglement) are formed between nodes in V_1 and V_2, that is, $J_i \in V_1$ and $J_j \in V_2$, but not within the same subset, resulting in independent two-vertex sets I_{set}, leading to an anti-HOM peak. Additionally, by introducing a third set of vertices V_3, representing a Jones matrix $J_k \in V_3$ nonorthogonal to $\{J_i, J_j\}$, for example, $J_k = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, weighted graphs can be realized (where the weighted edges are represented by the dashed lines in Fig. 1E). This weight, described by Eq. 9, is related to the probability of measuring coincident events. Because of the dual nature of these graphs, in the two-photon case, they can be interpreted from a metasurface design perspective as (for $\eta_{i,j} = 1$, Eq. 5)

$$\left| \langle J_i, J_j \rangle_{\mathrm{HS}} \right|^2 = \begin{cases} 0 \rightarrow \{J_i, J_j\} \in E\text{: HOM-type entanglement,} \\ \text{Otherwise} \rightarrow \text{partial or no entanglement} \end{cases} \tag{10}$$

Particularly, an edge between two vertices reflects a metasurface design constraint requiring the Hilbert-Schmidt inner product $\left| \langle J_i, J_j \rangle_{\mathrm{HS}} \right|^2$ of their respective Jones matrix transformations to be zero, leading to HOM-type maximal entanglement (Eq. 10). If this condition is not met, partial or no entanglement occurs, indicated by either a weighted edge or an independent set (I_{set}) (Fig. 1E), respectively (Eqs. 5 and 6). On the other hand, these graphs also visually grasp the structure of the corresponding Fock states. By examining the graph properties in Fig. 1D, we can derive a general expression for the Fock state corresponding to a complete bipartite graph $K_{n,m}(V,E)$, representing two photons. Let $V = V_1 \cup V_2$ with V_1 and V_2 being disjoint vertex sets of sizes n and m. We define $I_{\mathrm{set}} \subseteq V$ as the set of vertices for which $\{v_i, v_j\} \notin E$ if and only if $v_i, v_j \in I_{\mathrm{set}}$. The resulting Fock state can be partitioned into bunching and antibunching terms as

$$\left| K_{n,m} \right\rangle = \sum_{k=1}^{N_{\mathrm{bunching}}} \left(a_k^{\dagger} \right)^2 |0\rangle^{\otimes N_{\mathrm{bunching}}} + 2 \sum_{\sigma \in \forall I_{\mathrm{set}}}^{\frac{N_{\mathrm{antibunching}}}{2}} \sqrt{1 - A_{i,j}^{(\sigma)}}\, a_i^{\dagger} a_j^{\dagger} |0\rangle^{\otimes N_{\mathrm{bunching}}} \tag{11}$$

Here, $\hat{a}_k^{\dagger}$ represents the creation operator in mode k. The total number of bunching terms, N_{bunching}, is determined by the number of nodes in the graph, such that $N_{\mathrm{bunching}} = n + m$ (see Fig. 1, D and E). The total number of antibunching terms, $N_{\mathrm{antibunching}}$, is equal to twice the total number of two-vertex independent sets, given by $N_{\mathrm{antibunching}} = n(n - 1) + m(m - 1)$. Therefore, the total number of states is $N_{\mathrm{tot}} = n^2 + m^2$. Moreover, for a complete bipartite graph, the total number of suppressed coincidence events (*29*, *44*), corresponding to the graph's edges, is $N_{\mathrm{E}} = nm$. This result shows that complex quantum states—traditionally requiring cascaded interferometers—can be implemented in a single step using planar metasurfaces, greatly reducing system complexity.

Results and discussion

First, to demonstrate our concept, we designed and fabricated two amorphous silicon metasurfaces [see supplementary note 3 (*37*)] to realize quantum graph correlations analogous to those depicted in Fig. 1D [representing two-vertex linear and square HOM-type graphs; see supplementary note 2 (*37*) for detailed derivations]. The scanning electron micrograph is shown in Fig. 2A. Accordingly, the first metasurface, developed using the matrix Fourier optics approach, was engineered to have two diffraction orders: The +1 diffraction order encodes a diagonal (D) polarization analyzer J_{+1}, whereas the –1 order encodes an antidiagonal (A) analyzer J_{-1}. A schematic representation of the target metasurface operation is provided in Fig. 2B (bottom), while the expected output is based on Eq. 1.

To validate the functionality of the metasurface, we performed classical Mueller polarimetric measurements at a wavelength of 1550 nm and reconstructed the Jones matrix for each diffraction order (*45*). Initially, the expected polarization operations were confirmed through full-wave finite-difference time-domain simulations [see supplementary note 1 (*37*)], which provided reference Mueller matrices and diattenuations (polarization contrast). By sending a set of input Stokes vectors $\left\{ \vec{\mathbf{S}}_{\mathrm{in}} \right\}$ and collecting the output Stokes vectors $\left\{ \vec{\mathbf{S}}_{\mathrm{out}} \right\}$ through a full-Stokes polarimeter, we retrieved the 4×4 Mueller matrix $\mathbf{M}_q$ for each diffraction order [see supplementary note 4 (*37*) for details].

The first row of each Mueller matrix directly corresponds to the Stokes vector $\vec{\mathbf{S}}_q = [M_{01}, M_{02}, M_{03}]/M_{00}$ of the analyzed polarization

state. From the measured $\mathbf{M}_q$, we observed that the projective action of the metasurface corresponded to partial diattenuation $\left(\left|\vec{\mathbf{S}}_q\right| = \frac{\sqrt{M_{01}^2+M_{02}^2+M_{03}^2}}{M_{00}} \leq 1\right)$, as summarized in the histograms of Fig. 2B. To visualize the polarization transformations, we operated on input Stokes vectors uniformly sampling the Poincaré sphere (Fig. 2A, bottom) using $\{\mathbf{M}_q\}$ and plotted the resulting distorted distribution of the output states (Fig. 2B). These results indicate that the projections of the +1 (blue arrow) and −1 (orange arrow) diffraction orders are nearly orthogonal, representing diagonal (D) and antidiagonal (A) analyzers, respectively. Ideally, perfect linear analyzers would collapse all polarization inputs to a single point represented by the state $\vec{\mathbf{S}}_q$. Thus, we modeled the Hermitian component of the Jones matrix J_q^{H} as a partial elliptical diattenuator (*45*, *46*), given by

$$J_q^{\mathrm{H}} = \exp\left(\frac{\tanh^{-1}\left|\vec{\mathbf{S}}_q\right|}{2\left|\vec{\mathbf{S}}_q\right|}\left(M_{01}\sigma_z + M_{02}\sigma_x + M_{03}\sigma_y\right)\right) \quad (12)$$

where, $\sigma_{\{x,y,z\}}$ are the Pauli matrices (*45*, *46*). The second metasurface supports four diffraction orders acting as polarization analyzers, engineered to produce photon statistics and coincidence patterns characteristic of square HOM-type graphs (Fig. 1C). As schematically illustrated in Fig. 2C (bottom), the +1 and +2 orders were designed to operate as D analyzers $\{J_{+1}, J_{+2}\}$, while the −1 and −2 orders function as A analyzers $\{J_{-1}, J_{-2}\}$ (as shown in Fig. 2C, bottom). The experimentally determined projections of the four diffraction orders are plotted on the Poincaré spheres in Fig. 2C, revealing that within each analyzer set of arrows {blue, green} and {orange, purple}, the vectors are nearly aligned, while the two sets are nearly orthogonal, corroborating the intended design of the D and A analyzers, respectively.

Next, we experimentally demonstrated that a single-layer metasurface can generate path-entangled states analogous to those created by a conventional HWP and PBS, as illustrated in Fig. 1B (*35*). The experimental setup, conceptually shown in Fig. 3A, used a type II spontaneous parametric down-conversion (SPDC) source to produce photon pairs with orthogonal polarizations [horizontal (H) and vertical (V)] [supplementary note 5 (*37*)]. These photons were guided through a polarization Mach-Zehnder interferometer containing a delay line to control their relative time delay ($\Delta\tau$). Post-interferometer, photons were transformed using a HWP before being directed to either the reference PBS or the metasurface setup, selectable through a flip mirror (Fig. 3A). Output photons were collected by fiber-coupled collimators and detected using superconducting nanowire single-photon detectors (SNSPDs) for coincidence counting [supplementary note 6 (*37*)].

When cross-polarized photons (H and V) with HWP set to 0° traversed the metasurface, scanning $\Delta\tau$ revealed a HOM dip in the coincidence counts (Fig. 3B, left and far right), indicative of bosonic coalescence. To quantify the strength of quantum interference and entanglement facilitated by the metasurface, we compared the HOM dip visibility of the metasurface to that of the reference PBS. The HOM visibility is defined as $\mathrm{VIS} = \frac{|I_+ - I_-|}{I_+}$, where I_+ and I_- are the maximum and minimum coincidence counts, respectively. By rotating the HWP to 22.5°, we measured two-photon interference at the PBS, establishing an upper visibility limit of ~71% (*20*, *35*), whereas the metasurface achieved a visibility of ~61%. Both surpassed the classical visibility threshold of 50% (Fig. 3B, solid gray line) [supplementary note 7 (*37*)].

To rule out the influence of imperfections in the fabricated metasurface on the observed interference, we applied Eqs. 5 and 6 to predict the coincidence probabilities. In the ideal case of perfect diattenuators $\left(\left|\vec{\mathbf{S}}_q\right| = 1\right)$, the Hermitian component J_q^{H} is equivalent to the reduced density matrix ρ_q^{out} of the single photon detected at the qth diffraction order. Hence, one can directly examine orthogonality between these subspaces using $\langle J_i, J_j \rangle_{\mathrm{HS}}$ (Eq. 5) in terms of their corresponding Jones matrices.

