Can China's global infrastructure push really go green? p. 676

Tackling biased policing data wins NOMIS & *Science* Award p. 701

Freshwater sponge hints at ancestral cell types p. 717

# Science

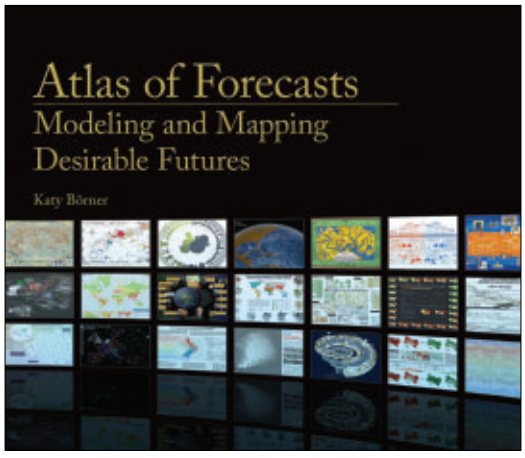
$15
5 NOVEMBER 2021
science.org

AAAS



# MARTIAN DELTA

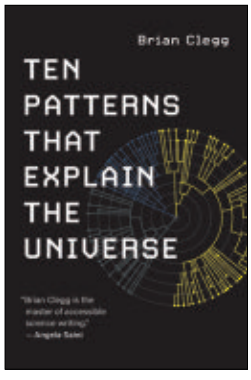
Perseverance rover investigates ancient hydrology p. 711




# New books from the MIT Press

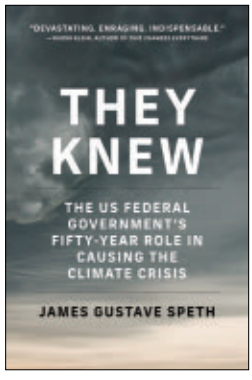
## Atlas of Forecasts

"[Börner's] sumptuous, detailed book tackles issues of error and bias head-on." ***—New Scientist***



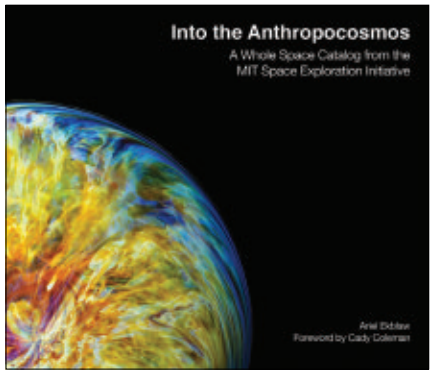

## Ten Patterns That Explain the Universe

"Using beautiful illustrations and illuminating diagrams, the book is a feast for the eyes." ***—Science***




## They Knew

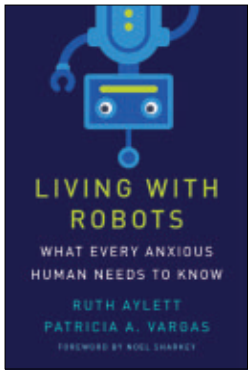
"A chilling description of the gulf between the safer course of action recommended by scientists and advisers, and the reality of federal policy." ***—Nature***




## Into the Anthropocosmos

"An insightful and comprehensive yet easy-to-understand summary of the complex science required to establish an interplanetary species."

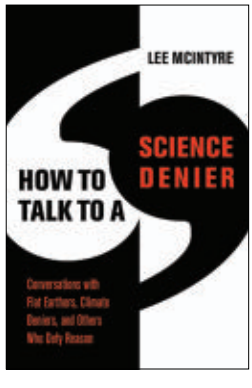
**—Scott Kelly, retired NASA astronaut; author of *Endurance: A Year in Space, a Lifetime of Discovery***




## Living with Robots

"A book by genuine experts and based on the latest science that shows us how robots will really feature in our lives."

**—Takanori Shibata, Chief Senior Research Science, AIST**




## How to Talk to a Science Denier

"Superb." ***—E&T Magazine***

**in bookstores & mitpress.mit.edu**

# CONTENTS



**674** A Chinese crew at work on the China-Laos Railway, part of the Belt and Road Initiative, in January 2020.

5 NOVEMBER 2021 • VOLUME 374 • ISSUE 6568

## NEWS

### IN BRIEF

**664** News at a glance

### IN DEPTH

**667 Methane removal seen as tool to slow warming**
As nations look to cut emissions of methane, some researchers want to pull it out of the air *By C. O'Grady*

**668 Pig kidney transplant obscures value of engineered animals**
Skin and nerves from pigs lacking immune-provoking sugar could help patients today *By E. Dolgin*

**669 Reactor mishap derails U.S. studies with neutron beams**
NIST's research reactor suffered little damage, but will be offline for months *By A. Cho*

**671 Max Planck director loses post after probe of misconduct**
Nicole Boivin remains on research staff, but major archaeological institute is left with interim director *By A. Curry*

**672 RNA drugs custommade for rare diseases face wider test**
With tempered expectations, more scientists try to accelerate "n-of-1" approach *By J. Kaiser*

**663 China Initiative spawns distrust—and activism**
Two surveys reveal fallout from U.S. investigations as Chinese faculty form campus groups *By J. Mervis*

### FEATURES

**674 A greener path**
China is making its Belt and Road Initiative more environmentally friendly. The massive infrastructure program could still cause ecological devastation *By D. Normile*

## INSIGHTS

### PERSPECTIVES

**680 Children and COVID-19 in schools**
The benefits of in-person schooling with mitigations in place outweigh the risks of COVID-19 for children *By S. N. Ladhani and the sKIDs Investigation Team*

**682 A helminth-induced antimicrobial protein**
Small proline-rich protein 2A regulates the intestinal microbiota *By N. Harris and L. Wickramasinghe*
RESEARCH ARTICLE p. 710

**684 Stabilizing synapses**
Adenosine fine-tunes the fate of nascent synapses in brain development *By D. Blum and L. V. Lopes*
RESEARCH ARTICLE p. 709

**685 How ammonia feeds and pollutes the world**
It is cheaper to cut ammonia emission now than to deal with its consequences later *By J. W. Erisman*
REPORT p. 758

**686 Midair transformations of aerosols**
Spontaneous reactions play an important rolein shaping the chemistry of aerosols *By M. F. Ruiz-Lopez*
REPORT p. 747

**688 A radical shift in air pollution**
Fifty years ago, Levy identified hydroxyl radicals as the driver of chemistry in the troposphere *By C. L. Heald and J. H. Kroll*

### POLICY FORUM

**690 National climate institutions complement targets and policies**
Institutions can affect coordination, consensus, and strategy *By N. K. Dubash* et al.

**693 Can updated climate pledges limit warming well below 2°C?**
Increased ambition and implementation are essential *By Y. Ou* et al.

### BOOKS *ET AL.*

**696 Embodying consciousness**
Bodies and feelings are foundational to conscious experience, argues a neuroscientist *By A. Gomez-Marin*

**697 Immunology meets the masses**
A YouTuber's vividly illustrated immune system explainer delights and informs *By D. M. Davis*

PHOTO: KAIKEO SAIYASANE/XINHUA NEWS AGENCY/GETTY IMAGES

# EMPOWER WITH EVIDENCE

AAAS | ANNUAL MEETING

Join us for the hybrid 2022 AAAS Annual Meeting, held in Philadelphia, and online, February 17-20.

**AAAS requires that all participants in its in-person meetings be fully vaccinated.**

## Be there for lectures from these renowned speakers during the hybrid #AAASmtg!



**Deepakraj M. Divan**
Director, Center for Distributed Energy, Georgia Tech



**Katalin Karikó**
Senior Vice President at BioNTech RNA Pharmaceuticals
Adjunct Professor, University of Pennsylvania School of Medicine



**Susan Amara**
President, AAAS
Chair, 2022 Annual Meeting



**John O. Dabiri**
Centennial Professor of Aeronautics and Mechanical Engineering, California Institute of Technology



**Andrea M. Ghez**
Professor of Physics and Astronomy, University of California, Los Angeles



**Katie Hinde**
Associate Professor, School of Human Evolution and Social Change, Arizona State University



**Nadya Mason**
Rosalyn S. Yalow Professor of Physics, University of Illinois at Urbana-Champaign



**Kate Starbird**
Associate Professor, University of Washington, Human Centered Design & Engineering



**Robin Nelson**
Associate Professor, School of Human Evolution and Social Change, Arizona State University



**Kyle Whyte**
George Willis Pack Professor, School for Environment and Sustainability, University of Michigan



**Audra J. Wolfe**
Writer and Editor

**aaas.org/meetings | #AAASmtg**

## LETTERS

**699 Protect China's karst cave habitats**
*By Y.-F. Duan* et al.

**699 Biodiversity loss due to more than climate change**
*By A. Dobson* et al.

**700 Science-informed salmon conservation**
*By M. S. Bank* et al.

## PRIZE ESSAY

**701 Revealing racial bias**
Causal inference can make sense of imperfect policing data
*By D. Knox*

# RESEARCH

## IN BRIEF

**704** From *Science* and other journals

## REVIEW

**707 Radiocarbon**
Radiocarbon: A key tracer for studying Earth's dynamo, climate system, carbon cycle, and Sun *T. J. Heaton* et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABD7096
PODCAST

**708 Membranes**
Subatomic species transport through atomically thin membranes: Present and future applications
*P. R. Kidambi* et al.
REVIEW SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABD7687

## RESEARCH ARTICLES

**709 Neurodevelopment**
Convergence of adenosine and GABA signaling for synapse stabilization during development *F. Gomez-Castro* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABK2055
PERSPECTIVE p. 684

**710 Immunology**
Small proline-rich protein 2A is a gut bactericidal protein deployed during helminth infection *Z. Hu* et al.
RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABE6723
PERSPECTIVE p. 682




**711 Martian geology**
Perseverance rover reveals an ancient delta-lake system and flood deposits at Jezero crater, Mars *N. Mangold* et al.

**717 Developmental biology**
Profiling cellular diversity in sponges informs animal cell type and nervous system evolution *J. M. Musser* et al.

**723 Biosynthesis**
Modular polyketide synthase contains two reaction chambers that operate asynchronously *S. R. Bagde* et al.
REPORT p. 729

## REPORTS

**729 Biosynthesis**
Mapping the catalytic conformations of an assembly-line polyketide synthase module *D. P. Cogan* et al.
RESEARCH ARTICLE p. 723

**734 Organic chemistry**
Carbonyl 1,2-transposition through triflate-mediated α-amination
*Z. Wu* et al.

**741 Organic chemistry**
Scalable Birch reduction with lithium and ethylenediamine in tetrahydrofuran
*J. Burrows* et al.

**747 Aerosol chemistry**
A surface-promoted redox reaction occurs spontaneously on solvating inorganic aerosol surfaces *X. Kong* et al.
PERSPECTIVE p. 686

**752 Organic chemistry**
Enantioselective catalytic 1,2-boronate rearrangements
*H. A. Sharma* et al.

**758 Air pollution**
Abating ammonia is more cost-effective than nitrogen oxides for mitigating $PM_{2.5}$ air pollution *B. Gu* et al.
PERSPECTIVE p. 685

**762 Neuroscience**
The glial framework reveals white matter fiber architecture in human and primate brains
*R. Schurr and A. A. Mezer*

**768 Hematopoiesis**
Resistance to inflammation underlies enhanced fitness in clonal hematopoiesis
*S. Avagyan* et al.

## DEPARTMENTS

**659 Editorial**
A new lane for science
*By Enola K. Proctor and Elvin Geng*

**782 Working Life**
Keeping science honest *By Julia Behnfeldt*

## ON THE COVER

Enhanced-color optical image of Jezero crater, Mars, taken by the Perseverance rover's Mastcam-Z instrument on 18 April 2021. The foreground flat-topped hill, informally named Kodiak butte, is 2.2 kilometers distant and 250 meters wide. It exposes ancient layered rocks that indicate gradual deposition of sediments in a river delta, followed by floods. See page 711. *Credit: NASA/JPL-Caltech/ASU/MSSS*




*Science* Staff ..... 658
*Science* Careers ..... 773

**SCIENCE (ISSN 0036-8075) is published weekly on Friday, 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 © 2021 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): $2148; Foreign postage extra: Air assist delivery: $98. 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.

PHOTO: HANS ENGBERS/SHUTTERSTOCK

# Science AAAS

**Editor-in-Chief** Holden Thorp, hthorp@aaas.org

**Executive Editor** Monica M. Bradford

**Editors, Research** Valda Vinson, Jake S. Yeston **Editor, Insights** Lisa D. Chong

**DEPUTY EDITORS** Julia Fahrenkamp-Uppenbrink (UK), Stella M. Hurtley (UK), Phillip D. Szuromi, Sacha Vignieri **SR. EDITORIAL FELLOW** Andrew M. Sugden (UK) **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Grocholski, Pamela J. Hines, Di Jiang, Priscilla N. Kelly, Marc S. Lavine (Canada), Yevgeniya Nusinovich, Ian S. Osborne (UK), Beverly A. Purnell, L. Bryan Ray, H. Jesse Smith, Keith T. Smith (UK), Jelena Stajic, Peter Stern (UK), Valerie B. Thompson, Brad Wible, Yuen Yiu, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Bianca Lopez, Seth Thomas Scanlon (UK), Yury V. Suleymanov **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmec **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Chris Filiatreau, Julia Haber-Katris, Nida Masiulis, Abigail Shashikanth, Suzanne M. White **SR. EDITORIAL COORDINATORS** Carolyn Kyle, Beverly Shields **EDITORIAL COORDINATORS** Aneera Dobbins, Joi S. Granger, Jeffrey Hearn, Lisa Johnson, Maryrose Madrid, Ope Martins, Shannon McMahon, Jerry Richardson, Hilary Stewart (UK), Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Alexander Kief, Ronmel Navas, Isabel Schnaidt, Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldock (UK)

**News Editor** Tim Appenzeller

**NEWS MANAGING EDITOR** John Travis **INTERNATIONAL EDITOR** Martin Enserink **DEPUTY NEWS EDITORS** Elizabeth Culotta, Lila Guterman, David Grimm, Eric Hand (Europe), David Malakoff **SR. CORRESPONDENTS** Daniel Clery (UK), Jon Cohen, Jeffrey Mervis, Elizabeth Pennisi **ASSOCIATE EDITORS** Jeffrey Brainard, Kelly Servick, Catherine Matacic **NEWS REPORTERS** Adrian Cho, Jennifer Couzin-Frankel, Jocelyn Kaiser, Rodrigo Pérez Ortega (Mexico City), Robert F. Service, Erik Stokstad, Paul Voosen, Meredith Wadman **INTERNS** Menna Ibrahim, Sofia Moutinho, Anil Oza, Alex Viveros **CONTRIBUTING CORRESPONDENTS** Warren Cornwall, Andrew Curry (Berlin), Ann Gibbons, Sam Kean, Eli Kintisch, Kai Kupferschmidt (Berlin), Andrew Lawler, Mitch Leslie, Eliot Marshall, Virginia Morell, Dennis Normile (Tokyo), Elisabeth Pain (Careers), Charles Piller, Michael Price, Tania Rabesandratana (Barcelona), Joshua Sokol, Emily Underwood, Gretchen Vogel (Berlin), Lizzie Wade (Mexico City) **CAREERS** Rachel Bernstein (Editor), Katie Langin (Associate Editor) **COPY EDITORS** Julia Cole (Senior Copy Editor), Morgan Everett, Cyra Master (Copy Chief) **ADMINISTRATIVE SUPPORT** Meagan Weiland

**Creative Director** Beth Rakouskas

**DESIGN MANAGING EDITOR** Marcy Atarod **INTERIM GRAPHICS MANAGING EDITOR** Chris Bickel **PHOTOGRAPHY MANAGING EDITOR** William Douthitt **WEB STRATEGY MANAGER** Kara Estelle-Powers **MULTIMEDIA MANAGING PRODUCER** Joel Goldberg **DESIGN EDITOR** Chrystal Smith **DESIGNER** Christina Aycock **GRAPHICS EDITOR** Nirja Desai **INTERACTIVE GRAPHICS EDITOR** Kelly Franklin **SENIOR GRAPHICS SPECIALISTS** Holly Bishop, Nathalie Cary **SENIOR SCIENTIFIC ILLUSTRATOR** Valerie Altounian **SCIENTIFIC ILLUSTRATOR** Ashley Mastin **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan **SOCIAL MEDIA STRATEGIST** Jessica Hubbard **SOCIAL MEDIA PRODUCER** Sabrina Jenkins **WEB DESIGNER** Jennie Pajerowski **SENIOR PODCAST PRODUCER** Sarah Crespi **VIDEO PRODUCER** Meagan Cantwell

**Chief Executive Officer and Executive Publisher** Sudip Parikh

**Publisher, Science Family of Journals** Bill Moran

**DIRECTOR, BUSINESS SYSTEMS AND FINANCIAL ANALYSIS** Randy Yi **DIRECTOR, BUSINESS OPERATIONS & ANALYSIS** Eric Knott **DIRECTOR OF ANALYTICS** Enrique Gonzales **MANAGER, BUSINESS OPERATIONS** Jessica Tierney **SENIOR BUSINESS ANALYST** Cory Lipman **FINANCIAL ANALYST** Isacco Fusi **ADVERTISING SYSTEM ADMINISTRATOR** Tina Burks **DIGITAL/PRINT STRATEGY MANAGER** Jason Hillman **SENIOR MANAGER, PUBLISHING AND CONTENT SYSTEMS** Marcus Spiegler **ASSISTANT MANAGER DIGITAL/PRINT** Rebecca Doshi **SENIOR CONTENT AND PUBLISHING SYSTEMS SPECIALIST** Jacob Hedrick **SENIOR CONTENT SPECIALISTS** Steve Forrester, Antoinette Hodal, Lori Murphy **PRODUCTION SPECIALIST** Kristin Wowk **DIGITAL PRODUCTION MANAGER** Lisa Stanford **CONTENT SPECIALIST** Kimberley Oster **ADVERTISING PRODUCTION OPERATIONS MANAGER** Deborah Tompkins **DESIGNER, CUSTOM PUBLISHING** Jeremy Huntsinger **SR. TRAFFIC ASSOCIATE** Christine Hall **SPECIAL PROJECTS ASSOCIATE** Sarah Dhere

**ASSOCIATE DIRECTOR, BUSINESS DEVELOPMENT** Justin Sawyers **GLOBAL MARKETING MANAGER** Allison Pritchard **DIGITAL MARKETING MANAGER** Aimee Aponte **JOURNALS MARKETING MANAGER** Shawana Arnold **MARKETING ASSOCIATES** Ashley Hylton, Mike Romano, Tori Velasquez, Jenna Voris, Justin Wood **SENIOR DESIGNER** Kim Huynh

**DIRECTOR AND SENIOR EDITOR, CUSTOM PUBLISHING** Sean Sanders **ASSISTANT EDITOR, CUSTOM PUBLISHING** Jackie Oberst

**DIRECTOR, PRODUCT & PUBLISHING DEVELOPMENT** Chris Reid **DIRECTOR, BUSINESS STRATEGY AND PORTFOLIO MANAGEMENT** Sarah Whalen **ASSOCIATE DIRECTOR, PRODUCT MANAGMENT** Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **PUBLISHING TECHNOLOGY MANAGER** Michael Di Natale **SR. PRODUCT ASSOCIATE** Robert Koepke **PRODUCT ASSOCIATE** Anne Mason **SPJ ASSOCIATE MANAGER** Samantha Bruno Fuller **SPJ ASSOCIATE** Casey Buchta

**DIRECTOR, INSTITUTIONAL LICENSING** Iquo Edim **MARKETING MANAGER** Kess Knight **BUSINESS DEVELOPMENT MANAGER** Rasmus Andersen **SENIOR INSTITUTIONAL LICENSING MANAGER** Ryan Rexroth **INSTITUTIONAL LICENSING MANAGER** Marco Castellan, Claudia Paulsen-Young **CUSTOMER SUCCESS MANAGER** Judy Lillibridge **SENIOR OPERATIONS ANALYST** Lana Guz **FULFILLMENT COORDINATOR** Melody Stringer

**DIRECTOR, GLOBAL SALES** Tracy Holmes **US EAST COAST AND MID WEST SALES** Stephanie O'Connor **US MID WEST, MID ATLANTIC AND SOUTH EAST SALES** Chris Hoag **US WEST COAST SALES** Lynne Stickrod **ASSOCIATE DIRECTOR, ROW** Roger Goncalves **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** Yoshimi Toda (Tokyo), Miyuki Tani (Osaka)

**DIRECTOR, COPYRIGHT, LICENSING AND SPECIAL PROJECTS** Emilie David **RIGHTS AND LICENSING COORDINATOR** Jessica Adams **RIGHTS AND PERMISSIONS ASSOCIATE** Elizabeth Sandler **LICENSING ASSOCIATE** Virginia Warren

---

**MAIN HEADQUARTERS**
Science/AAAS
1200 New York Ave. NW
Washington, DC 20005

**SCIENCE INTERNATIONAL**
Clarendon House
Clarendon Road
Cambridge, CB2 8FH, UK

**SCIENCE CHINA**
Room 1004, Culture Square
No. 59 Zhongguancun St.
Haidian District, Beijing, 100872

**SCIENCE JAPAN**
ASCA Corporation
Sibaura TY Bldg. 4F, 1-14-5
Shibaura Minato-ku
Tokyo, 108-0073 Japan

**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

**MEDIA CONTACTS**
scipak@aaas.org

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

**INSTITUTIONAL SALES AND SITE LICENSES**
science.org/librarian

**PRODUCT ADVERTISING & CUSTOM PUBLISHING**
advertising.science.org/
products-services
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



**AAAS BOARD OF DIRECTORS**
**CHAIR** Claire M. Fraser
**PRESIDENT** Susan G. Amara
**PRESIDENT-ELECT** Gilda A. Barabino
**TREASURER** Carolyn N. Ainslie
**CHIEF EXECUTIVE OFFICER** Sudip Parikh
**BOARD** Cynthia M. Beall, Rosina M. Bierbaum, Ann Bostrom, Janine Austin Clayton, Laura H. Greene, Kaye Husbands Fealing, Maria M. Klawe, Robert B. Millard, William D. Provine



*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 and comment, 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.



**BOARD OF REVIEWING EDITORS** (Statistics board members indicated with **(S)**)

**Erin Adams,** *U. of Chicago*
**Takuzo Aida,** *U. of Tokyo*
**Leslie Aiello,** *Wenner-Gren Fdn.*
**Deji Akinwande,** *UT Austin*
**Judith Allen,** *U. of Manchester*
**Marcella Alsan,** *Harvard U.*
**Sebastian Amigorena,** *Inst. Curie*
**James Analytis,** *UC Berkeley*
**Trevor Archer,** *NIEHS, NIH*
**Paola Arlotta,** *Harvard U.*
**David Awschalom,** *U. of Chicago*
**Delia Baldassarri,** *NYU*
**Nenad Ban,** *ETH Zürich*
**Nandita Basu,** *U. of Waterloo*
**Franz Bauer,** *Pontificia U. Católica de Chile*
**Ray H. Baughman,** *UT Dallas*
**Carlo Beenakker,** *Leiden U.*
**Yasmine Belkaid,** *NIAID, NIH*
**Philip Benfey,** *Duke U.*
**Kiros T. Berhane,** *Columbia U.*
**Joseph J. Berry,** *NREL*
**Alessandra Biffi,** *Harvard Med.*
**Chris Bowler,** *École Normale Supérieure*
**Ian Boyd,** *U. of St. Andrews*
**Emily Brodsky,** *UC Santa Cruz*
**Ron Brookmeyer,** *UCLA* **(S)**
**Christian Büchel,** *UKE Hamburg*
**Dennis Burton,** *Scripps Res.*
**Carter Tribley Butts,** *UC Irvine*
**György Buzsáki,** *NYU School of Med.*
**Mariana Byndloss,** *Vanderbilt U. Med. Ctr.*
**Annmarie Carlton,** *UC Irvine*
**Simon Cauchemez,** *Inst. Pasteur*
**Ling-Ling Chen,** *SIBCB, CAS*
**M. Keith Chen,** *UCLA*
**Zhijian Chen,** *UT Southwestern Med. Ctr.*
**Ib Chorkendorff,** *Denmark TU*
**Amander Clark,** *UCLA*
**James J. Collins,** *MIT*
**Robert Cook-Deegan,** *Arizona State U.*
**Virginia Cornish** *Columbia U.*
**Carolyn Coyne,** *Duke U.*
**Roberta Croce,** *VU Amsterdam*
**Ismaila Dabo,** *Penn State U.*
**Jeff L. Dangl,** *UNC*
**Chiara Daraio,** *Caltech*
**Nicolas Dauphas,** *U. of Chicago*
**Christian Davenport,** *U. of Michigan*
**Frans de Waal,** *Emory U.*
**Claude Desplan,** *NYU*
**Sandra Díaz,** *U. Nacional de Córdoba*
**Ulrike Diebold,** *TU Wien*
**Stefanie Dimmeler,** *Goethe-U. Frankfurt*
**Hong Ding,** *Inst. of Physics, CAS*
**Dennis Discher,** *UPenn*
**Jennifer A. Doudna,** *UC Berkeley*
**Ruth Drdla-Schutting,** *Med. U. Vienna*
**Raissa M. D'Souza,** *UC Davis*
**Bruce Dunn,** *UCLA*
**William Dunphy,** *Caltech*
**Scott Edwards,** *Harvard U.*
**Todd Ehlers,** *U. of Tübingen*
**Andrea Encalada,** *U. San Francisco de Quito*
**Nader Engheta,** *UPenn*
**Karen Ersche,** *U. of Cambridge*
**Beate Escher,** *UFZ & U. of Tübingen*
**Barry Everitt,** *U. of Cambridge*
**Vanessa Ezenwa,** *U. of Georgia*
**Michael Feuer,** *GWU*
**Toren Finkel,** *U. of Pitt. Med. Ctr.*
**Gwenn Flowers,** *Simon Fraser U.*
**Peter Fratzl,** *Max Planck Inst. Potsdam*
**Elaine Fuchs,** *Rockefeller U.*
**Jay Gallagher,** *U. of Wisconsin*
**Daniel Geschwind,** *UCLA*
**Ramon Gonzalez,** *U. of South Florida*
**Sandra González-Bailón,** *UPenn*
**Nicolas Gruber,** *ETH Zürich*
**Hua Guo,** *U. of New Mexico*
**Taekjip Ha,** *Johns Hopkins U.*
**Sharon Hammes-Schiffer,** *Yale U.*
**Wolf-Dietrich Hardt,** *ETH Zürich*
**Louise Harra,** *U. Coll. London*
**Jian He,** *Clemson U.*
**Carl-Philipp Heisenberg,** *IST Austria*
**Ykä Helariutta,** *U. of Cambridge*
**Janet G. Hering,** *Eawag*
**Christoph Hess,** *U. of Basel & U. of Cambridge*
**Heather Hickman,** *NIAID, NIH*
**Hans Hilgenkamp,** *U. of Twente*
**Kai-Uwe Hinrichs,** *U. of Bremen*
**Deirdre Hollingsworth,** *U. of Oxford*
**Randall Hulet,** *Rice U.*
**Auke Ijspeert,** *EPFL*
**Akiko Iwasaki,** *Yale U.*
**Stephen Jackson,** *USGS & U. of Arizona*
**Erich Jarvis,** *Rockefeller U.*
**Peter Jonas,** *IST Austria*
**Matt Kaeberlein,** *U. of Wash.*
**William Kaelin Jr.,** *Dana-Farber Cancer Inst.*
**Daniel Kammen,** *UC Berkeley*
**Kisuk Kang,** *Seoul Nat. U.*
**Sabine Kastner,** *Princeton U.*
**V. Narry Kim,** *Seoul Nat. U.*
**Robert Kingston,** *Harvard Med.*
**Nancy Knowlton,** *Smithsonian Institution*
**Etienne Koechlin,** *École Normale Supérieure*
**Alex L. Kolodkin,** *Johns Hopkins U.*
**Julija Krupic,** *U. of Cambridge*
**Paul Kubes,** *U. of Calgary*
**Gabriel Lander,** *Scripps Res.* **(S)**
**Mitchell A. Lazar,** *UPenn*
**Wendell Lim,** *UCSF*
**Luis Liz-Marzán,** *CIC biomaGUNE*
**Omar Lizardo,** *UCLA*
**Jonathan Losos,** *Wash. U. in St. Louis*
**Ke Lu,** *Inst. of Metal Res., CAS*
**Christian Lüscher,** *U. of Geneva*
**Jean Lynch-Stieglitz,** *Georgia Inst. of Tech.*
**David Lyons,** *U. of Edinburgh*
**Fabienne Mackay,** *QIMR Berghofer*
**Anne Magurran,** *U. of St. Andrews*
**Asifa Majid,** *U. of York*
**Oscar Marín,** *King's Coll. London*
**Charles Marshall,** *UC Berkeley*
**Christopher Marx,** *U. of Idaho*
**David Masopust,** *U. of Minnesota*
**Geraldine Masson,** *CNRS*
**Jason Matheny,** *Georgetown U.*
**Heidi McBride,** *McGill U.*
**C. Robertson McClung,** *Dartmouth*
**Rodrigo Medellín,** *U. Nacional Autónoma de México*
**C. Jessica Metcalf,** *Princeton U.*
**Baoxia Mi,** *UC Berkeley*
**Tom Misteli,** *NCI, NIH*
**Alison Motsinger-Reif,** *NIEHS, NIH* **(S)**
**Suresh Naidu,** *Columbia U.*
**Danielle Navarro,** *U. of New South Wales*
**Daniel Nettle,** *Newcastle U.*
**Daniel Neumark,** *UC Berkeley*
**Beatriz Noheda,** *U. of Groningen*
**Helga Nowotny,** *Vienna Sci. & Tech. Fund*
**Rachel O'Reilly,** *U. of Birmingham*
**Pilar Ossorio,** *U. of Wisconsin*
**Andrew Oswald,** *U. of Warwick*
**Isabella Pagano,** *Istituto Nazionale di Astrofisica*
**Elizabeth Levy Paluck,** *Princeton U.*
**Jane Parker,** *Max Planck Inst. Cologne*
**Giovanni Parmigiani,** *Dana-Farber Cancer Inst.* **(S)**
**Daniel Pauly,** *U. of British Columbia*
**Ana Pêgo,** *U. do Porto*
**Samuel Pfaff,** *Salk Inst.*
**Julie Pfeiffer,** *UT Southwestern Med. Ctr.*
**Philip Phillips,** *UIUC*
**Matthieu Piel,** *Inst. Curie*
**Kathrin Plath,** *UCLA*
**Martin Plenio,** *Ulm U.*
**Katherine Pollard,** *UCSF*
**Elvira Poloczanska,** *Alfred-Wegener-Inst.*
**Julia Pongratz,** *Ludwig Maximilians U.*
**Philippe Poulin,** *CNRS*
**Jonathan Pritchard,** *Stanford U.*
**Lei Stanley Qi,** *Stanford U.*
**Trevor Robbins,** *U. of Cambridge*
**Joeri Rogelj,** *Imperial Coll. London*
**Amy Rosenzweig,** *Northwestern U.*
**Mike Ryan,** *UT Austin*
**Miquel Salmeron,** *Lawrence Berkeley Nat. Lab*
**Nitin Samarth,** *Penn State U.*
**Erica Ollmann Saphire,** *La Jolla Inst.*
**Joachim Saur,** *U. zu Köln*
**Alexander Schier,** *Harvard U.*
**Wolfram Schlenker,** *Columbia U.*
**Susannah Scott,** *UC Santa Barbara*
**Anuj Shah,** *U. of Chicago*
**Vladimir Shalaev,** *Purdue U.*
**Jie Shan,** *Cornell U.*
**Beth Shapiro,** *UC Santa Cruz*
**Jay Shendure,** *U. of Wash.*
**Steve Sherwood,** *U. of New South Wales*
**Brian Shoichet,** *UCSF*
**Robert Siliciano,** *JHU School of Med.*
**Lucia Sivilotti,** *U. Coll. London*
**Alison Smith,** *John Innes Ctr.*
**Richard Smith,** *UNC* **(S)**
**Mark Smyth,** *QIMR Berghofer*
**John Speakman,** *U. of Aberdeen*
**Tara Spires-Jones,** *U. of Edinburgh*
**Allan C. Spradling,** *Carnegie Institution for Sci.*
**V. S. Subrahmanian,** *Northwestern U.*
**Ira Tabas,** *Columbia U.*
**Eriko Takano,** *U. of Manchester*
**Patrick Tan,** *Duke-NUS Med. School*
**Sarah Teichmann,** *Wellcome Sanger Inst.*
**Rocio Titiunik,** *Princeton U.*
**Shubha Tole,** *Tata Inst. of Fundamental Res.*
**Maria-Elena Torres Padilla,** *Helmholtz Zentrum München*
**Kimani Toussaint,** *Brown U.*
**Barbara Treutlein,** *ETH Zürich*
**Wim van der Putten,** *Netherlands Inst. of Ecology*
**Ivo Vankelecom,** *KU Leuven*
**Henrique Veiga-Fernandes,** *Champalimaud Fdn.*
**Reinhilde Veugelers,** *KU Leuven*
**Bert Vogelstein,** *Johns Hopkins U.*
**Julia Von Blume,** *Yale School of Med.*
**David Wallach,** *Weizmann Inst.*
**Jane-Ling Wang,** *UC Davis* **(S)**
**Jessica Ware,** *Amer. Mus. of Natural Hist.*
**David Waxman,** *Fudan U.*
**Chris Wikle,** *U. of Missouri* **(S)**
**Terrie Williams,** *UC Santa Cruz*
**Ian A. Wilson,** *Scripps Res.* **(S)**
**Hao Wu,** *Harvard U.*
**Wei Xie,** *Tsinghua U.*
**Yu Xie,** *Princeton U.*
**Jan Zaanen,** *Leiden U.*
**Kenneth Zaret,** *UPenn School of Med.*
**Bing Zhu,** *Inst. of Biophysics, CAS*
**Xiaowei Zhuang,** *Harvard U.*
**Maria Zuber,** *MIT*

EDITORIAL

# A new lane for science

A recent *Science* editorial on the social and political headwinds that have blunted, obfuscated, and confused public behavior in the United States' COVID-19 response cautioned both politicians who appoint themselves scientists and scientists—including virologists and epidemiologists—to stay in their lanes. The warning raises an important question: Should science add another lane?

Despite the remarkable development of safe and effective vaccines, only about two-thirds of Americans have received their first dose. Even nonmedical actions (social distancing and masking) supported by rigorous evidence are met with widespread indifference, resistance, and rage. Unfortunately, this number is the rule rather than the exception. Broadly, Americans receive about 55% of clinical interventions known to benefit their health.

To address this failing, science needs to add another lane—one called implementation research. Implementation scientists move beyond medication and device development and study how to facilitate their use by clinics, front-line health care providers, patients, communities, and policy-makers. Public health failures that could have been avoided, as well as successes attributable to this science, illustrate the importance of this work. The use of beta-blockers after myocardial infarction was proven to reduce mortality in a 1982 trial, but 15 years later, only 34% of Americans hospitalized with a heart attack were discharged with a prescription for these drugs, and the practice was not universally adopted until 2007. In HIV prevention research, numerous randomized trials in Africa found that adult male circumcision reduces HIV transmission. However, its use remained low despite the World Health Organization's endorsement and massive donor-funded work to scale up surgical capacity. A recent randomized trial in Tanzania showed that the engagement of pastors—who are influential community opinion leaders—boosted the acceptability and uptake of circumcision in men, demonstrating the value of social influence in implementation.

Similar research on how to expand the use of proven COVID-19 interventions is underway but must be scaled up substantially to address pressing questions: What strategies lead to vaccine acceptability, feasibility, fidelity, equity, scale-up, and spread? What social marketing messages are most effective? Who are the best opinion leaders? How can health systems overcome delays in identifying mildly ill outpatients eligible for monoclonal antibodies? Data are emerging about how to equip vaccine champions with the resources necessary to train others, build coalitions, and optimize organizations to administer vaccines as widely as possible. But more must be done, especially given the current politicized pandemic response and frayed social fabric.

Society needs a lane of science that studies rapid uptake of proven interventions. Questions pursued in implementation research require cross-disciplinary collaborations among scientists who understand communication, marketing, anthropology, economics, and social psychology—disciplines that have not historically interacted with one another.

> "COVID-19 has shown the world that 'knowing what to do' does not ensure 'doing what we know.'"

Three steps would contribute to a better pandemic response now and in the years ahead. The US National Institutes of Health (NIH) should create an Office of Implementation Research with funding that institutes must compete for, modeled on the Office of AIDS Research. The office would study emerging interventions and address obstacles to their use. Insights would guide health delivery, making learning-while-doing a standard. The office should support innovations that track rates of intervention use (vaccination and effective therapeutics) and capture the strategies leading to their uptake. And the NIH should support networks for implementation research, similar to the AIDS Clinical Trials Group. At least 10% of the NIH budget should be dedicated to this work. If this seems expensive, consider the costs of not taking these steps: Effective interventions that are not used optimally will fail to reap value from existing investments.

COVID-19 has shown the world that "knowing what to do" does not ensure "doing what we know." It demonstrates that intervention discovery is the start, not the end, of the scientific journey. There is no better time for science to establish a new lane, one devoted to ensuring that our nation's health discoveries are used to improve population health. The headwinds demand nothing less.

**–Enola K. Proctor and Elvin Geng**

**Enola K. Proctor**
is the Shanti K. Khinduka Distinguished Professsor Emerita, Brown School, Washington University in St. Louis, MO, USA.
ekp@wustl.edu

**Elvin Geng**
is a professor of Medicine in the Department of Medicine, Division of Infectious Diseases, Washington University in St. Louis, MO, USA.
elvin.geng@wustl.edu

10.1126/science.abn0184

# Enabling the next generation of pioneering interdisciplinary researchers

David Brückner is a physicist who shifted from investigating classical mechanics to studying migrating cells at the interface of physics and biology. His postdoc at the Institute of Science and Technology Austria (IST Austria) is unusual–he has two advisors, one of whom specializes in theoretical biophysics and neuroscience and one who works on experimental bioscience.

Christopher Currin is a neuroscientist with expertise in artificial intelligence (AI) and machine learning. His postdoc at IST Austria also involves a pair of mentors from disparate arenas. One of his principal investigators (PIs) focuses on computational neurotheory and neuroscience and the other on experimental neurogenetics and molecular neuroscience.

Both scientists are funded by a new fellowship program inaugurated by the NOMIS Foundation and situated at IST Austria that seeks to spark high-risk/high-reward, vanguard science by investing in early-career scholars whose work requires a merging of traditional disciplinary boundaries. Brückner and Currin agree: The likelihood that they would be able to do this type of inventive, cross-border, high-stakes research at any other institution is low, as is the chance they could convince a funding agency to finance this work. Their research is just too … risky.



Chris Currin, NOMIS fellow

But here on the outskirts of Vienna, nestled on the edge of the Viennese Alps, there is a new model for science erupting. At its core, it is a unique mentorship approach that unites investigators from diverse specialties with a distinct purpose: to transform science and the training of scientists for the betterment of humanity.

IST Austria, established in 2009, is led by computer scientist Thomas A. Henzinger. With only 67 professors and 250 students, it is youthful enough to be agile and creative in its pursuits but mature enough to take a distinctive approach to mobilize its scientific assets. "At IST Austria, all major disciplines of science come together on one campus," says Henzinger. "As a young institute with a single, multidisciplinary Ph.D. program, we provide a unique atmosphere to foster basic research across disciplines." There are no departments, but individual groups focus on different research topics. There are three essential reasons for the Institute's success: Typical boundaries between scientific fields are blurred, the emphasis is entirely on problem-solving, and cooperation between groups is encouraged.

In 2020, the Institute received a significant contribution to its investment in interdisciplinary investigations and early-career training through the NOMIS Fellowship Program at IST Austria, which furthers the same entrepreneurial spirit endemic to the Institute's culture by supporting postdocs who have radical ideas. With an appreciation for engendering scientific examination and discovery in an increasingly complex ecosystem, the program provides research support and mentoring for future trailblazers in interdisciplinary science.

IST Austria is the right place for these innovative funding mechanisms. One of the founding principles of the Institute is interdisciplinarity. In tailoring this fellowship to train postdoctoral scientists in interdisciplinary research, IST Austria and the NOMIS Foundation are paving the way for future thought leaders to navigate science's most complex grand challenges, those which require creative insight and tools to converge at the intersection of many fields.

## Crossing disciplinary borders

Each year, two postdocs who have proven themselves to be flexible, blue-sky thinkers are chosen for the fellowship program. Their proposals must illustrate how they will leverage and coalesce seemingly incongruent subject areas into a viable research project to be completed over a 3-year period. In fact, one of the unique aspects of the program is that fellows are required to have two PIs as mentors, to inspire fresh, synergistic collaborations between diverse groups.

"There is a lot of conversation and communication between experimental, computational, and clinical groups," says Currin, emphasizing that his research would not be conceivable without this dialogue. The postdoc, who is also an AI consultant and entrepreneur, is aiming to broaden our understanding of how neurons work. He analyzes neuronal electrical signals to better understand the mechanisms that cause epilepsy by using spike sorting, a technique that helps him identify synchronous activity in the brain. "Epilepsy is characterized by highly synchronous nerve cell activity, so we are looking to control the synchronicity of these neurons," he says.

One of Currin's mentors is Gaia Novarino, a neurobiologist and translational neurogeneticist who investigates the molecular and physiological cellular basis of brain disorders, especially autism spectrum disorders and epilepsies. His other mentor is Tim Vogels, whose group builds models and theories about the computational mechanics of learning, computing, and executing thought. "Chris is using in vitro models of human cells to study the electrical properties of cells and neurons," says Novarino. "He is purely theoretical. He is applying computational techniques to understand the data we are acquiring from those in vitro neuronal networks." Currin meets with Novarino and Vogel frequently as they pursue a shared research agenda. The mentorship plays a vital role in building Currin's innovation and leadership capabilities. For example, by having diverse mentors, he is exposed to different avenues of communicating and problem-solving. "The mentorship process starts with clarifying ways to communicate with each other," says Novarino. "With people from different backgrounds, you have to find a common language to explain goals and share possible solutions." The fellow is exposed to and becomes fluent in that common language–an essential step in the partnership.

And as Novarino describes it, it is a win-win alliance: The postdoc gains critical skills and leadership opportunities, and the mentors get a unique chance to stretch their perspectives and delve into problem-solving in new domains. "I learn every day," she says, "for example, how to use



Gašper Tkačik, professor of biophysics and information theory

mathematical models or computational tools to analyze the data in an effective and different way. It is mutually beneficial."

Gašper Tkačik, one of Brückner's mentors, agrees. "Our hope is that David will be the person to bring us together," he says. "It is not just a shared research project. It is a legacy that the groups will stay connected. You can have a collaboration that is purely transactional–one project of interest to two groups, where you coauthor a paper, and then you finish. Then there are longer-term collaborations where the two PIs get to be friends and colleagues; there is a shared responsibility and social life."

Brückner's ambitious research agenda is already exceeding expectations. Endeavoring to elucidate cellular decision-making processes–i.e., how cells determine their fate to become certain tissues–Brückner is applying tools from information theory. "There are existing frameworks with which to think about these embryos, and theories as to how to make decisions given noisy input like mechanical and chemical information," he says. "The challenge is to draw on theories from information processing and apply it to biological systems."

"There are not many groups that are trying to do this work David is envisioning," says Tkačik. "Edouard is a specialist in migrating cells and we are good at the technical aspects of modeling." It's a perfect marriage.

## The freedom to explore

Freedom is a theme here. The NOMIS Fellows are given liberty from many barriers prominent in the typical postdoc journey. For example, fellows are not tethered to an expected outcome for their proposal; instead, "I can spend time thinking about how I would solve a problem," says Currin. "Not having to worry about the result first means I can make sure the method is done better. This lessens the burden." Adds Brückner: "What's special is they give me permission to explore in any direction that makes sense. In other proposals, you have to do exactly what is written in the proposal. But high-risk/high-reward doesn't work that way–you may start on one thing and end up somewhere else. So that freedom is a huge deal."

There is freedom to fail here, too, with the knowledge that every setback is a potential step forward for human understanding and discovery. "Sometimes we have to ask those big scientific questions without knowing the clinical value or relevance yet," says Currin. "Freedom to fail is a luxury in life and certainly in science, but here it is OK. I feel free to change things if they are not going well."

The mentors honor this freedom, too. "I view the NOMIS fellowship as an opportunity, or mostly an obligation, for mentors to give more freedom to fellows," says Tkačik. "They have been chosen by high standards, so there is trust implicit in this. It is not just about writing a proposal and us putting our stamp on it. The fellows should be able to take our advice and suggestions in different ways, but they should be in the driver's seat."



Gaia Novarino, professor of neuroscience, investigating genetic causes of neurodevelopmental disorders



David Brückner, NOMIS fellow

> "Science is based on diversity on so many different levels. So being able to increase the range of thinking and the number of scientists to bridge the gap across disciplines is essential to thinking in different ways."

## Mentorship from beginning to end

At the core of IST Austria's success is its substantial emphasis on mentorship. "By definition, the work I do would not be possible without the experimentalists, and they in turn would not be able to do their work without the computational side," says Currin. "You need both."

The NOMIS fellowship program "brings together people who typically would not interact and trusts that they will develop something special," says Novarino. "Facilitating the interactions between different groups of people is very important. In other fellowship schemes, this typically wouldn't happen."

Diversification of science (and science funding) is a critical component here. "We don't want more of the same, with everyone thinking the same," she adds. "This is not good science. Science is based on diversity on so many different levels. So being able to increase the range of thinking and the number of scientists to bridge the gap across disciplines is essential to thinking in different ways."

"A problem doesn't know about disciplines and predetermined theories. It is just there and needs to be solved," says Brückner. "Bringing together these disciplines allows us to solve problems in new ways that might not have been done if we stuck with one way of doing it."

"It's great to cure cancer, but that's not why we do these things. There are parts of the research that are curiosity driven, where we don't know what the absolute next step is," says Tkačik. "We should be fair to ourselves in that some research is about making clearly planned, hypothesis-driven next steps. And applying for funding at any level with that knowledge is hard. Granting agencies are not very happy if you ask for funding only because a problem is interesting, but there are no promises of what to do next," says Tkačik. "It is good to have a full spectrum of institutions that support research at all levels of 'fuzziness.' For that reason, NOMIS–which funds high-risk research that provides flexibility to guide your work without it being fully planned out–is needed."

## An energetic ecosystem of innovation

The benefits of the research ecosystem at IST Austria are plentiful, say the fellows and their PIs. The weekly seminar is a cornucopia of interdisciplinary delights, including a special series devoted to theoretical neuroscience topics. "It's never been the case that I've had too many talks to go to–until now!" jokes Currin. "I have felt very welcome and empowered here, able to focus on the research and not the administrative stuff that takes away from your mental capacity to focus. Here I feel my needs are taken care of and I am able to engage with my community."

IST Austria and the NOMIS Foundation see this pilot as a potential global model, something that is sorely needed. "The program is great," says Novarino. "NOMIS gives opportunities to these early-career scientists to develop directions that could not survive in more classical, conservative programs. The facility and ability to explore in new ways creates the possibility of developing a new class of scientists who can find other space and initiatives and missions."

"I wish there were more opportunities like the NOMIS fellowships," says Tkačik. "The really complex questions of today, such as vaccine hesitancy, climate change, or the genetic basis of complex diseases, require us to step out of a disciplinary field and to collaborate, but that is not enough. They require a new breed of scientist. Fellowships like these get people to where they can work at the interface of fields and have the time and freedom to succeed."

NOMIS FOUNDATION | FELLOWS

IST AUSTRIA
Institute of Science and Technology

## NOMIS Foundation fellowships

NOMIS fellowships support emerging, early-career researchers who demonstrate the potential to lead groundbreaking, high-risk basic research. The fellowships provide postdocs with the freedom to explore big unanswered questions and work on high-risk interdisciplinary research with the support of outstanding scientific mentorship. The NOMIS Fellowship Program at IST Austria appoints two fellowship positions each year. More information is available at ist.ac.at/en/education/postdocs/nomis-fellowship.

NOMIS & Science
YOUNG EXPLORER AWARD

## NOMIS & *Science* Young Explorer Award

In 2021, the NOMIS Foundation and *Science*/AAAS launched the NOMIS & *Science* Young Explorer Award, which recognizes bold early-career researchers who ask fundamental questions at the intersection of the life and social sciences. The focus is on researchers who take risks to address relevant and exciting questions with creative approaches. The annual call for applications will open in the spring. For more information, visit www.science.org/content/page/nomis-science-young-explorer-award.

# NEWS

> "It's not a very happy place. We are getting quite concerned about this."
>
> **David Frost,** top U.K. negotiator for Brexit, in *The Guardian*, about his country's impasse with the European Union over participation by U.K. scientists in the EU Horizon Europe funding program.

## IN BRIEF

*Edited by* **Jeffrey Brainard**

FUNDING

## Biden plan offers smaller science boost

The U.S. Congress could begin to vote as early as this week on President Joe Biden's Build Back Better plan, a $1.75 trillion package that includes hefty boosts for some federal research agencies. Many of the sums for science in the plan's latest version—which also supports efforts to deal with climate change and enhance social welfare—are far smaller than those included in an earlier one. A proposed $10 billion increase for the Department of Energy's Office of Science for 19 major projects, including new instruments at national laboratories, shrank to just $985 million over 5 years, all targeted at fusion energy, in the current plan. The National Science Foundation, once slated to get $11 billion, instead would receive $3.5 billion through 2028. Congress has yet to finish work on separate bills that set spending for science in 2022.

**How select agencies and programs fared**

| AGENCY/PROGRAM | AMOUNT FOR RESEARCH ($ MILLIONS) | PURPOSE |
|---|---|---|
| National Science Foundation | 3500 | Includes $1.5 billion for new technology directorate |
| Department of Energy Office of Science | 985 | Studies of fusion energy, low-dose radiation |
| National Institute of Standards and Technology | 340 | Advanced manufacturing, worker training, and wildfire studies |
| National Oceanic and Atmospheric Administration | 100 | Climate research grants |
| NASA | 85 | Earth science studies involving climate change |
| National Institutes of Health | 75 | Support for minority-serving institutions |

## Antidepressant fights COVID-19

THERAPEUTICS | An inexpensive antidepressant commonly used to treat obsessive-compulsive disorder significantly decreased the risk of COVID-19 patients becoming hospitalized, conclude authors of a large trial published last week. The randomized, controlled study examined fluvoxamine in 1500 symptomatic, unvaccinated people in Brazil with at least one preexisting condition that made them more likely to develop severe COVID-19. Patients who received a fluvoxamine pill twice a day starting within 7 days of developing symptoms were 32% less likely to be hospitalized or need prolonged observation in an emergency room than those in a placebo group, the authors reported in *The Lancet Global Health*. In patients who took the pills for at least 80% of the recommended 10-day course, the risk of hospitalization was 66% less than in those in the placebo group, and just one of the treated patients died, compared with 12 who received the placebo. Although fluvoxamine's mechanism of action in the COVID-19 context is uncertain, the drug is a selective serotonin reuptake inhibitor that has antiviral and anti-inflammatory effects; it also has antiplatelet activity that could potentially reduce the widespread clotting that can complicate severe COVID-19.

## Mood disorders qualify for shots

COVID-19 | Tens of millions of additional U.S. adults became eligible for COVID-19 vaccine booster shots last month, when the U.S. Centers for Disease Control and Prevention (CDC) added depression, schizophrenia, and other mental health disorders to a list of conditions that put people at higher risk of severe COVID-19. About 50 million U.S. adults suffer from mental health conditions, the government estimates. CDC cited two meta-analyses published in *JAMA Psychiatry* in July and October. The first found a 38% increased risk of COVID-19 mortality in people with any mental health disorder and a 67% increased risk in people with schizophrenia or bipolar disorder. The second, which analyzed studies involving 19 million people, found a higher (31%) risk of COVID-19 hospitalization and mortality (51%) among people with mood disorders.

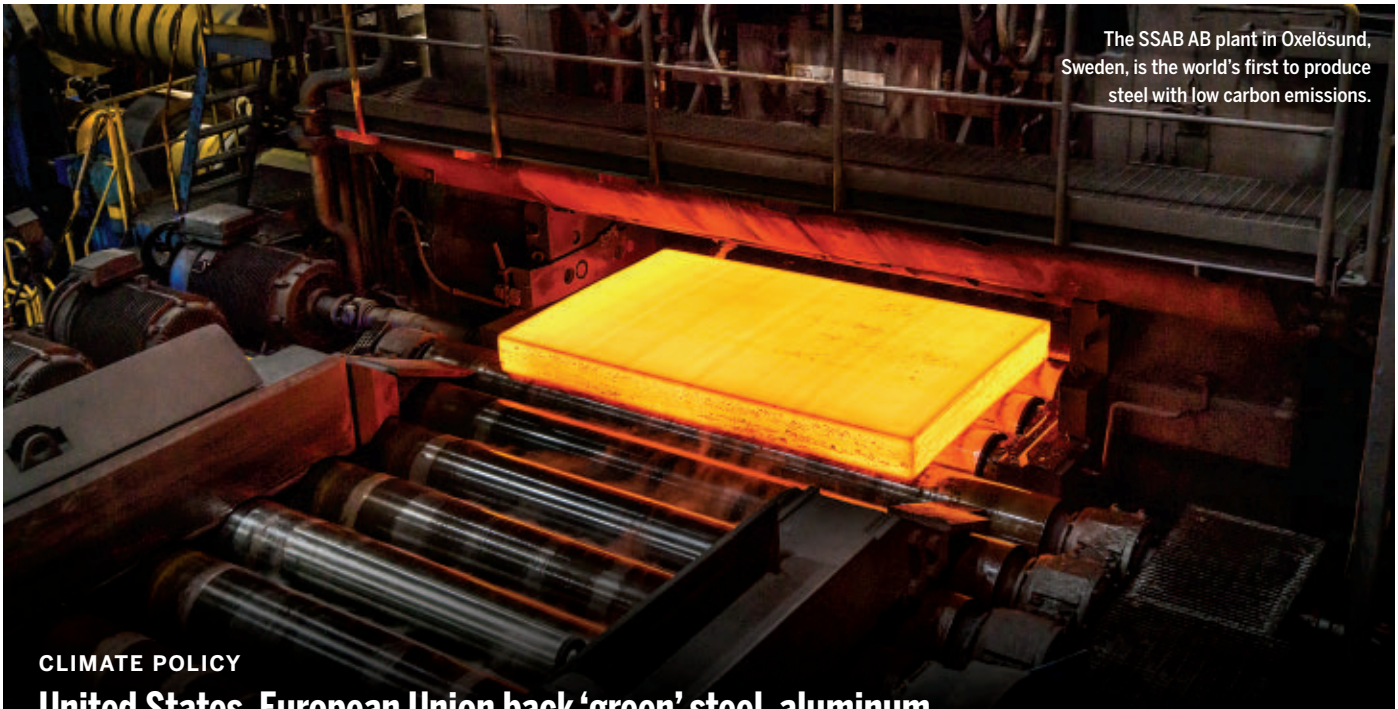
## Ukraine to get treasures back

COLLECTIONS | A disputed set of archaeological treasures loaned from museums in Crimea should be returned to the Ukrainian government, a Dutch court ruled on 26 October. Crimea was part of Ukraine until March 2014, when Russia annexed it. At the time, the University of Amsterdam's Allard Pierson Museum was holding the objects—including jewels from Chersonesus, a Greek colony founded some 2500 years ago in present-day Sevastopol, and a gold helmet from the



An Amsterdam museum must return to Ukraine this silver pin and other treasures on loan since 2014.

PHOTO: ALLARD PIERSON MUSEUM



The SSAB AB plant in Oxelösund, Sweden, is the world's first to produce steel with low carbon emissions.

PHOTO: JAN LINDBLAD JR./SSAB

CLIMATE POLICY

## United States, European Union back 'green' steel, aluminum

On the eve of the U.N. climate summit in Glasgow, U.K., the United States and European Union this week said they will negotiate a deal that expands methods to reduce carbon emissions from the manufacturing of steel and aluminum. Collectively, steel and aluminum production account for 10% of all global carbon releases. The pact also calls for an immediate truce in a trade war between the two parties over those products. The emissions deal, to be negotiated during the next 2 years, aims to restrict imports of metals produced with "dirty" methods, targeting especially those from China, which are produced largely by burning coal. The agreement is also expected to contain incentives to reduce emissions by U.S. and EU producers of those metals.

Scythians, a nomadic people living on the region's steppes around the same time. The museums wanted them back, but the Ukrainian government, which owned the museums until the annexation, claimed the objects as well. The Allard Pierson Museum says it has "no opinion" on the outcome, but is "pleased that a verdict has been reached."

## Indonesia OKs Novavax vaccine

COVID-19 | Indonesia's government this week authorized Novavax's vaccine, making it the first protein-based COVID-19 vaccine to win authorization anywhere, the company said. The small biotechnology company and its partner, the Serum Institute of India, asked the World Health Organization for emergency use listing in September, but no decision has been issued. U.S.-based Novavax says it expects to ask the U.S. Food and Drug Administration for authorization by December, after it struggled to meet the agency's vaccine purity standards. Twenty-seven percent of people in Indonesia, the fourth most populous nation, had been fully vaccinated against COVID-19 by 31 October, according to *Our World in Data*. Since 27 October, Novavax has also applied for authorizations in Australia, Canada, and the United Kingdom.

## Co-author wins retraction bid

PUBLISHING | A journal retracted a controversial article about plant biodiversity last week after more than 1 year of effort by a co-author turned whistleblower who sought to expose a range of data anomalies attributed to the paper's senior author. The paper, in *Biodiversity and Conservation*, purported to show a method of distinguishing plant species by comparing snippets of DNA—known as DNA barcoding—is more cost-effective than traditional plant-survey methods. Ken Thompson, a postdoctoral fellow at Stanford University, worked on the paper as an undergraduate in 2014 at the University of Guelph (UG) with Steven Newmaster, a botany professor there. But he eventually asked the journal and UG to investigate the paper—his first to be published—asserting in part that the data bore a dubious similarity to separate data collected at the Canadian Centre for DNA Barcoding, also located at UG. The retraction notice in *Biodiversity and Conservation* corroborated several of Thompson's concerns. Newmaster, it added, "has not responded to any correspondence from the Editor or publisher about this retraction." Newmaster did not respond to a request for comment from *Science*. A UG spokesperson says the university "is not in a position to comment on a decision made by a journal."

## NIH yanks diversity policy

EQUITY | The U.S. National Institutes of Health (NIH) last week canceled a policy meant to help bridge a funding gap between Black and white scientists. A May notice from three institutes offered to highlight applicants from groups underrepresented in science, including people of color and people from disadvantaged backgrounds, which could boost a proposal close to the threshold to win funding. But on 25 October, NIH's Office of the Director rescinded the notice to promote "clarity in communications." NIH said institutes already had leeway to consider "diverse scientific perspectives"

when picking proposals. Last week, NIH also released an agencywide notice that encourages applications from underrepresented groups, but the applications won't be tagged as such. Outside scientists voiced disappointment, noting that program officers won't know which applications come from scientists from underrepresented groups. The now-rescinded policy came after NIH Director Francis Collins apologized in March for "structural racism" in science. Data through 2020 show a long-standing gap between white and Black investigators' success rates when applying for research grants had narrowed only slightly.

## MIT food computer paper pulled

PUBLISHING | The journal *PLOS ONE* last week retracted a controversial 2019 paper co-authored by scientists at Massachusetts Institute of Technology's (MIT's) Media Lab about a "food computer" it suggested could revolutionize agriculture by identifying optimum growing conditions. The retraction notice cited "concerns about the reliability of the article's results and conclusions" and said the latter were not supported by experimental data. The retraction followed the filing in August of a wrongful termination lawsuit against MIT by Babak Babakinejad, a scientist who ran the food computer project. He says he was fired after he complained about the paper to MIT officials. MIT's news office ended up widely promoting it. This week the institute declined to comment on the pending litigation, but wrote in an emailed statement: "Ensuring an accurate research record is important to MIT, and we have updated our MIT News article about this paper to reflect the retraction." The retraction adds to a string of troubles for the Media Lab, whose former director, Joi Ito, resigned in 2019.

## Gene therapy gets a boost

BIOMEDICINE | A new U.S. public-private partnership is putting up $75 million over 5 years to speed the development of genetic treatments for rare inherited diseases. The Bespoke Gene Therapy Consortium, announced on 27 October, will focus on adeno-associated viruses (AAVs), a widely used vehicle to introduce new genes into cells. Researchers will study the basic biology of AAV gene delivery, and how to streamline manufacturing of the viruses and regulatory steps such as toxicology testing. The consortium will also launch four to six clinical trials. Participants include the National Institutes of Health, which will contribute just over half the funding; the Food and Drug Administration; 10 companies; and five nonprofit groups, including the Foundation for the National Institutes of Health, which will manage the project. The United States has only approved gene therapies for two inherited diseases—spinal muscular atrophy and a blindness disorder.



Sabertoothed tigers needed healthy hips to leap on and take down big prey.

ILLUSTRATION: LA BREA TAR PITS

ANIMAL BEHAVIOR

## Did social bonds help disabled sabertooth survive?

Rendered lame by hip dysplasia, a sabertoothed cat couldn't have hunted, dooming it to a premature death—unless it got a helping hand. Last week, a research team described in *Scientific Reports* evidence for just that, supporting previous work indicating these big cats lived in social groups. The team used computed tomography imaging to create 3D models of a fossil pelvis and right femur from an adult sabertoothed tiger that died more than 11,000 years ago and was found in California's La Brea tar pits. They discovered the cat's hip joint had been afflicted by hip dysplasia from a young age; common in modern pet dogs and cats, the condition can cause hip dislocations. The researchers suggest other sabertoothed cats must have shared their kills with their injured companion. *Smilodon* outnumber herbivores at the La Brea pits, leading scientists to speculate that they hunted in packs.

BY THE NUMBERS

**4%**

Share of global funding for climate change research that goes to topics in Africa, even though the United Nations deems the continent the most vulnerable to climate impacts. African institutions receive less than one-quarter of the money. (*Climate and Development*)

**$1.8 billion**

Funding obtained by 18 private fusion energy companies, according to the first annual report about their industry. Two-thirds of the 23 companies featured were founded in or after 2010. (Fusion Industry Association, U.K. Atomic Energy Authority)

**SCIENCE.ORG/NEWS**
Read more news from *Science* online.



Flaring burns off excess methane in oil fields, preventing the potent greenhouse gas from accumulating in the air. Some researchers want to remove methane in other ways.

PHOTO: ORJAN F. ELLINGVAG/CORBIS/GETTY IMAGES

CLIMATE CHANGE

# Methane removal seen as tool to slow warming

## As nations look to cut emissions of the gas, some researchers want to pull it out of the air

*By* **Cathleen O'Grady**, *in Glasgow, U.K.*

Methane, long eclipsed by carbon dioxide ($CO_2$) as a villain in climate change, is having its moment in the spotlight. The United Nations climate summit opened here this week with faltering global commitments to reducing $CO_2$, but more than 100 nations have agreed to cut their methane emissions by one-third by 2030. Officials hope cracking down on the potent greenhouse gas will help limit global warming to a 1.5°C rise, the goal established by the 2015 Paris agreement.

Now, some researchers say they want to not only put less methane into the atmosphere, but also actively pull it out. At a side event at the summit, researchers with the advocacy group Methane Action argued that so-called negative emissions technologies—alongside every trick in the book to reduce emissions—could restore methane to preindustrial levels and trim an estimated 0.4°C to 0.6°C of warming. With a limited time window to reach the Paris target, "Methane is really the only lever we have to buy a few decades and shave off peak temperatures by a few tenths of a degree," says Rob Jackson, an earth scientist at Stanford University and a Methane Action board member.

Methane is a ripe target for climate action because it traps 84 times more heat than the same amount of $CO_2$ over 20 years. It has more than doubled since preindustrial times and accounts for about half the 1.1°C of global warming to date. Big offenders include leaks from coal mines and oil and gas operations, burps and farts from cattle, and landfills, where methane-spewing microbes consume organic matter. Emissions from natural sources such as wetlands are picking up, too, as the planet warms.

Thankfully, methane is short-lived, breaking down into water and $CO_2$ in about 10 years. That means methane cuts would quickly translate into lower atmospheric levels.

Ideas to remove methane would "enhance, or speed up, this natural oxidation of methane," Jackson says. A fleeting, naturally occurring compound, the hydroxyl radical, defangs most atmospheric methane, but about 2% is destroyed by reactions with chlorine, produced when the Sun shines on salt spray in the ocean, says Maarten van Herpen, scientific adviser to Methane Action. Lab experiments show that adding iron to sprays of seawater can boost production of both chlorine and hydroxyl and significantly speed up methane destruction.

Van Herpen and his colleagues want to test the concept with a handy, real-world source of iron at sea. Marine fuel oil contains iron additives; by studying the iron, chlorine, and methane composition in the plumes from ship smokestacks, the team hopes to provide a proof of principle for the technology. Deployment could rely on these smokestacks, or even purpose-built chimneys, to emit the ideal iron-salt aerosol mix. Emitting just 10% of the amount of iron that naturally reaches the atmosphere in mineral dust could be enough to restore methane to its preindustrial level, says Matthew Johnson, a chemist at the University of Copenhagen.

Another option is to deploy a catalyst across large areas. One candidate is titanium dioxide paint, which in the presence of ultraviolet light can oxidize methane in air that blows across its surface. Buildings or even wind turbines could be coated with these "photocatalyst" paints, Jackson says.

But the technologies face many challenges. For starters, methane only exists at levels of 2 parts per million, whereas $CO_2$ has now surpassed levels of 400 ppm. To make a difference, methane removal technologies would have to be deployed on a mammoth scale. "It's like pulling a needle from a haystack over and over again," Jackson says. The dream might be to remove the 2 billion to 3 billion tons that would restore preindustrial levels, Jackson says, but a more modest goal might be removing 100 million tons per year to offset persistent agricultural emissions, or new emissions produced by a warming Arctic.

Funding is also lacking, says Daphne Wysham, CEO of Methane Action. Although technologies to remove carbon from the atmosphere have received billions of dollars in subsidies and incentives, methane removal has gotten next to nothing, she says. "Right now, it is so unknown in policy circles." The group hopes to persuade diplomats at the summit to approve a declaration that calls on governments to commit resources to methane removal.

And some researchers and environmentalists are wary of monkeying with global atmospheric chemistry. Lofting vast amounts of iron-salt particles above the ocean is reminiscent of long-standing schemes for cooling Earth by creating a reflective haze in the stratosphere, an idea some have derided as a potentially dangerous form of "geoengineering." But Jackson rejects that label. Unlike solar geoengineering, the iron-salt method would just speed up a chemical reaction that occurs naturally, he says. And the particles are short-lived, so the scheme could be easily halted. The group is planning an independent safety study before going ahead.

Methane removal is "a fascinating idea," says University of Washington, Seattle, atmospheric chemist Alex Turner, "as long as those methods don't lead to destruction of ecosystems or harm to the environment." But he would prefer to target the small number of actors in agriculture and the fossil fuel industry that are responsible for a hefty portion of methane emissions.