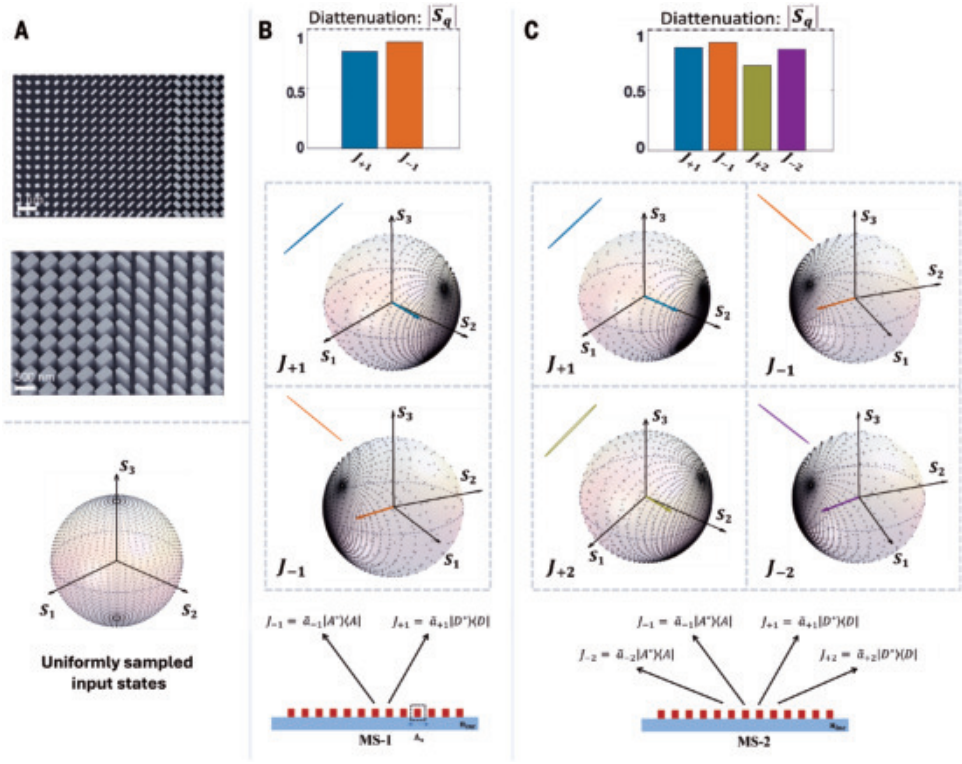


Fig. 2. Metasurface designs and classical polarimetric characterization. (**A**) (Top) Scanning electron microscopy images of the fabricated devices. (Bottom) Representation of the input states uniformly sampled over the Poincaré sphere. (**B**) Design and characterization of the two-order metasurface. (Bottom) Schematic of the first metasurface (MS-1), encoding two diffraction orders: +1 order acts as a diagonal (D) polarization analyzer, and the −1 order acts as an antidiagonal (A) analyzer. (Top) Experimental confirmation of the metasurface's functionality via classical Mueller polarimetry. Results are plotted on the Poincaré spheres, with gray dots indicating the concentration of the output polarization states for a uniformly distributed set of input states. The blue arrow represents the +1 diffraction order (D analyzer), and the orange arrow represents the −1 order (A analyzer). Both arrows correspond to the analyzed state $\vec{\mathbf{S}}_q$, while the polarization ellipse insets highlight their orientation. Histogram displays the associated diattenuation $\left|\vec{\mathbf{S}}_q\right|$. (**C**) Design and characterization of the four-order metasurface. (Bottom) Schematic of the second metasurface (MS-2), engineered to have four diffraction orders, functioning as polarization analyzers. The +1 and +2 diffraction orders are designed as D analyzers, while the −1 and −2 orders act as A analyzers. (Top) Experimental characterization of the four diffraction orders on the Poincaré spheres, with corresponding diattenuation values. Blue and green arrows correspond to the +1 and +2 orders (D analyzers), while the orange and purple arrows represent the −1 and −2 orders (A analyzers).

However, for partial diattenuators $\left(\left|\vec{\mathbf{S}}_q\right|<1\right)$, J_q^{H} no longer equals ρ_q^{out}; instead, $\rho_q^{\mathrm{out}}=J_q^{\mathrm{H}}\rho_r^{\mathrm{in}}J_q^{\mathrm{H}\dagger}$ (*38*, *41*), where ρ_r^{in} is the single-photon reduced density matrix of the two-photon input. For two orthogonally polarized photons (i.e., H and V) impinging simultaneously on the metasurface, we have, $\rho_r^{\mathrm{in}}=\frac{1}{2}(|\mathrm{H}\rangle\langle\mathrm{H}|+|\mathrm{V}\rangle\langle\mathrm{V}|)=\frac{1}{2}\mathbf{I}$, yielding $\rho_q^{\mathrm{out}}=\frac{1}{2}J_qJ_q^{\mathrm{H}\dagger}$. In this scenario, a photon detected at the qth order may originate from either H or V with equal probability. Since ρ_q^{out} can be approximated from the classically characterized Jones matrices, we therefore refer to it as the effective Jones matrix J_q^{eff}

$$J_q^{\mathrm{eff}}=\frac{1}{2}J_q^{\mathrm{H}}J_q^{\mathrm{H}\dagger} \tag{13}$$

To compute the predicted coincidence probabilities $P_{c(i,j)}$ under partial diattenuation (Eqs. 12 and 13), the Hilbert-Schmidt inner product in Eq. 5 is replaced by $\left\langle J_i^{\mathrm{eff}},J_j^{\mathrm{eff}}\right\rangle_{\mathrm{HS}}$, reflecting the overlap between single-photon density matrices at output modes i and j. Thus, by substituting the reconstructed effective Jones matrices (Eq. 13) into Eq. 5 and setting $\chi=0.71$ in $\gamma_{1,2}(\Delta\tau)$ (Eq. 6)—corresponding to the maximum attainable visibility in the reference measurement—we theoretically predict both HOM and anti-HOM behaviors (Fig. 3B, solid red curves). These predictions closely match the experimental data (black dots for the metasurface, and purple dots for PBS) averaged over five second-order correlation $g^{(2)}$ measurements, with error bars indicating the standard deviation.

Figure 3B (middle and right) compares the behavior of the beam splitter and the metasurface with HWP set to 0° and 22.5°. When HWP is rotated to 22.5°, aligning the photon polarizations with the metasurface's projective analyzers, a flat line is theoretically expected in the correlation measurement (in contrast to a dip seen with a PBS). However, the experimental data for the metasurface (Fig. 3B, middle) showed a slight dip because of the partial diattenuation of the analyzers implemented in both diffraction orders. By incorporating the classically characterized J_q^{eff} into Eq. 5, assuming D- and A-polarized two-photon input, we modeled the coincidence behavior with a fidelity of 99.95% (Fig. 3B, middle, solid red curve). See supplementary note 8 (*37*) for fidelity calculations.

To further validate our approach, we experimentally reconstructed the square-HOM graph (Fig. 1D) by performing six two-fold coincidence measurements with HWP at 0°, ensuring one photon was H-polarized and the other V-polarized. These measurements correspond to four edges ($N_{\mathrm{E}}=4$) and two independent sets in the graph. According to Eqs. 5 and 9, edges should yield HOM dips, whereas independent sets should produce anti-HOM peaks as the path length difference varies. Indeed, the diffraction order pairs {(+1, −1), (+1, −2), (−1, +2), (−2, +2)} exhibited HOM dips exceeding the classical limit (Fig. 3C), where black dots represent the average of five $g^{(2)}$ measurements, and error bars indicate standard deviation. From a graph-theoretic perspective, these orders define edge coordinates, whereas from a metasurface design standpoint, they act as nearly orthogonal projectors satisfying Eqs. 5 and 6, as shown on the Poincaré sphere (Fig. 2C).

Conversely, the pairs {(+1, +2), (−1, −2)} displayed anti-HOM peaks (Fig. 3C), consistent with Eq. 5, and had a smaller angular difference on the Poincaré sphere, indicating nonorthogonal Jones matrices. Notably, the observed anti-HOM peak is attributed to the anticoalescence (antibunching) of bosons (*47*).

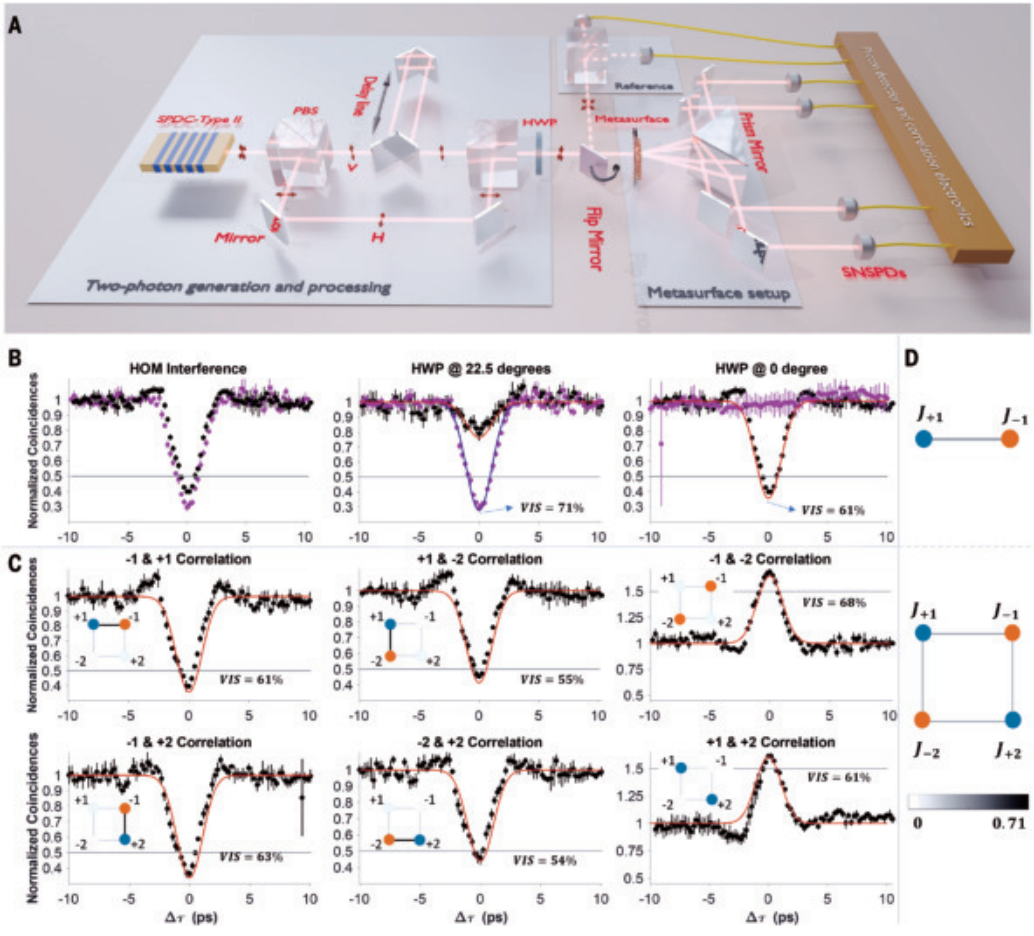


Fig. 3. Experimental demonstration of metasurface-enabled quantum multiport interferometry. (**A**) Schematic of the setup. (1) Two-photon generation and processing: Photon pairs with orthogonal polarizations (H and V) from a type II SPDC source pass through a polarization Mach-Zehnder interferometer introducing a delay $\Delta\tau$. (2) Reference: A PBS provides a benchmark for maximum HOM visibility. (3) Metasurface: Selected output channels are fiber-coupled and sent to SNSPDs for coincidence measurements. (**B**) Measured second-order correlation $g^{(2)}$ versus $\Delta\tau$. (Left) A HOM dip is observed for both the two-order metasurface (black) and PBS (purple), albeit with different visibilities (VIS). (Middle) With HWP at 22.5°, the metasurface correlation remains nearly flat, whereas the PBS dip reaches 71%. (Right) Setting HWP to 0° reverses this behavior, yielding a 61% dip for the metasurface and a flat PBS response; red curves indicate theoretical predictions based on Eq. 5. (**C**) Four-order metasurface coincidence measurements reveal four HOM dips and two anti-HOM peaks, corresponding to edges and independent sets, respectively, in the quantum graph insets. The solid red curves represent the theoretically predicted behavior. (**D**) Reconstructed two- and four-order quantum-correlation graphs, where each edge's color (normalized to the reference PBS) indicates the two-photon visibility, confirm the theoretically predicted interference patterns.