Euan Nisbet, an earth scientist at Royal Holloway University of London, says methane removal efforts ought to focus on sites where the gas is already concentrated, like "monster" landfills, which could be capped with a layer of soil packed with methane-guzzling microbes. Cow barns and coal mine shafts could also be suitable targets for removal technologies, he says.

Jackson says methane removal would be futile if carbon and methane emissions aren't curtailed. But removal is urgent, too, he says. "It has to be deployed in the next 20 to 30 years to really make a difference." ■



An engineered pig's kidney was attached last month to the circulation of a brain-dead person.

PHOTO: JOE CARROTTA/NYU LANGONE HEALTH

BIOMEDICINE

# Pig kidney transplant obscures value of engineered animals

## Skin and nerves from pigs lacking immune-provoking sugar could help patients today

*By* **Elie Dolgin**

Though the study leader heralded an operation as providing "new hope for an unlimited supply of organs," last month's much-publicized transplant of a kidney from a genetically modified pig to a brain-dead human recipient did little to usher in cross-species organ swaps, others say. Still, the engineered pigs could find less flashy uses today: as donors for skin and nerve grafts, providers of long-lasting heart valves, and sources of allergy-free meat.

The transplant study, performed at New York University (NYU), showed the human immune system doesn't immediately reject an organ from a pig engineered to lack alpha-gal, a sugar molecule that provokes a frenzied immune response in people. The porcine kidney filtered waste from the blood and produced urine for at least 54 hours. After that, doctors shut off the recipient's ventilator and ended the experiment.

But those findings are "seriously not a surprise," says Wayne Hawthorne, a transplant scientist at Australia's Westmead Institute for Medical Research and president of the International Xenotransplantation Association, a society for researchers who study using animal organs and cells to treat human disease. Monkey experiments had already shown that pig kidneys like the one the brain-dead woman received usually last up to 1 week with no problems. *Science* could not reach NYU surgeon Robert Montgomery, but at a press conference last month, he defended the study, noting "many other examples of preclinical primate studies that have not translated well into what happens in humans."

Lacking just one sugar gene, the organs aren't a viable option for xenotransplants, which need to last for years. Other sugar molecules on porcine cells have similar immune-triggering effects, and experts say they, too, should be genetically purged. Plus, several human genes may need to get added to the pig genome to further temper the body's immune reaction. But the pigs, engineered at a biotech company now called Revivicor, do have promise for other medical applications with less chance of immune rejection. "Just taking out one sugar, you can make a big difference," says Thomas Platts-Mills, a clinical immunologist at the University of Virginia.

In the late 2000s, for example, Platts-Mills showed tick bites could provoke an unusual allergic response to alpha-gal. People with this "alpha-gal syndrome" (AGS) develop severe reactions to red meat and many medical byproducts of animal husbandry, including the blood thinner heparin (made from pig intestines) and implantable devices such as bioprosthetic heart valves from cows or pigs. (The antibodies that cause these allergic reactions, elicited only by tick bites, differ from those that attack transplanted pig organs, which everyone carries.)

Revivicor had won regulatory approval for its engineered pigs as a source of meat in December 2020, and although pork chops and sausage patties from the swine aren't yet commercially available, the company has begun to send free samples to people like Amber Shifflett of Charlotte Hall, Maryland, who has AGS and spent months avoiding red meat. On a Saturday in mid-October, Shifflett, 30, cooked an allergy-safe ham steak for breakfast. Then she had another ham steak for dinner. "I was nervous at first," Shifflett says. Would her chest tighten, as it usually did after eating meat? "I was completely fine." Now, she says, "I just want to savor every last bite."

Beyond meat, Scott Commins, an allergist at the University of North Carolina, Chapel Hill, who treats AGS, hopes to offer his patients allergy-free pharmaceuticals. "Bacon is nice," Commins says, "but I really believe that there is a broader use of the animals medically that could arguably be more important." During the pandemic, for example, a safe version of heparin would have benefited AGS patients who developed blood clotting issues from COVID-19.

Some medical products from alpha-gal–free pigs could find uses beyond AGS patients. Replacement heart valves, for instance, are increasingly being made from cow and pig tissues rather than mechanical alternatives. But in part because of immune attack, these bioprosthetic valves deteriorate and must be replaced after 10 to 15 years. To slow that attack, animal-derived valves are stripped of cellular material and chemically treated to mask immune-stimulating remnants. But according to pediatric heart surgeon Joseph Turek of Duke University, alpha-gal remains detectable at "actually pretty astounding" levels—as he, Commins, and others reported in *The Journal of Thoracic and Cardiovascular Surgery* in April. An alpha-gal–free option might last twice as long as other bioprosthetic valves, estimates Turek, whose work is funded in part by Revivicor.

The alpha-gal–free pigs could also find use as skin donors for burn victims and as a source of neurons to treat peripheral nerve injuries. A company called Alexis Bio (formerly XenoTherapeutics) has already tested skin from the pigs on six people with third-degree burns. Ordinarily, doctors rely on human skin from dead people as a temporary dressing, until patients can get a graft of their own skin. But cadaveric skin is expensive and often in short supply. Engineered pig skin seems to function just as well to help wounds heal for up to 9 days, the longest duration evaluated so far.

"The most striking thing [about the pig skin] is that there is nothing striking," says trial investigator Jeremy Goverman, a burn surgeon at Massachusetts General Hospital (MGH). Because the graft only stays on the body for a week or two, the immune system doesn't have time to reject it. "The use of it is intended to be short term," says Paul Holzer, Alexis Bio's CEO.

His company and another startup, Axonova Medical, are also developing nerve grafts from the same sorts of pigs. These could supplant the current go-to approach for nerve repair surgery: harvesting nerves from somewhere else on a person's body, an option that is not always possible and can leave people with sensory problems.

Axonova research director Kritika Katiyar and her colleagues start with neurons from Revivicor pig embryos and grow them into nerves up to 5 centimeters long. Having already repaired damaged leg nerves in rats and facial nerves in pigs, the company plans to move on to monkey studies next. Meanwhile, MGH plastic surgeon Curtis Cetrulo, Holzer, and colleagues reported in September using leg nerves from pigs missing alpha-gal to repair damaged arm nerves in rhesus macaques. In both teams' studies, over about 6 months, recipient animals' own nerve cells replaced the pig tissue—eliminating the risk of immune rejection. "It looks like the host tissue has completely taken over, and the graft has been removed from the body naturally," Katiyar says.

Researchers such as University of Miami xenotransplantation scientist Christopher Burlak await the day when larger organ transplants from pigs—kidneys, livers, or hearts—help saves lives. The NYU study was a "first step in a long road," he says. Companies are already taking the next ones, designing pig organ donors with three or more genes removed and up to nine added human genes. None of those pigs' organs has been tested in people. "The next big advance in the field is going to be real transplants," Burlak says, "not the short-term evaluation of xenografts in dead people." ■

Elie Dolgin is a science journalist based in Somerville, Massachusetts.

MATERIALS SCIENCE

# Reactor mishap derails U.S. studies with neutron beams

## NIST's research reactor suffered little damage, but will be offline for months

*By* **Adrian Cho**

An accident in February has shut down a diminutive nuclear reactor at the National Institute of Standards and Technology (NIST)—and has cost the United States, at least temporarily, almost half of its capacity to study materials with beams of neutrons. In the incident, a fuel rod in the 52-year-old reactor at the NIST Center for Neutron Research (NCNR) in Gaithersburg, Maryland, overheated and partially melted, releasing a small amount of radiation. The public was never in danger, NIST says. But the reactor won't restart until April 2022 at the earliest, leaving thousands of users scrambling to find beamtime elsewhere.

"We fully intend to restart the reactor," says Robert Dimeo, director of NCNR. "We're only going to do it when we're confident that it is safe to do so."

In the meantime, the shutdown is "a huge problem," says Robert Birgeneau, a condensed matter physicist at the University of California (UC), Berkeley, whose team has used the neutron source to study exotic iron-based superconductors. Michael Hore, a polymer physicist at Case Western Reserve University, says the shutdown will set one of his projects back 1 or 2 years. "It's not like we can go somewhere else," he says. "There are only so many instruments and they're all oversubscribed as it is."

Liberated when atomic nuclei split, neutrons can probe materials in ways x-rays cannot. Whereas x-rays interact with both the electrons and the atomic nuclei in a sample, uncharged neutrons bounce off only the nuclei, providing a complementary probe of a material's atomic-scale structure. They can also penetrate materials that x-rays cannot, enabling researchers to image the interiors of big objects

such as a running engine or a steel girder. Because neutrons act like little magnets, they can reveal the atomic-scale patterns of magnetism within materials.

Hore uses neutrons to probe the dynamics of polymers, the chainlike molecules in plastics and many biological materials. By replacing hydrogen in a specific part of a molecule with deuterium—which neutrons are more likely to bounce off—researchers can track that bit of the molecule in a jumble of similar material, Hore says. Neutrons also do less damage to delicate samples than x-rays, says Tonya Kuhl, a chemical engineer at UC Davis. "You can just blast biological samples with neutrons and it's not an issue," she says.

The NCNR reactor generates just 20 megawatts of heat—less than 1% as much as a typical power reactor—and is the smallest of the three main neutron sources in the United States. The other two are the 85-megawatt High Flux Isotope Reactor (HFIR) at the Department of Energy's (DOE's) Oak Ridge National Laboratory and the $1.4 billion Spallation Neutron Source (SNS), also at Oak Ridge, which fires protons from an accelerator into a target to blast out pulses of neutrons.

Nevertheless, NCNR's 29 spectrometers, imagers, and other instruments nearly equal the total at HFIR and SNS together, and the lab serves more than 2600 researchers each year. Scientists credit the lab's staff of 194 with its success. "In some odd way, having a lower power reactor has forced them to be more creative," Birgeneau says.

On 3 February, research at NCNR came to an abrupt halt. At 8 a.m. that Wednesday, operators began to restart the reactor after a refueling stop, according to a report NIST submitted to the Nuclear Regulatory Commission (NRC) last month. At 9:07 a.m., the reactor's power plummeted from 15 megawatts to 7 megawatts. Within 1 minute, monitors sensed radiation in the reactor's reinforced concrete confinement building. At 9:09 a.m., automated systems shut down the reactor and sealed the building.

Recognizing that the reactor's uranium fuel may have been damaged, the operator immediately issued an alert, the most urgent alarm a research facility can sound, says Scott Burnell, an NRC spokesperson. Melting fuel is rare at research reactors, says Dale Klein, a nuclear engineer at the University of Texas, Austin, who chaired NRC from 2006 to 2009. "I used to be the director of our research reactor," he says, "and my worst nightmare was having a fuel failure."

Still, the safety systems worked. The 10 staff then in the confinement building received a radiation dose roughly equal to that from a computerized tomography scan, NIST's website says. Only trace amounts of three radioactive isotopes escaped the building, and radiation at the boundary of NIST's 2.34-square-kilometer campus never increased above background levels, according to the website.

NIST traced the accident's roots to a mistake in refueling the reactor a month earlier. Typically, operators replace the oldest four of 30 rodlike fuel elements and rearrange the others. Working by feel, they use a special tool to lock each element in place by twisting it into a spring-loaded latching mechanism, and an inexperienced crew failed to secure one element. Circulating cooling water then lifted it out of position, impeding the flow around it. When the reactor restarted, the element overheated and part of its aluminum cladding melted.



A small research reactor in the Maryland suburbs has been a mainstay of neutron beam studies for a half-century.

PHOTO: RUSH CAPPELLETTI

Staff turnover and institutional culture played a role in the accident, according to the NIST report. In 2011, nine of NCNR's 21 reactor operators had more than 20 years of experience. Now, just three of 22 do. NIST also relied too much on hands-on training and too little on explicit procedures to guide operators, the report says. "We didn't make that transition effectively from a skills-based workforce to a knowledge-based workforce," Dimeo says.

The reactor suffered no damage beyond the overheated fuel element, Dimeo says. Workers have removed all but three elements and now plan to clear debris from the core and purify the deuterated cooling water, he says, which means the reactor cannot restart before April. However, NCNR needs NRC's explicit permission to restart it, and lab officials must convince the commission they have eliminated the causes of the accident, which could take longer, Burnell says. Still, he says, "The agency absolutely understands the importance of the facility, and we are going to do the most prompt and thorough review we can."

A radiation release led to the loss of another neutron source 2 decades ago. In 1996, researchers found tritium in the groundwater near the reactor-based source at DOE's Brookhaven National Laboratory. Although the leak was small and confined to the Brookhaven campus, public outcry led DOE to permanently shutter the reactor in 1999.

NIST hopes openness will forestall a similar outcry. It held a virtual meeting with residents of the suburbs surrounding the lab on 10 February, says Jennifer Huergo, a NIST spokesperson. NIST is also posting its communications with NRC to its website and updating residents by email, Huergo says. "As soon as we get questions, we respond as quickly as we can."

The shutdown comes as the United States has lost the lead in neutron resources. DOE shut down an accelerator-based neutron source at Argonne National Laboratory in 2008 and, in 2015, stopped supporting basic research at an accelerator-based source still running at Los Alamos National Laboratory. "Europe has us beat by a factor of three, Asia has us beat by a factor of two," Kuhl says. "It'd be catastrophic to lose more capabilities."

Even before the accident, researchers fretted about the reactor-based neutron sources. Both NCNR's reactor and 55-year-old HFIR run on "highly enriched" fuel, in which more than 90% of the uranium is the fissile isotope uranium-235. In principle, such fuel could be fashioned into a nuclear bomb. As part of U.S. antiproliferation efforts, both reactors are supposed to be retrofitted to use a fuel containing less than 20% uranium-235 when it becomes available, perhaps in the 2030s. In recent years, expert panels have also advised NIST and DOE to plan to replace the aging reactors.

Planning needs to start now because it will take 10 to 20 years to build a new reactor, Birgeneau says. NCNR's reactor is licensed through 2029, and Dimeo envisions renewing its license even as NIST explores replacing it. "From our perspective," he says, "neutrons are here to stay at NIST." ■



Archaeologist Nicole Boivin no longer directs research at a Max Planck institute.

PHOTO: SVEN DÖRING/LAIF/REDUX

ARCHAEOLOGY

# Max Planck director loses post after probe of misconduct

## Nicole Boivin remains on research staff, but major archaeological institute is left with interim director

*By* **Andrew Curry**

Archaeologist Nicole Boivin has been removed from her position as a director at the Max Planck Institute for the Science of Human History (MPI-SHH) in Jena, Germany, after an internal investigation found evidence of workplace misconduct and bullying, sources inside the institute say.

According to a Max Planck Society (MPG) spokesperson, on 23 October, President Martin Stratmann notified Boivin she was no longer director of archaeological research at MPI-SHH, a premier institute in the study of the human past. Last week, in an all-staff meeting in the institute's airy, glass-walled library, MPG Vice President of Human Sciences Ulman Lindenberger spoke to about 100 employees, group leaders, and postdocs to announce the decision and answer questions about what would happen next, according to people who were present.

Lindenberger could not say whether MPI-SHH would continue its investigations into the human past or shift focus entirely. "That is the subject of deliberations in the governing bodies of the Max Planck Society," spokesperson Christina Beck wrote in an email to *Science*. The assembled staff were told permanent staffers would keep their contracts and those on temporary contracts would be allowed to finish them out.

Several current and former employees contacted by *Science* voiced support for Boivin, but multiple former staffers endorsed the investigation's findings and her demotion.

Boivin remains a member of MPG and will now run a small research group at the institute, separate from the archaeology department, according to an email sent to staffers on 30 October by the institute's acting head, chemical ecologist Jonathan Gershenzon.

Launched in 2014, MPI-SHH was supposed to marry archaeogenetics, archaeology, and linguistics to explore the recent human past. Boivin came on in 2016 as founding director of archaeology, managing a budget in the millions of euros that is one of the field's largest. The department focused on applying multidisciplinary approaches to sometimes overlooked landscapes, exploring how human migration, farming, and cultural practices have altered ecosystems in Africa and Asia. Recent work includes the earliest evidence for milk consumption in Africa and human impacts on ancient tropical forests.

But in 2020, the archaeogenetics and linguistics departments decamped for the Max Planck Institute for Evolutionary Anthropology in nearby Leipzig, in what was rumored to be an acrimonious split between Boivin and the two other MPI-SHH directors, Johannes Krause and Russell Gray. That left Boivin as sole director at MPI-SHH.

Her removal last week followed a commission's lengthy investigation, conducted by Ulrich Sieber, a lawyer and director emeritus of the Max Planck Institute for the Study of Crime, Security and Law, and Silja Vöneky, a law professor at the University of Freiburg, into allegations of bullying and other "violations of good scientific practice," Beck says. Former staffers told *Science* the complaints included taking credit for others' ideas.

Investigators interviewed more than 50 current and former staffers and scrutinized hundreds of email chains, according to people involved. Sieber recommended in January that Boivin be removed from her position, Beck says.

Several researchers say the probe got it right. "The commission was thorough and has documented a really systematic culture of mistreatment," says University of Colorado, Boulder, archaeozoologist William Taylor, who worked in Boivin's department between 2017 and 2019. "The findings accurately reflect a pretty severely dysfunctional environment that was exploiting junior scientists."

In a statement emailed to *Science* last week, Boivin pointed to positive evaluations of the institute by two recent external review boards. "Over the past 5 years I have built up an internationally renowned Department of Archaeology," she wrote.

In an email to staff on 24 October, Boivin called her demotion "abrupt and shocking." But last weekend she agreed to the change, sources at the institute say.

At least three other MPG directors have faced accusations of bullying since 2018. One resigned her directorship, one was removed, and one remains a director.

The society has also contended with accusations of institutional misogyny. Fewer than one-quarter of its directors are women.

Many staffers remained in the dark about the investigation until a few months ago, when Boivin mentioned it in a Zoom call. "For me this came out of the blue," says a researcher who has worked in Boivin's department since 2017 and asked to stay anonymous to protect his relationships in the institute. "I haven't experienced any bullying or harassment … I still don't know exactly what the accusations are, and considering they were so serious they should have been made more public."

What comes next for MPI-SHH remains unclear. It may wind down after some staff serve out 2-year contracts, or one or more new directors may be chosen.

Senior researchers have been meeting with representatives from MPG to make sure the turmoil "does not adversely affect the students and postdocs," says Robert Spengler, an ethnobotanist at the institute. "In the short term we will continue conducting top-notch research—as we always have." ■

With reporting by Ann Gibbons and Gretchen Vogel.

BIOMEDICINE

# RNA drugs custommade for rare diseases face wider test

## With tempered expectations, more scientists try to accelerate "n-of-1" approach

*By* **Jocelyn Kaiser**

Three years ago, neurologist Tim Yu stunned colleagues at the American Society of Human Genetics (ASHG) annual meeting with news that in just 10 months, his team at Boston Children's Hospital had developed, tested, and administered a drug tailored to the unique genetic mutation causing a 7-year-old girl's deadly brain disease. The story of milasen (named for the girl, Mila) raised hopes for more such personalized treatments made from short strands of RNA, known as antisense oligonucleotides (ASOs), that can overcome a gene's flaws.

At this year's ASHG meeting, held online last month, Yu shared a sobering update: Mila died earlier this year—tragic evidence that for neurodegenerative diseases, "it's critical to intervene early," he told a session.

But Yu's group and others are now extending the customized ASO approach, known as "n of 1" because the drug is designed for just one or a few patients. A foundation has formed to provide the bespoke RNAs for free. The U.S. Food and Drug Administration (FDA) is developing guidance for testing, manufacturing, and monitoring the outcomes of the personalized treatments. And at the ASHG meeting, Yu balanced Mila's outcome with a more hopeful tale of another young girl who was given a custom ASO and has seen her disease stabilize for now.

Developing a treatment for a genetic disease is often a multiyear quest. But after sequencing Mila's genome, Yu's team realized she had an extra stretch of DNA in a gene called *CLN7* that disrupted the messenger RNA (mRNA), or protein-building instructions, made from the gene. Yu realized this could potentially be overcome by quickly synthesizing an ASO that would stick to the RNA during its processing into a mature mRNA, hiding the error so that Mila's cells would make protein correctly.

Yu's team raced to get milasen through toxicological testing and FDA approval under a compassionate use protocol for life-threatening conditions. Mila was already blind and could not walk without help when the drug was first injected into her spinal fluid in early 2018. Although it helped her—she had fewer seizures and regained some muscle strength—her condition eventually worsened. But her story led many more families to seek help from Yu, who estimates that 10% to 15% of people with inherited diseases have a genetic error similar to Mila's that can be treated with an ASO customized to match the flaw.

Among them was Ipek Kuzu, born in March 2017 with ataxia-telangiectasia (A-T), a neurodegenerative condition caused by a defect in a DNA repair gene. Ipek began to receive spinal infusions of a tailored ASO in January 2020, when she only had mild symptoms, such as a speech delay. Nearly 2 years later, Ipek has had no worsening of her problems. However, A-T's more severe symptoms don't appear until about age 5. "It's still too early to declare success," Yu said at the meeting.



Ipek Kuzu, age 4, receives a custommade RNA drug for her ataxia-telangiectasia.

PHOTO: MEHMET KUZU

Ipek's father, Mehmet Kuzu, is realistic: If Ipek doesn't need a wheelchair until her late teens, instead of the usual age of 9 or 10, "it will be a very big difference," he says. Other children might also benefit. Yu has come across four A-T children with the same mutation as Ipek. "We are a little bit guilty of a misnomer by calling these drugs individualized," he said in his online talk.

Similarly, the biotech company Ionis Pharmaceuticals in 2019 rushed through regulatory steps to give 26-year-old Jaci Hermstad a custom ASO for her rare inherited early-onset form of amyotrophic lateral sclerosis (ALS). Seriously ill when treatment began, she died in May 2020. However, 10 other ALS patients with the same mutation have started on the ASO, jacifusen, and some have had signs of clinical improvement, says neurologist Neil Shneider of Columbia University, who works with the company.

Ionis launched a 64-patient clinical trial in April to test jacifusen compared with a placebo. Extending a custom ASO's use to other patients helps defray the staggering cost of safety testing the molecule in animals—typically more than $1 million.

Early treatment with a custom ASO may be less crucial for nonprogressive genetic disorders. Rush University Medical Center neurologist Elizabeth Berry-Kravis leads an ASO trial for Angelman syndrome, in which children develop intellectual disabilities. The drug binds to the mRNA for a protein that turns off the paternal copy of gene called *UBE3A* so the gene is activated and makes up for problems with the mother's version.

In all five participants, her team saw striking changes, including acquiring new words and sleeping better, she said at the ASHG meeting. Although younger children improved more, teens made gains, too. "The neurons are all still there, they're just not wired right," Berry-Kravis said.

The trial was halted 1 year ago because inflammation around the injection site caused temporary leg weakness in patients, but it has restarted outside the United States with a lower drug dose. Such side effects are one potential obstacle with ASOs delivered to the spinal cord, retired Ionis CEO Stanley Crooke cautioned at the ASHG meeting.

Still, in January 2020 he created the n-Lorem Foundation to accelerate tailored ASOs for "ultrarare" disease mutations—shared by no more than 30 people, too few to interest most drug companies. It is working with more than 35 families and hopes to treat the first patient, a boy with a developmental disorder caused by an uncommon mutation, early next year.

Academic groups in several countries are also working on customized ASOs. Yu cautions that as researchers move forward with more n-of-1 ASO treatments, they must "avoid raising expectations too quickly." But even if the outcome is uncertain, the risk is worth taking, Mehmet Kuzu says. With his daughter Ipek's disease, "If nothing is done, we know what will happen." ■

U.S. NATIONAL SECURITY

# China Initiative spawns distrust—and activism

## Two surveys reveal fallout from U.S. investigations as Chinese faculty form campus groups

*By* **Jeffrey Mervis**

Scientists of Chinese heritage working in the United States are making the case that the U.S. government's 3-year-old China Initiative and rising anti-Asian violence threaten not only their research, personal safety, and civil rights, but also the academic research enterprise.

Two recent surveys document what had previously been anecdotal evidence of the impact of the China Initiative, which has led to federal investigations of hundreds of U.S. researchers and criminal charges against more than two dozen, mostly of Chinese origin. Several scientists have been convicted of failing to disclose foreign funding ties and defrauding the government.

The surveys find that disturbingly large percentages of respondents have abandoned scientific collaborations with their counterparts in China, thought twice about applying for federal funding, and said no to Chinese graduate students and visiting scholars asking to work in their labs (see chart, right). Significant numbers also fear being monitored by government agencies and are weighing leaving the United States. One survey polled members of the University of Michigan Association of Chinese Professors (UM-ACP); the other examined a national sample of both Chinese and non-Chinese academic scientists.

This week, UM-ACP and similar groups at a dozen other U.S. universities wrote to President Joe Biden asking him to end the initiative, compensate those whose careers have been derailed by unproven allegations, and narrow the scope of future investigations to focus on genuine threats to national security.

Speaking out doesn't come naturally to many members of the groups writing to Biden, its leaders say. "Chinese American scientists traditionally don't do that, but the time has come," says Duxin Sun, a professor of pharmacy and president of UM-ACP.

In July, UM-ACP polled its 370 members, all UM faculty. Nearly 20% of the 123 respondents said they had been the subject of an investigation connected to research collaborations in China. They were contacted by university officers in 15 cases, federal agents in nine cases, and a federal research agency in one case. (Some faculty faced scrutiny from more than one entity.) To preserve anonymity, the survey didn't ask for details about the investigations.

The survey also revealed a broader sense of angst among Chinese faculty at UM. Nearly two-thirds said they "don't feel safe" as a Chinese academic for a variety of reasons, and 42% feel their university would not defend them if they were the target of a federal investigation. Some 71% believe their concerns "are largely invisible" to university administrators, and 89% believe Asian Americans are underrepresented in top-level posts. "I'm grateful to the UM-ACP for identifying faculty concerns about policies, practices, and climate that impact their success at UM and across the country," says Susan Collins, UM's provost, adding that "the strength of the university rests on the excellence of our faculty."

A second survey tallied responses from nearly 2000 U.S. scientists at 83 of the country's largest research universities. Almost half were of Chinese heritage. Sponsored by the Committee of 100 (C-100), a group of prominent Chinese Americans from all sectors, it found that Chinese researchers are four times more likely to fear government surveillance than their non-Chinese colleagues and to have canceled projects with researchers in China for fear those activities could make them a target. "We anticipated some evidence of [Chinese scientists feeling subject to] racial profiling, but not how stark these differences would be," says the study's lead author, education researcher Jenny Lee of the University of Arizona.

C-100 President Zhengyu Huang, an investment banker and philanthropist, notes that discrimination against Asians has deep roots in U.S. society. "Historically, the government has played an explicit role in upholding racism and racial profiling," Huang asserts. "And even when it is implicit, the effect is the same: Lives are ruined."

Critics of the China Initiative believe terminating it would help end that ugly legacy. In recent weeks, faculty at prominent U.S. universities, including Stanford University; the University of California, Berkeley; UM; and Princeton University, have called on Biden to find "an alternative response to the challenges posed by" U.S.-Chinese relations. This week's joint letter from the collection of Chinese academic organizations supports that move, but also emphasizes the human toll the initiative has taken on its members, their families, and the larger scientific community.

The emergence of these groups represents a reluctant coming of age for Chinese American scientists like Rongwei Yang. "I personally have little interest in becoming political," says Yang, a mathematician who leads the 130-member University at Albany Association of Chinese Professionals, formed in May. "However, like many of my peers, I worry that the China Initiative may inadvertently harm the U.S. academic enterprise in the long run. So what we are doing is more than safeguarding our own rights." ■

### Feeling threatened

At the University of Michigan (UM), 64% of 123 Chinese faculty responding to a survey said the China Initiative and U.S.-Chinese tensions have made them feel unsafe; 33% said the climate has caused them to consider not applying for U.S. grants. A survey of nearly 2000 scientists at 83 large universities found Chinese and non-Chinese scientists respond differently to issues raised by the China Initiative.




CREDITS: (GRAPHIC) K. FRANKLIN/*SCIENCE*; (DATA) J. LEE/UNIVERSITY OF ARIZONA AND X. LI/COMMITTEE OF 100, RACIAL PROFILING AMONG SCIENTISTS OF CHINESE DESCENT AND CONSEQUENCES FOR THE U.S. SCIENTIFIC COMMUNITY; A. LIN AND D. SUN/UNIVERSITY OF MICHIGAN ASSOCIATION OF CHINESE PROFESSORS

# A GREENER PATH

China is making its Belt and Road Initiative more environmentally friendly. The massive infrastructure program could still cause ecological devastation

*By* **Dennis Normile**




The Belt and Road Initiative provides many developing countries with much-needed infrastructure, including this highway under construction in Sri Lanka.

PHOTO: PAULA BRONSTEIN/GETTY IMAGES



The 40 wind turbines dotting the rolling grassland 9 kilometers south of Zhanatas, an impoverished industrial town in southern Kazakhstan, are monuments to change. Each as tall as a 50-story building, the towers dominate the arid landscape. Completed in June, the $160 million Zhanatas Wind Farm has a capacity of 100 megawatts (MW), making it one of the most powerful wind farms in central Asia. It's a milestone in Kazakhstan's quest to boost reliance on renewables and achieve carbon neutrality by 2060. The country has 300 years' worth of coal, but isn't building a single new plant to burn it.

The project stands out for another reason as well: It is one of the biggest renewable energy projects built in the region under China's Belt and Road Initiative (BRI), a colossal infrastructure plan that has alarmed environmental advocates.

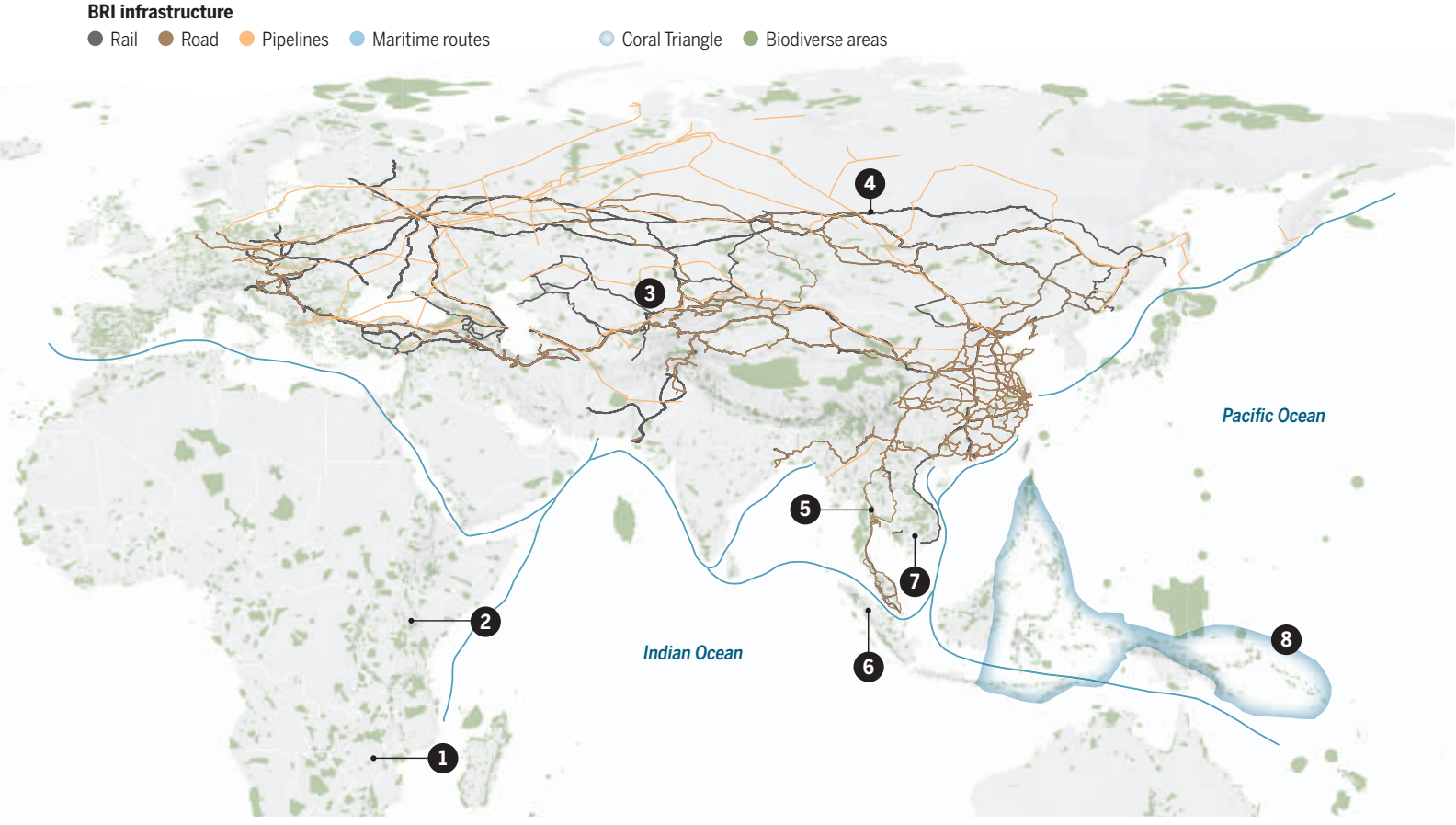
Launched in 2013, BRI links China to markets and sources of raw materials around the world while stimulating economic growth in developing countries. It is providing much-needed power plants, roads, ports, and railways in Asia, Africa, and Latin America. But the "infrastructure tsunami" also threatens to "open a Pandora's box of environmental crises, including large-scale deforestation, habitat fragmentation, wildlife poaching, water pollution and greenhouse gas emissions," ecologist William Laurance of James Cook University, Cairns, wrote in *The Conversation* in 2017.

China countered such criticism by putting an environmentally friendly face on BRI. In June 2017, its Ministry of Ecology and Environment issued "Guidelines on Promoting Green Belt and Road," and President Xi Jinping himself said in April 2019 that BRI "aims to promote green development." That same month, China established a BRI International Green Development Coalition, which brings together government agencies, nongovernmental organizations, and think tanks to study ways to make BRI investments environmentally sustainable.

Two years after Xi's pronouncement, the initiative has a greener cast. Chinese-funded coal power is in retreat, as environmental concerns dovetail with economic forces. Of the 52 coal-fired power plants with Chinese backing announced in BRI countries between 2014 and 2020, only one has gone into operation. Another seven are under construction and 11 are in planning; the remaining 33 have been shelved or canceled, according to a June report by Christoph Nedopil, a development economist at Fudan University in Shanghai. Confirming

## A mixed picture

The "infrastructure tsunami" planned under China's Belt and Road Initiative (BRI) poses risks to biodiversity hot spots throughout Asia, Europe, and Africa. But investments in coal are increasingly giving way to renewable energy.




**1 Coal on the way out?**
A planned coal power plant in **Zimbabwe** is off the table. All told, 33 of 52 Chinese-financed coal plants announced between 2014 and 2020 have been shelved or canceled.

**2 Here comes the Sun**
A Chinese contractor completed **Kenya's** first solar power station in 2020, 1 year after a major Chinese bank scrapped plans to finance a coal power plant in the country.

**3 Windmills on the steppe**
The 100-megawatt **Zhanatas Wind Farm** in southern **Kazakhstan**, completed in June, is an example of BRI's recent investment in renewable energy.

**4 Migratory hurdles**
East-west roads and rail lines across Eurasia threaten to block the north-south migrations of saiga antelopes, Mongolian gazelles, and kulans, a species of wild ass.

**5 Ecosystems at risk**
BRI road and rail links encroach on two major protected areas in **Thailand** and one in Laos. All host endangered species.

**6 Orangutans threatened**
The **Batang Toru hydropower project** on **northern Sumatra** in **Indonesia** threatens the Tapanuli orangutan, the world's rarest ape. The Bank of China has withdrawn its financial support.

**7 Great river dammed**
Studies suggest **Cambodia's Lower Sesan 2 Dam** will have a major impact on aquatic diversity and fisheries in the **Mekong**. More than two dozen other Mekong basin dams are planned under BRI.

**8 Peril at sea**
Expanded shipping lanes through the **Coral Triangle**, a global center of marine biodiversity, threaten reefs, mangroves, and seagrass beds.

CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) KEY BIODIVERSITY AREAS PARTNERSHIP; CORAL TRIANGLE ATLAS; A. HUGHES, *CONSERVATION BIOLOGY*, 33:4, 883 (2019)

what appeared to be de facto policy, Xi told the United Nations General Assembly on 21 September that China would stop building coal-fired plants abroad altogether. (Coal power continues to boom at home, however.)

The shift away from coal "occurred more rapidly than many observers, including me, expected; this is cause for optimism about the greening of BRI projects in the developing world," says Tyler Harlan, a geographer at Loyola Marymount University in Los Angeles.

Meanwhile, renewable energy projects like Zhanatas are burgeoning. Solar and wind power are increasingly cost competitive, and like other developing countries, "Kazakhstan is aiming to decarbonize to preserve the environment," says Kuanysh Baltabayev, executive director of EcoJer, an association of regional environmental initiatives in Kazakhstan. Chinese companies are studying two more renewable projects there, one wind and one solar. And the shift toward renewables is seen throughout the BRI countries. Total BRI investments in solar, wind, and hydroelectric power surpassed spending on fossil fuel plants in 2020, Nedopil reported.

But Harlan and others caution that carbon dioxide is only one part of BRI's environmental cost. The initiative includes the construction of hundreds of dams, some of which fragment aquatic communities and block migratory fish species, Laurance says. And "there are massive road-building projects in the works" under BRI, adds ecologist Alex Lechner of Monash University, Indonesia. These damage local ecologies and pose a major risk to global biodiversity, a study published in 2020 showed, while also increasing emissions and reducing the extent of forests, which

act as carbon sinks. Although many countries require environmental impact assessments before construction begins, they are often not public and rarely up to international standards.

BRI's growing investment in renewable power "is certainly laudable," Laurance concludes. But the constellation of projects still includes "an awful lot that will be environmentally destructive."

**DEFINING BRI CAN** be difficult. Xi first outlined the Silk Road Economic Belt and 21st-Century Maritime Silk Road Development Strategy during visits to Kazakhstan and Indonesia in the fall of 2013, 6 months after becoming president. Within months, Chinese officials were describing a "One Belt, One Road" plan—later renamed BRI—for overland roads and railways and shipping lanes connecting ports in China and elsewhere in Asia to Europe and East Africa. It envisioned linking some 60 countries, home to half the world's population.

Soon, the BRI label was also used for pipelines, power plants, cement factories, and industrial facilities backed by Chinese private and government lenders anywhere in the world. There are even proposed projects in the Arctic (the Polar Silk Road), and a space and cyberspace initiative (the Belt and Road Space Information Corridor). More than 130 countries have signed BRI memoranda of understanding with China, which pledge cooperation on infrastructure and other projects, according to the Belt and Road Portal website.

Analysts typically consider a project part of BRI if it has Chinese backing and is located in a country that has signed a BRI memorandum. Still, "rather than being a coherent, geopolitically-driven grand strategy, BRI is an extremely loose, indeterminate scheme, driven primarily by competing domestic interests," Lee Jones, of Queen Mary University of London, and Jinghan Zeng, of Lancaster University, wrote in *Third World Quarterly* in 2019. "The BRI has become an all-encompassing slogan that domestic actors [use to] brand their projects," Zeng added in an email to *Science*. That makes its cost elusive; estimates run anywhere from $1 trillion to $8 trillion, though many consider the upper sum unrealistic.

The initiative's amorphous character also means "reliable, accurate maps of the BRI are almost nonexistent," Alice Hughes, a conservation biologist at the Xishuangbanna Tropical Botanical Garden of the Chinese Academy of Sciences, wrote in a 2019 *Conservation Biology* paper. Hughes filled the gap herself by producing maps, now widely cited, of BRI road and rail routes (see map, p. 678).

Western countries, wary of China's increasing influence, have warned that some projects are debt traps. Tiny Montenegro, in the Balkans, owes $1 billion for an unfinished highway that the country cannot afford and many say it does not need, for example. Chinese lenders could seize land in Montenegro that was put up as collateral. Yet many developing countries do need the infrastructure. To sustain growth, eradicate poverty, and respond to climate change, Asian countries—including China—will need to invest $26 trillion in infrastructure between 2016 and 2030, or $1.7 trillion per year, according to a 2017 study by the Asian Development Bank. With regional spending in 25 Asian countries totaling only $881 billion in 2015, BRI is partly filling the gap.



A Belt and Road Initiative project upgraded the port of Mombasa, Kenya, to handle increasing trade.

PHOTO: LUIS TATO/BLOOMBERG/GETTY IMAGES

At the plan's outset, there was little concern for the environment, and China's large banks and engineering firms rushed to expand their coal businesses. In 2015, more than 46% of BRI investment went into coal plants, according to Nedopil's report. By the end of the decade, "people started to realize there were a lot of problematic projects in the pipeline," says Li Shuo, a climate policy adviser to Greenpeace East Asia. Decreasing renewable electricity costs and pressure to cut emissions dampened coal's allure.

For example, the state-owned Industrial and Commercial Bank of China (ICBC) has apparently lost interest in a $1.2 billion, 1050-MW coal-fired plant under construction near the town of Lamu, Kenya, after a group named deCOALonize won a court order halting work on the plant in 2019. (Meanwhile, a Chinese contractor completed Kenya's first solar power station, a 50-MW, $130 million installation northeast of Nairobi, in 2020.) ICBC also recently withdrew from a $3 billion, 2800-MW coal-fired plant in Zimbabwe, according to news reports, and Egypt and Bangladesh both shelved plans for new coal plants financed by China.

The course change is essential to limiting global warming. If BRI countries followed historic patterns of fossil fuel use rising in tandem with economic growth, their emissions "would be enough to induce nearly 3 degrees of warming even if the rest of the world takes 2-degree compliant action," warned a 2019 study by economist Ma Jun of Tsinghua University in Beijing and Simon Zadek of the Finance for Biodiversity Initiative, a U.K. nongovernmental organization.

China's about-face won't be enough, however: Despite the country's image as the world's foremost promoter of coal, Japanese and Western private institutions financed 87% of new coal power capacity outside China that went online or into development between 2013 and mid-2019, according to a July report from Xinyue Ma and Kevin Gallagher of the Global Development Policy Center at Boston University. The pair called on the G-20 grouping of the world's largest economies "to commit to limiting all overseas fossil fuel financing, starting with overseas coal finance." Climate advocates are also seeking incentives for countries to retire existing coal power plants early so they aren't spewing carbon dioxide for another 4 to 5 decades.

**BUT BRI'S MAIN** alternative for coal, hydropower, has its own problems. Dams disturb rivers' natural flows and can degrade riverine ecosystems and block migratory fish from spawning and feeding grounds. And dam construction often requires access roads later used for illegal logging and wildlife poaching.

Most of BRI's renewable spending has gone into hydroelectric power: It received 35% of total energy sector investments in 2020, versus 23% for solar and wind. (Overall energy investments have gone down in recent years [see graphic, p. 680].) Spending is skewed in part because the solar and wind industries are mostly private. China's major



The 55-megawatt Garissa Solar Power Station, built northeast of Nairobi, Kenya, exemplifies the Belt and Road Initiative's shift in focus, from coal to renewable power.

dam builders, in contrast, are state-owned enterprises supported by state-owned banks, which puts them in a better position to win financial and political support, Harlan notes.

BRI's hydropower investments are already having a disastrous impact on the Mekong River Basin—the world's most biodiverse river system after the Amazon. Rising in China, the Mekong runs through five countries before reaching the South China Sea. It supports the world's largest inland fishery, providing a significant source of protein for the 65 million people living in the 800,000-square-kilometer basin.

The river system already has more than 400 dams, 126 of them in China itself, that "reduce the mightiness of the Mekong's natural flow," says Brian Eyler, a water security expert at the Stimson Center in Washington, D.C. The dams suppress the Mekong's unique wet season flooding, which is key to fishery productivity and the riverside agriculture that tens of millions rely on.

Environmental advocates point to Cambodia's Lower Sesan 2 Dam as an example of the influence a single hydropower plant can have. The contract to build the 400-MW dam, Cambodia's largest, was signed in 2012, but was later brought under the BRI umbrella. Completed in 2018, the dam is located just downstream from the confluence of the Sesan and Srepok rivers, two Mekong tributaries that are key migratory pathways for hundreds of fish species, Eyler says. Seventy-five meters tall, it stretches 6 kilometers across a wide valley to block the meandering, shallow river. The flooding it caused displaced 5000 mostly Indigenous peoples and ethnic minorities from six villages.

A 2012 study in the *Proceedings of the National Academy of Sciences* by Guy Ziv, now at the University of Leeds, and colleagues warned that the dam could cut the entire Mekong's fish population by 9.3%—"a huge reduction caused by a single dam," Eyler says. A 2008 impact assessment commissioned by Vietnam Electricity estimated the dam would shrink the annual incomes of 300,000 fishers and farmers living near the river by millions of dollars.

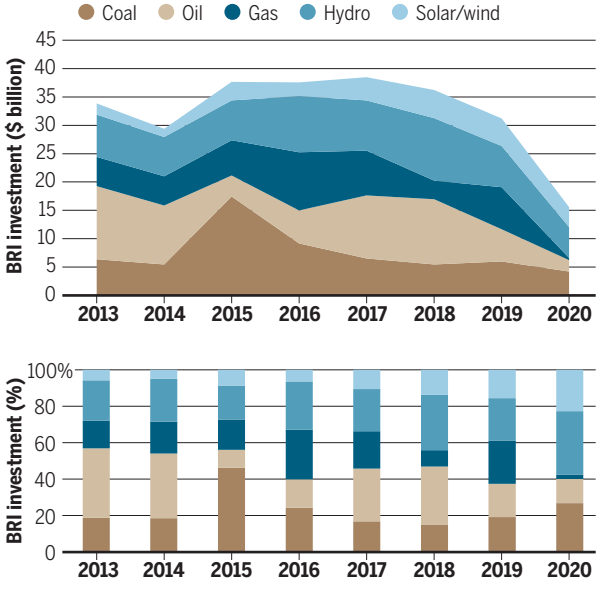
Peer-reviewed studies of the dam's impact have yet to appear, but fishers have reported that catches in 2019 and 2020 were down 80% compared with previous years, Eyler says. And a 10 August report by Human Rights Watch says those affected have not received compensation.

Chinese companies are full or part owners of another big dam already built in Cambodia and 18 more in neighboring Laos. And more are on the way: Chinese interests have eight Mekong dams under construction and another 34 under study for locations scattered throughout the watershed. (Developers and investors from all of the Mekong basin countries as well as from the United States, South Korea, and Japan are also planning Mekong dams.)

But just as with coal-fired power plants, the financial calculus is changing. Eyler says the Mekong dams were envisioned 20 years ago when many saw hydroelectric power as a clean alternative and solar and wind power were still costly. Now, some dam developers are having trouble finding buyers for their power because other types of renewable energy are becoming cheaper. This makes arranging financing difficult. "You can't get money from the bank if you don't have a power purchase agreement," Eyler says. In March 2020, poor economic prospects led Cambodia to delay

## Renewables up, coal down

Hydroelectric, solar, and wind power make up a growing share of the Belt and Road Initiative's (BRI's) energy investments. The pandemic dampened all BRI activity in 2020.




CREDITS: (PHOTO) XIE HAN/XINHUA/ALAMY LIVE NEWS; (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) C. NEDOPIL WANG, *COAL PHASE-OUT IN THE BELT AND ROAD INITIATIVE (BRI): AN ANALYSIS OF CHINESE-BACKED COAL POWER FROM 2014-2020*, GREEN BRI CENTER, INTERNATIONAL INSTITUTE OF GREEN FINANCE, BEIJING (2021)

construction of two planned mainstream Mekong dams for 10 years. (A month later, a 60-MW solar farm owned by a Chinese-Cambodian joint venture was connected to the national grid.)

And in March 2019, the Bank of China posted a statement on its website promising to "very carefully" evaluate the impact of the $1.68 billion Batang Toru hydropower project on northern Sumatra in Indonesia, which it planned to finance. That project alarmed conservationists because it could destroy the last remaining habitat of the Tapanuli orangutan, a frizzy-haired species considered the world's rarest ape (*Science*, 13 September 2019, p. 1064). The bank withdrew from the project in summer 2020, according to government reports. Batang Toru may go ahead without Chinese support, but the planned start in 2022 may be delayed until 2025.

By 2030, megadams "will likely be obsolete," Eyler concludes. But unlike coal-fired power plants, existing dams are unlikely to be retired early. The only course of action "is to replace future dams with solar and wind," he says.

**ROADS, RAILWAYS**, and pipelines—linear infrastructure—may pose BRI's biggest threat to ecological systems and biodiversity. These ribbons of development stretching for hundreds or thousands of kilometers cause habitat loss and fragmentation, sometimes isolating wildlife populations in pockets too small to survive. They can disrupt migratory routes, cause roadkill, and open pristine areas to illegal logging and poaching. They have secondary impacts as well, such as the construction of concrete plants and quarries for limestone, sand, and gravel, causing further ecological destruction.

Although many routes are under construction, few have been completed, so conservationists are still rushing to assess the threats. A team led by Lechner mapped planned new BRI roads and railways, along with upgrades to existing routes, through Southeast Asia, assessing overlaps with protected areas, key biodiversity areas, and the ranges of threatened species. The team also looked at shipping routes that might benefit from BRI's new or expanded ports throughout the Coral Triangle, a hot spot for marine biodiversity in the waters surrounding Philippines, Malaysia, Indonesia, and Papua New Guinea. For each, they assessed proximity to marine protected areas and reefs, mangroves, and seagrass beds.

They found that land and marine routes now in planning or under construction will cut directly through 21 protected areas and encroach on the 25-kilometer buffer zones of nearly 30% of the region's protected and key biodiversity areas. In Thailand, for example, BRI routes encroach on the Dong Phayayen–Khao Yai Forest Complex—a UNESCO World Heritage site home to 800 animal species—and the Khao Sok National Park–Khlong Saeng Wildlife Sanctuary, which hosts a relic of the oldest remaining rainforest in the world. In Laos, BRI threatens a refuge for endangered species of cats, deer, gibbons, and bears. Other BRI routes pose a risk to such iconic animals as the large-antlered muntjac, the Malayan tapir, the white-handed gibbon, the Sumatra serow, and the critically endangered Edwards's pheasant.



A hydropower project on northern Sumatra in Indonesia threatens the endangered Tapanuli orangutan.

PHOTO: ANDREW WALMSLEY/ALAMY STOCK PHOTO

"In Southeast Asia, BRI is increasingly seen as a threat to the already imperiled biodiversity," the team wrote in an August 2020 paper in *Biological Conservation*.

The same is true elsewhere, according to Hughes's *Conservation Biology* paper. She looked at 170,126 kilometers of roads and 80,451 kilometers of railways planned to cross Eurasia and Africa and plotted their proximity to key biodiversity and protected areas. "The BRI portends a new and significant threat to biodiversity globally," she concluded.

There are glimmers of hope for mitigating the devastation, Hughes found. Carefully planned forest planting, for example, could connect forest patches isolated by new construction or extend ecological corridors. Such strategies would create new carbon sinks as well. Identifying animal migration routes could pinpoint the best places to build fences to keep animals off highways and forested overpasses so wildlife can cross them. Seeking to burnish its environmental credentials, China has recently advocated steps to protect biodiversity at home (*Science*, 15 October, p. 249). Conservationists hope it will also foster a network of protected areas and wildlife corridors across Eurasia to limit BRI's damage.

But any effective mitigation must rely on rigorous environmental impact assessments done with international oversight, conservationists say. BRI projects have tended to follow the assessment requirements of the host countries, which vary greatly depending on each country's economic development and capacity to study and protect its biodiversity. (Impact assessments for dams are often limited to effects in the dam's immediate vicinity and rarely consider the cumulative impacts of multiple dams, a January review in the journal *Water* found.) Whether commissioned by the developers or the host country government, environmental impact assessments are not always made public, and some are little more than fig leaves, biodiversity advocates say.

Banks could help, Hughes says. "Many Chinese and Asian banks are starting to consider green finance and developing green financial policies," she notes. They could require comprehensive impact assessments before opening their coffers. But time is short. "Many of the routes are under construction now, with opening dates set between 2022 and 2024," Hughes warns.

The window of opportunity for further greening BRI is closing. ■

With reporting by Bian Huihui and Aiya Kuchukova.



PHOTO: WOLFGANG RATTAY/REUTERS

VIEWPOINT: COVID-19

# Children and COVID-19 in schools

## The benefits of in-person schooling with mitigations in place outweigh the risks of COVID-19 for children

*By* **Shamez N. Ladhani**[1,2]
*and* **the sKIDs Investigation Team**

Children have a very low risk of severe or fatal COVID-19 but, early in the pandemic, uncertainties about their role in virus transmission led most countries to close educational settings as part of national lockdowns to control the spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In many countries, including parts of North America and the Asia-Pacific regions, schools remained closed to in-person teaching for more than a year. School closures affect not only the education of children, but also their social and emotional well-being, and limit access to social and welfare services, school meals, and school-based immunizations. Therefore, understanding the role of children in SARS-CoV-2 infection and transmission, and identifying effective interventions to mitigate risks, are critical for keeping schools safe for staff, students, and their families.

Reports early in the pandemic in 2020 indicated that children represented only a tiny fraction of total cases, hospitalizations, and deaths due to COVID-19 and invariably developed mild, transient, and self-limiting illnesses. Case fatality rates are strongly associated with age, ranging from 0.1% in <20-year-olds but remaining below 1% among <50-year-olds and then rising exponentially to 35% among ≥80-year-olds (*1*). Moreover, household contact studies

A teacher prepares COVID-19 rapid tests for students at a secondary school. These are one of several mitigation measures to ensure in-person schooling is safe.

suggested that children were less likely to become infected or transmit SARS-CoV-2 to others. Many of these studies, however, involved polymerase chain reaction (PCR) testing of symptomatic contacts and therefore missed asymptomatic and mild, transient infections, which are more common in children than adults. Nonetheless, educational settings were closed in many countries as part of national lockdown mitigation strategies, and the net effect was a rapid decline in cases across all age groups, although the specific contribution of school closures compared to all the other lockdown measures implemented at the same time was difficult to estimate.

As community cases fell during national lockdowns, discussions about reopening educational settings were highly divisive among parents, educationalists, clinicians, politicians, and policy-makers. Some countries such as Iceland and Sweden had successfully kept their preschools (age 1 to 6 years) and schools (age 7 to 16 years) open throughout the pandemic. The UK, like many other countries, opted to partially reopen preschools and some primary and secondary school (age 5 to 16 years) for in-person teaching from 1 June 2020. Strict physical distancing and infection control measures were implemented, including smaller class sizes and clustering staff and students into bubbles, with face coverings recommended for staff but not for students. National surveillance conducted by Public Health England (PHE) identified very few infections and outbreaks during the 6 weeks until mid-July 2020, with most cases affecting staff and limited transmission among students (*2*). Notably, the risk of SARS-CoV-2 infection and outbreaks in educational settings correlates strongly with local community infection rates (*2*, *3*), which is expected because higher community rates allow more opportunities for virus introductions into schools.

In June 2020, PHE also initiated active surveillance in primary schools where staff and students provided blood samples to test for SARS-CoV-2 antibodies as a robust marker of prior infection, irrespective of symptom status or duration of infection (*4*). SARS-CoV-2 antibody positivity rates were similar in students and staff and comparable to local community seroprevalence, which is consistent with similar school serosurveillance studies performed in other high-income countries. This is important because it indicated that children were as likely to be infected and develop antibody responses against SARS-CoV-2 as adults. Additionally, only 14% and 31% of seropositive primary and secondary school students, respectively, reported COVID-19 symptoms compared to 53% and 77% of staff, respectively, confirming the higher proportion of asymptomatic infections with younger age (*4*). After 6 weeks at school, only 4 out of 603 (0.7%) 4- to 12-year-old students and 1 out of 1015 (0.1%) staff sampled in schools across England seroconverted, further confirming very low risk of infection following partial reopening of primary schools, albeit when community infection rates were also low.

As occurred in many countries, the decision to fully reopen schools in late 2020 led to further concerns about the difficulties in implementing mitigations and physical distancing with full attendance in classrooms with poor ventilation and lack of air quality monitoring, using carbon dioxide detectors, for example. Although the low risk of severe COVID-19 in children was well-recognized, there were increasing reports of pediatric multisystem inflammatory syndrome (MIS-C), a rare immune complication in ~0.03% of ≤20-year-olds that arises 2 to 4 weeks after SARS-CoV-2 exposure. There was also increasing recognition of Long Covid which, although lacking a formal case definition, comprises persistent, intermittent, relapsing, and new-onset symptoms lasting several weeks to months after SARS-CoV-2 infection. Additionally, the personal risk, due to age and comorbidities, of education staff being infected was a concern.

The UK experience of full reopening of in-person school teaching was similar to those of other countries, whereby cases in students increased but lagged behind and did not exceed local community infection rates in adults (*3*) (supplementary fig. S1). In November 2020, however, the UK imposed a second national lockdown because of increasing cases, but schools remained open. Cases declined first in adults and then in children who continued to attend full-time education. From late November, there was a rapid increase in cases due to the more transmissible Alpha (B.1.1.7) variant. This was first observed among secondary school–aged children because they were still attending school, whereas adults remained in lockdown.

One of the main concerns with the full reopening of schools was the risk of in-school transmissions leading to large outbreaks, with onward transmission to household members and potentially into the wider community. In England, PHE surveillance found that most schools did not experience a COVID-19 outbreak. When they did occur, more than half involved only two cases before they were controlled, most likely because of mitigation measures: frequent sanitization, physical distancing, class bubbles, staff masking, and rapid testing with lateral flow devices and isolation of cases and their bubble contacts. PHE surveillance during September–October 2020 identified more frequent outbreaks in secondary schools (10%) compared to primary schools (2%) or nurseries (0.3%) (*5*). Although data were collected over different periods, infection rates in nursery staff (24.3%) during November 2020 were significantly higher than in primary (9.8%) or secondary school (4.0%) teaching staff during September–October 2020. By contrast, infection rates in children across the three educational settings remained low (2.3, 0.8, 1.2%, respectively), especially when the index case was also a child (*5*). The higher infection rates in nurseries may be explained by difficulties in maintaining physical distancing owing to smaller premises and constant adult care needed by infants and toddlers.

These investigations, however, did not provide sufficient granularity to distinguish new introductions of the virus from the community compared with in-school transmission. But recent US studies with active contact tracing found that very low rates of secondary infections occurred inside educational premises in K–12 (5 to 18 years) schools. In North Carolina, surveillance of 11 school districts with >90,000 students and staff attending in-person schooling for 9 weeks found 773 community-acquired SARS-CoV-2 infections, whereas contact tracing found only 32 additional infections acquired in school (*6*). Among 17 rural Wisconsin schools, COVID-19 incidence among 4876 students and 654 staff during 31 August–29 November 2020 was lower (3453 out of 100,000) than in the county overall (5466 out of 100,000) (*7*). Of the 191 cases in students and staff, only seven (3.7%) were linked to in-school spread (*7*). In New York City, COVID-19 prevalence in public schools was similar to or less than community prevalence during 9 October–18 December 2020 (*8*). Additionally, of 36,423 school-based close contacts, only 191 (0.5%) subsequently tested positive for SARS-CoV-2, and the likely index case was an adult for 78.0% of secondary cases (*8*). Taken together, despite the inevitable biases and limitations of such studies, the risk of acquiring SARS-CoV-2 in school remains very low when appropriate mitigation measures are in place, although recom-

[1]Immunisation and Countermeasures Division, Public Health England, London, UK. [2]Paediatric Infectious Diseases Research Group, St. George's University of London, London, UK. The members of the sKIDs Investigation Team are provided in the supplementary materials. Email: shamez.ladhani@phe.gov.uk

mendations for face coverings or masks for students have varied from none in the UK to being mandated from 2 years of age in the US. In England, community infection rates in adults were generally higher than rates in children, except when adults were in national lockdown while children continued in-person schooling in November–December 2020 and March–April 2021.

A key question about school safety, however, relates to asymptomatic infections and their role in transmission. In adult institutional settings, especially care homes, by the time an outbreak was identified, a high proportion of residents and staff had already been infected, with 30 to 60% infected by the end of the outbreak. Reassuringly, antibody seroprevalence studies in schools have consistently reported similar or lower seroconversion and seropositivity rates in students and staff compared to local community rates, most likely contributed by the mitigations in educational settings (*4*, *9*, *10*). Reasons for lower transmission by children compared to adults remain uncertain but include higher rates of asymptomatic infection, smaller lung capacity, and shorter duration of viral shedding than in adults.

That infection and outbreaks occur in educational settings is undeniable and likely reflects a combination of factors, including local community infection rates and inadequate in-school mitigations. Additionally, where wider testing has been performed, transmission to household members of staff and students can also occur (*11*), highlighting the importance of early and rapid contact tracing and isolation to control virus transmission. The number of mitigation measures, rather than individual measures, are important in preventing in-school transmission (*12*). Although the risk of severe COVID-19 in children and their households is low (*13*, *14*), concerns remain about Long Covid in children, which can be debilitating. The risk of Long Covid has been difficult to estimate because of the wide range of symptoms and varied surveillance methodology, but ranges between 2 and 14% in children compared to 20 to 30% in adults.

The emergence of more transmissible variants of concern, first Alpha and now Delta (B.1.617.2), poses additional challenges for educational settings. Increasing infection rates in unvaccinated populations, including children, result in widespread educational disruption because of cases and class bubbles self-isolating at home. In the UK, >10% of students were absent from school in July 2021 when the Delta variant was circulating. Notably, although the number of children hospitalized with COVID-19 increased with the Alpha and Delta waves, hospitalization rates in children with confirmed COVID-19 remained stable, suggesting that these variants are not associated with more severe disease in children. The recent licensure of highly effective vaccines, however, is already changing the landscape of the pandemic and should allow schools to reopen safely worldwide. Most high-income countries are now offering COVID-19 vaccines to children aged 12 years and older even though their personal risk of severe disease remains very low. Vaccinating adults, however, will protect staff and household members against COVID-19, and the attendant lower community infection rates should confer indirect (herd) protection that will help reduce infection risk in children (*15*).

In their recent updates on the role of schools in transmission and COVID-19 in children, the European Centre for Disease Prevention and Control (ECDC) and US Centers for Disease Control and Prevention (CDC) have reported that, when adequate mitigation measures are in place, transmission in schools is usually similar to or lower than community transmission. Therefore, although school closures may contribute to reducing transmission, by themselves, they would be inadequate in preventing community transmission and, consequently, the benefits of in-person schooling outweigh the risks, especially in countries where adults are fully vaccinated against COVID-19. The focus now must be on assessing and implementing evidence-based mitigations to reduce the risk of infection in schools not only to protect students and staff but also to minimize disruption to education and access to additional services. ■

## REFERENCES AND NOTES


1. C. Signorelli, A. Odone, *Int. J. Public Health* **65**, 1435 (2020).
2. S. Ismail *et al.*, *Lancet Infect. Dis.* **S1473**, 30882 (2020).
3. A. A. Mensah *et al.*, *J. Infect.* **82**, 67 (2021).
4. S. N. Ladhani *et al.*, *Lancet Child Adolesc. Health* (2021).
5. F. Aiano *et al.*, *Lancet Europe* 10.1016/j.lanepe.2021.100120 (2021).
6. K. O. Zimmerman *et al.*, *Pediatrics* **147**, e2020048090 (2021).
7. A. Falk *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **70**, 136 (2021).
8. J. K. Varma *et al.*, *Pediatrics* **147**, e2021050605 (2021).
9. S. N. Ladhani *et al.*, *EClinicalMedicine* **37**, 100948 (2021).
10. Office for National Statistics, https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/covid19schoolsinfectionsurveyengland/round5englandmay2021
11. M. Okarska-Napierała, J. Mańdziuk, E. Kuchar, *Emerg. Infect. Dis.* **27**, 317 (2021).
12. J. Lessler *et al.*, *Science* **372**, 1092 (2021).
13. H. Forbes *et al.*, *BMJ* **372**, 628 (2021).
14. R. Wood *et al.*, *Arch. Dis. Child.* archdischild-2021-321604 (2021).
15. R. J. Harris *et al.*, *N. Engl. J. Med.* **385**, 759 (2021).


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj2042

10.1126/science.abj2042

IMMUNOLOGY

# A helminth-induced antimicrobial protein

## Small proline-rich protein 2A regulates the intestinal microbiota

*By* **Nicola Harris** *and* **Lakshanie Wickramasinghe**

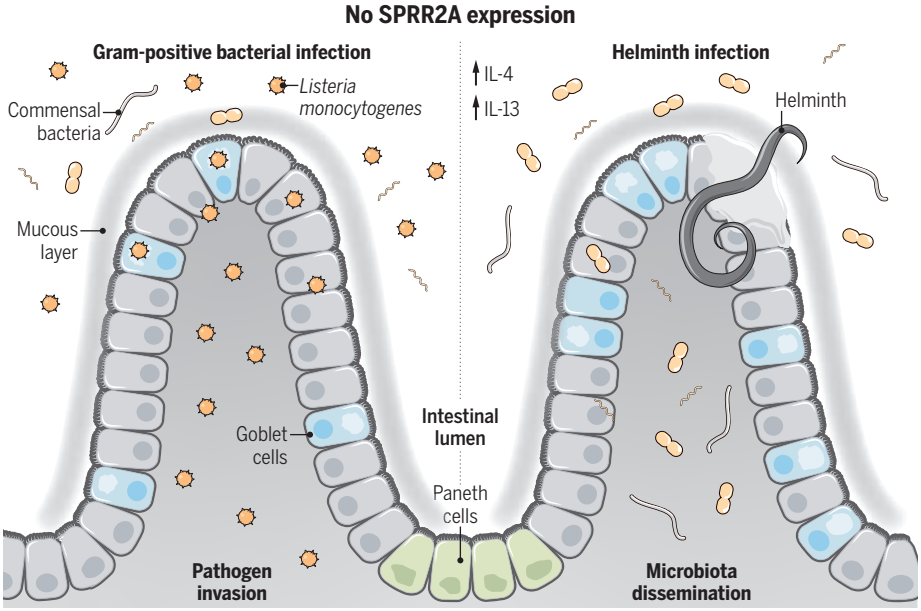
Antimicrobial proteins (AMPs) constitute the first line of host defense against invading microorganisms. Both immune and nonimmune cells can deploy a diverse array of AMPs to combat specific pathogens (*1*). This forms part of an evolutionarily conserved mechanism within the host to rapidly kill and neutralize infectious microbes that breach barrier tissues, including the skin, lungs, and intestines (*2*). Not only are AMPs constitutively expressed, but they can also be induced by an inflammatory stimulus (*1*). On page 710 of this issue, Hu *et al.* (*3*) identify small proline-rich protein 2A (SPRR2A) as a distinct AMP with antibacterial properties against Gram-positive, but not Gram-negative, species. SPRR2A maintains intestinal barrier integrity during homeostasis and limits infection with the Gram-positive bacterial pathogen *Listeria monocytogenes*. Intriguingly, SPRR2A also plays a distinct role in protecting the intestines from bacterial invasion during helminth (parasitic worm) infection.