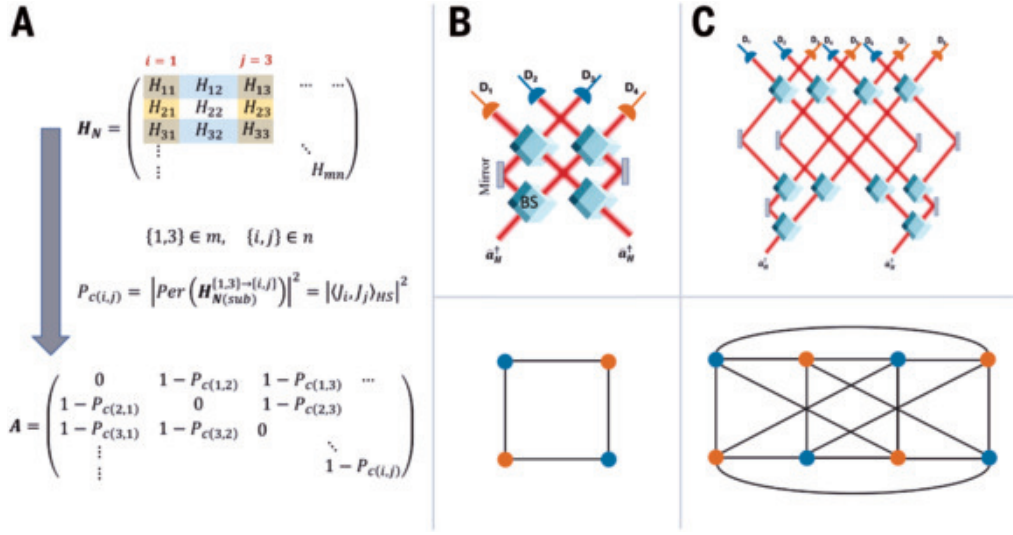


Fig. 4. Graph representations of truncated Hadamard interferometers. (**A**) Conceptual framework for mapping a multiport Hadamard interferometer of order N, described by $\boldsymbol{H}_N$, onto a single-layer metasurface defined by an adjacency matrix $\boldsymbol{A}$. The coincidence probability $P_{c(i,j)}$ is described as the permanent of the submatrix transformation $\left|\mathrm{Per}\left(\boldsymbol{H}^{\{1,3\}\to\{i,j\}}_{(N)\mathrm{sub}}\right)\right|^2$, corresponding to the scattering amplitude of two input photons entering from orthogonal modes k_1 and k_3 and exiting from i and j output channels. A single-layer metasurface can replicate any higher-dimensional Hadamard transformation, with its design extracted from the corresponding complete bipartite graph. Therefore, the output quantum correlations would satisfy the equality: $P_{c(i,j)} = \left|\mathrm{Per}\left(\boldsymbol{H}^{\{1,3\}\to\{i,j\}}_{(N)\mathrm{sub}}\right)\right|^2 = \left|\langle J_i, J_j\rangle_{\mathrm{HS}}\right|^2$. (**B** and **C**) (Top) Implementations of truncated Hadamard interferometers of the second and third order, constructed using 50:50 nonpolarizing beam splitters, mirrors, and single-photon detectors D_n, when two indistinguishable photons impinge on these interferometers. (Bottom) Fock graph states corresponding to Hadamard interferometers with 2^N output ports.

Specifically, the transformation implemented by the four-channel metasurface leads to a balanced gain and loss in probability, governed by Eq. 5, as a function of temporal distinguishability. This occurs because, when twofold correlations are measured by selecting outputs from only two diffraction orders out of four, the remaining output modes act as correlated loss channels. When the photons are temporally indistinguishable, specific output configurations directed to loss channels are suppressed through quantum interference, enhancing coincidences in the measured modes, as per Eq. 5, rendering the system non-Hermitian (*29*).

By calculating the effective Jones matrix J_q^{eff}, we precisely characterized how the four-order metasurface transforms single-photon reduced density matrices then substituted J_q^{eff} into Eq. 5 to obtain theoretical predictions (Fig. 3C, solid red line) that match the experimental data with an average fidelity of 99.97%. The reconstructed quantum-correlation graphs for two- and four-order metasurfaces are shown in Fig. 3D, where edge color reflects the normalized two-photon visibility relative to the PBS reference.

Metasurface-enabled higher-order Hadamard interferometry

Our theoretical framework and the experimental results of Fig. 3 reveal a broader implication stemming from the dual nature of these quantum graphs: The correlation outcomes of existing LQO networks can be represented as graphs of Fock states and then reinterpreted as metasurface designs. Notably, we demonstrate that metasurfaces can greatly simplify the design of an entire class of interferometers—Hadamard interferometers (*11*, *48*)—which traditionally require a complex network of beam splitters. The transformation of a Hadamard interferometer with 2^N output ports is governed by a Hadamard matrix $\boldsymbol{H}_N$ of order N (Fig. 4A)

$$\boldsymbol{H}_N = (\boldsymbol{H}_1)^{\otimes N}, \text{ where } \boldsymbol{H}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{14}$$

An example of a simple optical system that performs a Hadamard transformation of order one $\boldsymbol{H}_1$, is a 50:50 nonpolarizing beam splitter. In the two-photon case, this transformation leads to the well-known HOM interference, as illustrated in Fig. 1A. A truncated Hadamard interferometer of order N with two input photons and 2^N output ports normally requires $N_{\mathrm{BS}} = 3 \times 2^{N-1} - 2$ balanced nonpolarizing beam splitters, where N_{BS} scales exponentially with the degree N of the Hadamard transformation . In this case, the coincidence probability $\mathrm{P}_{c(i,j)}$ is given by the permanent (Per) of the 2×2 submatrix transformation $\left|\mathrm{Per}\left(\boldsymbol{H}^{\{1,3\}\to\{i,j\}}_{(N)\mathrm{sub}}\right)\right|^2$ describing the scattering amplitudes of two indistinguishable photons injected (*30*, *49*, *50*), for example (Fig. 4A), into two orthogonal spatial modes $\{1,3\} \in m$, and detected at output modes $\{i,j\} \in n$ (*26*, *30*), as per the example shown in Fig. 4A, would have permanent [also see supplementary note 9 (*37*)]

$$\left|\mathrm{Per}\left(\boldsymbol{H}^{\{1,3\}\to\{i,j\}}_{(N)\mathrm{sub}}\right)\right|^2 = \left|\mathrm{Per}\begin{pmatrix} H_{1i} & H_{1j} \\ H_{3i} & H_{3j} \end{pmatrix}\right|^2 = \left|H_{1i}H_{3j} + H_{1j}H_{3i}\right|^2 \tag{15}$$

Instances of truncated Hadamard interferometers of order 2 and 3 constructed using 50:50 nonpolarizing beam splitters, mirrors, and single-photon detectors (D_n) are shown in Fig. 4, B and C (top). These interferometers are configured to distribute the input photons to a set of 2^N output ports corresponding to two orthogonal Hadamard modes. By applying the graph-theoretic approach described earlier, we can cast the pairwise correlations of such interferometers into a graph, where, according to Eq. 9, $A_{i,j} \propto 1 - \left(P_{c(i,j)} = \left|\mathrm{Per}\left(\boldsymbol{H}^{\{1,3\}\to\{i,j\}}_{(N)\mathrm{sub}}\right)\right|^2\right)$. As depicted in Fig. 4, B and C (bottom), we find that these graphs consistently correspond to complete bipartite graphs for any $\boldsymbol{H}_N$, reflecting the mutual orthogonality of the rows and columns in the Hadamard matrix.

However, by using the metasurface design protocol outlined earlier, we show that any higher-order truncated Hadamard interference involving two photons can be efficiently implemented using a single-layer metasurface, as demonstrated both theoretically and experimentally in Figs. 1 and 3, respectively. Thus, the metasurface condition ensures (Fig. 4A)

$$P_{c(i,j)} = \left|\mathrm{Per}\left(\boldsymbol{H}^{\{1,3\}\to\{i,j\}}_{(N)\mathrm{sub}}\right)\right|^2 = \left|\left\langle J_i, J_j\right\rangle_{\mathrm{HS}}\right|^2 \quad (16)$$

As such, each output port—represented as a node in the corresponding graph (Fig. 4, B and C)—is replaced by a metasurface matrix-encoded diffraction order corresponding to a set of orthogonal analyzers.

Concluding remarks

Our generalized framework for quantum metasurface–based multiport interferometry demonstrates that Jones matrix–encoded diffraction orders can replicate high-dimensional linear optical transformations that would otherwise require complex beam-splitter networks—such as higher order Hadamard interferometers (Figs. 3 and 4). Central to our approach is the Hilbert-Schmidt space generalization, which reveals that quantum correlations between diffraction orders are directly proportional to the overlap of their respective Jones matrix transformations. This allows us to precisely predict twofold correlations and disentangle the effect of input photons' distinguishability from the metasurface's transformation.

Moreover, our graph-theoretic approach—where diffraction orders map to nodes and photon-correlation events map to edges—captures the experimentally observed multiphoton interference, encompassing both HOM (bunching) and anti-HOM (antibunching) phenomena. Our method can be naturally extended to other degrees of freedom, the study of quantum walks of correlated photons and other quantum information tasks (*15*, *51*) within a single metasurface layer. Our results open previously unexplored possibilities for high-dimensional quantum interference in loss-tolerant photonic architectures and advance the development of compact, scalable quantum optical systems [see supplementary notes 10 and 11 (*37*)].