In vitro experiments by Hu *et al.* revealed that SPRR2A exerts its antimicrobial properties by binding to negatively charged lipids on bacterial membranes, effectively killing these microorganisms. The authors found that SPRR2A is expressed by intestinal epithelial goblet and Paneth cells in response to the bacterial colonization of germ-free mice. The generation of SPRR2A-deficient mice revealed an expansion of Gram-positive bacteria within the small, but not large, intestinal lumen and mucus layers, consistent with its selective antimicrobial activity against Gram-positive bacterial species. The optimal activ-

Department of Immunology and Pathology, Central Clinical School, Monash University, Melbourne, VIC, Australia. Email: nicola.harris@monash.edu

## An antimicrobial protein with dual roles

Paneth cells and goblet cells in the host intestinal epithelium secrete small proline-rich protein 2A (SPRR2A) to prevent luminal bacteria from invading tissue during homeostasis and helminth infection. SPRR2A production is promoted by pathogenic bacterial colonization or by interleukin-4 (IL-4) and/or IL-13 production. When SPRR2A is absent, bacteria invade intestinal tissue during Gram-positive bacterial infection and helminth infection, causing inflammation.




ity of SPRR2A, as determined in vitro, includes low salt concentrations and an acidic pH. These are the exact conditions found in the small intestine, tying together the authors in vitro and in vivo observations.

The production of SPRR2A is also elicited by the type 2 cytokines interleukin-4 (IL-4) and IL-13, which are typically produced in response to helminth infection as part of a protective type 2 immune response. Given the number of known AMPs, it is not surprising that an additional family member has been discovered. But the finding that SPRR2A production can be enhanced by IL-4 and IL-13 renders SPRR2A distinct among the known AMPs.

The relevance of this cytokine-mediated induction of SPRR2A is demonstrated when mice genetically deficient for SPRR2A are exposed to the helminth *Heligmosomoides polygyrus*, a model of mammalian intestinal helminth infection. SPRR2A does not affect the ability of *H. polygyrus* to infect the host but instead protects infected mice from bacterial invasion across the epithelial barrier following its damage by the helminth (see the figure). In a normal situation, the host responds to this damage through an IL-4 and IL-13–dependent increase in SPRR2A expression by the epithelial cells. This in turn leads to the elimination of Gram-positive bacteria in close proximity to the epithelium, thereby preventing their translocation across the damaged epithelium and into the underlying tissues. Although the consequences of bacterial entry into the tissue following helminth infection were not examined, it is likely this would result in increased intestinal, and possibly even systemic, inflammation. This suggests that there is an innate drive for the host to heighten its defenses against bacterial infection during intestinal helminth infection. One major benefit of this response is the ability to curb the onset of a concurrent helminth and bacterial infection that could overwhelm the host immune response. Notably, this reveals that a close evolutionary relationship exists between the mammalian host, helminth, and intestinal bacteria.

A plethora of recent studies reveal that helminths can shape the intestinal bacterial microbiota (*4*), with helminth–bacterial cross-talk able to modulate host resistance against future infection with the same parasite (*5*) and the development of inflammatory diseases such as allergy (*6*) and inflammatory bowel disease (IBD) (*7*). Hu *et al.* also found *SPRR2A* gene expression in biopsies taken from the large intestine of IBD patients. However, whether SPRR2A plays a protective or pathological role in IBD remains unknown. Helminth-associated changes to microbial composition have been repeatedly linked to host production of the type 2 cytokines IL-4 and IL-13 (*4*), yet the exact mechanisms remain unclear. Considering the findings of Hu *et al.*, it is likely that SPRR2A represents one such mechanism, and there are likely to be more.

Hu *et al.* found that two common residents of the mammalian intestine, helminths and bacteria, can independently lead to expression of SPRR2A by the intestinal epithelium. This raises the question of whether other common intestinal pathogens, such as viruses or protozoa, also drive SPRR2A expression. If so, further studies may reveal a role for SPRR2A in preventing bacterial invasion during such infections—akin to its role in helminth infection. The authors also noted *SPRR2A* gene expression in healthy esophagus and bladder, two other sites that commonly interface with microbes from the environment. Although the implications of this finding are not clear, it is compelling to consider that while the epithelia of these tissues differ markedly in terms of structure, they must both constantly adapt to physical alterations such as the mechanical distension caused by food or urine. In keeping with this, SPRR proteins have previously been postulated to contribute to the mechanical properties of the skin epithelium (*8*). Whether SPRR2A plays a role in tissue adaption to mechanical stress and/or resistance to bacterial infection in the esophagus and bladder deserves further attention. This is especially relevant to the bladder, with urinary tract infections by the Gram-positive bacteria *Staphylococcus saprophyticus*, *Enterococcus faecalis,* and *Streptococcus agalactiae* common in certain demographic groups, such as pregnant women and the elderly (*9*). Future studies exploring the specific role of SPRR2A across different diseases and in different groups may reveal additional functions of SPRR2A.

Some of the most important immunological discoveries have been made from research based on careful observation of natural immune phenomena that take place during health, disease, or infection. The findings of Hu *et al.* highlight the potential for discovery of unknown pathways through evaluation of the complex interplay that occurs during transkingdom interactions in the mammalian intestine. ■

GRAPHIC: N. DESAI/*SCIENCE*


### REFERENCES AND NOTES

1. L. J. Zhang, R. L. Gallo, *Curr. Biol.* **26**, R14 (2016).
2. S. Mukherjee, L. V. Hooper, *Immunity* **42**, 28 (2015).
3. Z. Hu *et al.*, *Science* **374**, eabe6723 (2021).
4. A. Rapin, N. L. Harris, *Trends Immunol.* **39**, 724 (2018).
5. E. C. White *et al.*, *Sci. Adv.* **4**, aap7399 (2018).
6. M. M. Zaiss *et al.*, *Immunity* **43**, 998 (2015).
7. D. Ramanan *et al.*, *Science* **352**, 608 (2016).
8. E. Candi, R. Schmidt, G. Melino, *Nat. Rev. Mol. Cell Biol.* **6**, 328 (2005).
9. K. A. Kline, A. L. Lewis, *Microbiol. Spectr.* **4**, UTI-0012 (2021).



### ACKNOWLEDGMENTS

N.L.H. is supported by a National Health and Medical Research Council (NHMRC) of Australia SRF-B fellowship.

10.1126/science.abm3876

NEUROSCIENCE

# Stabilizing synapses

## Adenosine fine-tunes the fate of nascent synapses in brain development

By **David Blum**[1] and **Luísa V. Lopes**[2]

Synaptogenesis is a dynamic process during normal brain development. An overabundance of synapses is initially formed. Synaptic connectivity is then refined through the elimination of synapses. The essential process of removing the irrelevant synapses during this period is called synaptic pruning and is a process regulated by neuronal activity (*1*), commonly referred to as the "use it or lose it" principle. This rapid period of synaptogenesis early in life plays a vital role in network wiring and affects learning, memory, and brain adaptation later in life. On page 709 of this issue, Gomez-Castro *et al.* (*2*) provide a deeper look into this fundamental mechanism, showing that adenosine has a key role in the activity-dependent stabilization of inhibitory γ-aminobutyric acid (GABA) synapses.

The nucleoside adenosine is a building block of any cell, is crucial as a direct and indirect component of RNA and DNA, and is an important methyl donor for epigenomic regulation; adenosine phosphorylation to form adenosine diphosphate (ADP) and adenosine triphosphate (ATP) is directly involved in cellular metabolism. Adenosine action on its four G protein–coupled receptors ($A_1R$, $A_{2A}R$, $A_{2B}R$, and $A_3R$) is also an essential process in intercellular communication. In the brain, extracellular adenosine increases as a function of neuronal activity. It can be formed either from the breakdown of ATP coreleased with various neurotransmitters upon neuronal stimulation or as a result of intracellular catabolic activity. It then activates pre- or postsynaptic adenosine receptors and modulates neuronal excitability through, for example, feedback control of neurotransmitter release and neuronal activity or fine-tuning of long-term potentiation (*3*–*5*). These processes allow for the proper establishment of synaptic communication and the control of excitability, which affects specific functions such as learning and memory.

Whether adenosine affects neuronal development and synapse maturation has been overlooked. Most knowledge was inferred from studies using caffeine administration (an adenosine receptor blocker) during pregnancy. Offspring of dams (pregnant mice) exposed to high doses of caffeine displayed an $A_{2A}R$-dependent delayed migration and insertion of GABA neurons into circuits (*6*), leading to increased cortical and hippocampal network excitability and susceptibility to seizures in adulthood (*6*, *7*). This ultimately impairs cognition in adulthood (*6*, *8*) but also accelerates memory impairment in a model of tauopathy (a form of neurodegeneration) (*9*). More recently, a role of $A_{2A}R$ in the radial migration of cortical projection neurons was described in mice (*10*). This is an important developmental process by which cortical neurons travel from their birthplace to their final position in the circuit. However, the role of adenosine during synaptogenesis remained largely unknown.

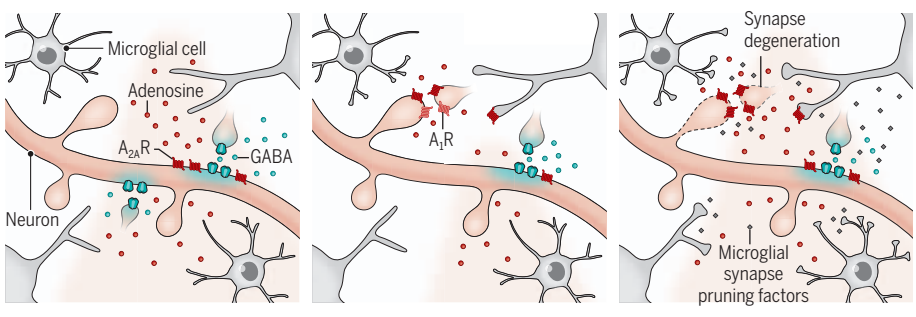
Gomez-Castro *et al.* uncover a role of adenosine in the crucial period for GABAergic synaptic stabilization in the mouse hippocampus. They describe a transient peak in extracellular adenosine, originating from activity-dependent ATP breakdown, and $A_{2A}R$ expression at postnatal days 4 to 9, coinciding with GABA receptor peak at postnatal days 7 to 9 (*11*). They show that the activation of $A_{2A}R$ by activity-dependent adenosine formation is crucial for maintaining GABAergic synapses, whereas receptor ablation in neurons results in synapse elimination. This mechanism is not only important for shaping circuits at these developmental stages, it is crucial for memory formation in adults. The authors elegantly demonstrate the downstream involvement of the calcium–calmodulin–adenylyl cyclase pathway, protein kinase A activation, and subsequent phosphorylation of gephyrin—a central protein that ultimately anchors, clusters, and stabilizes GABA receptors at inhibitory synapses. This establishes adenosine as an important signal to detect activation of GABAergic synapses and favor their stabilization during development. Impaired production of activity-formed adenosine therefore favors elimination of weak synapses.

Whether such developmental control is more broadly engaged in the stabilization or even pruning of other types of synapses remains to be addressed. GABA is known to be excitatory during early developmental stages in a large variety of tissues and species, providing the initial excitatory drive in the absence of glutamatergic synapses (*12*). Perhaps the mechanism described by Gomez-Castro *et al.* has a broader effect on excitatory synaptogenesis and may constitute a fundamental aspect of early synapse stabilization.

Moreover, although the study of Gomez-Castro *et al.* provides clear evidence of the involvement of neuronal $A_{2A}R$ activation,

### Allostatic regulation of synapse fate by adenosine

Adenosine A2A receptor ($A_{2A}R$) signaling controls synaptic allostasis (adaptation to challenges) during development and potentially in the aged brain, which modulates processes crucial for network and cognitive functions.




Enhanced adenosine tone and $A_{2A}R$ expression allow γ-aminobutyric acid (GABA)–ergic synapse stabilization, allowing proper network wiring.

Adenosine sets physiological plasticity by acting on pre- and postsynaptic adenosine receptors.

Enhanced adenosine tone and $A_{2A}R$ expression affect neuron-glial dialog to promote synaptic dysfunction or degeneration.

[1]Lille Neuroscience & Cognition, Inserm UMR-S1172, Alzheimer & Tauopathies, LabEx DISTALZ, Lille Cedex, France. [2]Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal. Email: david.blum@inserm.fr; lvlopes@medicina.ulisboa.pt

GRAPHIC: KELLIE HOLOSKI/*SCIENCE*

whether microglia are also involved in the process remains unknown. Indeed, microglial cells, which also express adenosine receptors, play an essential role in synaptic pruning (*13*). Motile microglial processes make brief contacts with synaptic structures to sense immature synaptic activity and physically remove the defective synapses (*13*). Therefore, an interesting matter for future study would be to understand whether adenosine, acting on $A_{2A}R$, may serve as a coincidence activity detector to coordinate the roles of neurons and microglia in this pruning process.

Imbalances in the adenosine pathway (i.e., enhanced extracellular amounts of adenosine and $A_{2A}R$ up-regulation) have also been demonstrated in the aged brain as well as in epilepsy, neurodegenerative, and psychiatric disorders. All of these physiological and pathological situations involve functional and morphological synapse remodeling. Compelling evidence demonstrates that the synaptic and neuronal up-regulation of $A_{2A}R$ in aging and neurodegeneration is instrumental to the decline of synaptic function and degeneration, presumably through neuron-glial dialog (*14*, *15*). The mechanism uncovered for synaptogenesis by Gomez-Castro *et al.* may have broader implications—e.g., that of adenosine possibly playing a more general role as an activity detector, regulating synaptic dynamics and presumably synaptic loss, in the aged and diseased brain (see the figure). Adenosine and $A_{2A}R$ would then regulate the fate of particular synapses in extreme stages of the brain life cycle. This supports a prime role for adenosine in synaptic allostasis (i.e., during brain adaptation to challenges), the controlling of brain network wiring, and cognitive function. ■

## REFERENCES AND NOTES


1. S. Cohen-Cory *et al.*, *Science* **298**, 770 (2002).
2. F. Gomez-Castro *et al.*, *Science* **374**, abk2055 (2021).
3. D. Boison, *Drug News Perspect.* **20**, 607 (2007).
4. A. Badimon *et al.*, *Nature* **586**, 417 (2020).
5. N. Rebola, R. Lujan, R. A. Cunha, C. Mulle, *Neuron* **57**, 121 (2008).
6. C. G. Silva *et al.*, *Sci. Transl. Med.* **5**, 197ra104 (2013).
7. W. Fazeli *et al.*, *Exp. Neurol.* **295**, 88 (2017).
8. Y. Li *et al.*, *Physiol. Res.* **67**, 975 (2018).
9. S. Zappettini *et al.*, *Front. Cell. Neurosci.* **13**, 438 (2019).
10. S. Alçada-Morais *et al.*, *Cereb. Cortex* bhab188 (2021).
11. U. Ådén, E. Herlenius, L.-Q. Tang, B. B. Fredholm, *Pediatr. Res.* **48**, 177 (2000).
12. H. Gozlan, Y. Ben-Ari, *Cereb. Cortex* **13**, 684 (2003).
13. R. C. Paolicelli *et al.*, *Science* **333**, 1456 (2011).
14. M. Temido-Ferreira *et al.*, *Mol. Psychiatry* **25**, 1876 (2020).
15. K. Carvalho *et al.*, *Brain* **142**, 3636 (2019).


## ACKNOWLEDGMENTS


D.B. and L.V.L. thank their teams at Inserm and the Instituto de Medicina Molecular (IMM) as well as H. Carvalho Pinheiro (IMM, Lisbon) for help with the figure.

10.1126/science.abm3902




Manure injection is a technology that uses less mineral fertilizer and reduces ammonia emission.

PHOTO: HANS ENGBERS/SHUTTERSTOCK

POLLUTION

# How ammonia feeds and pollutes the world

## It is cheaper to cut ammonia emission now than to deal with its consequences later

*By* **Jan Willem Erisman**

Ammonia, with the chemical formula $NH_3$, is a common ingredient in many industrial and agricultural applications and plays a pivotal role in producing the fertilizers needed to produce enough food for 7.9 billion people. However, the agricultural use of ammonia also negatively affects the environment, resulting in the loss of biodiversity and the pollution of water, air, and soil (*1*). Ammonia contributes to the formation of nitrous oxide, a powerful greenhouse gas, worsening the problem in most scenarios but also curbing it somewhat through cooling by aiding the formation of particulate matter and clouds (*2*). On page 758 of this issue, Gu *et al.* (*3*) show that the economic cost associated with the loss of human life—owing to ammonia's contribution to air pollution alone—far outweighs the economic cost to curb ammonia emission.

When released into the air, ammonia bonds with nitrogen oxides and sulfur dioxide and forms particulates less than 2.5 μm in diameter, known as $PM_{2.5}$. These particulates can lead to premature death when inhaled. The World Health Organization (WHO) recently recognized air pollution as the single biggest environmental threat to human health and reassessed the impact of $PM_{2.5}$ on overall air-quality (*4*). Gu *et al.* report the annual cost of human life due to nitrogen-related air pollution to be 23.3 million years. For the sake of comparison with the monetary cost to curb nitrogen emission, this toll in human life can be interpreted as an economic cost of roughly US$420 billion. Based on these numbers, the authors present a quantitative argument for the reduction of ammonia emission. According to their analysis, the abatement cost for ammonia is only a fraction of the corresponding human life cost, as well as only 10% of what it would take to cut out an equivalent amount of nitrogen oxides (*3*).

Ammonia belongs to a group of compounds known as reactive nitrogen ($N_r$), which is the basis for the formation of amino acids—the building blocks of life. These compounds are available in nature to a limited extent. They can be released during volcanic eruptions and lightning and can be formed by the activity of nitro-


Institute for Environmental Sciences, Leiden University, Netherlands. Email: j.w.erisman@cml.leidenuniv.nl

gen-fixing bacteria, which is essential for the growth of plants and thus for the entire food chain, including humans. The ingenuity with which nature deals with the limited availability of these compounds has, over several millennia, led to the development of many natural processes and organisms and rich biodiversity.

The rapid economic development of the global economy over the past century has led to enormous growth in $N_r$ production (*5*). Because of this artificial abundance of $N_r$ in natural and agricultural ecosystems, some natural processes have become unnecessary, resulting in biodiversity loss, coastal eutrophication, acidification and eutrophication of nature areas, human health impacts, and stratospheric ozone depletion. A single molecule of $N_r$ can be transferred from one system to another and contributes over time to all the different negative effects listed above, in what is known as the "cascade effect" (*6*).

Although much effort has been dedicated to the reduction of nitrogen oxide ($NO_x$) emission, Gu *et al.* argue that ammonia can provide a cheaper pathway for $N_r$ reduction. They estimate the abatement cost of $NO_x$ emission at US$16.0 per kilogram of nitrogen reduced, compared with US$1.5 per kilogram of nitrogen reduced by way of ammonia, and this can be achieved by using relatively low-tech solutions such as the reduction of fertilizer use and better management of manure storage and usage. Moreover, policies targeting $NO_x$ reduction alone may not adequately slow down the increasing concentration of $N_r$ in ecosystems and the atmosphere (*3*, *4*), and ammonia reduction may be the extra tool necessary to achieve these targets.

Ammonia from animal manure in livestock production alone contributes roughly 60% of the global ammonia emission and is increasing because of the rising demand for animal products (*7*). Global emission of ammonia has doubled since the 1960s. Ammonia has long been considered a local problem in intensively cultivated areas with concentrated livestock production, such as the Netherlands and Denmark. But, as satellites with infrared spectrometers continue to provide data on ammonia concentration in the atmosphere, the focus of studies assessing ammonia emission has since shifted to the global stage (*8*, *9*). Studies using these new data have indicated an overall underestimation of both anthropogenic and natural ammonia emission in the current data (*10*). These reports indicate that current modeling likely underestimates ammonia pollution, especially in high-intensity agricultural areas. These areas, being relatively close to human populations, may also amplify the negative effects on public health.

Although the Netherlands and Denmark reduced their ammonia emission by 60% over the past three decades (*11*), further reductions are necessary owing to the ruling of the European Court in 2018 that requires $N_r$ emission to be reduced (*12*). Based on this ruling, the State Council of the Netherlands has since judged the Dutch nitrogen policy to be ineffective, resulting in what has become known as the "nitrogen crisis" (*11*, *12*). The council ruling drastically limits the availability of $N_r$ emission permits across all sectors until enough $N_r$ reductions have been achieved within the nation, and the economic loss has been estimated to be 14 billion euros (*12*). The nitrogen crisis in the Netherlands has shown that there needs to be a better balance between long-term sustainability of economy alongside ecology with respect to $N_r$ emission.

The contributions from excess $N_r$ to climate change, water pollution, biodiversity loss, and landscape quality need to be addressed. This will require changes that may disrupt the economy of food production and may require incentives and provisions to help farmers adapt (*11*). The ongoing economic shock in the Netherlands caused by the nitrogen crisis shows that nitrogen pollution can be quite expensive to fix if industrial, and especially agricultural, developments continue to progress unsustainably. Gu *et al.* show that the global burden of disease associated with nitrogen air pollution exposure is estimated to cause millions of deaths and lost years of healthy life annually and, together with economic setbacks, will cost far more to fix tomorrow than it would if action is taken today. ■

## REFERENCES AND NOTES


1. M. A. Sutton *et al.*, *Philos. Trans. R. Soc. London Ser. A* **378**, 20190315 (2020).
2. Intergovernmental Panel on Climate Change (IPCC), "Summary for policymakers" in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte et al., Eds. (Cambridge Univ. Press, 2021); www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_SPM.pdf.
3. B. Gu *et al.*, *Science* **374**, 758 (2021).
4. World Health Organization (WHO), "WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide, and carbon monoxide" (WHO, 2021).
5. J. N. Galloway, A. Bleeker, J. W. Erisman, *Annu. Rev. Environ. Resour.* **46**, 745 (2021).
6. J. N. Galloway *et al.*, *Bioscience* **53**, 341 (2003).
7. A. Uwizeye *et al.*, *Nat. Food* **1**, 437 (2020).
8. J. X. Warner *et al.*, *Geophys. Res. Lett.* **44**, 2875 (2017).
9. M. Van Damme *et al.*, *Environ. Res. Lett.* **16**, 55017 (2021).
10. E. Dammers *et al.*, *Atmos. Chem. Phys.* **19**, 12261 (2019).
11. J. W. Erisman, *One Earth* **4**, 15 (2021).
12. E. Stokstad, *Science* **366**, 1180 (2019).


10.1126/science.abm3492

ATMOSPHERIC SCIENCE

# Midair transformations of aerosols

## Spontaneous reactions play an important role in shaping the chemistry of aerosols

*By* **Manuel F. Ruiz-Lopez**

The presence of atmospheric aerosols, which are tiny solid or liquid particles suspended in the air, have a major impact on the global climate and public health. A wide variety of natural sources and human activities produce aerosols, ranging from volcanic eruptions to coal burning, which means that the composition of aerosols also varies greatly. Some of these aerosols can absorb or react with water vapor and other gaseous compounds while airborne. They influence the atmospheric chemical cycles (*1*, *2*) and must be considered for improving current air quality and climate models. On page 747 of this issue, Kong *et al.* (*3*) describe an aerosol surface–driven reaction of great atmospheric relevance and suggest potential applications for the chemistry. They discovered that ammonium sulfate crystals exposed to water vapor would spontaneously form elemental sulfur and nitrogen gas on their surface at a much greater rate than expected (see the figure).

> **"...ammonium sulfate crystals exposed to water vapor would spontaneously form elemental sulfur and nitrogen gas on their surface at a much greater rate than expected."**

Ammonium sulfate is an aerosol that can act as a seed for the formation of cirrus

Laboratoire de Physique et Chimie Théoriques, UMR CNRS 7019, University of Lorraine, CNRS, BP 70239, 54506 Vandoeuvre-lès-Nancy, France

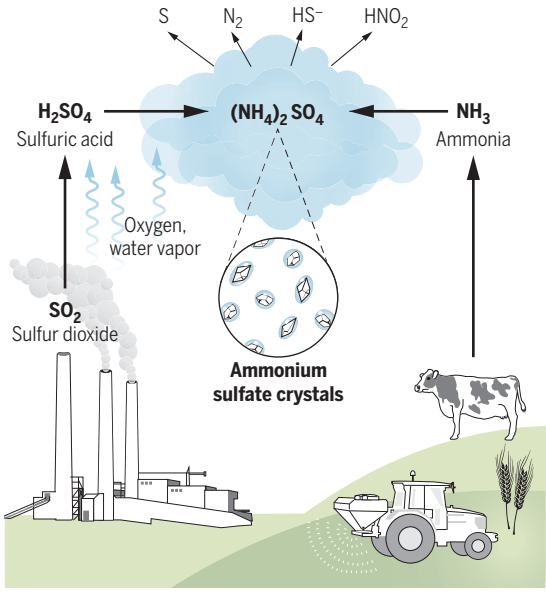
clouds (*4*). Although natural sources can produce ammonium sulfate aerosols, most of them are produced by human activities. In the troposphere, the lowest layer of the atmosphere, the sulfur dioxide emitted from fossil fuel combustion dissolves in water droplets and is transformed into sulfuric acid. Precipitation of these droplets to the ground creates acid rain. Sulfuric acid in the troposphere can also react with ammonia emitted from agricultural activities and produces ammonium sulfate salt, either in the form of aqueous solutions or crystallized particles, depending on air humidity and temperature. Kong *et al.* studied the behavior of ammonium sulfate crystals exposed to varying concentrations of water vapor with an experimental technique called ambient pressure x-ray photoelectron spectroscopy. They measured the binding energy of the inner electrons in sulfur and nitrogen atoms on the crystal surface, which conveys the chemical information on what kind of compounds these atoms are forming.

In dry conditions, only the expected signatures for sulfate and ammonium ions are observed. However, when the relative humidity increases, the ammonium sulfate salts begin to dilute and react over the surface, leading to additional signatures for elemental sulfur, bisulfide ions, nitrous acid, and ammonia, which was unexpected. If the relative humidity is further increased to the deliquescence threshold of 78%, the salts gradually dissolve completely in water. During this transformation, there is a transient period in which elemental sulfur dominates the spectra and nitrogen species signals vanish. As the dilution continues, however, the sulfate, ammonium, and ammonia signals reappear, and a new steady state is reached. The assignment of the signals and important aspects of the underlying transformation mechanisms have been guided and supported by quantum chemistry and molecular dynamics calculations.

Kong *et al.* conclude that spontaneous sulfate-reducing ammonium oxidation reaction is at play as the salt surface is being diluted, which does not happen when the salt is completely dry or completely diluted. The final products of this reaction are elemental sulfur and nitrogen gas. The ease with which the reaction occurs in the aqueous layer at the vapor-crystal interface is notable because it is an unfavorable reaction under ordinary conditions. For instance, biocatalysts are typically needed to stimulate the reaction in wastewater treatment processes.

## How ammonium sulfate crystals break down upon hydrated surface catalysis

When exposed to water vapor, ammonium sulfate crystals spontaneously form elemental sulfur, nitrogen gas, bisulfide ions, and nitrous acid. This unexpected reaction has implications for the long-term effects of ammonia and sulfur dioxide emissions.




GRAPHIC: N. CARY/*SCIENCE*

The spontaneous oxidation reaction implies that, in humid air, highly oxidized sulfur compounds, such as sulfate, can be converted to less oxidized compounds, such as elemental sulfur or bisulfide. This chemistry was previously not expected to occur in the atmosphere, which is considered an oxidizing medium. The counterintuitive discovery is likely to change the current understanding of the sulfur cycle. In addition, the authors also identified nitrous acid as an intermediate product of this process. Nitrous acid is a key precursor to the extremely reactive hydroxyl radical sometimes known as the "atmosphere's detergent." The presence of the spontaneous reactions on sulfate ammonium could help explain why the concentration of atmospheric hydroxyl radicals is underestimated in current models.

The findings also have broad implications for the "on water catalysis" phenomenon (*5*), where the rate of chemical reactions is accelerated at water surfaces and aqueous interfaces. The work by Kong *et al.* provides new perspectives on the current understanding of these accelerated reactions (*6*). Though still incompletely understood, "on water catalysis" has been shown to play a role in several chemical and photochemical processes on atmospheric aqueous aerosols *(7–9)*. In addition, because the interfaces formed by water droplets coated with an organic film share similarities with cell membranes, "on water catalysis" may have played an important part in the origination of life on Earth (*10*). In terms of applications, the effect of "on water catalysis" has been exploited in organic synthesis that uses dispersed interface-rich systems such as oil-in-water microemulsions (*11*) or sprayed aqueous microdroplets (*12*). For instance, certain reactions, such as the oxidation of water to form hydrogen peroxide, have been identified to proceed spontaneously "on water" even if they do not occur "in water" (*13*). These findings open new perspectives for the development of greener synthetic techniques by reducing the need for organic solvents.

There is no simple answer to why some unfavorable chemical reactions are boosted at aqueous interfaces. Some fundamental molecular properties, such as the oxidation potential (*14*) or the acidity (*15*), change when the molecules are adsorbed at an aqueous interface. The local electric fields associated with the structural asymmetry of the interface have been pointed out as a key for explaining these effects. The alignment and confinement of the reagents, or the large surface-to-volume ratio exhibited by microscopic systems, have also been suggested as important factors to be considered for the effect. These issues are being intensely debated in the literature (*6*), and future research should aim at clarifying their respective contributions. In this regard, the parallel study of selected reactions with different experimental techniques and elaborated theoretical molecular simulations would be most valuable. ■

### REFERENCES AND NOTES


1. A. R. Ravishankara, *Science* **276**, 1058 (1997).
2. M. O. Andreae, P. J. Crutzen, *Science* **276**, 1052 (1997).
3. X. Kong *et al.*, *Science* **374**, 747 (2021).
4. J. P. Abbatt *et al.*, *Science* **313**, 1770 (2006).
5. S. Narayan *et al.*, *Angew. Chem. Int. Ed.* **44**, 3275 (2005).
6. M. F. Ruiz-Lopez, J. S. Francisco, M. T. C. Martins-Costa, J. M. Anglada, *Nat. Rev. Chem.* **4**, 459 (2020).
7. E. M. Knipping *et al.*, *Science* **288**, 301 (2000).
8. C. George, M. Ammann, B. D'Anna, D. J. Donaldson, S. A. Nizkorodov, *Chem. Rev.* **115**, 4218 (2015).
9. J. M. Anglada, M. T. C. Martins-Costa, J. S. Francisco, M. F. Ruiz-López, *J. Am. Chem. Soc.* **142**, 16140 (2020).
10. C. M. Dobson, G. B. Ellison, A. F. Tuck, V. Vaida, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11864 (2000).
11. S. Serrano- Luginbühl, K. Ruiz-Mirazo, R. Ostaszewski, F. Gallou, P. Walde, *Nat. Rev. Chem.* **2**, 306 (2018).
12. X. Yan *et al.*, *Angew. Chem. Int. Ed.* **55**, 12960 (2016).
13. J. K. Lee *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19294 (2019).
14. M. T. C. Martins-Costa, J. M. Anglada, J. S. Francisco, M. F. Ruiz-Lopez, *Angew. Chem. Int. Ed.* **51**, 5413 (2012).
15. H. Mishra *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18679 (2012).


10.1126/science.abl8914

LANDMARK: ATMOSPHERIC CHEMISTRY

# A radical shift in air pollution

## Fifty years ago, Levy identified hydroxyl radicals as the driver of chemistry in the troposphere

By **Colette L. Heald**[1] *and* **Jesse H. Kroll**[2]

The lower atmosphere, or troposphere, contains ~90% of the mass of the atmosphere, including the air that we breathe at the surface of Earth. Fifty years ago, during a time when scientists were focused on understanding the chemistry of the ozone layer and urban smog, the troposphere was thought to be largely chemically inert and simply a receptacle for emissions and for gases transported from the stratosphere above. In 1971 in *Science* (*1*), Levy discovered that hydroxyl (OH), a highly reactive radical, could be formed efficiently in the lower atmosphere. This revelation shifted the paradigm of atmospheric chemistry and shapes how we view both air quality and climate forcing today (see the figure).

Prior to 1971, only a few regions of the atmosphere were thought to be chemically active, driven by ultraviolet (UV) radiation from the Sun. In 1930, Chapman (*2*) proposed that ozone ($O_3$) was generated in the stratosphere from the UV photolysis of $O_2$. This process forms O atoms that add to $O_2$ to make $O_3$. Subsequent work by other researchers found that the photolytic reactions of trace stratospheric species form a host of free radicals (such as OH, $HO_2$, NO, and $NO_2$) that take part in a complex network of radical-radical and radical-molecule reactions. The net result of this photochemistry is the stratospheric ozone layer, which shields living creatures at the surface from the damaging effects of UV radiation.

However, the absorption of UV radiation by the ozone layer implies diminished photochemistry in the underlying troposphere. As a result, it was believed that most of the troposphere was essentially inert, involving only slow atmospheric reactions initiated by ozone (descending from the stratosphere) or the occasional O atom. Anthropogenic air pollutants could therefore accumulate in the troposphere; for example, there were concerns that carbon monoxide (CO) emitted from vehicles might reach toxic levels and become a widespread public health threat.

The main exception to this view of the inert troposphere was the polluted urban atmosphere. The late 1940s saw the emergence of "Los Angeles smog," a noxious mix of gases and particles leading to eye irritation, damage to plants, and reduced visibility. Studies (*3, 4*) quickly revealed that this pollution included ozone (which at ground level is a major respiratory irritant and phytotoxin) and other oxidants, and moreover was photochemical in nature, formed by the irradiation of hydrocarbons and nitrogen oxides (both emitted in high quantities from vehicles). Researchers focused on understanding this photochemistry, to predict and ultimately mitigate smog formation (*5*).

By the late 1960s, there were indications that the troposphere was more photochemically active than was generally assumed. The atmospheric lifetime of CO was found to be surprisingly short (*6, 7*), indicating a large unknown reactive sink. Although it was understood that urban $O_3$ was formed from the near-UV photolysis of $NO_2$, its fast formation, which required the rapid oxidation of hydrocarbons and conversion of NO to $NO_2$, could not be explained by existing mechanisms (*8*). These gaps provided hints that the atmospheric oxidation of trace species (CO, hydrocarbons, and other pollutants) was much faster than was appreciated, and it was suggested that highly reactive OH radicals might be responsible (*7, 8*). However, there was no known viable tropospheric source of OH.

Levy solved this conundrum, by showing that the hydroxyl radical (OH) could be efficiently formed throughout the troposphere (what he referred to as the "normal atmosphere") by the UV photolysis of ozone. The process involves two steps. In the first, photolysis forms $O(^1D)$ (electronically excited O atoms) caused by absorption of a photon. In the second, $O(^1D)$ reacts with water vapor:

$$O_3 + \text{photon} \rightarrow O_2 + O(^1D) \quad (1)$$
$$O(^1D) + H_2O \rightarrow OH + OH \quad (2)$$

These same reactions were known to produce OH in the stratosphere, but they had been assumed not to occur in the troposphere because of the lack of high-energy UV photons (wavelengths of <320 nm) required for reaction 1. However, Levy pointed out that the very small fraction of low-wavelength UV that did make it through the ozone layer, combined with the high water-vapor concentrations in the lower atmosphere, was sufficient to lead to substantial OH generation throughout the troposphere.

The impact of the troposphere being a highly reactive environment was immediate and profound. The ubiquity of the OH radical explained the short chemical lifetimes of CO (and other pollutants), as well as the rapid formation of urban ozone pollution. It also led to many new discoveries that are now central to our understanding of the lower atmosphere. For example, within just a few years, it was shown that the oxidation of methane by OH had an enormous impact on the global atmosphere, not only serving as the main source of CO (*9*) but also leading to the formation of ozone throughout the entire troposphere (*10, 11*), and not just the stratosphere or polluted urban areas, as previously thought. This realization also placed a new focus on trace compounds that do not react with OH (such as $CO_2$, CFCs, and $N_2O$). Because these species could accumulate in the atmosphere, they have outsized impacts on stratospheric ozone depletion, climate, or both.

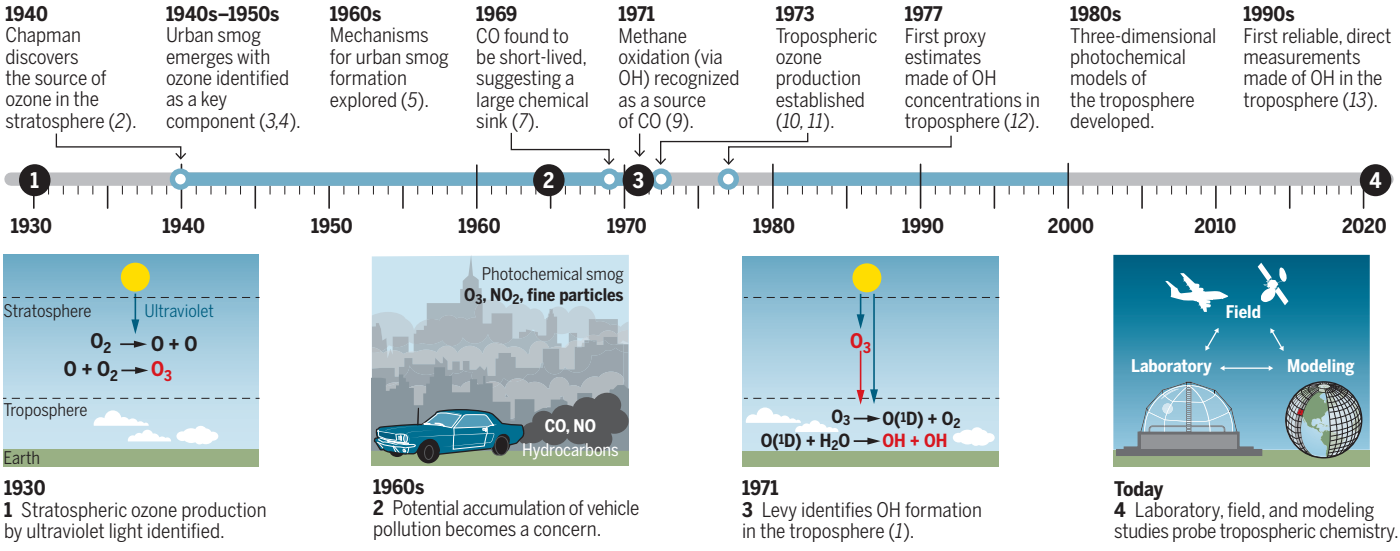
It would be many more years before atmospheric OH levels were measured. The high reactivity of OH ensures that its atmospheric concentration is exceedingly small—below parts per trillion by volume—and posed a daunting measurement challenge. Measurements of long-lived halocarbons (whose main sinks were oxidation by OH) provided the first estimates of OH concentrations (*12*). Direct, spectroscopic detection of OH with laser-induced fluorescence proved to be highly challenging, and it would be decades before OH could be detected reliably in the troposphere (*13*). Currently, OH measurements remain challenging and relatively sparse, but it is generally accepted that global OH has a concentration on the order of $1.2 \times 10^6$ molecules $cm^{-3}$ (*14*), in reasonable agreement with the early estimates by Levy.

The production of OH in the troposphere underpins our modern understanding of the chemistry of the troposphere, and particularly the relationships between emissions, air quality, and anthropogenic

[1]Department of Civil and Environmental Engineering and Department of Earth, Atmospheric, and Planetary Sciences, Massachussets Institute of Technology, Cambridge, MA 02139, USA. [2]Department of Civil and Environmental Engineering and Department of Chemical Engineering, Massachussets Institute of Technology, Cambridge, MA 02139, USA. Email: heald@mit.edu; jhkroll@mit.edu

## A century of atmospheric chemistry

The timeline shows key discoveries in tropospheric chemistry and its effects on our understanding of pollution. In 1971, Levy showed that hydroxyl radical (OH) could be formed efficiently throughout the troposphere, when it had been thought that such radicals could only be generated in the stratosphere. The ubiquity of tropospheric OH gives rise to complex networks of chemical reactions that control the levels of air pollutants and climate forcers. The ongoing study of this chemistry, from 3D models, laboratory investigations, and fieldwork, continues to inform our understanding of tropospheric pollutants and efforts to mitigate their effects.




**1930**
**1** Stratospheric ozone production by ultraviolet light identified.

**1960s**
**2** Potential accumulation of vehicle pollution becomes a concern.

**1971**
**3** Levy identifies OH formation in the troposphere (*1*).

**Today**
**4** Laboratory, field, and modeling studies probe tropospheric chemistry.

climate change. To illustrate this point, consider a hypothetical world without tropospheric OH. Leading air pollutants of concern, which are typically regulated for their deleterious effects on human health, would shift toward primary (that is, directly emitted) species. For example, as feared in the 1960s, the present-day tropospheric burden of CO would soar, putting a large fraction of the global population at risk for CO poisoning. Similarly, concentrations of sulfur dioxide ($SO_2$) and nitrogen oxides ($NO_x$) would rise along with attendant exposures to $SO_2$ and $NO_2$ that increase the risk of respiratory disease.

At the same time, concentrations of secondary pollutants would fall. In a troposphere without OH, there would be little to no tropospheric production of $O_3$ and thus no urban $O_3$ air quality concerns, largely alleviating the health burden of $O_3$. Fine particulate matter ($PM_{2.5}$, particles less than 2.5 μm in diameter), which leads to the premature deaths of millions of people every year, would also decrease sharply because of lower formation rates of secondary sulfate, nitrate, and organic aerosols. Air quality regulation, both policy and the deployment of control technology, would be far simpler, because pollutants of concern could be moderated directly (and linearly) through emissions cuts.

Similarly, Earth's climate and our understanding of anthropogenic climate forcing would also be greatly altered in a world without tropospheric OH. Atmospheric concentrations of methane, a leading greenhouse gas, would surge. However, given that a substantial fraction of methane sources are natural, methane would have accumulated in (and warmed) the preindustrial atmosphere, such that anthropogenic additions of methane would contribute more modestly to climate forcing. Coupled with the lack of anthropogenic $O_3$ production in the troposphere and the greatly diminished anthropogenic source of $PM_{2.5}$, there would be much less climate forcing from what are currently referred to as short-lived climate forcers (SLCFs).

It is also probable that a troposphere without OH would produce any number of new environmental concerns. Longer-lived tropospheric gases could be transported upward to the stratosphere, likely exacerbating the destruction of the ozone layer. Concentrations of harmful pollutants known to be removed by OH could increase rapidly and require new air pollution policies. The production of OH in the troposphere, identified 50 years ago by Levy, precludes this counterfactual of a simple but markedly different troposphere.

In 1971, the work of Levy showed that, counter to prevailing assumptions at the time, the lower atmosphere is not inert, but rather is a complex chemical reactor. This revelation is what necessitated and created the field of tropospheric chemistry. The work of the last 50 years has revealed generations of oxidative chemistry in the troposphere, highlighted the prominence of secondary pollutants, identified intricate interactions between natural and anthropogenic emissions, and elucidated the nuanced response of air pollutants and SLCFs to regulatory policies and a changing climate. It would be convenient but unrealistic (and even counterproductive) to assume that air-pollution policy can be developed with a knowledge of emissions alone. The challenge initiated by Levy is to understand how these emissions evolve in the reactive troposphere. ■

### REFERENCES AND NOTES


1. H. Levy 2nd, *Science* **173**, 141 (1971).
2. S. Chapman, *Mem. R. Meteorol. Soc.* **3**, 103 (1930).
3. A. J. Haagen-Smit, *Ind. Eng. Chem.* **44**, 1342 (1952).
4. A. J. Haagen-Smit, C. Bradley, M. Fox, *Ind. Eng. Chem.* **45**, 2086 (1953).
5. P. A. Leighton, Photochemistry of Air Pollution, in *Physical Chemistry: A Series of Monographs*, E. Hutchinson, P. Van Rysselberghe, Eds. (Academic Press, 1961), vol. 9.
6. E. Robinson, R. C. Robbins, "Sources, Abundance, and Fate of Gaseous Atmospheric Pollutants" (Stanford Research Institute, 1968).
7. B. Weinstock, *Science* **166**, 224 (1969).
8. J. Heicklen, K. Westberg, N. Cohen, "The Conversion of NO to $NO_2$ in Polluted Atmospheres" (The Pennsylvania State University, Center for Air Environment Studies, 1969).
9. J. C. McConnell, M. B. McElroy, S. C. Wofsy, *Nature* **233**, 187 (1971).
10. W. Chameides, J. C. G. Walker, *J. Geophys. Res.* **78**, 8751 (1973).
11. P. Crutzen, *Pure Appl. Geophys.* **106**, 1385 (1973).
12. H. B. Singh, *Geophys. Res. Lett.* **4**, 101 (1977).
13. P. S. Stevens, J. H. Mather, W. H. Brune, *J. Geophys. Res. Atmos.* **99**, 3543 (1994).
14. M. Rigby *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5373 (2017).



### ACKNOWLEDGMENTS

We thank H. Levy for his insight and feedback.

10.1126/science.abl5978


GRAPHIC: N. CARY AND C. SMITH/*SCIENCE*

POLICY FORUM

CLIMATE POLICY

# National climate institutions complement targets and policies

## Institutions can affect coordination, consensus, and strategy

By **Navroz K. Dubash[1], Aditya Valiathan Pillai[1], Christian Flachsland[2], Kathryn Harrison[3], Kathryn Hochstetler[4], Matthew Lockwood[5], Robert MacNeil[6], Matto Mildenberger[7], Matthew Paterson[8], Fei Teng[9], Emily Tyler[10]**

Discussions about climate mitigation tend to focus on the ambition of emission reduction targets or the prevalence, design, and stringency of climate policies. However, targets are more likely to translate to near-term action when backed by institutional machinery that guides policy development and implementation. Institutions also mediate the political interests that are often barriers to implementing targets and policies. Yet the study of domestic climate institutions is in its infancy, compared with the study of targets and policies. Existing governance literatures document the spread of climate laws (*1*, *2*) and how climate policy-making depends on domestic political institutions (*3*–*5*). Yet these literatures shed less light on how states organize themselves internally to address climate change. To address this question, drawing on empirical case material summarized in table S1, we propose a systematic framework for the study of climate institutions. We lay out definitional categories for climate institutions, analyze how states address three core climate governance challenges—coordination, building consensus, and strategy development—and draw attention to how institutions and national political contexts influence and shape each other. Acontextual "best practice" notions of climate institutions are less useful than an understanding of how institutions evolve over time through interaction with national politics.

### CATEGORIES OF INSTITUTIONALIZATION

Climate institutions shape and are shaped by national politics, and this process is a leading determinant of climate policy. Recent literature suggests that distributional conflict within countries is a more important determinant of climate policy than considerations of free-riding across countries (*6*). Consequently, the interplay between institutions and politics also shapes what countries put forward as their Nationally Determined Contributions (NDCs) under the Paris Agreement. Institutions determine who sits at the table, how decisions are made and by whom, and the cognitive priors carried into decision-making. Over time, successful climate institutions create long-term dynamics that favor more ambitious climate policy by restructuring decision-making and interest-group politics.

For our purposes, "institutions" are defined as "…the formal or informal procedures, routines, norms and conventions embedded in the organisational structure of the polity or economy" [(*7*), p. 938], including laws, organizations in government, and interdepartmental coordination processes. In this article, our focus is on national and formal institutions, while acknowledging that informal practices and institutions at regional and national scales and nonstate organizations of various kinds play a substantial governance role.

Even this limited subset of national climate institutions are challenging to define because they take a wide array of forms. For example, in the UK, a Climate Change Act requires 5-year carbon budgets, on the basis of advice from a credible analytical body. In India, a climate action plan led to the National Solar Mission that provided a political arena for solar energy interests to counter fossil fuel interests within bureaucratic negotiations. By contrast, in Brazil, institutional restructuring that reduced deforestation-linked emissions was driven entirely by nonclimate considerations, suggesting that institutional shifts outside the climate sphere are relevant to climate outcomes. The definition of climate institutions cannot be limited to an assessment of institutional mandates but must also include those that bring climate cobenefits.

Rather than seeking sharp boundaries, we focus on the institutionalization of climate concerns as a process, which proceeds at different rates and in different ways across countries, with the possibility of both advancement and retreat. This process is usefully mapped by using three categories of institutionalization. First, the deepest form of institutionalization involves the creation of new, "purpose-built" climate institutions—laws, organizations, or processes—primarily and intentionally focused on climate change. Second, in many cases, institutionalization happens through intentional "layering" (*8*) of climate obligations on existing institutions, which involves repurposing of existing institutional structures. Third, climate change–related outcomes, positive or negative, may result from "latent" institutions that are not intentionally aimed at climate mitigation.

This last category has a very wide scope and is consequently not included in table S1. However, its inclusion as a category signals that any part of state activity could be salient to climate change and that over time, these institutions could be made more intentionally climate oriented through layering or substituted through purpose-built institutions. The study of climate institutions needs to consider all three categories. They serve as analytical ideal types through which to track the process of institutionalization, which is shaped in path-dependent ways by national political context, interactions with international processes, and existing bureaucratic norms (*9*).

### THREE PROBLEMS OF GOVERNANCE

We suggest that national climate institutions are usefully assessed against three problems of governance that are particularly salient to climate change. As an all-of-economy and all-of-society problem, climate solutions require rapid and large-scale social and economic transformations that have the potential to create economic losers who may oppose change. Examples of institutional efforts to address these challenges—drawing on the authors' recent study of Australia, Brazil, China, Germany, India, South Africa, the UK, and the United States (*10*)—are discussed below and further summarized in table S1.

#### Coordination challenges

The scale and scope of the mitigation challenge requires coordination across policy domains such as energy, industry, agriculture, and urbanization through "climate

[1]Centre for Policy Research, New Delhi, India. [2]Centre for Sustainability, Hertie School, Berlin, Germany. [3]Department of Political Science, University of British Columbia, Vancouver, BC, Canada. [4]Department of International Development, London School of Economics and Political Science, London, UK. [5]Science Policy Research Unit, University of Sussex, Sussex, UK. [6]Department of Government and International Relations, University of Sydney, Sydney, NSW, Australia. [7]Department of Political Science, University of California Santa Barbara, Santa Barbara, CA, USA. [8]School of Social Sciences, University of Manchester, Manchester, UK. [9]Institute of Energy Environment and Economy, Tsinghua University, Beijing, China. [10]African Climate and Development Institute, University of Cape Town, Cape Town, South Africa. Email: ndubash@gmail.com



A worker cleans solar cells in a housing complex in Kolkata, India. A climate action plan in India led to the National Solar Mission, providing a political arena for solar energy interests.

PHOTO: PARTH SANYAL/REUTERS

policy integration" (*11*), and "mainstreaming" of climate policy into related policy arenas, as well as the supportive tasks of monitoring and assessing progress. Because each domain also includes established and competing interests, this is not only a technical challenge but also a political one.

States have developed four broad responses to the coordination challenge. The default option is to layer climate coordination responsibilities onto existing environmental ministries, but this approach has proved to be relatively ineffective. For example, in India, the environment ministry was too weak to direct other, more powerful ministries. In Germany, sectoral governance structures, such as long-standing corporatist approaches in the transport sector, proved more enduring than attempts at coordination by the environment department, albeit without a formal mandate. An important exception is the United States, where the deep capacity of the Environmental Protection Agency was deployed with positive effect once it was explicitly mandated to do so.

Second, countries have reorganized their bureaucracies to allocate climate change responsibilities. In the UK, bureaucratic reorganization brought environment together with industrial strategy in a new department. Consistent with its view of climate change as a developmental challenge, China embedded climate coordination in its powerful National Development and Reform Commission from 2007 to 2018, with formal authority to steer line ministries, before shifting responsibility to the Ministry of Ecology and Environment.

Third, many countries have developed new interministerial coordination committees, with mixed success. They have had limited effect in Brazil on the politically heavyweight sectors of forests and agriculture and in South Africa on electricity and mining. In India, coordination was successful when endowed with the authority of the prime minister's office through a special envoy, who was able to drive ministerial cooperation. However, these coordination capacities have often proved insufficiently durable. The Indian envoy did not endure interministerial politics. In the United States or Australia, executive branch coordination structures have emerged and vanished with electoral outcomes. Coordination in China is more stable because it draws on the institutional device of National Leading Groups, which are used for a range of cross-ministerial coordination.

A fourth approach is creating climate laws, subsuming the coordination challenge within a framework of target setting and compliance, as in the UK, Germany, and briefly, Australia. Yet coordination capacity remains weak even within these frameworks. The UK Climate Change Act lacks an explicit mechanism to coordinate across sectors, leading to uneven performance—aggressive mitigation in electricity but limited in transport, for example. In Australia, coordination was implicitly assumed to occur through the setting of a carbon price, but the Climate Change Act was short-lived. More long-lasting were specialized law-backed agencies to promote renewable energy, which were politically resilient perhaps because they distributed benefits rather than allocating costs.

### Building consensus

Institutions are needed to mediate low- and high-carbon interests and ideally build consensus toward low-carbon futures. That climate mitigation efforts may create losers now and winners later complicates political mediation (*12*). Conventional interest groups, such as labor and capital, may not line up neatly on one or another side of policies (*5*). In this complex political landscape, institutions can provide credible information and analysis that shifts politics, creates new narratives, and creates credible spaces for bargaining among interests.

The UK Climate Change Committee illustrates how providing credible information can help build consensus around mitigation; the Committee's analysis is invoked when setting carbon budgets, serving as a ratchet to prevent backsliding. Australia has established a Climate Change Authority with a mandate to provide independent information, but a similarly salutary effect on policymaking has not been documented, although whether this is due to the effectiveness of the institution or the intractability of Australian climate politics is unclear.

Deliberative bodies can help change climate narratives, opening new political opportunities and configurations. The Brazilian Climate Change Forum has, since 2000, formulated and injected proposals into public discourse, including a net zero target. However, its ability to influence climate-friendly political administrations is greater than its ability to buffer more climate-skeptical administrations. In South Africa, the National Economic Development and Labour

Commission has injected a "just transition" narrative, with the promise of realigning the politics of climate policy.

Institutional changes can create feedback effects that support consensus or that undermine it (*13*). In Brazil, a plan to limit Amazonian deforestation, in part through greater investment in the capacity of the environment ministry, led to swift pushback by mobilized "ruralists" who stood to lose from stronger environmental enforcement. Yet it also activated large agribusiness who sought market access tied to stricter environmental controls, creating a positive feedback effect. An Indian solar mission as part of a climate action plan provided space for solar interests in the policy process.

### Strategy development

Deep decarbonization requires transformational rather than incremental change. This requires, among other things, the ability to define sectoral transformative pathways and develop NDCs that are backed by institutional capacity to establish targets and policies, monitor progress, and calibrate them over time.

Many countries may not have established the domestic institutions needed to implement the implied approach of the Paris Agreement. They may have mechanisms to set targets, but these mechanisms are frequently ad hoc and not rooted in domestic climate processes aimed at long-term strategic transformation. For example, the US approach oscillates with political control of the White House. Other countries have created ad hoc processes run by the head of government, such as the presidency in Brazil and the prime minister's offices in India and Australia. Although these efforts may be informed by a long-term strategy as articulated in an NDC, they are not institutionally embedded and frequently lack the benefit of consistent analysis, review, and structured opportunities to revise strategies.

UK, Germany, and China are examples of exceptions, where purpose-built mechanisms and responsibilities for strategy development, monitoring, and iterating targets exist, which are also law-backed in the first two cases. The UK approach is built around setting 5-year budgets, backed by the analytical capacity of the Climate Change Committee; Germany sets annual sectoral targets with a mandated review and policy adjustment process; and China's approach relies on a politically powerful national leading group and 5-year plans, backed by a climate change department.

### INTERACTION WITH NATIONAL CONTEXT

Climate institutions are underdeveloped in part because countries do not start with a blank slate or a free hand in creating institutions. Instead, interests and ideas around climate change, operating in the context of existing institutions, interact to generate distinct "varieties of climate governance" (*9*). Although there are several factors at play, cross-country comparative analysis suggests two variables that are particularly useful in mapping these varieties: the extent of political polarization around climate change, and the narrative frame through which climate politics are debated. Political consensus around climate action supports the creation of purpose-built institutions, whereas such institutions are hard to support when climate politics are polarized. Institutions are differently formed if they are informed by a primary political narrative of greenhouse gas reductions versus by a narrative that embeds mitigation within other pressing domestic goals, such as job creation, energy security, or green growth.

**"The way forward lies in recognizing that more limited climate institutions consistent with domestic political context can still help shift political narratives and create conditions for future strategic institutions."**

Low political polarization with mitigation as a dominant objective is most likely to create the conditions for new purpose-built institutions created to strategically chart and achieve a mitigation course over time. The UK and, since passage of its 2019 Climate Change Act, Germany, fall in this category, with explicit legally backed emission reduction mandates. These conditions hold in China, too; polarization is low under conditions of "authoritarian environmentalism," and China has explicitly embraced a mitigation and carbon neutrality narrative.

Political contestation around mitigation-centric narratives may lead to emergent climate institutions, but these are likely to be unstable and subject to political shifts. In Australia and the United States, efforts to build mitigation-focused institutions have foundered, as evidenced by the ebb and flow of such institutions as coordinating offices in the US White House and the Australian Climate Change Commission. Where politics are contested, the specifics of domestic political institutions matter to climate outcomes. For example, multiple veto points in the US political system led to the Senate blocking climate legislation, but the Supreme Court provided an alternative route to confer the power to regulate greenhouse gases on the Environmental Protection Agency through layering. In Australia, despite conflictual politics, relatively stable renewable energy promotion organizations were created that avoided conflict by sharing benefits rather than imposing costs.

In some countries, climate mitigation debates are subordinated and linked to other, local, more politically salient narratives. Where these are tied to deeply contested sectors, such as electricity in South Africa and forests in Brazil that have great mitigation salience, progress on mitigation may occur serendipitously through shifts in sectoral politics and result in latent climate institutions. Examples include shifts in forest governance in Brazil in the 2000s or renewable energy in South Africa in the 2010s. Under such conditions, climate outcomes may depend on strengthening sectoral institutions.

Where mitigation is embedded within other objectives, and political polarization is low because climate politics operates below the radar, opportunistic climate institutions may emerge. Opportunistic institutions are designed to maximize instances in which mitigation agendas can be made complementary to other agendas, such as energy security and air pollution, as in India. Because there is little domestic impetus for emergent climate institutions, they may be more likely to result from international pressures. But for the same reason, these institutions are frequently unstable, as in India.

Those seeking a more active state role in accelerating climate action face a conundrum. Strategic climate institutions capable of driving economy-wide low-carbon transformation require the stringent conditions of national political consensus around a mitigation agenda; building such institutions under conditions of political contestation and/or embedded climate narratives risk instability and failure. Other contexts are conducive only to more limited climate institution–building.

The way forward lies in recognizing that more limited climate institutions consistent with domestic political context can still help shift political narratives and create conditions for future strategic institutions. In India, the creation of sectoral "missions" for solar and energy efficiency have led to concrete policy gains but have additionally contributed to

greater political acceptance of mitigation actions, albeit within a cobenefits framing. In Brazil, the success of a forest policy, enabled in part by institutional strengthening of the forest bureaucracy, led to creation of climate-focused bodies. Deliberative institutions in South Africa are engaged in creation of a just transition narrative around climate change, which may provide more favorable ground for strategic institution building than would a narrow mitigation framing.

Over time, as climate politics in countries shift, in part through the ratcheting effect that deeper institutionalization of climate change can provide, conditions for mitigation-centric institutions may emerge. The emergence of the German Climate Change Act reflects greater political convergence around climate change, spurred also by public mobilization. Even after these conditions are reached, climate institutions are needed to preserve hard-won political gains. The UK Climate Change Committee's analyses have also been credited with supporting formulation of the country's net-zero target.

These examples suggest that countries develop climate institutions in context-dependent ways. Universal best-practice formulations are of less use than an approach based on understanding how climate politics and climate institutions shape each other. Although choices may therefore be constrained, there is scope for building a virtuous cycle from seemingly modest initial institutions. Both the analytical and normative agenda of climate policy requires further attention to the role of climate institutions. ■


### REFERENCES AND NOTES

1. A. Averchenkova, S. Fankhauser, M. Nachmany, *Trends in Climate Change Legislation* (Edward Elgar Publishing, 2017).
2. G. Iacobuta, N. K. Dubash, P. Upadhyaya, M. Deribe, N. Höhne, *Clim. Policy* **18**, 1114 (2018).
3. K. Harrison, L. M. Sundstrom, *Global Commons, Domestic Decisions: The Comparative Politics of Climate Change* (MIT Press, 2010).
4. J. Meckling, J. Nahm, *Governance* (Oxford) **31**, 741 (2018).
5. M. Mildenberger, *Carbon Captured: How Business and Labor Control Climate Politics* (MIT Press, 2020).
6. M. Aklin, M. Mildenberger, *Glob. Environ. Polit.* **20**, 4 (2020).
7. P. A. Hall, R. C. R. Taylor, *Polit. Stud.* **44**, 936 (1996).
8. J. Mahoney, K. A. Thelen, Eds., *Explaining Institutional Change: Ambiguity, Agency, and Power* (Cambridge University Press, 2010).
9. N. K. Dubash, *Env. Polit.* **30** (sup1), 1 (2021).
10. Varieties of Climate Governance, *Env. Polit.* **30** (suppl. 1), 1 (2021).
11. J. J. L. Candel, R. Biesbroek, *Policy Sci.* **49**, 211 (2016).
12. J. J. Finnegan, *APSA Preprints* 10.33774/apsa-2019-6ft75 (2019).
13. M. Lockwood, C. Kuzemko, C. Mitchell, R. Hoggett, *Environ. Plan. C Polit. Space* **35**, 312 (2017).



### ACKNOWLEDGMENTS

The authors acknowledge financial support from the John D. and Catherine T. MacArthur Foundation that enabled two project workshops (grant 19-1905-153965-CLS).


### SUPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm1157

10.1126/science.abm1157

CLIMATE POLICY

# Can updated climate pledges limit warming well below 2°C?

## Increased ambition and implementation are essential

*By* **Yang Ou[1], Gokul Iyer[1], Leon Clarke[2], Jae Edmonds[1], Allen A. Fawcett[4], Nathan Hultman[2,3], James R. McFarland[4], Matthew Binsted[1], Ryna Cui[2], Claire Fyson[5], Andreas Geiges[5], Sofia Gonzales-Zuñiga[6], Matthew J. Gidden[5,7], Niklas Höhne[8,9], Louise Jeffery[6], Takeshi Kuramochi[8,10], Jared Lewis[11,12,13], Malte Meinshausen[11,12,13], Zebedee Nicholls[11,12,13], Pralit Patel[1], Shaun Ragnauth[4], Joeri Rogelj[7,14], Stephanie Waldhoff[1], Sha Yu[1], Haewon McJeon[1]**

As part of the 2015 Paris Agreement, countries agreed to regularly revisit and enhance their national climate strategies and, every 5 years, to offer new emissions targets in the form of nationally determined contributions (NDCs) (*1*). This year's 26th Conference of Parties provides a waypoint in this updating process as countries have been offering enhanced or completely new NDCs (*2*, *3*) (henceforth, updated pledges) (*4*). We find that compared with the 2015 pledges, the updated pledges suggest a strengthening of ambition through 2030. By calculating probabilistic temperature outcomes over the 21st century for five emissions scenarios (see the figure and table S1), we find that the updated pledges provide a stronger near-term foundation to deliver on the long-term goals of the Paris Agreement of reducing the probability of the worst levels of temperature change this century and increasing the likelihood of limiting temperature change to well below 2°C.

As of 30 September of this year, 121 parties covering 52% of global emissions have submitted updated pledges (*5*). The updated pledges, in most cases, target near-term reductions—through 2025, 2030, or 2035 —and, hence, they could play a critical role in achieving the Paris Agreement's longer-term goals (*1*). Any assessment of the long-term temperature implications of the updated pledges will need to account for changes in our scientific understanding of the climate system, including uncertainties related to the effect of global emissions on global mean surface temperature change. Multiple lines of evidence now suggest a narrower range for climate sensitivity—defined as the steady-state global temperature increase for a doubling of carbon dioxide ($CO_2$)—compared with past assessments. Notably, new evidence suggests higher likelihoods for climate sensitivity values near the middle part of the range identified in previous literature and lower likelihoods for values in the lower and higher parts of the range (*6*).

Several recent policy, societal, and technological developments could also affect future emissions trajectories and, hence, temperature outcomes. Foremost, 33 countries representing 32% of global emissions have communicated official long-term strategies (LTSs) under the United Nations Framework Convention on Climate Change (UNFCCC) or unofficial long-term objectives that outline emission reduction strategies through the midcentury (*7*), and 52 parties representing 54% of global emissions have communicated net-zero emissions targets (*8*). Other developments since 2015—such as the economic slowdown driven by the COVID-19 pandemic and subsequent recovery (*9*), the enactment of new sectoral policies and regulations across the globe (*10*), and technological advances driving rapidly declining costs of renewable energy, electric vehicles, and other clean technologies (see table S2)—have all affected near-term emission trajectories with potential long-term implications.

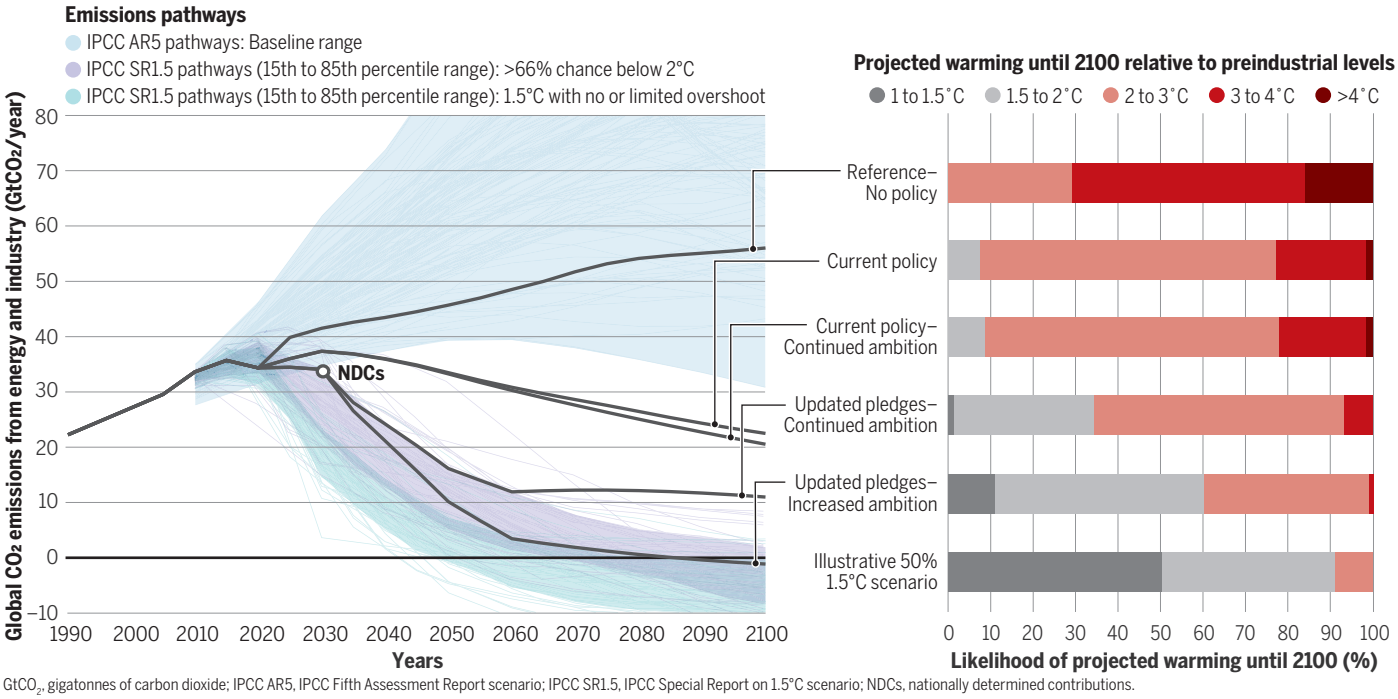
Our analysis accounts for these recent developments. In doing so, this study provides new and updated insights over previous studies that have examined the long-term temperature outcomes of the original 2015 NDCs (see materials and methods, section S1). We also note that our analysis is intentionally designed to facilitate easy comparison with Fawcett *et al.* (*11*)—a key study by a similar team of authors—that evaluated the long-term probabilistic temperature outcomes of the original 2015 pledges using previous versions of the models used in the present study. Comparing with Fawcett *et al.* allows us to quantify how much of the new insights obtained in this study can be attributed to an improved understanding of climate science versus updated emissions trajectories.