REFERENCES AND NOTES

1. J. L. O'Brien, *Science* **318**, 1567–1570 (2007).
2. J. O'Brien, A. Furusawa, J. Vučković, *Nat. Photonics* **3**, 687–695 (2009).
3. H.-S. Zhong *et al.*, *Science* **370**, 1460–1463 (2020).
4. C. H. Bennett, D. P. DiVincenzo, *Nature* **404**, 247–255 (2000).
5. B. Vest *et al.*, *Science* **356**, 1373–1376 (2017).
6. N. Spagnolo *et al.*, *Phys. Rev. Lett.* **111**, 130503 (2013).
7. C. K. Hong, Z. Y. Ou, L. Mandel, *Phys. Rev. Lett.* **59**, 2044–2046 (1987).
8. F. Bouchard *et al.*, *Rep. Prog. Phys.* **84**, 012402 (2021).
9. E. Knill, R. Laflamme, G. J. Milburn, *Nature* **409**, 46–52 (2001).
10. M. Reck, A. Zeilinger, H. J. Bernstein, P. Bertani, *Phys. Rev. Lett.* **73**, 58–61 (1994).
11. R. Barak, Y. Ben-Aryeh, *J. Opt. Soc. Am. B* **24**, 231–240 (2007).
12. D. J. Brod *et al.*, *Adv. Photonics* **1**, 034001 (2019).
13. W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, I. A. Walmsley, *Optica* **3**, 1460–1465 (2016).
14. N. Yu *et al.*, *Science* **334**, 333–337 (2011).
15. A. H. Dorrah, F. Capasso, *Science* **376**, eabi6860 (2022).
16. M. Khorasaninejad *et al.*, *Science* **352**, 1190–1194 (2016).
17. A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, B. Luk'yanchuk, *Science* **354**, aag2472 (2016).
18. N. A. Rubin *et al.*, *Science* **365**, eaax1839 (2019).
19. K. An *et al.*, *Nat. Commun.* **15**, 3944 (2024).
20. P. Georgi *et al.*, *Light Sci. Appl.* **8**, 70 (2019).
21. L. Li *et al.*, *Science* **368**, 1487–1490 (2020).
22. A. S. Solntsev, G. S. Agarwal, Y. S. Kivshar, *Nat. Photonics* **15**, 327–336 (2021).
23. T. Stav *et al.*, *Science* **361**, 1101–1104 (2018).
24. K. Wang *et al.*, *Science* **361**, 1104–1108 (2018).
25. T. Santiago-Cruz *et al.*, *Science* **377**, 991–995 (2022).
26. J. B. Spring *et al.*, *Science* **339**, 798–801 (2013).
27. N. Spagnolo *et al.*, *Nat. Photonics* **8**, 615–620 (2014).
28. J. Carolan *et al.*, *Science* **349**, 711–716 (2015).
29. M. C. Tichy, M. Tiersch, F. de Melo, F. Mintert, A. Buchleitner, *Phys. Rev. Lett.* **104**, 220405 (2010).
30. S. Aaronson, A. Arkhipov, in *STOC '11: Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing* (Association for Computing Machinery, 2011), pp. 333–342.
31. J. A. Bondy, U. S. R. Murty, *Graph Theory with Applications* (Macmillan London, 1976).
32. P. Walther *et al.*, *Nature* **434**, 169–176 (2005).
33. S. Bartolucci *et al.*, *Nat. Commun.* **14**, 912 (2023).
34. X. Gu, M. Erhard, A. Zeilinger, M. Krenn, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4147–4155 (2019).
35. W. P. Grice, I. A. Walmsley, *Phys. Rev. A* **56**, 1627–1634 (1997).
36. N. A. Rubin, A. Zaidi, A. H. Dorrah, Z. Shi, F. Capasso, *Sci. Adv.* **7**, eabg7488 (2021).
37. See supplementary materials.
38. B. C. Hall, *Quantum Theory for Mathematicians*, vol. 267 in Graduate Texts in Mathematics (Springer New York, 2013).
39. M. Keyl, R. F. Werner, in *Quantum Information: From Foundations to Quantum Technology Applications*, D. Bruß, G. Leuchs, Eds. (Wiley, 2016), chap. 5, pp. 75–90.
40. I. Bengtsson, K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge Univ. Press, 2006).
41. M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge Univ. Press, 2010).
42. C. Drago, A. M. Bran'czyk, *Can. J. Phys.* **102**, 411–421 (2024).
43. M. Tillmann *et al.*, *Phys. Rev. X* **5**, 041015 (2015).
44. A. Crespi, *Phys. Rev. A* **91**, 013811 (2015).
45. R. A. Chipman, W.-S. T. Lam, G. Young, *Polarized Light and Optical Systems* (CRC Press, ed. 1, 2018).
46. N. A. Rubin, Z. Shi, F. Capasso, *Adv. Opt. Photonics* **13**, 836–970 (2021).
47. Q. Li *et al.*, *Nat. Photonics* **15**, 267–271 (2021).
48. U. Chabaud, E. Diamanti, D. Markham, E. Kashefi, A. Joux, *Phys. Rev. A* **98**, 062318 (2018).
49. E. R. Caianiello, *Nuovo Cim.* **10**, 1634–1652 (1953).
50. M. A. Broome *et al.*, *Science* **339**, 794–798 (2013).
51. A. Peruzzo *et al.*, *Science* **329**, 1500–1503 (2010).

ACKNOWLEDGMENTS

We thank N. A. Rubin, L. Li, E. Hu, and M. Abobeih for helpful discussions; C. Jin and C. J. Xin for assistance in maintaining the superconducting nanowire single-photon detectors; K. Yang and R. Jiang for providing the pump laser; J. Oh, Z. Sun, and D. Cassara for advice on nanofabrication. All individuals acknowledged are affiliated with Harvard University. **Funding:** K.M.A.Y. acknowledges partial financial support from the Harvard Quantum Initiative–Amazon Web Service Generation Q Fellowship Program. F.C. acknowledges support from the Air Force Office of Scientific Research under award no. FA9550-21-1-0312. This work was performed at the Harvard University Center for Nanoscale Systems (CNS), a member of the National Nanotechnology Coordinated Infrastructure Network (NNCI), supported by the National Science Foundation under award no. ECCS-2025158. **Author contributions:** K.M.A.Y. developed the theory, designed and simulated the device, and conceived of the experiment. M.D. fabricated the metasurfaces and acquired the scanning electron microscopy images. K.M.A.Y. and M.D. built the measurement setup, performed the experiments, and analyzed the data. K.M.A.Y., M.D., and M.Y. contributed to operating and collecting data from the superconducting nanowire single-photon detectors. K.M.A.Y., F.C., and M.D. wrote the manuscript. N.S., M.L., and F.C. supervised the research. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adw8404
Materials and Methods; Supplementary Text; Figs. S1 to S5

Submitted 21 February 2025; accepted 2 June 2025

10.1126/science.adw8404

PHONON PHYSICS

Atom-by-atom imaging of moiré phasons with electron ptychography

Yichao Zhang[1,2,3], Ballal Ahammed[2,4], Sang Hyun Bae[1,2], Chia-Hao Lee[1,2], Jeffrey Huang[1,2], Mohammad Abir Hossain[2,4], Tawfiqur Rakib[2,4,5], Arend M. van der Zande[1,2,4,5], Elif Ertekin[2,4,5], Pinshane Y. Huang[1,2,5]*

Twisted two-dimensional materials exhibit distinctive vibrational modes called moiré phonons, which arise from the moiré superlattice. Here, we demonstrate atom-by-atom imaging of phasons, an ultrasoft class of moiré phonons in twisted bilayer tungsten diselenide (WSe_2). Using ultrahigh-resolution (<15 picometers) electron ptychography, we imaged the size and shape of each atom to extract time-averaged vibrational amplitudes as a function of twist angle and position. We observed several signature properties of moiré phasons, such as increased vibrational amplitudes at solitons and AA-stacked regions. By correlating experiments with molecular dynamics simulations and lattice dynamics calculations, we show that phasons dominate the thermal vibrations in low-angle twisted bilayers. These results represent a powerful route to image thermal vibrations at atomic resolution, unlocking experimental studies of a thus far hidden branch of moiré phonon physics.

Twisted van der Waals bilayers exhibit distinctive vibrational modes, broadly defined as moiré phonons (*1*–*5*) that arise from the moiré superlattice. These modes tune with interlayer twist angle and couple strongly to important physical properties of twisted van der Waals materials, ranging from charge transport to unconventional superconductivity (*3*–*19*). Experimental studies of moiré phonons have focused mainly on the optical modes of the moiré superlattice, such as the interlayer breathing and shear modes with frequencies of tens of wave numbers (we refer to these as "higher-frequency moiré phonons") that can be accessed with low–wave number Raman spectroscopy (*20*, *21*). However, two-dimensional (2D) moiré materials are predicted (*3*–*5*) to host a second class of moiré phonons, which have not been directly observed. These modes, which we refer to as phasons, are the lowest energy acoustic or acoustic-like modes resulting from the moiré superlattice.

At small twist angles of a few degrees, van der Waals bilayers reconstruct into moiré superlattices (*22*–*24*) containing quasi-periodic regions of rotationally aligned regions (such as AA or AB), separated by a network of stacking faults often referred to as solitons (Fig. 1, A to C). In these low-twist angle moirés, the solitons form a soft elastic network that hosts a set of vibrational modes that are not present in the individual monolayers. Just as phasons generally represent excitations in the phase of incommensurate modulations, the moiré phason modes are the in-plane vibrational modes of the incommensurate modulations represented by the soliton network. In real space, the moiré phasons correspond to localized interlayer sliding at the soliton domain walls and translations of the moiré superlattice (*3*–*5*).

Moiré phasons have been referred to as ultrasoft shear modes or ultrasoft moiré phonons. These terms are often used interchangeably, but phasons are more precisely defined as vibrational modes of incommensurate phases that resemble acoustic modes with zero-energy excitations at infinitely long wavelengths (*25*, *26*). We follow the convention of referring to the ultrasoft shear modes as phasons regardless of their precise twist angle and commensurability (*27*) because real moiré superlattices are typically disordered across length scales of tens of nanometers and above (*28*–*32*), including in our samples. For example, the triangular AB regions in Fig. 1A vary in their size and symmetry, indicating that the sample is not perfectly periodic across the ~30-nm field of view.

Although moiré phasons have been discussed in several theoretical and computational works (*3*–*5*, *9*, *25*, *27*, *33*–*35*), they have only been detected indirectly through electron-phason coupling (*36*). These phasons are extremely difficult to access experimentally because their predicted frequencies and energies are very low, on the order of 0.1 cm^{-1} or 0.01 meV (*5*), placing them out of reach of conventional spectroscopic techniques. Nonetheless, moiré phasons have already been implicated in diverse and important phenomena. They have been predicted to underlie superlubricity and disorder in the moiré lattice (*27*), alter charge transport (*33*, *37*) and exciton diffusion (*9*), dominate thermal conductivity (*37*) and specific heat at low temperatures (*5*), and couple strongly to the emergent electronic properties at low-angle twists (*4*, *27*, *36*, *38*).

The challenges in measuring moiré phasons point to a gap in experimental techniques for probing thermal vibrations. Existing diffraction-based measurements, such as x-ray or neutron scattering, lack the sensitivity and atomic-scale spatial resolution needed to directly visualize phasons. Conversely, spectroscopic techniques such as Raman spectroscopy or monochromated electron energy-loss spectroscopy can either access sufficiently low frequencies to detect moiré phonon modes (*1*, *21*) or operate at atomic-scale resolution (*39*–*42*), but not both.

Rapid advances in segmented detectors (*43*), 4D scanning transmission electron microscopy (STEM), and electron ptychography (*44*) offer a potential new route for measuring atomic vibrations. Electron ptychography can achieve spatial resolutions of tens of picometers (*44*–*53*), sufficiently high that thermal vibrations measurably blur atomic images (*44*). These advances indicate the potential for electron ptychography to serve as a real-space, atomically resolved method for imaging thermal vibrations, with the potential to newly access low-frequency, anisotropic, and spatially localized modes that are beyond the reach of present methods.

Achieving <15-pm resolution with multislice electron ptychography

Figure 1 shows how we used multislice electron ptychography (MEP) to achieve the picometer-scale spatial resolution required to image thermal vibrations in twisted bilayer WSe_2. We fabricated a WSe_2 bilayer (see materials and methods) with a 1.7° interlayer twist angle (0° is defined as parallel or AB stacking). Next, we imaged the sample with aberration-corrected annular dark-field STEM (ADF-STEM) to visualize the moiré superlattice and locate a region of interest. Figure 1, A to C, shows the resulting images and corresponding atomic structures, which reveal the reconstructed moiré superlattice and its network of soliton stacking faults.