### EMISSIONS SCENARIOS

We developed our emissions scenarios (see the figure, left) using the Global Change Analysis Model (GCAM; see materials and methods, section S2). Our scenarios include the impacts of COVID-19 on the economy and the latest technological trends, includ-

## Emissions pathways and temperature probabilities

Five scenarios modeled using the Global Change Analysis Model (GCAM; materials and methods, section S2) are shown (left). Thick gray lines reflect global $CO_2$ emissions from fossil fuel production and utilization and from industrial processes. The scenarios represent deterministic pathways that share all assumptions other than those about the stringency of emissions mitigation policies through 2030 and beyond (materials and methods, section S4). The open circle shows emissions in 2030 with new and updated pledges (*4*). Shaded areas show ranges for baseline, 2°C, and 1.5°C pathways per the Intergovernmental Panel on Climate Change (IPCC) (*15*). Faint lines behind the IPCC ranges represent emissions trajectories that determine the range. Likelihoods of different levels of increase in global mean surface temperature change through 2100 relative to preindustrial levels are shown (right). Temperature outcomes are based on the full suite of greenhouse gas, aerosol, and short-lived species emissions generated by GCAM. The illustrative 50% scenario corresponds to an emission pathway (fig. S1) that achieves a 50% chance of limiting 2100 warming below 1.5°C.




$GtCO_2$, gigatonnes of carbon dioxide; IPCC AR5, IPCC Fifth Assessment Report scenario; IPCC SR1.5, IPCC Special Report on 1.5°C scenario; NDCs, nationally determined contributions.

ing rapidly declining costs of photovoltaic cells, the expansion of electric vehicles, and the availability of $CO_2$ removal measures such as bioenergy in conjunction with carbon capture and storage, afforestation, and direct air capture (see materials and methods, sections S2 and S4; and table S2).

The Reference–No policy, Current policy, and Current policy–Continued ambition scenarios serve as counterfactuals against which to compare scenarios with the updated pledges. The Reference–No policy scenario assumes no new emission mitigation policy through 2100. Compared with Fawcett *et al.* (*11*), the Reference–No policy scenario in this study results in lower emissions throughout the century (see fig. S1). This is primarily driven by policy and technological developments since 2015, such as reduced coal power investment and rapid cost reductions in renewable energy. The Current policy and Current policy–Continued ambition scenarios assume that countries implement current sectoral policies and national climate mitigation measures through 2030 (*5*). Examples of these policies include increased deployment of renewable energy, improved energy efficiency, and additional investment in public transportation.

Beyond 2030, the Current policy scenario assumes that countries continue to decarbonize their economies at the same annual decarbonization rate—defined as the annual rate of improvement in $CO_2$ emissions per unit of gross domestic product (GDP)—that was required to implement the policies between 2015 and 2030. By contrast, the Current policy–Continued ambition scenario assumes a minimum decarbonization rate of 2% per year beyond 2030. In other words, this scenario assumes that countries continue to decarbonize their economies at the same rate as that required to implement the current policies between 2015 and 2030 or increase to 2% per year if the decarbonization rate is below this level.

The Updated pledges–Continued ambition and Updated pledges–Increased ambition scenarios assume that countries achieve the updated pledges through 2030 as stated. Collectively, the updated pledges result in lower $CO_2$ emissions from energy and industry in 2030 by 15% relative to the 2015 pledges (which corresponds to a 15% increase relative to 2005), suggesting a strengthening of ambition since the original 2015 submissions (see fig. S1). Beyond 2030, both scenarios assume that countries without official UNFCCC LTSs achieve the same level of decarbonization rate as the rate between 2015 and 2030 or a minimum rate if their decarbonization rate is below this minimum rate. Countries with LTSs are assumed to follow their LTSs until the target year, followed by the path defined by the minimum decarbonization rate between 2015 and their LTS target year.

Countries with net-zero pledges are assumed to achieve net-zero emissions in the target year and then continue to keep their

[1]Joint Global Change Research Institute, Pacific Northwest National Laboratory and University of Maryland, College Park, MD, USA. [2]Center for Global Sustainability, School of Public Policy, University of Maryland, College Park, MD, USA. [3]US Department of State, Washington, DC 20520, USA. [4]US Environmental Protection Agency, Washington, DC, USA. [5]Climate Analytics, Berlin, Germany. [6]NewClimate Institute, Berlin, Germany. [7]International Institute for Applied Systems Analysis, Laxenburg, Austria. [8]NewClimate Institute, Cologne, Germany. [9]Environmental Systems Analysis Group, Wageningen University and Research, Wageningen, Netherlands. [10]Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, Netherlands. [11]Australian-German Climate and Energy College, The University of Melbourne, Parkville, Victoria, Australia. [12]School of Geography, Earth and Atmospheric Sciences, The University of Melbourne, Parkville, Victoria, Australia. [13]Climate Resource, Northcote, Victoria, Australia. [14]Grantham Institute, Imperial College London, London, UK. Email: hmcjeon@pnnl.gov

GRAPHIC: K. FRANKLIN/*SCIENCE*

emissions constant beyond that year. The two scenarios differ in their assumptions about the minimum decarbonization rate: Whereas the Updated pledges–Continued ambition scenario assumes a 2% minimum decarbonization rate, the Updated pledges–Increased ambition scenario assumes a 5% rate. The 2 and 5% decarbonization rate assumptions have been used previously and are consistent with the average and high rates observed historically (*11*).

We note that our emissions scenarios are consistent with most scenarios in the literature (see the figure). However, many scenarios in the literature result in net-negative emissions in the second half of the century. By contrast, our scenarios assume that countries reduce emissions at most to net-zero because no country has committed to net-negative emissions thus far.

## TEMPERATURE OUTCOMES

We estimate probabilistic temperature outcomes for the emission scenarios using a simple climate model [Model for the Assessment of Greenhouse Gas Induced Climate Change (MAGICC)] in a probabilistic setup (see materials and methods, section S3) that accounts for the latest understanding of climate uncertainties (*12*) (see the figure, right). Because emission trajectories in the Updated pledges scenarios are lower than those in the Reference–No Policy and Current policy scenarios, they result in lower probabilities of temperature change exceeding 4°C in 2100 and higher probabilities of limiting temperature change below 2°C. Under the Updated pledges–Continued ambition scenario, the probability of global mean surface temperature changes exceeding 4°C is virtually eliminated (compared with a ~10% chance with the 2015 NDCs) and the probabilities of limiting temperature change to below 2°C and 1.5°C increase to 34 and 1.5%, respectively, compared with about 8 and 0% with the 2015 pledges. By contrast, if countries scale up mitigation efforts as in the Updated pledges–Increased ambition scenario, not only is the probability of exceeding 4°C virtually eliminated but the probabilities of limiting temperature change below 2°C and 1.5°C also increase to 60 and 11% respectively (compared with 31 and 5% with the 2015 pledges).

## THE ROLE OF IMPROVED CLIMATE SCIENCE

To understand the role of improved climate science since the original Fawcett *et al.* study (*11*), we performed additional simulations by combining the emissions scenarios in this study and those from Fawcett *et al.* with the versions of MAGICC used in the two studies (see fig. S4 and table S1). The version of MAGICC used in this study accounts for the improved understanding of climate science (see materials and methods, section S3). Therefore, our additional simulations help us understand how much of the probabilistic temperature outcomes obtained in this study can be attributed to the improved understanding of climate uncertainties versus updates to emissions trajectories between the two studies.

Results from the additional simulations suggest that most of the differences in probabilistic temperature outcomes below 2°C between Fawcett *et al.* (*11*) and this study can be attributed to the updated emissions trajectories, which are in turn driven by the updated pledges, LTSs, net-zero pledges, and other recent societal, policy, and technological trends. For example, under the Increased ambition scenarios, the updated emission trajectories contribute to 27 percentage points out of a total improvement of about 29 percentage points in the probabilities of limiting temperature change to below 2°C, with two percentage points being contributed by an improved climate science understanding of the narrowed climate sensitivity ranges (see table S3).

## AMBITION NEEDS IMPLEMENTATION

Our analysis is based on a single integrated assessment model driven by a set of assumptions about future population, GDP, technological change, and resources. Other models and assumptions could result in different quantitative results. However, the core qualitative insight of this analysis would remain: if implemented as stated, updated pledges offered by nations in the past year would lay a stronger near-term foundation for reaching global climate goals compared with the 2015 pledges. This underscores that the process for countries to revisit their national climate strategies and offer new targets has generated a notable increase in ambition over the past 6 years. This increase in ambition could spawn even higher ambition in the future: As countries implement updated pledges, the widespread development and deployment of low-carbon technologies, expansion of policy instruments across all sectors and levels of government, and enhanced international cooperation could drive down costs and enable more cost-effective and stronger mitigation in the long term (*13*). Nevertheless, to the extent that countries continue to have opportunities to further enhance their ambition—both before and after 2030— this would provide additional and needed ambition to substantially improve climate outcomes and increase the likelihood to limit warming to less than 1.5°C.

Ultimately, realizing the long-term climate benefits described in this study will require putting words into action by implementing these newer and enhanced targets. This could entail substantial political, financial, institutional, social, and behavioral challenges (*14*). Surmounting these challenges with policies and measures and successfully achieving the targets would substantially reduce global climate risks, essentially avoiding the worst outcomes from climate change and substantially improving the probability of limiting global warming to well below 2°C. ■

## REFERENCES AND NOTES


1. UNFCCC, The Paris Agreement (2021); https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement.
2. NDC Registry, The latest submissions (2021); https://www4.unfccc.int/sites/ndcstaging/Pages/LatestSubmissions.aspx.
3. The White House, Leaders summit on climate summary of proceedings (2021); www.whitehouse.gov/briefing-room/statements-releases/2021/04/23/leaders-summit-on-climate-summary-of-proceedings/.
4. UNFCCC, Glasgow Climate Change Conference (2021); https://unfccc.int/process-and-meetings/conferences/glasgow-climate-change-conference.
5. Climate Action Tracker, CAT climate target update tracker (2021); https://climateactiontracker.org/climate-target-update-tracker/.
6. S. Sherwood *et al.*, *Rev. Geophys.* **58**, e2019RG000678 (2020).
7. Climate Watch, Explore long-term strategies (LTS) (2021); www.climatewatchdata.org/lts-explore.
8. Climate Watch, Net-zero tracker (2021); www.climatewatchdata.org/net-zero-tracker.
9. International Energy Agency, Economic impacts of Covid-19 (2021); www.iea.org/reports/global-energy-review-2021/economic-impacts-of-covid-19.
10. J. Setzer, R. Byrnes, "Global trends in climate change litigation: 2020 snapshot" (Policy report, Grantham Research Institute on Climate Change and the Environment and Centre for Climate Change Economics and Policy, London School of Economics and Political Science, 2020).
11. A. A. Fawcett *et al.*, *Science* **350**, 1168 (2015).
12. Z. Nicholls *et al.*, *Earth's Future* **9**, e2020EF001900 (2021).
13. G. Iyer, L. Clarke, J. Edmonds, N. Hultman, H. McJeon, *Energy Policy* **86**, 493 (2015).
14. D. Cullenward, D. G. Victor, *Making Climate Policy Work* (Polity Press, 2020).
15. J. Rogelj *et al.*, in "Global warming of 1.5°C: An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty," V. Masson-Delmotte *et al.*, Eds. (Intergovernmental Panel on Climate Change, 2018).


## ACKNOWLEDGMENTS


Y.O. and G.I. contributed equally to this work. The views and opinions expressed in this paper are those of the authors alone and do not necessarily state or reflect those of the Environmental Protection Agency, the Department of State, or the US government, and no official endorsement should be inferred. Current policy projections and update pledges used in this analysis were derived by Climate Action Tracker, with thanks to all project team members who contributed. **Funding:** This work was funded by US Environmental Protection Agency IAA DW-089-92460001 (Y.O., G.I., M.B., J.E., P.P., S.Y., S.W., and H.M.) and German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety grant 16_I_291_Global_A_CAT (C.F., A.G., S.G.-Z., M.J.G., N.H., L.J., and T.K.). **Data and materials availability:** GCAM is an open-source community model available at https://github.com/JGCRI/gcam-core/releases. Source code and data associated with this analysis are available at https://doi.org/10.5281/zenodo.5553171.


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abl8976


10.1126/science.abl8976

NEUROSCIENCE

# Embodying consciousness

## Bodies and feelings are foundational to conscious experience, argues a neuroscientist

*By* **Alex Gomez-Marin**



The brain and body are inextricably linked in the manifestation of consciousness, maintains Antonio Damasio.

PHOTO: SCIEPRO/ISTOCK.COM

Most neuroscientists conceive of consciousness as a phenomenon that emerges somewhere in the brain. In *Feeling and Knowing: Making Minds Conscious*, neuroscientist Antonio Damasio presents a different view, one that centers on the body and feeling. The conscious mind, he argues, is not only in the head but also in the heart, the gut, and indeed the rest of the body. The brain may be necessary for consciousness, he concedes, but it is not sufficient.

Noting that the entirety of one's nervous system is situated within one's body, Damasio argues that "feelings are not conventional perceptions of the body but rather *hybrids*, at home in both body and brain." In the intimate exchange between neurons and flesh, he seeks to bridge Descartes's split.



**Feeling and Knowing: Making Minds Conscious**
*Antonio Damasio*
Pantheon, 2021. 256 pp.

To elucidate human consciousness, Damasio advances piecemeal, following the steps of evolution to unfurl the natural history of the mind. His narration gives primacy to a third-person chronological account rather than a first-person phenomenological one—a framework that is not new but is powerful. Sensing comes first, he proposes, followed by mind and then feeling, which in turn opens the way to consciousness.

All living organisms, from bacteria to Bach, are capable of basic sensing, allowing them to survive and thrive in Earth's ever-changing environments. The simplest of life-forms can perform wonders. They solve basic problems with "intelligent but unminded competences," Damasio claims, never holding their own homeostatic processes as "images" in mind. This latter ability requires feelings, he argues, by which we "*experience in mind* a process that clearly takes place in the *physical realm of the body*."

Feelings let a mind know, by direct contact, about the state of the body in which it is situated. In the encounter, the object perceived and the perceiving subject literally meet. Feelings and emotions are related but different. Whereas emotions are physical reactions of the body (dry mouth, goose bumps), feelings involve the mental awareness that accompanies such changes (fear, joy). Emotions are public, feelings private.

Damasio is not particularly interested in neural correlates of thoughts or feelings. Nor does he spend much time on brain structures and their putative role as the neural basis of consciousness. What makes us conscious of the images that brain regions fabricate, he insists, is "the *addition of knowledge certifying the ownership of those images*" as provided by "homeostatic feelings."

The thesis of the book is this: Only by establishing a scientific understanding of feeling can we begin to understand what it means to be conscious. The chief corollary: The brain alone does not experience make.

One might feel underwhelmed by such an apparently modest exposition, especially when compared with the expectations raised by the so-called "hard problem" of consciousness (why and how do physical processes in the brain give rise to conscious experience?). Damasio disagrees with the "in the brain" part of the formulation. And yet, even when spread to the body, the question still looks rather insoluble, if not ill-posed.

For Damasio, there is no mystery. There is only the problem that consciousness has been deemed a separate, special phenomenon. Consciousness, in his estimation, is biological business as usual. As he puts it, consciousness "calls for a rearrangement of the furniture of mind" but not for a rehabilitation of the house of science or its philosophical foundations.

In a section titled "The Problem of Consciousness," Damasio loses his measured temper and balanced erudition. He rushes to ridicule panpsychism, the position that consciousness pervades the cosmos. "Why stop at living things," he asks, referring to the work of two prominent consciousness researchers who, he says, believe that "the universe and all the stones in it" might be conscious as well. Indeed, the recent revival of the idea that mind is already in matter, rather than a tardy emanation from it, has opened a Pandora's box among consciousness researchers, angering some die-hard materialists while leading others out of conceptual dead ends (*1*).

While he rejects human exceptionalism and brain centrism, Damasio remains neuron-centric on the question of consciousness: Brains need bodies, but feelings need neurons. Thus, life-forms without nervous systems cannot be conscious. Such a veto is not only conservative but also confusing, as he postulates that plants cannot have feelings, while flirting with the idea that machines may.

Altogether, Damasio's concise, precise, and lucid prose effectively conveys the core insight he has distilled over decades (*2*): that affect—encompassing emotions, feelings, motivations, and moods—is central to understanding what we do, how we think, and who we are. ■

The reviewer is at the Instituto de Neurociencias, Consejo Superior de Investigaciones Científicas–Universidad Miguel Hernández de Elche, Alicante, Spain. Email: agomezmarin@gmail.com

### REFERENCES AND NOTES

1. P. Goff, *Galileo's Error: Foundations for a New Science of Consciousness* (Pantheon, 2019).
2. A. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (Putnam, 1994).

10.1126/science.abm6378

IMMUNOLOGY

# Immunology meets the masses

## A YouTuber's vividly illustrated immune system explainer delights and informs

*By* **Daniel M. Davis**

In 1882, in a private laboratory in Sicily, Ukrainian microbiologist Ilya Mechnikov pierced starfish larvae with a thorn from a nearby bush and watched as some of the organism's cells moved toward the injury. The cells, he hypothesized, were moving there to engulf any disease-causing microbes that might dare to enter the wound. The cells Mechnikov observed were later named phagocytes—"phago" is Greek for "eating" and "cyte" is Greek for "cell"—and the cell best able to "eat" microbes came to be called the macrophage ("the big eater"). For this discovery, Mechnikov won a Nobel Prize in 1908.

Fast-forward a little more than a century to today, when immune system enthusiasts can celebrate the microbiologist's visionary breakthrough with a macrophage enamel pin (*1*) or a poster-sized illustration of the "immune system battlefield" (*2*). Both products are available on the retail site affiliated with the popular YouTube channel Kurzgesagt (German for "in a nutshell"), which posts easy-to-understand animations on a variety of scientific subjects, from black holes to dinosaurs, for an audience of more than 16 million subscribers. With little formal science education, the site's creator, Philipp Dettmer, has become one of the world's most influential science communicators, thanks in great part to the clear scripts and vivid art and design of the site's videos.

The nitty gritty of the immune system is of particular interest to Dettmer, who was diagnosed with cancer at the age of 32. Knowing that it would be difficult to capture the complexity of this system in a 10-minute animation, he has written a book, *Immune*, that charts the vast array of different immune cells, what they do, the way they communicate with each other, and what happens when things go wrong. The book's colorful illustrations are reminiscent of those that appear in Kurzgesagt's video explainers and are among the best I have seen in any scientific text.

Throughout the book, Dettmer uses wonderful, original metaphors to explain immune function, which build—layer upon layer—as the details of this complex system unfold. In chapter 17, for example, he likens the diversity of immune recognition to planning a fantastical dinner party for millions of guests, each of whom is very particular about what they like. From basic kitchen ingredients, which represent gene segments, he shows how 3,577,446 different dinners can be made.

Along the way, Dettmer covers many profound immunological principles, such as how a body is able to detect foreign entities—from bacteria to transplanted organs—and how it can recognize a virus that has only recently come into existence. Over the course of two chapters, with the help of a vivid illustration of the B cell "career," likening the activation process to online encryption, he lucidly describes how antibodies are produced by B cells and are selected to target a particular pathogen. "B Cells are very dangerous cells," he writes, "so they need a strict two-factor authentication to be truly activated."

Some experts might be uncomfortable with the colorful language Dettmer regularly uses. Receptor proteins on cell surfaces are, in Dettmer's phrasing, "the noses of your cells," and when they activate a macrophage, it gets "angry," sending out "octopus-like arms" to engulf bacteria. In another passage, he likens different types of T cells to the classes and subclasses of characters used in the role-playing game Dungeons & Dragons. But make no mistake, Dettmer's chatty and accessible tone is intentional, helping readers to process an enormous amount of complex information. The outcome is reminiscent of Bill Bryson's *The Body*—fast-paced and full of wonder.

There is relatively little in Dettmer's book about the people who shaped our understanding of the immune system or about the experiments they performed to glean this information. There is a place for such stories, but this book's narrative is driven instead by the timeline of an immune response, freeing Dettmer from the jargon and precision that might otherwise have weighed down the text.

Antibodies, T cells, and B cells were once rarely discussed outside of research labs and scientific talks, but thanks to COVID-19, everyone, it seems, has become an immune system enthusiast. Now, more than ever, there is a hunger for details about this complex biological process, and Dettmer's book is the feast we have been waiting for. Although he could not have anticipated how his research would be made accessible to the public more than a century later, I cannot help but think Mechnikov would be delighted at Dettmer's efforts. ■

The reviewer is at the Lydia Becker Institute of Immunology and Inflammation, The University of Manchester, Manchester M13 9PL, UK, and the author of *The Secret Body* (Princeton Univ. Press, 2021) and *The Beautiful Cure* (Penguin Random House, 2018). Email: daniel.davis@manchester.ac.uk

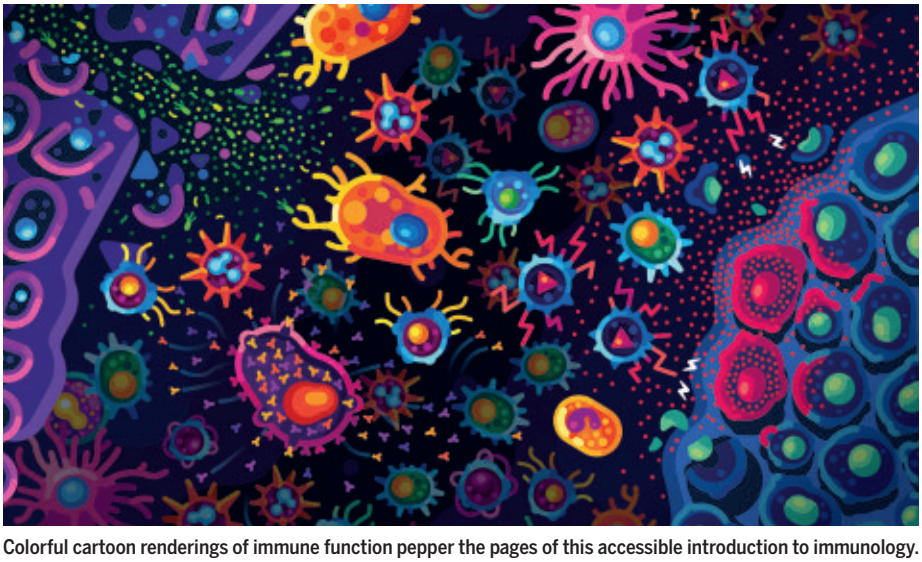


IMAGE: PHILIP LAIBACHER

Colorful cartoon renderings of immune function pepper the pages of this accessible introduction to immunology.

**Immune: A Journey into the Mysterious System That Keeps You Alive**
*Philipp Dettmer*
Random House, 2021.
368 pp.




### REFERENCES AND NOTES

1. Macrophage Enamel Pin, https://shop-us.kurzgesagt.org/products/macrophage-enamel-pin.
2. Immune System Battlefield Poster, https://shop-us.kurzgesagt.org/products/immune-battlefield-poster.

10.1126/science.abm0134

Gordon Research Conferences

# Advancing the Frontiers of Science and a Commitment to Diversity

## Did you know that GRC offers two diversity fellowship programs for eligible applicants?

### The Carl Storm Underrepresented Minority Fellowship (CSURM)

This program was established by the GRC Board of Trustees to honor Director Emeritus Carl Storm's many years of service to GRC and to support his commitment to increasing diversity at the conferences. This fellowship program provides an opportunity for URM graduate students, postdocs, faculty and scientists to receive funding assistance to attend their first GRC.

### The Carl Storm International Diversity Fellowship

This program was established by the GRC Board of Trustees to honor Director Emeritus Carl Storm's many years of service to GRC and to support his commitment to increasing diversity at the conferences. This fellowship program allows conference chairs to nominate graduate students and postdocs from China, Africa, India, Central America, South America, Russia, and Ukraine. Funding is available on a first-come, first-served basis.







**Apply now for 2022 conferences!**
By completing a brief form, eligible applicants can obtain additional information about these funding opportunities!



**Visit www.grc.org today!**



China's karst cave ecosystems, home to myriad unique species, remain inadequately protected.

PHOTO: MA XIPING/XINHUA/GETTY IMAGES

*Edited by* **Jennifer Sills**

# Protect China's karst cave habitats

China's 1.9 million-km$^2$ karst area is one of the most threatened biodiversity hotspots in the world (*1*). Karst caves have unique habitats with dim light, limited nutrient supply, high humidity, and relatively low temperature fluctuation (*2*), and most of the organisms that live in the caves have very narrow distributions and small populations. Although small parts of the karst cave region in South China were designated a United Nations Educational, Scientific, and Cultural Organization (UNESCO) World Heritage Site in 2007 (*3*) and 2014 (*4*), most karst caves in China do not fall within the current system of national natural reserves. Without effective protection, starting with the designation of the caves as national reserves and classification of species as endangered, the many species that live only in karst caves may not survive.

The karst caves are home to entire genera of some organisms, as well as recently discovered species and likely countless species yet to be discovered. Almost all of the 49 known species of the golden-line barbel genus (*Sinocyclocheilus*) in the cyprinid fish family (Cyprinidae) live in or around one or a few caves (*5*). More than 100 newly described plant species are endemic to karst caves, including *Chiritopsis* spp. and *Petrocosmea* spp. in Gesneriaceae and *Polystichum* spp. in Dryopteridaceae (*6*–*8*). Despite their small populations and limited habitats, many of these species have not been listed by the International Union for Conservation of Nature (IUCN) as Critically Endangered because assessments are outdated or have not been done at all.

Karst cave organisms are highly adapted to the environments and restricted in their ranges. Increased human exploitation of natural resources for herbal remedies, mining activities (*8*), tourism (*9*), fishing, and agricultural activities has led to habitat and vegetation degradation in karst caves, putting endemic and rare species at risk of extinction. With no charismatic flagship species, karst cave ecosystems receive little governmental support. No governmental agency or nongovernmental organization focuses on these ecosystems (*10*).

Protecting karst cave species requires urgent and coordinated action from the government and the public. China should pass legislation to protect the region and make a government body, such as the National Forestry and Grassland Administration, accountable for success. The legislation should include the designation of a karst national reserve that includes both caves in the UNESCO world heritage sites and those outside their borders. A national reserve would draw funding to manage and restore habitats and place limits on nearby human activities. In addition, the IUCN should assess or update previous assessments of karst cave species. China should raise public awareness to ensure that biological species in the karst region are added to the latest list of national key protected wild organisms (*11*). We must take action to protect the karst cave ecosystems if we want a chance to mitigate biodiversity loss.

**Yi-Fan Duan[1], Meng Li[1], Ke-Wang Xu[1], Liang Zhang[2], Li-Bing Zhang[3,4]***

[1]Co-Innovation Center for Sustainable Forestry in Southern China & Key Laboratory of National Forestry and Grassland Administration on Subtropical Forest Biodiversity Conservation, College of Biology and the Environment, Nanjing Forestry University, Nanjing 210037, China. [2]Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China. [3]Missouri Botanical Garden, St. Louis, MO 63110, USA. [4]Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China.
*Corresponding author.
Email: libing.zhang@mobot.org

## REFERENCES AND NOTES

1. D. X. Yuan, *Karstology of China* (Geological Publishing House, Beijing, 1994) [in Chinese].
2. H. Ren, *et al.*, *Ecol. Indic.* **120**, 106947 (2021).
3. UNESCO World Heritage Convention, "South China Karst" (2007); https://whc.unesco.org/en/list/1248/.
4. UNESCO World Heritage Convention, "Twenty-six new properties added to World Heritage List at Doha meeting" (2014); https://whc.unesco.org/en/news/1162/.
5. Y. H. Zhao, C. G. Zhang, *Endemic fishes of Sinocyclocheilus (*Cypriniformes: Cprinidae*) in China: Species Diversity, Cave Adaptation, Systematics and Zoogeography* (Science Press, Beijing, 2009) [in Chinese].
6. W.-B. Xu, Y. Liu, H.-S. Gao, *Novon* **19**, 559 (2009).
7. M. Q. Han *et al.*, *Kew Bull.* **74**, 22 (2019).
8. Y. F. Duan, N. T. Lu, L. Zhang, L.-B. Zhang, *Phytotaxa* **313**, 296 (2017).
9. A. K. Monro, N. Bystriakova, L. Fu, F. Wen, Y. Wei, *PLOS One* **13**, e0190801 (2018).
10. T. Whitten, *J. Appl. Ecol.* **46**, 520 (2009).
11. Ministry of Agriculture and Rural Affairs, National Forestry and Grassland Administration of China, "List of national key protected wild plants" (2021); www.forestry.gov.cn/main/5461/20210908/162515850572900.html [in Chinese].

10.1126/science.abm5389

# Biodiversity loss due to more than climate change

With the UN Climate Change Conference (COP26) taking place this week, climate change is on everyone's radar. Unfortunately, the myopic lens of climate change has both diminished and distorted our understanding of the biodiversity crisis (*1*). The conservation

Where Science Gets Social.

AAAS.ORG/COMMUNITY

AAAS' Member Community is a one-stop destination for scientists and STEM enthusiasts alike. It's "Where Science Gets Social": a community where facts matter, ideas are big and there's always a reason to come hang out, share, discuss and explore.

Member COMMUNITY AAAS

of biodiversity should be approached as a solution to the climate crisis, rather than an ancillary victim. Losses of biodiversity are increasingly corrosive, and they are driven by traditional threats that dwarf the potential impacts of climate change (*2*).

The principal drivers of species extinctions vary by taxonomic group and location (*3*, *4*), but evidence shows that conventional anthropogenic factors (habitat loss, land-use change, and deforestation) are consistently more important than the contemporary effects of climate change (*3*–*5*). Land use–driven loss of biodiversity will have the most pronounced impact in the Tropics and Temperate Zones, where biodiversity is greatest and land-use change most rapid. In marked contrast, the impacts of climate change will initially be felt most sharply in polar regions, where biodiversity is unique but much less abundant (*5*). At current rates of land-use change, it is unlikely that much biodiversity will be left by the time climate change severely affects the Tropics (*5*).

Paradoxically, tropical forests and savannahs are the only economically viable buffer against climate change (*6*, *7*). To save biodiversity and reverse climate change, we must tackle habitat destruction and exploitation of species rather than believing that reducing fossil fuel use will solve both problems.

**Andrew Dobson[1,2]*, Zeke Rowe[3], Joel Berger[4,5], Philippa Wholey[3], Tim Caro[3,6]**
[1]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA. [2]Santa Fe Institute, Santa Fe, NM 87501, USA. [3]School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK. [4]Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523, USA. [5]Wildlife Conservation Society, Bronx, NY 10460, USA. [6]Center for Population Biology, University of California, Davis, CA 95616, USA.
*Corresponding author.
Email: dobber@princeton.edu

### REFERENCES AND NOTES

1. P. Legagneux *et al.*, *Front. Ecol. Evol.* **5**, 175 (2018).
2. D. Tilman *et al.*, *Nature* **546**, 73 (2017).
3. S. Ducatez, R. Shine, *Conserv. Lett.* **10**, 186 (2017).
4. Ł. Tracewski *et al.*, *Conserv. Biol.* **30**, 1070 (2016).
5. W. Jetz, D. Wilcove, A. P. Dobson, *PLOS Biol.* **5**, e157 (2007).
6. G. B. Bonan, *Science* **320**, 1444 (2008).
7. C. Kremen *et al.*, *Science* **288**, 1828 (2000).

10.1126/science.abm6216

## Science-informed salmon conservation strategies

Human population growth and activities over the past century have broadly affected marine fish biodiversity, with several species declining or near extinction (*1*). Many of the world's salmon species are particularly vulnerable because they rely on a diverse array of habitats for survival and reproduction (*2*), which compromises their ability to adapt to environmental changes (*3*). Salmon populations throughout the world face unprecedented threats to their survival and viability, including habitat degradation, climate change, pathogens, illegal trade, and overfishing (*4*–*9*). Protecting these fish and their ecosystems will require scientific strategies and technology on both land and sea.

Global assessments can help to map the intact Northern Hemisphere rivers that serve as stronghold habitats for salmon, better informing where whole watershed-scale conservation investments should be made. Aquatic telemetry acoustic tags and environmental sensors can also improve our understanding of salmon migration patterns and habitat use (*10*). Better information about the habitats on which salmon depend throughout their life cycle and changes to their distribution resulting from drought, food web fluctuations, and river fragmentation will enable more effective and comprehensive conservation efforts (*2*, *4*). Targeted high-throughput molecular screening of infective agents may aid in addressing potential fish health problems arising from existing and unknown pathogens (*6*, *11*). Genotype technology can help enforce legal trade by identifying fish that are being sold illegally, which in turn could mitigate overexploitation of sensitive fisheries (*2*, *4*, *9*).