Figure 1D shows an image of the bilayer WSe_2 obtained with MEP (*44*, *54*), which calculates an object from a series of diffraction patterns acquired as an atomic-scale electron beam is rastered across a sample. In an ideal case, the reconstructed phase object is proportional to the electrostatic potential, which is blurred by the thermal vibration of

[1]Department of Materials Science and Engineering, University of Illinois Urbana-Champaign, Urbana, IL, USA. [2]The Grainger College of Engineering, University of Illinois Urbana-Champaign, Urbana, IL, USA. [3]Department of Materials Science and Engineering, University of Maryland, College Park, MD, USA. [4]Department of Mechanical Science and Engineering, University of Illinois Urbana-Champaign, Urbana, IL, USA. [5]Materials Research Laboratory, University of Illinois Urbana-Champaign, Urbana, IL, USA. *Corresponding author. Email: pyhuang@illinois.edu

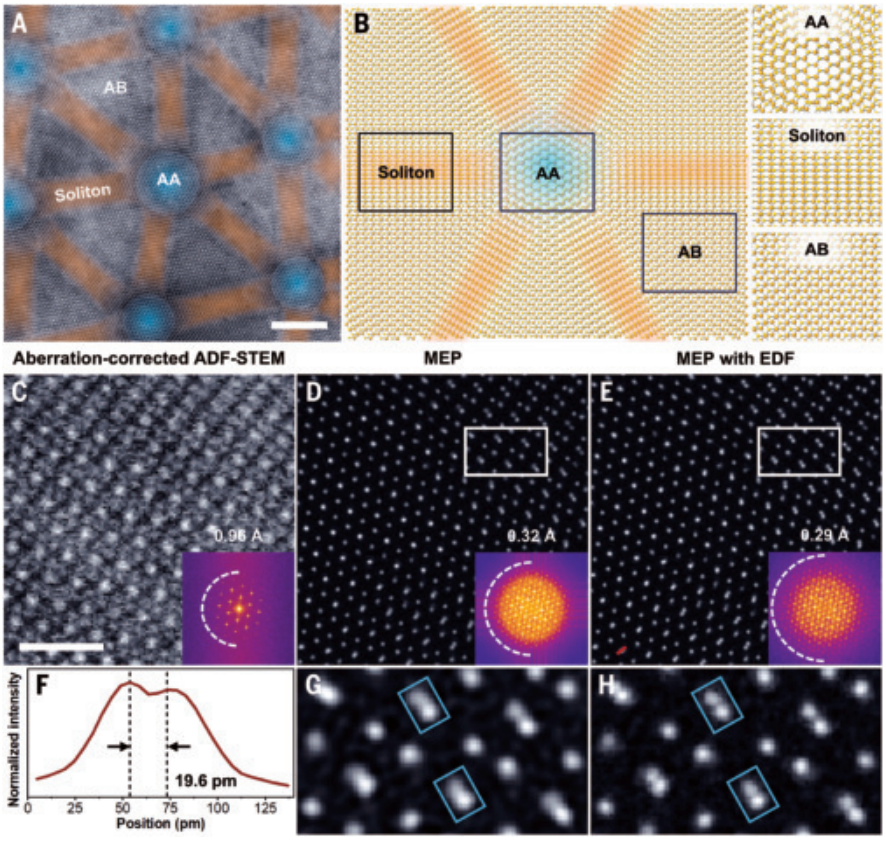


Fig. 1. Picometer-scale resolution achieved with MEP. (**A**) Aberration-corrected ADF-STEM image of 1.7°-twisted bilayer WSe_2 with soliton (orange) and AA stacked (blue) regions shaded. Scale bar is 5 nm. (**B**) Atomic structure of 1.7°-twisted bilayer WSe_2. The insets magnify representative AA, soliton, and AB regions. W atoms are orange, and Se atoms are yellow. (**C** to **E**) Images of soliton regions in 1.7°-twisted bilayer WSe_2 obtained by (C) aberration-corrected ADF-STEM, (D) MEP, and (E) MEP with EDF. The insets of (C) to (E) are their corresponding Fourier transforms. Information transfer limits (0.96 Å for aberration-corrected ADF-STEM, 0.32 Å for MEP, and 0.29 Å for MEP with EDF) are labeled with dashed white arcs. Scale bars are 5 Å for (C) to (E). The red line at the bottom left of (E) marks the position of the line profile in (F). (**F**) Line profile showing a resolvable projected atomic spacing of 19.6 pm. (**G** and **H**) Magnified ptychographic phase images from areas outlined by white boxes in (D) and (E), respectively. Blue rectangles label two pairs of atoms unresolvable with MEP but resolvable using MEP with EDF.

atoms during the few-second image acquisition. MEP has received intense recent interest because it has achieved the highest spatial resolution of any real-space imaging method (*44*, *55*). With MEP (Fig. 1D), we obtained exceptionally clear images of each atomic site (see fig. S1 and materials and methods). For example, the Fourier transform of the ptychography image in Fig. 1D displayed peaks indicating 0.32-Å information transfer, triple that of the aberration-corrected STEM image in Fig. 1C (0.96 Å).

We further improved the 2D spatial resolution of MEP by using an extended depth of field (EDF) algorithm (*56*, *57*) to extract the 2D projection in Fig. 1E (see individual slices in figs. S2 to S4 and materials and methods). EDF algorithms are commonly used to extract all-in-focus image projections from focal series (*57*); here, we adapted this method to remove background signals from MEP depth slices that did not contain atoms. We applied this method to all ptychography images unless otherwise specified. Using this method, we improved the information transfer even further to <0.29 Å. This high resolution allowed us to resolve detailed atomic structures, even in atom pairs that were blurred together in conventional MEP projections (blue boxes in Fig. 1, G and H).

As a second measure of resolution, we determined the minimum projected atomic spacings resolvable with a visible intensity dip between the atoms, in analogy to the Abbe limit. We observed atomic spacings of 19.6 pm (Fig. 1F) and 14.7 pm (fig. S5). These values approached calculated room temperature in-plane vibrational amplitudes of 5 to 6 pm for monolayer WSe_2 (*58*), indicating that the vibrational amplitudes might be accessible with ptychography.

Visualization of moiré phasons in a soliton domain wall

In Fig. 2, we leveraged the extremely high spatial resolution of ptychography to measure anisotropic thermal vibrations in a 2D moiré superlattice. We acquired and reconstructed an MEP image of 1.7°-twisted bilayer WSe_2 at room temperature, in a region containing a soliton and AB domains (Fig. 2A). The position-averaged convergent beam electron diffraction patterns from this region indicate that the region is primarily oriented perpendicular to the electron beam (see fig. S6). As shown in Fig. 2, B and C, we fit 2D Gaussians to each atomic site in the image and obtained their in-plane short- and long-axis Gaussian root mean square (RMS) widths, ellipticity, and long-axis orientation (see materials and methods). These data are akin to atomic anisotropic displacement parameters (*59*), also called thermal ellipsoids, obtained from the Debye-Waller factor in x-ray diffraction. Unlike x-ray measurements, however, our method is atomically resolved rather than spatially averaged, making it ideal for resolving spatially localized modes such as moiré phonons and phasons.

Figure 2, D and E, shows an atom-by-atom map and corresponding polar plot displaying the magnitude and orientation of elliptical 2D Gaussian fits of the single W atoms (see fig. S7 and materials and methods). We analyzed single W atoms so that local shear or tilt would not influence the shape and size of the atoms (see materials and methods). We also excluded atoms that were too close to their nearest neighbors in projection. These steps ensured that the extracted shape and size of each W atom was due to thermal vibration rather than projection artifact.

These data show that within the soliton, W atoms had increased long-axis lengths and aligned parallel to the soliton direction (−36° in Fig. 2E), whereas in the AB region, their orientations were shorter with more randomly distributed orientations. We obtained similar results in a second region in the same sample (fig. S8) and in a corresponding analysis of Se columns (fig. S9). As a test of robustness, we repeated our processing methods on conventional MEP projection images without applying EDF and obtained similar results (fig. S10). The large amplitude vibrations, their strong ellipticity, and their localization to the solitons are unusual features that are consistent with the predictions for moiré phasons (*3*–*5*).

We performed harmonic lattice dynamics calculations (see materials and methods) to visualize the phason modes and compared them with our experimental data. Figure 2, F and G, and movie S1 show the lowest energy eigenmode of a 1.7° moiré superlattice, represented as displacement vectors for the two W layers (see fig. S11 for the eigenvector maps of all atoms). In this phason mode, the motion was highly localized to the unstable stacking regions: the solitons (orange) and

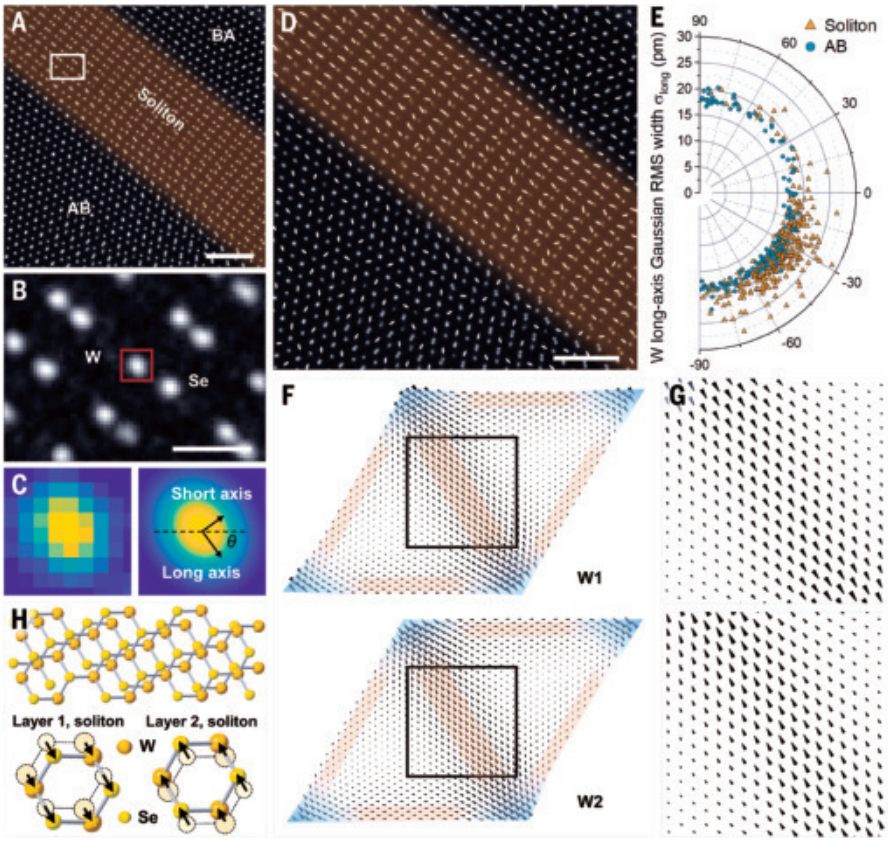


Fig. 2. Direct imaging of anisotropic phason modes at a soliton. (**A**) MEP image of 1.7°-twisted bilayer WSe_2. Orange shading marks the soliton region. Scale bar is 1 nm. (**B**) Magnified image from the white rectangle in (A) showing ellipticity of atoms. The ellipticity magnitude is calculated as $\sigma_{long}/\sigma_{short}$. Scale bar is 2 Å. (**C**) Magnified image (left) and 2D Gaussian fit (right) of a single W atom labeled by the red box in (B). (**D**) Vector field map of anisotropic phonon vibration overlaid on (A). White arrows indicate the relative magnitude and direction of the ellipticity of all single W atoms, magnified for visibility. Scale bar is 1 nm. (**E**) Polar plot of single W atom long-axis Gaussian RMS width σ_{long}. (**F**) Eigenvector maps of the lowest energy phonon mode of a 1.7°-twisted bilayer WSe_2 supercell. This mode is one of two phason or ultrasoft shear modes (see fig. S11 for the second mode). The top and bottom panels, respectively, show the displacement of W atoms in the top and bottom layers. The phason mode is localized to the AA (blue) and soliton (orange) regions. The size of the black arrows are magnified for visibility. (**G**) Magnified eigenvector maps from the black boxes in (F). (**H**) Schematic illustrations of atomic displacements in the soliton due to the phason mode. Shown at the top is the atomic structure of the soliton; shown at the bottom are the magnified in-plane atomic displacements in the different layers.

their connection points, the AA sites (blue). Within the solitons, the W and Se atoms in each monolayer displaced in parallel, whereas the two WSe_2 layers displaced antiparallel with respect to one another (Fig. 2H). The second-lowest-energy phason mode, nearly degenerate at the gamma point, contained similar motions for the symmetry-equivalent solitons in the moiré superlattice (fig. S11 and movie S2). The dispersion relations of these modes are shown in fig. S12 and convergence tests in fig. S13 and are discussed in the supplementary text. Our experimental ellipticity maps in Fig. 2D were qualitatively consistent with the anisotropic thermal vibration expected to arise if the phason modes dominated the vibrational amplitudes. These data indicate that our ptychography methods can directly image the moiré phason modes and their impact on thermal vibrations.

Thermal vibrations across the moiré superlattice

In Fig. 3, we show how phasons and higher-frequency moiré phonons affected the vibrational motions and atomic profiles in MEP, considering all thermally activated modes. We studied a 2.45°-twisted bilayer WSe_2, as shown in the MEP image in Fig. 3A. Our experimental measurements were conducted at room temperature, and so although we expected the vibrational amplitudes to be dominated by the lowest frequency and thus highest amplitude modes, multiple phonon modes should be populated according to a Boltzmann distribution.

We conducted molecular dynamics (MD) simulations to understand the total atomic vibration amplitudes from all thermally activated phonon modes at room temperature (see materials and methods). First, we used MD to calculate the atomic trajectories at 300 K. The atomic coordinates at each time step ($\Delta t = 0.5$ fs for a total of 3 ns, corresponding to ~30 periods of the lowest frequency modes) populated a probability density function (PDF) for each atom. Figure 3B shows these probability maps and corresponding 2D Gaussian fits for three W atoms in the AA, AB, and soliton regions. These PDFs revealed key trends in the real-space distribution of moiré phonon modes. The total vibrational amplitudes were isotropic in AA and AB regions (approximately equal short- and long-axis amplitudes with $\sigma_{AA} > \sigma_{AB}$) and anisotropic in the soliton, consistent with the observations reported in Fig. 2.

These simulations also provided a series of predicted signatures for the expected in-plane vibrational profiles resulting from phasons. These are (i) anisotropic vibrational amplitudes within the solitons, resulting in increased long-axis amplitudes; (ii) orientational alignment of anisotropy in the soliton in the direction parallel to the soliton; and (iii) increased, isotropic vibrational amplitudes at AA sites. Having previously established (i) and (ii), we next examined (iii), the vibrational amplitudes at AA sites and how they compared with other regions of the moiré superlattice.

We visualized the varying vibrational amplitudes in real space, comparing simulation and experiment. Figure 3, C and E, displays experimental and MD maps showing the short-axis amplitudes of the thermal ellipsoids for individual W atoms, comparing similar fields of view for a 2.45°-twisted superlattice. We focused first on the short axes of the W atoms because we expected them to be more robust to potential experimental artifacts, such as errors in 2D Gaussian fitting.

To produce the experimental map in Fig. 3C, we acquired, reconstructed, and fitted 2D Gaussians to each W atom in MEP data as described previously. Next, we isolated the thermal vibration amplitudes from the ptychographic atom profiles by deconvolving a blurring factor to remove nonthermal contributions to the atom size (see fig. S14 for undeconvolved short-axis Gaussian RMS width). To calculate this blurring factor, we used ptychography simulations to model the contributions from the finite width of the atomic potential and the finite resolution of the ptychographic reconstruction [see fig. S15 and materials and methods for details of the ptychography simulations and deconvolution, as well as extraction of thermal root mean square displacement (RMSD) from MD simulation results]. Figure S16 shows that the precise magnitude of the blurring factor does not change the qualitative spatial trends of the vibrational amplitudes, but it does

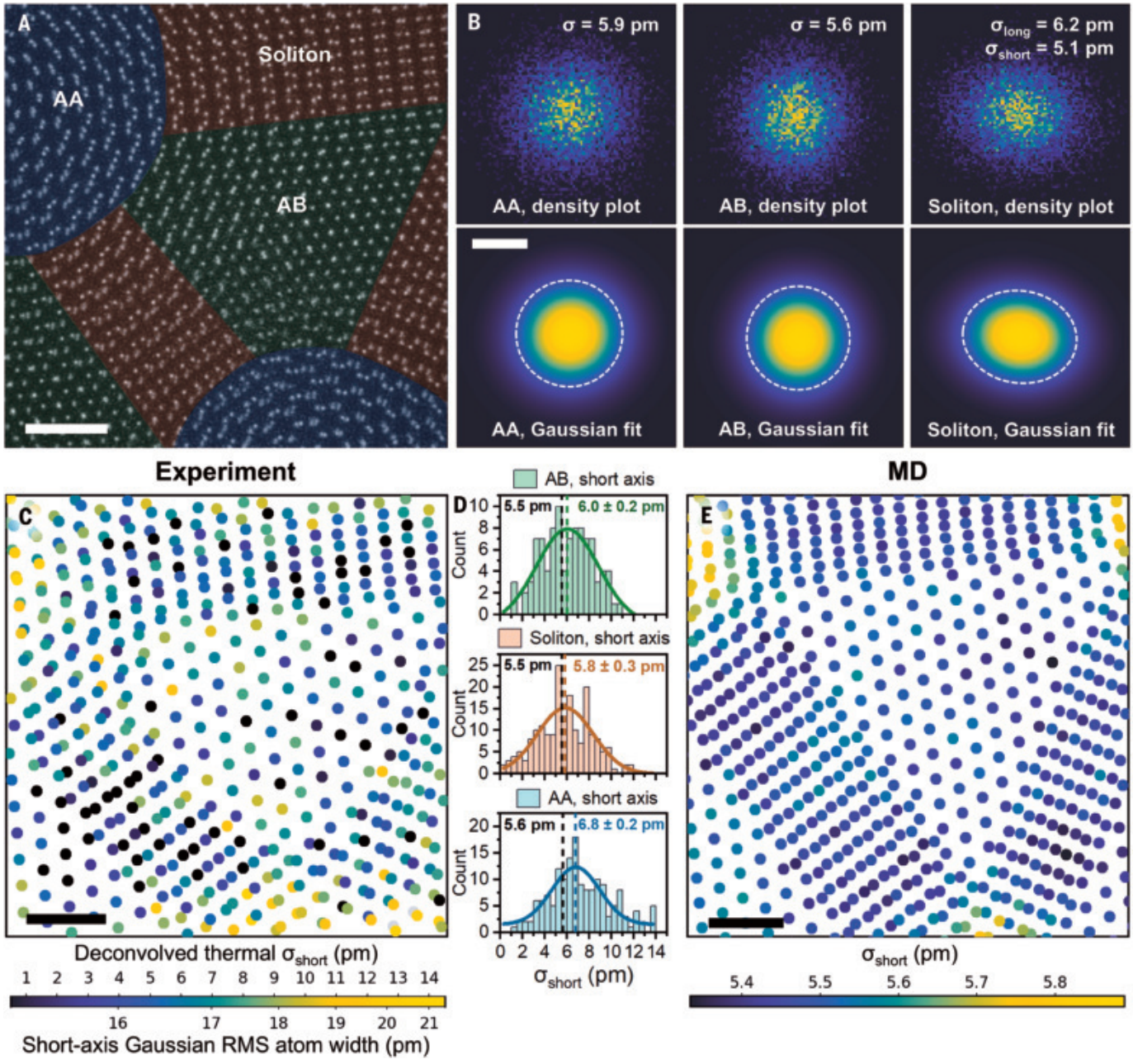


Fig. 3. Experimental and MD comparison of single W atom sizes. (**A**) MEP image of 2.45°-twisted bilayer WSe_2. Soliton, AA, and AB regions are labeled with orange, blue, and green, respectively. (**B**) Extraction of thermal vibration amplitude in 2.45°-twisted bilayer WSe_2 from MD simulations. The top row shows probability density function plots for single W atoms in AA (left), AB (middle), and soliton (right) regions. The bottom row shows corresponding 2D Gaussian fits, where σ is the Gaussian RMS width. See materials and methods for details on generating the PDFs and Gaussian fits. Dashed circles and ellipses mark full width at quarter maximum of fitted Gaussians. (**C** and **E**) Experimental (C) and MD (E) short-axis thermal vibration amplitudes σ_{short} of single W atoms. The bottom axis of the color bar in (C) labels the corresponding short-axis Gaussian RMS atom width before deconvolution. (**D**) Histograms of experimentally measured single W short-axis thermal vibration amplitude in regions AB (top), soliton (middle), and AA (bottom) regions, respectively. Bell curves label Gaussian fits. Vertical lines label centers of Gaussian fits for experiment (colored) and MD (black). Errors are standard errors of the mean. Scale bars are 1 nm for (A), (C), and (E). Scale bar is 10 pm for all panels in (B).

shift the quantitative values for extracted thermal vibrations (see materials and methods).

We obtained notable agreement between the experimental data in Fig. 3C and the MD simulations in Fig. 3E (see full supercell shown in fig. S17 and phason contribution in fig. S18). Both showed the highest amplitudes near the AA regions, followed by the AB and soliton regions. The large vibrational amplitudes at AA sites could be understood intuitively: As discussed previously, the AA sites are nodes of the elastic network of solitons, where vibrations are expected to be the largest. In addition, atoms in the AA sites sit at local maxima in the interlayer potential energy landscape, a property that should amplify atomic vibrations

We further confirmed that our experimental signals arose from phasons and higher-frequency moiré phonons by duplicating our analyses on simulated ptychography data with varying phonon symmetries. First, we conducted ptychography simulations on WSe_2

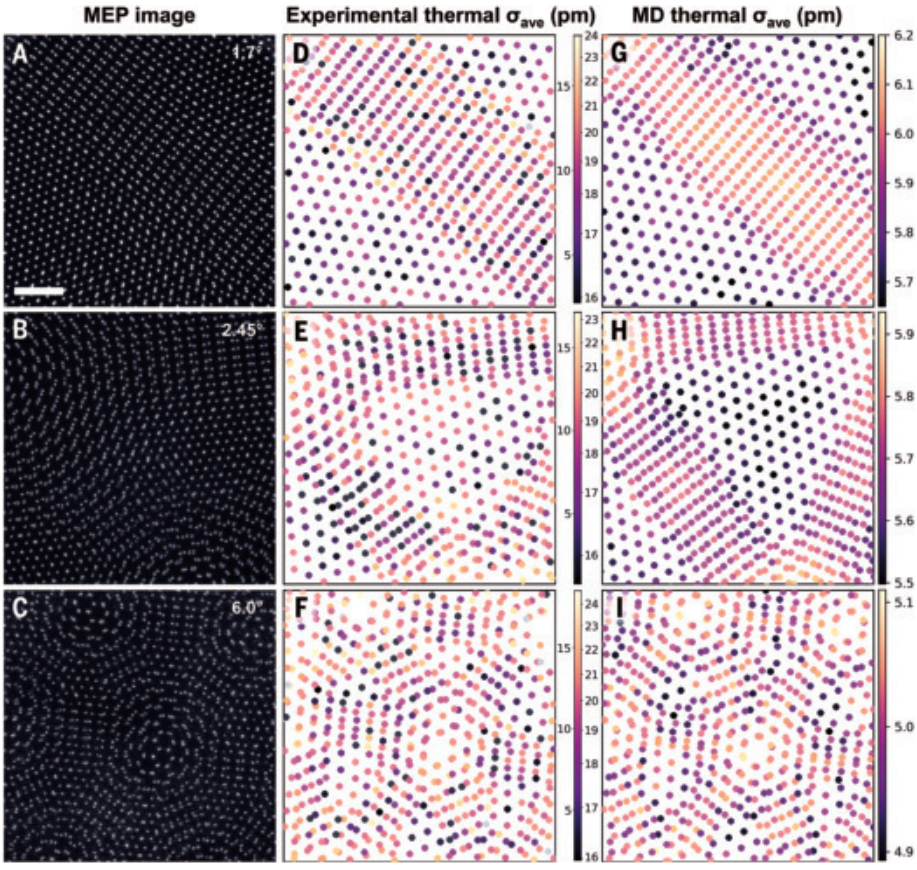


Fig. 4. Evolution of phasons with twist angle. (**A** to **C**) MEP images of 1.7°-, 2.45°-, and 6.0°-twisted bilayer WSe_2. (**D** to **F**) Experimental, deconvolved average thermal vibration amplitudes σ_{ave} (color bar left axis) and average Gaussian RMS atom width (color bar right axis) of single W atoms in (A) to (C), respectively. (**G** to **I**) MD simulated average RMSD of single W atoms of 1.7°-, 2.45°-, and 6.0°-twisted bilayer WSe_2, respectively. Scale bars are 1 nm for all panels.

bilayers with isotropic frozen phonons, as used in standard STEM simulations (*60*). In the isotropic frozen phonon model, atoms were given random displacements according to a Gaussian probability distribution with a specified RMSD. Because these simulations lacked the detailed spatial and frequency behavior of real phonon modes, they were useful controls to validate phonon detection.

We compared these "isotropic phonon" ptychography simulations against a more sophisticated model containing phonons calculated using MD, which included moiré phonons and phasons as well as conventional lattice phonons (see figs. S19 to S21 and materials and methods). As shown in fig. S19, ptychography simulations with MD phonons replicated our experiment, exhibiting increased short-axis Gaussian RMS width near AA sites, whereas ptychography simulations with isotropic phonons did not. This correspondence ruled out ptychography and Gaussian fitting artifacts as the source of our signals, indicating that our measured spatial distribution of atom sizes stemmed directly from local variations in phonon vibration amplitude associated with the local stacking orders within the moiré superlattice.

Next, we examined ptychography's ability to quantify the thermal vibration amplitude σ and its spatial variations. Figure 3D shows histograms comparing the vibrational amplitudes in different regions of the moiré superlattice for single W atoms, extracted from an experimental ptychography reconstruction. Using these distributions, we calculated mean W short-axis amplitudes of $\sigma_{AB} = 6.0 \pm 0.2$ pm in the AB regions, $\sigma_{soliton} = 5.8 \pm 0.3$ pm, and $\sigma_{AA} = 6.8 \pm 0.2$ pm. These reported errors are the standard error of the mean, which we used to assess our ability to determine shifts in σ values within a single image. In addition, we estimated systematic errors of ±1 to 2 pm in σ for each image arising from uncertainties in the deconvolution of nonthermal contributions.

For comparison, an analysis based on MD coordinates in a similar field of view yielded $\sigma_{AB} = 5.5$ pm, $\sigma_{soliton} = 5.5$ pm, and $\sigma_{AA} = 5.6$ pm. These values were also consistent with reported values of the room-temperature, in-plane RMSD of W atoms (5.5 pm) calculated from first principles (*58*). Note that we obtained these averages from the image shown in Fig. 3A, which contained regions that approached AA stacking but did not reach the AA centers. These data show that although the experimental amplitude measurements for single atoms are noisy and exhibit large spread, their mean values are sufficiently precise to show key trends, in this case, an increase in σ_{short} near AA-stacked regions. More generally, we found that σ decreases as the 2D projected distance increases between the nearest W atoms (for example, going from AA to AB stacking) in both simulations and experiment (see fig. S22 and supplementary text).

Evolution of moiré phasons with twist angle

We investigated the evolution of moiré phason modes with twist angle and superlattice structure. Figure 4, A to C, shows ptychography images of 1.7°-, 2.45°-, and 6.0°-twisted bilayers. These twisted bilayer samples spanned the different major regimes of moiré superlattices and their phonon behaviors, from strongly reconstructed (<2°) to a transitional regime (2° to 6°) and finally near-rigid structures (>6°), where minimal lattice reconstructions occurred (*5*, *21*). To highlight the evolution of the phason modes, we plotted the average thermal vibration amplitude, σ_{ave}, for each twist angle from experiment (Fig. 4, D to F) and MD (Fig. 4, G to I). We used σ_{ave} for each atom, rather than σ_{short}, as plotted in Fig. 3, to better capture the anisotropic phasons (see comparison of ellipticity magnitudes in fig. S23). To generate these plots, we fit 2D Gaussians to single W atoms in each image, isolated the thermal contributions (see materials and methods), and then calculated a geometric average of the long- and short-axis thermal vibration amplitudes.

We observed localized, enhanced vibrations at the solitons and their connection points in both experiment (Fig. 4, D and E) and MD (Fig. 4, G and H). These trends were visible directly in the maps of σ_{ave}. For example, for the 1.7° twist, Fig. 4, D and G, shows an increase in σ_{ave} in the soliton compared with the surrounding AB region. For the 2.45° bilayer (Fig. 4, E and H), the most visible increases in σ_{ave} were observed near the AA stacked regions. In the 6.0° bilayer, the regional changes in vibration amplitude became more difficult to distinguish visually. There were likely two sources of the lack of visual trends in the 6.0° data. First, our MD simulations predicted that the regional differences in the vibration amplitudes were smaller for larger twist angles (Fig. 4, G to I). Second, for larger twist angles, the moiré unit cells shrank and the distinct stacking regions became smaller and more tightly spaced, making regional differences harder to discern.

Overall, these data match well with predictions for twist-angle dependent changes in moiré phasons and with a broader understanding of how lattice resconstructions affected the interlayer interactions and moiré phonons. In the strongly reconstructed regime (<2°), the lattice reconstructions dominate, producing solitons and their corresponding phasons (*3*–*5*). In the transitional regime, where lattice reconstructions decreased (*21*, *35*), the ultrasoft modes are predicted to gradually lose their localization to the soliton boundaries (*5*). For large twist angles, as the lattice reconstruction effects vanish, the structure and phonon behavior eventually approach that of two rigid WSe_2 layers (*5*, *21*, *61*).

Summary and conclusions

We developed an atom-by-atom method to image thermal vibrations using electron ptychography and used it to reveal moiré phasons in twisted bilayer WSe_2. By doing so, we validated the signature predictions of ultrasoft moiré phasons: the emergence of spatially localized, anisotropic vibrations at the solitons and AA sites in low-angle moirés, features that were thought to be out of reach of experiments. Our work showed that phasons dominated the thermal vibrations of reconstructed moiré structures, adding to the growing body of evidence that these ultrasoft modes play a key role in the properties of low-angle 2D moiré materials. More broadly, our electron ptychography methods introduce a powerful way to measure spatially nonuniform vibrational modes, a capability that should prove impactful for understanding the vibrational properties of defects and interfaces.

REFERENCES AND NOTES

1. M.-L. Lin *et al.*, *ACS Nano* **12**, 8770–8780 (2018).
2. N. Suri, C. Wang, Y. Zhang, D. Xiao, *Nano Lett.* **21**, 10026–10031 (2021).
3. M. Koshino, Y.-W. Son, *Phys. Rev. B* **100**, 075416 (2019).
4. H. Ochoa, *Phys. Rev. B* **100**, 155426 (2019).
5. I. Maity, M. H. Naik, P. K. Maiti, H. R. Krishnamurthy, M. Jain, *Phys. Rev. Res.* **2**, 013335 (2020).
6. M. Mandal *et al.*, *Phys. Rev. B* **108**, 115439 (2023).
7. C. Jin *et al.*, *Nature* **567**, 76–80 (2019).
8. K. Tran *et al.*, *Nature* **567**, 71–75 (2019).
9. A. Rossi *et al.*, *ACS Nano* **18**, 18202–18210 (2024).
10. Y. Xu *et al.*, *Nature* **587**, 214–218 (2020).
11. X. Huang *et al.*, *Nat. Phys.* **17**, 715–719 (2021).
12. L. Balents, C. R. Dean, D. K. Efetov, A. F. Young, *Nat. Phys.* **16**, 725–733 (2020).
13. Y. Xia *et al.*, *Nature* **637**, 833–838 (2025).
14. Y. Cao *et al.*, *Nature* **556**, 43–50 (2018).
15. F. Wu, A. H. MacDonald, I. Martin, *Phys. Rev. Lett.* **121**, 257001 (2018).
16. B. Lian, Z. Wang, B. A. Bernevig, *Phys. Rev. Lett.* **122**, 257002 (2019).
17. X. Liu, R. Peng, Z. Sun, J. Liu, *Nano Lett.* **22**, 7791–7797 (2022).
18. Z. Li, J.-M. Lai, J. Zhang, *J. Semicond.* **44**, 011902 (2023).
19. R. M. Fernandes, J. W. F. Venderbos, *Sci. Adv.* **6**, eaba8834 (2020).
20. W. Ren, J. Chen, G. Zhang, *Appl. Phys. Lett.* **121**, 140501 (2022).
21. J. Quan *et al.*, *Nat. Mater.* **20**, 1100–1105 (2021).
22. H. Yoo *et al.*, *Nat. Mater.* **18**, 448–453 (2019).
23. A. Weston *et al.*, *Nat. Nanotechnol.* **15**, 592–597 (2020).
24. J. S. Alden *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11256–11260 (2013).
25. Q. Gao, E. Khalaf, *Phys. Rev. B* **106**, 075420 (2022).
26. P. M. Chaikin, T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge Univ. Press, 1995).
27. H. Ochoa, R. M. Fernandes, *Phys. Rev. Lett.* **128**, 065901 (2022).
28. Y. Zhang *et al.*, *Sci. Adv.* **10**, eadk1874 (2024).
29. D. Halbertal, S. Shabani, A. N. Passupathy, D. N. Basov, *ACS Nano* **16**, 1471–1476 (2022).
30. T. Benschop *et al.*, *Phys. Rev. Res.* **3**, 013153 (2021).
31. L. J. McGilly *et al.*, *Nat. Nanotechnol.* **15**, 580–584 (2020).
32. N. P. Kazmierczak *et al.*, *Nat. Mater.* **20**, 956–963 (2021).
33. I. Maity, A. A. Mostofi, J. Lischner, *Nano Lett.* **23**, 4870–4875 (2023).
34. R. Samajdar, Y. Teng, M. S. Scheurer, *Phys. Rev. B* **106**, L201403 (2022).
35. J. Z. Lu, Z. Zhu, M. Angeli, D. T. Larson, E. Kaxiras, *Phys. Rev. B* **106**, 144305 (2022).
36. J. Birkbeck *et al.*, *Nature* **641**, 345–351 (2025).
37. H. Ochoa, R. M. Fernandes, *Phys. Rev. B* **108**, 075168 (2023).
38. S. Carr, S. Fang, E. Kaxiras, *Nat. Rev. Mater.* **5**, 748–763 (2020).
39. F. S. Hage, D. M. Kepaptsoglou, Q. M. Ramasse, L. J. Allen, *Phys. Rev. Lett.* **122**, 016103 (2019).
40. B. Haas *et al.*, *Nano Lett.* **23**, 5975–5980 (2023).
41. M. Xu *et al.*, *Nat. Mater.* **22**, 612–618 (2023).
42. O. L. Krivanek *et al.*, *Ultramicroscopy* **203**, 60–67 (2019).
43. K. Tabata *et al.*, *Small Sci.* **4**, 2300254 (2024).
44. Z. Chen *et al.*, *Science* **372**, 826–831 (2021).
45. Y. Jiang *et al.*, *Nature* **559**, 343–349 (2018).
46. H. Yang *et al.*, *Ultramicroscopy* **180**, 173–179 (2017).
47. P. Wang, F. Zhang, S. Gao, M. Zhang, A. I. Kirkland, *Sci. Rep.* **7**, 2857 (2017).
48. Z. Chen *et al.*, *Nat. Commun.* **11**, 2994 (2020).
49. H. Sha *et al.*, *Nat. Commun.* **14**, 162 (2023).
50. H. Sha *et al.*, *Sci. Adv.* **9**, eadf1151 (2023).
51. Z. Dong *et al.*, *J. Am. Chem. Soc.* **145**, 6628–6632 (2023).
52. H. Zhang *et al.*, *Science* **380**, 633–638 (2023).
53. A. Gladyshev, B. Haas, T. M. Boland, P. Rez, C. T. Koch, arXiv:2309.12017 [physics.atom-ph] (2023).
54. A. M. Maiden, M. J. Humphry, J. M. Rodenburg, *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **29**, 1606–1614 (2012).
55. H. Sha, J. Cui, R. Yu, *Sci. Adv.* **8**, eabn2275 (2022).
56. R. Hovden, H. L. Xin, D. A. Muller, *Microsc. Microanal.* **17**, 75–80 (2011).
57. B. Forster, D. Van De Ville, J. Berent, D. Sage, M. Unser, *Microsc. Res. Tech.* **65**, 33–42 (2004).
58. I.-C. Tung *et al.*, *Nat. Photonics* **13**, 425–430 (2019).
59. W. Hummel, A. Raselli, H.-B. Burgi, *Acta Cryst.* **B46**, 683–692 (1990).
60. R. F. Loane, P. Xu, J. Silcox, *Acta Cryst.* **A47**, 267–278 (1991).
61. S. Huang *et al.*, *Nano Lett.* **16**, 1435–1444 (2016).
62. Y. Zhang *et al.*, Data: Atom-by-atom imaging of moiré phasons with electron ptychography. Zenodo (2025); https://doi.org/10.5281/zenodo.15283332.

ACKNOWLEDGMENTS

Electron microscopy work was carried out in the Materials Research Laboratory Central Research Facilities at the University of Illinois and the NSF-MRSEC shared facilities supported by award nos. DMR-1720633 and DMR-2309037. This work used NCSA Delta GPU at the National Center for Supercomputing Applications at the University of Illinois through allocation MAT240032 from the National Science Foundation–supported Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program. Molecular dynamics and harmonic lattice dynamics calculations were performed on PSC Bridges-2 at the Pittsburgh Supercomputing Center through allocation MAT220011 from the same ACCESS program, which is supported by National Science Foundation grant nos. 2138259, 2138286, 2138307, 2137603, and 2138296. Additional computational resources were provided by the Illinois Campus Cluster operated by the Illinois Campus Cluster Program. We thank L. Wang for developing the 2D Gaussian fitting procedure and R. Fernandes and I. Maity for discussions of the theory of phasons. **Funding:** This work was funded by National Science Foundation MRSEC awards DMR-1720633 and DMR-2309037 (P.Y.H., A.M.v.d.Z., E.E.), Air Force Office of Scientific Research PECASE award AF FA9550-20-1-0302 (P.Y.H.), and a Packard Foundation Fellowship (P.Y.H.). **Author contributions:** Conceptualization: P.Y.H.; Methodology: Y.Z., C.-H.L., B.A., T.R., S.H.B., M.A.H.; Investigation: Y.Z., B.A.; Software: Y.Z., B.A., J.H., T.R.; Formal analysis: Y.Z., B.A.; Visualization: Y.Z., B.A.; Resources: S.H.B., M.A.H.; Funding acquisition: P.Y.H., E.E., A.M.v.d.Z.; Project administration: P.Y.H.; Supervision: P.Y.H.; Writing – original draft: Y.Z., P.Y.H., B.A.; Writing – review & editing: Y.Z., B.A., S.H.B., C.-H.L., J.H., M.A.H., T.R., A.M.v.d.Z., E.E., P.Y.H. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available on Zenodo (*62*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adw7751
Materials and Methods; Supplementary Text; Figs. S1 to S23; Tables S1 to S5; References (*63*–*79*); Movies S1 and S2

Submitted 14 February 2025; accepted 6 June 2025

10.1126/science.adw7751

myIDP ScienceCareers
INDIVIDUAL DEVELOPMENT PLAN

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.
- 20 career paths with a prediction of which ones best fit your skills and interests.
- A tool for setting strategic goals with optional reminders to keep you on track.
- Articles and resources to guide you through the process.
- Options to save materials online and print them for further review and discussion.
- A certificate of completion for users that finish myIDP.

Start planning your future today!
myIDP.sciencecareers.org

In partnership with:

FASEB Federation of American Societies for Experimental Biology
UMASS University of Massachusetts Medical School
UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND

Running out of road

Na Zhao

It took just 30 minutes for my carefully constructed future to crumble. First came the email, at 9:48 a.m. on 3 July. "Due to university budget constraints, we must rescind your job offer," it read. The tenure-track position I had been weighing for weeks had vanished. Do not panic, I commanded myself. I had a backup option: to stay at my current institution, which despite a hiring freeze had stretched to also offer me a tenure-track position, though with only a small startup package. But a few minutes later, I was on a scheduled call with the program officer for my grant from the National Institutes of Health (NIH). He gently explained that I couldn't activate the award unless my institution committed much more funding of its own. I sat frozen until day care called: My 2-year-old had bumped his head and I needed to sign an accident report at pickup. That mundane errand became an anchor in a fast-moving river, reminding me there were still dinners to cook, bedtime stories to read, a family to nurture, a life to live.

Twenty months earlier, I had been full of naïve optimism as I started the 2023–24 job season. But 200 hours of applications yielded only two screening interviews and zero offers. Search committees were polite but blunt: "We're looking for external funding." So, I threw myself into grant writing. By midsummer 2024 I held a fundable score on an NIH K award, intended to help early-career investigators transition into independent roles. My calendar lit up: screening calls with several institutions, and two on-site visits. Success, I told myself, is a numbers game. File enough grants, book enough flights, and one door will stay open.

Then the political winds shifted. In February, a proposed cut to NIH indirect costs spooked university administrators nationwide. University C, which had planned to invite me for a second visit, froze hiring. University R went silent. Only University A pressed ahead. I delivered a Zoom seminar in March, a chalk talk in April, and received a job offer in May.

But ecstasy quickly gave way to anxiety when my husband, who works in biotech, couldn't find a job in the same city. June blurred into spreadsheets and insomnia. Should we live apart so he could keep his current job? Perhaps, but who would the kids live with? I'd be launching a new lab; it couldn't be me. Yet I'm their mother, and I want to kiss them good night every single day until they won't let me. How about one parent with each child, on a rotating schedule? Too chaotic. Could my husband work in my lab? Aside from nepotism rules, I'd also like to stay married. I told University A I needed more time to decide.

If I stayed at my current institution, how much preliminary data could I generate on the limited startup budget? I feared a hopeless loop: no data, no new grants, no money to generate data. University A granted me a 1-week extension. But before I had a chance to decide, Congress passed the new spending bill, and I received that email saying the offer was canceled.

I allowed myself 1 day to mourn, rage-scribbling in my notebook, eating ice cream for dinner, calling friends who had survived academic limbo. Their stories shared a common theme: Careers look linear only in hindsight. Offers vanish, grants get triaged, political tides turn. We keep going. But exactly how many detours can a junior scientist afford before running out of road?

That sleepless night, a flood in Texas claimed more than 130 lives, many of them children. Faced with real tragedy, my hardship shrank. A tenure-track position is not life or death, and most setbacks are temporary (at least I hope!). Write one more grant even when the last review stings. Send one more email asking for an opening. Accept that "yes" can morph into "no" overnight, and sometimes back again.

I still don't know whether my institution can piece together a patchwork startup big enough to activate the K grant; the department chair promises to "see what we can do." I don't know when I will unlock a lab door bearing my name. I can't predict the next hiring cycle or the next election. But I can keep doing the science I love while I still have a bench. Between mouse surgeries and bedtime stories, I'll write the next proposal. That, I've decided, is what resilience really means. □

Na Zhao is a research assistant professor at Baylor College of Medicine.

ILLUSTRATION: ROBERT NEUBECKER

Who's the top employer for 2024?

Science 2024 TOP EMPLOYER

Science Careers' annual survey reveals the top companies in biotech & pharma voted on by *Science* readers.

Explore these highly-rated employers in our new interactive experience:
sciencecareers.org/topemployers