Together, conducting global research and applying technology to monitoring (*6*, *9*, *10*) and enforcement could help identify and address local combinations of threats. If successful, the protection of vulnerable salmon species would serve as a model for similarly complex species that require specialized conservation efforts (*12*).

**Michael S. Bank[1,2]*, Christian Sonne[3], Sophia V. Hansson[3,4], Matthias C. Rillig[5,6]**
[1]Institute of Marine Research, Bergen, Norway. [2]University of Massachusetts Amherst, Amherst, MA, USA. [3]Aarhus University, Roskilde, Denmark. [4]Laboratoire Ecologie Fonctionnelle et Environnement, Centre National de la Recherche Scientifique, Toulouse, France. [5]Institute of Biology, Freie Universität Berlin, 14195 Berlin, Germany. [6]Berlin-Brandenburg Institute of Advanced Biodiversity Research, 14195 Berlin, Germany.
*Corresponding author. Email: michael.bank@hi.no

### REFERENCES AND NOTES

1. H. Y. Yan *et al.*, *Sci. Adv.* **7**, eabb6026 (2021).
2. L. G. Crozier *et al.*, *Evol. Appl.* **1**, 252 (2008).
3. R. F. Sage, *Glob. Change Biol.* **26**, 3 (2020).
4. S. Wilson, "California's disappearing salmon," *Washington Post* (2021).
5. S. Castle, "As wild salmon decline, Norway pressures its giant fish farms," *The New York Times* (2017).
6. G. J. Mordecai *et al.*, *Sci. Adv.* **7**, eabe2592 (2021).
7. O. Torrissen *et al.*, *J. Fish Dis.* **36**, 171 (2013).
8. M. Krkosek *et al.*, *Science* **318**, 1772 (2007).
9. R. Ebersole, "Why you might not be getting the salmon you paid for," *National Geographic* (2021).
10. N. E. Hussey *et al.*, *Science* **348**, 1221 (2015).
11. A. W. Bateman, *Sci. Rep.* **11**, 3466 (2021).
12. C. S. Sonne *et al.*, *Science* **372**, 1271 (2021).

10.1126/science.abm4703




Body cameras are seen as a new means of documenting police-civilian interactions, but their use—and, at times, convenient disuse—can reflect bias as well.

PHOTO: SEBASTIAN KAHNERT/PICTURE ALLIANCE/GETTY IMAGES

PRIZE ESSAY

SOCIAL SCIENCES

# Revealing racial bias

## Causal inference can make sense of imperfect policing data

By **Dean Knox**

For decades, high-profile incidents of excessive force against minorities have fueled allegations of abusive policing in the United States and demands for reform. Yet one of the main drivers of today's policing crisis remains unchanged: massive racial disparities in law enforcement.

Courts and city councils struggle to measure the severity of racial bias in policing, let alone to identify the means to address such bias. Solutions are difficult to identify because the policing data landscape is fraught with inconsistent record-keeping and incomplete, task-specific datasets. In examining the dizzying array of analytic approaches used in this context, my colleagues and I found many to be mutually incompatible or even misleading, producing contradictory results and impeding knowledge accumulation (*1*–*3*). Making use of formal statistical frameworks for drawing causal inferences (*4*, *5*)—that is, reliable conclusions about how and why events occur, given explicitly stated assumptions and observed data—we have shown the importance of measuring and accounting for the long chain of events from officer deployment to contact, detainment, and violence.

Policing presents substantial challenges for statistical analysis. For almost 100 years, police agencies have been the sole source of data on police-civilian interactions (*6*). Administrative datasets only document incidents that were required to be reported—historically, violent and/or property crimes. Agencies now increasingly also report stops, frisks, arrests, and uses of force against civilians. Still, only a smattering of interactions are documented.

Why does this matter? The application of off-the-shelf statistical methods to "datasets of convenience"—that is, datasets focused solely on obtainable information, without considering what variables or observations might not be obtainable—often leads to fragile conclusions hinging on implausible or unstated assumptions (*7*). Similar challenges arise when analyzing datasets acquired through open record requests (*1*) or labor-intensive crowdsourcing (*2*).

An increasingly important subfield of statistics and computer science, causal inference, aims to address this issue. Causal inference focuses on a deceptively simple question: Where do our data come from? From this starting point, we can build frameworks (*4*, *5*, *8*, *9*) for analyzing datasets contaminated by inaccuracies, selective reporting, and omitted variables. Rather than ignoring imperfections in our data, we ask what the range of possible interpretations is and what new information must be collected to further narrow this list.

In disciplines facing similar challenges, such as medicine, where patients rarely

Operations, Information and Decisions Department, University of Pennsylvania, Philadelphia, PA 19014, USA. Email: dcknox@upenn.edu

NOMIS & Science YOUNG EXPLORER AWARD

visit doctors unless something has gone wrong, the causal-inference toolkit has proven invaluable for academics and policy-makers (*10*). Yet in policing research, careful causal analysis remains the exception, not the rule.

Failure to account for unobserved causal processes has frequently undermined our understanding of policing. Decades of research have analyzed police detainment records to compare treatment of white and minority detainees (*11*, *12*). Scholars frequently conclude that there is surprisingly little evidence of discrimination. However, our work shows that without accounting for an officer's initial decision to detain, an act that may itself reflect bias, analysts have little hope of recovering accurate estimates of discrimination (*1*).

This initial decision is relevant because minorities may be detained in situations where white individuals would not be detained, for example, jaywalking encounters. By contrast, individuals engaged in behaviors like assault will typically be detained irrespective of race. Comparisons of minority detainments to white detainments therefore fail to achieve the apples-to-apples conditions needed to demonstrate disparate treatment (*7*). Studies that ignore the underlying processes inherent in the formation of these datasets can thus sharply underestimate discrimination in police violence or mask it entirely.

To aid others studying policing datasets, we have developed techniques for estimating bounds—best- and worst-case levels of discrimination—consistent with the available, imperfect data. Bounding approaches consider the ways that unobserved phenomena may manifest in a dataset, for example, the decision to detain; thus, they can yield conclusions that are robust to those unobserved parameters. Similar causal bounding approaches are widely used in other domains, like epidemiology (*13*), but have appeared in policing only recently (*3*).

Crowdsourced datasets of police killings (*14*, *15*) increasingly provide an objective check on officer self-reports. Yet our work shows that challenges arise even when studying civilian-collected datasets, because most nonviolent incidents are not recorded. The consequences of this lack of data were recently highlighted by a flawed, high-profile study that erroneously concluded "if anything, [we] found anti-White disparities" in police violence (*16*). The study, which went on to dismiss widely proposed police diversification reforms, bolstered antireformers (*17*), but our analysis showed that its claims were mathematically baseless, and the paper was ultimately retracted (*2*, *18*).

Our own study of diversity in the Chicago Police Department illustrates the difficulty of conducting research in this area and how causal reasoning can be of aid. Using data on 2.9 million officer shifts, we found that officers from marginalized groups engage in substantially less violence, particularly toward minority civilians. These officers also engage in less discretionary enforcement for minor offenses (*19*). Such an analysis is feasible thanks to detailed patrol records that let us compare officers rotating through similar places at similar times with similar numbers and types of encounters. When simulating typical data constraints in prior work, which relied on enforcement records alone, we found that a failure to account for uneventful shifts could lead to inaccurate inferences based on an incomplete dataset.

As concerns about policing have exploded, policy-makers and the general public are turning to academics to make sense of its complexities. However, a literature filled with contradictory results has often led scholars to say that "we simply do not know" (*20*). Our causal analyses of common research methods reveal the roots of this confusion. More importantly, they suggest a path forward as the discipline grows.

We are now developing a system for reconstructing encounter timelines from body-worn camera footage, building on my prior work on conversation and vocal-tone analysis in audio data (*21*, *22*) and incorporating our collaborators' expertise in computer vision (*23*, *24*). We continue to develop new causal methods for enabling the systematic evaluation of officers and agencies by fusing existing, imperfect datasets (*25*). Above all, we demonstrate that a shared language—a coherent causal framework—is needed to evaluate evidence, adjudicate contradictory claims, and accumulate knowledge in this critically important domain (*3*). ■



**WINNER**

**Dean Knox**

Dean Knox received his undergraduate degree from the University of Illinois at Urbana-Champaign and a PhD from the Massachusetts Institute of Technology. After completing his postdoctoral fellowship at Microsoft Research, he joined the Politics Department at Princeton University as an assistant professor in 2018. Dean is currently an assistant professor at the Operations, Information and Decisions Department at the Wharton School of the University of Pennsylvania, where he co-founded the Research on Policing Reform and Accountability group with Jonathan Mummolo. His research develops statistical methods for analyzing imperfect social science data.



**FINALIST**

**Geoffrey Supran**

Geoffrey Supran received his undergraduate degree from Trinity College, University of Cambridge, and a PhD from the Massachusetts Institute of Technology (MIT). After completing joint postdoctoral fellowships at MIT and Harvard University, Geoffrey became a research fellow in the Department of the History of Science at Harvard in 2019 and also Director of Climate Accountability Communication at the Climate Science Social Network in 2020. His research focus is the quantitative historical analysis of climate change disinformation and propaganda by fossil fuel interests.

**science.org/doi/10.1126/science.abm3434**

**REFERENCES AND NOTES**

1. D. Knox, W. Lowe, J. Mummolo, *Am. Polit. Sci. Rev.* **114**, 619 (2020).
2. D. Knox, J. Mummolo, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1261 (2020).
3. D. Knox, J. Mummolo, *J. Polit. Inst. Polit. Econ.* 2020, **341** (2020).
4. J. Pearl, *Biometrika* **82**, 669 (1995).
5. D. B. Rubin, *J. Educ. Psychol.* **66**, 688 (1974).
6. International Association of Chiefs of Police, Uniform Crime Reports (1930).
7. D. Knox, W. Lowe, J. Mummolo, SSRN 3940802 (2020).
8. P. W. Holland, *J. Am. Stat. Assoc.* **81**, 945 (1986).
9. J. Neyman, "Sur les applications de la theorie des probabilites aux experiences agricoles: Essai des principes," thesis, Univ. of Warsaw (1923).
10. J. M. Robins, *Math. Model.* **7**, 1393 (1986).
11. R. G. Fryer Jr., *J. Polit. Econ.* **127**, 1210 (2019).
12. D. A. Smith, C. A. Visher, L. A. Davidson, *J. Crim. Law Criminol.* **75**, 234 (1984).
13. C. E. Frangakis, D. B. Rubin, *Biometrics* **58**, 21 (2002).
14. D. B. Burghart, Fatal encounters. Database (2012); https://fatalencounters.org/.
15. *Washington Post*, Fatal force. Database (2015); www.washingtonpost.com/graphics/investigations/police-shootings-database/.
16. D. J. Johnson, T. Tress, N. Burkel, C. Taylor, J. Cesario, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 15877 (2019).
17. H. Mac Donald, Written testimony before the Committee on the Judiciary of the United States House of Representatives, Oversight Hearing on Policing Practices, 19 September 2019; https://docs.house.gov/meetings/JU/JU00/20190919/109952/HHRG-116-JU00-Wstate-MacDonaldH-20190919.pdf.
18. D. Knox, J. Mummolo, "Prominent claims that policing is not racially biased rest on flawed science," Medium (2020); https://medium.com/@jon.mummolo/prominent-claims-that-policing-is-not-racially-biased-rest-on-flawed-science-6f66535dc7e5.
19. B. A. Ba, D. Knox, J. Mummolo, R. Rivera, *Science* **371**, 696 (2021).
20. D. A. Sklansky, *J. Crim. Law Criminol.* **96**, 12091244 (2005).
21. D. Knox, C. Lucas, *Am. Polit. Sci. Rev.* **115**, 649 (2021).
22. S. A. Mehr *et al.*, *Science* **366**, eaax0868 (2019).
23. O. Russakovsky *et al.*, *Int. J. Comput. Vis.* **115**, 211 (2015).
24. S. Yeung *et al.*, *Int. J. Comput. Vis.* **126**, 375 (2018).
25. G. Duarte, N. Finkelstein, D. Knox, J. Mummolo, I. Shpitser, arXiv:2109.13471 (2021).

10.1126/science.abm3432

PHOTOS: (TOP TO BOTTOM) COURTESY OF DEAN KNOX; STEPHANIE MITCHELL/HARVARD UNIVERSITY

PRIZE ESSAY



**FINALIST**
**Geoffrey Supran**
Geoffrey Supran received his undergraduate degree from Trinity College, University of Cambridge, and a PhD from the Massachusetts Institute of Technology (MIT). After completing joint postdoctoral fellowships at MIT and Harvard University, Geoffrey became a research fellow in the Department of the History of Science at Harvard in 2019 and also Director of Climate Accountability Communication at the Climate Science Social Network in 2020. His research focus is the quantitative historical analysis of climate change disinformation and propaganda by fossil fuel interests.
science.org/doi/10.1126/science.abm3434

SOCIAL SCIENCES

# Fueling their own climate narrative

## Using techniques from big data to decode Big Oil's climate change propaganda

*By* **Geoffrey Supran**

On Earth Day, 2021, New York City sued ExxonMobil and other firms "for systematically and intentionally deceiving New Yorkers" about climate change with "false advertising and greenwashing" (*1*). New York joins dozens of cities and states in accusing the fossil fuel industry of endangering the public by knowingly denying the science of climate change for decades (*2*). Increasingly, these legal challenges seek to hold fossil fuel companies accountable not just for outright climate denial (*3*), but for subtler forms of "deceptive marketing" that have served to delay action on climate change (*4*).

Naomi Oreskes and I recently published the first quantitative assessment of how ExxonMobil has used language to subtly yet systematically shape public discourse in misleading ways (*5*). By nature and by design, climate "delayism" is more insidious than denialism. To uncover these subtler tactics while also minimizing subjective judgments of language, our research brings statistical techniques from the field of computational linguistics to bear upon a discipline traditionally dominated by qualitative approaches.

Our findings are clear: ExxonMobil has used—and continues to use—rhetoric that mimics the tobacco industry in three key ways: (i) publicly downplaying the reality and seriousness of climate change, (ii) presenting fossil fuel dominance as reasonable and inevitable, and (iii) shifting responsibility for climate change away from itself and onto consumers.

How did we arrive at our conclusions? We analyzed 212 documents published between 1972 and 2019, including all internal company memos that led to allegations against ExxonMobil and peer-reviewed publications offered by the company in response. We also examined all of Mobil and ExxonMobil's paid "advertorials" concerning climate change published in *The New York Times*, as well as all of ExxonMobil's flagship climate reports. We processed the text of these documents using a machine learning model to identify recurring topics. We then statistically compared how often terms and topics appeared in one set of documents versus another—a translation and expansion of techniques previously used to study tobacco industry documents (*6*).

Our analysis showed that ExxonMobil's advertorials and flagship reports emphasized certain terms and topics while avoiding others, creating misleading public narratives about climate change. One example can be found in the rhetoric of "risk," which Exxon and Mobil systematically introduced as a modifier to "climate change" after their merger in 1999. By definition, a risk is something that may or may not happen. In characterizing climate change as a risk, ExxonMobil implied that it was a possibility, but not necessarily a reality—even after climate scientists had demonstrated that it was already underway. One ExxonMobil manager saw this as "an effort" by then-CEO Rex Tillerson "to carefully reset the corporation's profile on climate positions so that it would be… less exposed" (*7*). It was also a reprisal of the tobacco industry playbook, wherein the public was told that smoking was "a risk factor for certain diseases" but "not a proven cause" (*8*). To this day, ExxonMobil, Chevron, and ConocoPhillips continue to call human-caused climate change a "risk" rather than a reality (*9*–*11*).

ExxonMobil's public-facing documents also statistically overused terms that shifted responsibility for climate change from the producer to the consumer. In private, the company willingly recognized climate change as a "fossil fuel" problem caused by "fossil fuel combustion"—i.e., by the company's products. In public, however, the company overwhelmingly presented climate change as caused by the "energy demand" of "consumers"; a problem to be solved by "energy efficiency" (*5*). ExxonMobil's advertorials referred only to "greenhouse gas emissions," while, just as often, their scientists more-specifically labeled them "fossil fuel emissions" and "fossil fuel $CO_2$."

Using an inductive method called frame analysis (*12*) to further examine the advertorials, we found that ExxonMobil conflates

NOMIS & Science YOUNG EXPLORER AWARD

Department of the History of Science, Harvard University, Cambridge, MA 02139, USA. Email: gjsupran@fas.harvard.edu

PHOTO: STEPHANIE MITCHELL/HARVARD UNIVERSITY

demand for energy with an indefinite need for fossil fuels; casting itself as an impartial supplier. The benefits of fossil fuels are presented as profound and explicit, whereas their consequences are an ambiguous climate "risk."

Consumer demand is one aspect of the climate problem and its solution. However, the disproportionate fixation of ExxonMobil, a fossil fuel supplier, on demand-side responsibility is reminiscent of the tobacco industry's well-documented effort "to diminish its own responsibility…by casting itself as a kind of neutral innocent, buffeted by the forces of consumer demand" (*8*).

Today, narratives focused on personal climate responsibility are pervasive in shaping the agenda of many scholars, policy-makers, and members of the public. Yale University, for instance, refuses to divest from fossil fuel companies because "consumption of fossil fuels, not production, is the root of the climate change problem" (*13*). The Republican Party's 2020 climate agenda was premised on the idea that "fossil fuels aren't the enemy. It's emissions" (*14*). Even the Paris Agreement does not mention fossil fuels (*15*).

The fossil fuel industry is priming us to see ourselves as "consumers first and citizens second" (*16*). Experiments indicate that messages framed in terms of individual behavior decrease people's willingness to take both personal and collective climate actions (*17*).

Will Big Oil continue to follow in Big Tobacco's footsteps (*18*)? Already, ExxonMobil's public relations focus on meeting the demand of its consumers has begun to be repackaged by its lawyers to exonerate the company and blame its customers. In 2018, a lawyer defending ExxonMobil and other companies against a climate lawsuit argued that it is not "the production and extraction of oil that is driving these emissions. It's the energy use" (*19*). "It's the way people are living their lives." Fossil fuel firms could also invoke legal defenses that cite their prolific "risk" rhetoric as evidence that the public knew about and voluntarily assumed the risk of climate change—just as tobacco companies argued that their customers had been warned about the risks of smoking.

Both in the court of law and the court of public opinion, we must learn from the history of the tobacco industry to protect the future climate. It took 20 years for Big Tobacco's blame game to falter (*20*). We cannot afford to wait that long to stop smoking our planet. ■

## REFERENCES AND NOTES


1. L. Ly, "New York City sues Big Oil for allegedly misleading consumers about climate change," *CNN* (2021); https://perma.cc/M8QQ-JSU6.
2. Columbia Law School, Arnold&Porter, "U.S. Climate Change Litigation" (2021); https://perma.cc/ZHZ3-5HTX.
3. J. Cook *et al*., "America misled: How the fossil fuel industry deliberately misled Americans about climate change" (George Mason University Center for Climate Change Communication, 2019); https://perma.cc/7942-6FWJ.
4. Commonwealth of Massachusetts v. Exxon Mobil Corporation, 1984CV03333 (2020); https://perma.cc/4KLX-X9PZ.
5. G. Supran, N. Oreskes, *One Earth* **4**, 696 (2021).
6. S. Risi, R. N. Proctor, *Tob. Control* **29**, e41 (2020).
7. S. Coll, *Private Empire: ExxonMobil and American Power* (Penguin Books, 2012).
8. R. N. Proctor, *Golden Holocaust: Origins of the Cigarette Catastrophe and the Case for Abolition* (Univ. California Press, 2012).
9. ExxonMobil Corp, "Climate change" (corporate.exxonmobil.com); https://perma.cc/E42Q-MJU6.
10. Chevron Corp, "Climate change" (chevron.com); https://perma.cc/6H9L-N4GE.
11. ConocoPhillips, "Managing climate-related risks" (conocophillips.com); https://perma.cc/QSP7-DEVF.
12. B. Van Gorp, in *Doing News Framing Analysis: Empirical and Theoretical Perspectives*, P. D'Angelo, J. A. Kuypers, Eds. (Routledge, 2009), pp. 84–109.
13. Yale University Investments Office, "Climate change"; https://perma.cc/9UVU-YK8F.
14. J. Siegel, "How House Republicans won over conservatives to gain consensus on a climate agenda," *Washington Examiner* (2020); https://perma.cc/HF83-EMGD.
15. G. Piggot, P. Erickson, H. van Asselt, M. Lazarus, *Clim. Policy* **18**, 1189 (2018).
16. M. F. Maniates, *Glob. Environ. Polit.* **1**, 31 (2001).
17. R. Palm, T. Bolsen, J. T. Kingsland, *Weather Clim. Soc.* **12**, 827 (2020).
18. L. C. Friedman *et al*., *Am. J. Public Health* **105**, 250 (2015).
19. City of Oakland v. BP P.L.C., 18-16663, Transcript of proceedings (2018); https://perma.cc/EJ4Y-HDQV.
20. P. Mejia *et al*., *Am. J. Public Health* **104**, 1048 (2014).


10.1126/science.abm3434




## Pushing the Boundaries of Knowledge

As AAAS's first multidisciplinary, open access journal, *Science Advances* publishes research that reflects the selectivity of high impact, innovative research you expect from the *Science* family of journals, published in an open access format to serve a vast and growing global audience. Check out the latest findings or learn how to submit your research: **ScienceAdvances.org**

Science Advances
AAAS

**GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS**

# RESEARCH



## IN *SCIENCE* JOURNALS

*Edited by* **Michael Funk**

HEMATOPOIESIS

### Colorful clones in the blood

Stem cells in regenerating tissues such as the blood can acquire mutations that enable a growth advantage, increasing the chance of developing cancer. It is unclear how such diverse mutations promote clonal fitness. Avagyan *et al.* generated a platform in zebrafish to label clones with unique hues while inducing mutations in genes implicated in human blood disorders. Mutations in some genes caused clones to expand over time, resulting in clonal dominance. Progenitors in the dominant clone expressed anti-inflammatory factors to resist the inflammatory environment produced by their own mature progeny, leading to a self-perpetuating cycle promoting clonal fitness. Targeting these resistance pathways may be used to abate clonal hematopoiesis and prevent its associated pathology. —BAP

*Science*, aba9304, this issue p. 768

Microscopy image of multicolored blood cells, genetically labeled with fluorescent proteins, that can track clonal cell lineage

CREDIT: AVAGYAN *ET AL.*

NEUROSCIENCE

### How to quantify local axonal orientations

Mapping the axonal trajectories of the brain's white matter at cellular resolution is a long-standing goal of neuroscience. However, existing methods for mapping the axons are either limited to animal studies or require highly specialized equipment for data acquisition and processing. Nissl staining identifies cell nuclei and has been used extensively to investigate parcellations of the cortical gray matter, but the white matter has largely been neglected with this technique. Schurr and Mezer now show that Nissl staining, together with structure tensor analysis, can be used to study white matter architecture and the organization of the glial cell framework around axons over the whole brain. This technique greatly advances our knowledge regarding the organization of glial cells and the fine-grained organization of axonal projections in the brain. —PRS

*Science*, abj7960, this issue p. 762

BIOSYNTHESIS

### Big molecules build small

Actinomycete bacteria are prolific producers of bioactive small molecules such as polyketide antibiotics. These molecules are built by the addition of short carbon units to a growing, protein-tethered chain, either iteratively as in fatty acid synthesis or in a modular fashion by a hand-off from one distinct enzyme complex to the next. Bagde *et al*. and Cogan *et al*. report structures of polyketide synthase modules in action, taking advantage of antibody stabilization of one of the domains. Both groups visualized multiple conformational states and an asymmetric arrangement of domains, providing insight into how these molecular assembly machines transfer substrates from one active site to another. —MAF

*Science*, abi8532, abi8358, this issue p. 723, this issue p. 729

ORGANIC CHEMISTRY

### Easy aryl reductions

The Birch reduction has been widely used for more than half a century to achieve partial reduction of aryl rings by alkali metals at just two diametrically opposed carbon sites. However, the conditions require condensation of caustic gaseous ammonia. A variation developed soon afterward by Benkeser used safer liquid ethylene diamine but was prone to overreduction. By diluting ethylene diamine in tetrahydrofuran solvent, Burrows *et al*. now obtain selectivities comparable to Birch conditions but without the need for condensed ammonia. —JSY

*Science*, abk3099, this issue p. 741

EVOLUTIONARY BIOLOGY

### The variable tempo of early evolution

The Ediacaran Period, 635 to 539 million years ago, was marked by the rise of complex life on Earth. Yang *et al.* found that Ediacaran evolution was not slow and steady but rather marked by intervals of rapid diversification interspersed with dramatic environmental change. The authors present geochronologic data that better define the tempo of evolution and how changes in that tempo relate to changes in the global carbon cycle as evidenced by carbon isotopic measurements of carbonate rocks in the Ediacaran

stratigraphy. The process was cyclic: Specific assemblages of complex organisms persisted stably for tens of millions of years in each cycle, after which new assemblages arose over much shorter time intervals. Present data suggest that these bursts of diversification may have been caused by rapid environmental changes associated with changes in the carbon cycle. The causes for these episodes of environmental change remain poorly understood. —KVH

*Sci. Adv.* 10.1126/sciadv.abi9643 (2021).

DEVELOPMENTAL BIOLOGY

## Sponges and evolutionary origins

Sponges represent our distant animal relatives. They do not have a nervous system but do have a simple body for filter feeding. Surveying the cell types in the freshwater sponge *Spongilla lacustris*, Musser *et al*. found that many genes important in synaptic communication are expressed in cells of the small digestive chambers. They found secretory machinery characteristic of the presynapse in small multipolar cells contacting all other cells and also the receptive apparatus of the postsynapse in the choanocytes that generate water flow and digest microbial food. These results suggest that the first directed communication in animals may have evolved to regulate feeding, serving as a starting point on the long path toward nervous system evolution. —BAP

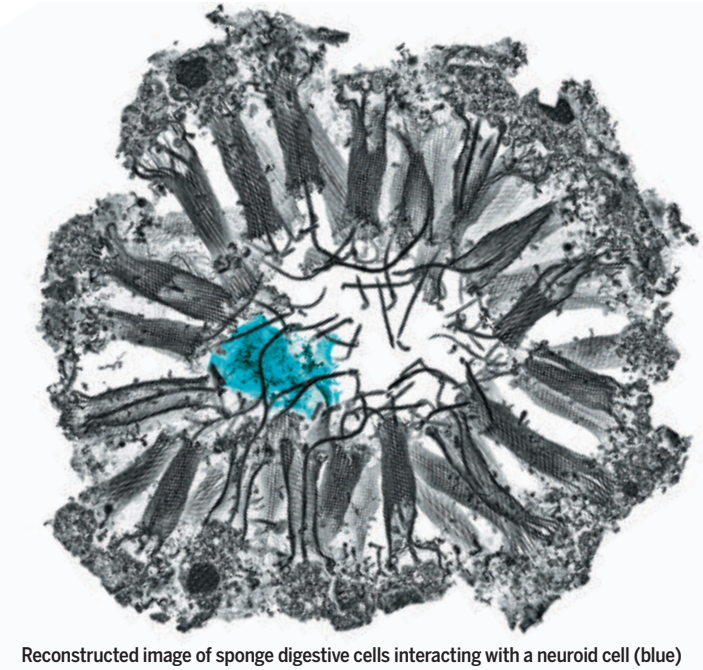
*Science*, abj2949, this issue p. 717

ALZHEIMER'S DISEASE

## A path to more sleep and less amyloid

Recent data have demonstrated a link between sleep disturbances and Alzheimer's disease (AD). However, our understanding of the mechanisms mediating sleep disturbances in AD remains incomplete. Jagirdar *et al.* show that the thalamic reticular nucleus (TRN) plays a main role in determining the effects of sleep on hallmarks of AD in mice and possibly in patients. TRN activity was reduced in a mouse model of AD, resulting in reduced slow-wave sleep. Its activation restored this sleep and reduced amyloid plaques in multiple brain areas. In patients with AD, markers of TRN activation correlated with disease stage. Targeting TRN activation might have therapeutic effects in AD. —MM

*Sci. Transl. Med.* **13**, eabh4284 (2021).

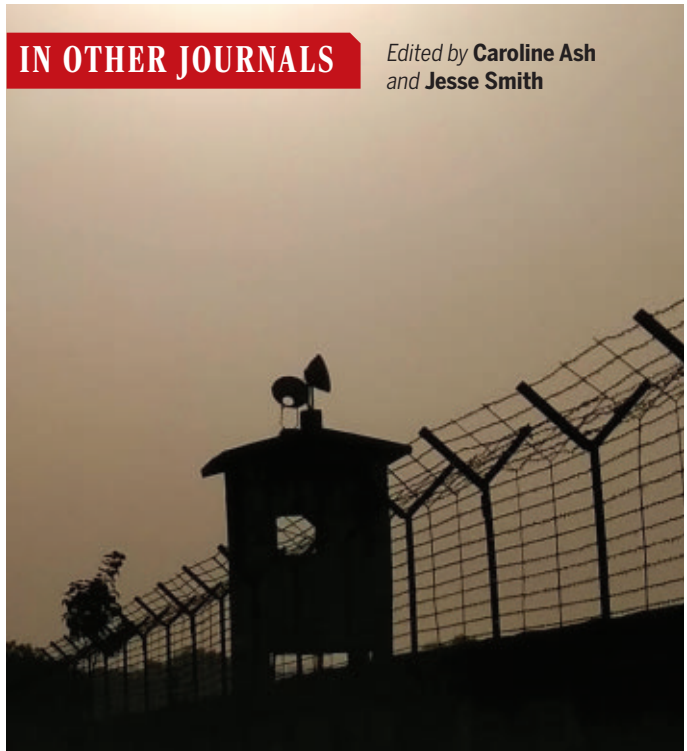


CREDITS (LEFT TO RIGHT): MUSSER ET AL.; JACOB MUSSER, GIULIA MIZZON, CONSTANTIN PAPE, NICOLE SCHIEBER, CREATIVE TEAM/EMBL

**Reconstructed image of sponge digestive cells interacting with a neuroid cell (blue)**

# IN OTHER JOURNALS

*Edited by* **Caroline Ash** *and* **Jesse Smith**



IMPRISONMENT

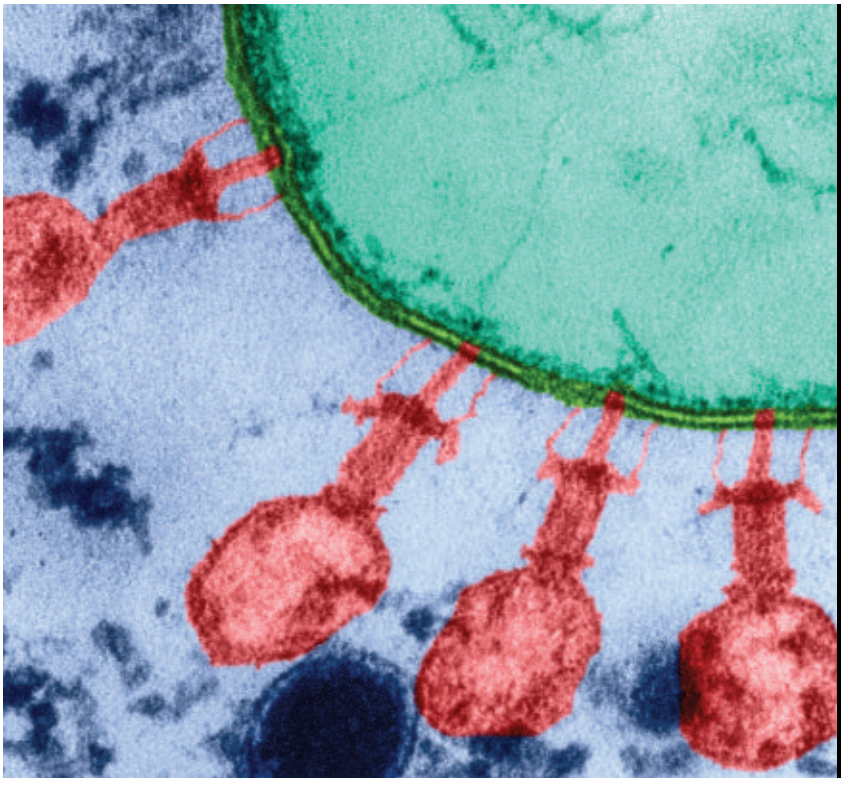
## Public health benefits of release

Although the United States is a world leader in science and technology, it also takes the top position for incarcerating its population, and jails are a hotbed for infectious disease transmission. Those jails that house pretrial detainees and others serving less than 1 year are a source of disease, particularly for disadvantaged communities that suffer disproportionate policing. Using data from jails in 1605 U.S. counties from January to December 2020, Reinhart *et al.* found that avoiding incarceration for nonviolent and petty offenses not only reduced jail populations by 80%, but also reduced the daily case growth rates of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in surrounding communities by 2%. By comparison, mask mandates are estimated to reduce daily case growth rates by 2.5%. —BAP *JAMA Netw. Open.* **4**, e2123405 (2021).

HUMAN GENETICS

## OAS1 links dementia and COVID-19

Dementia increases an individual's risk of severe COVID-19. Alzheimer's disease is characterized by extensive neuroinflammatory changes, amyloid-beta deposits, and tau tangles. Oligoadenylate synthetase 1 (OAS1) mediates innate immune responses through interferon and is a known risk factor for the development of Alzheimer's disease. Magusali *et al.* examined genetic variants around the gene OAS1 to investigate the correlation between severe COVID-19 and dementia. The authors found that a haplotype near OAS1 containing four linked genetic variants down-regulates the expression of OAS1. Two of the variants are associated with the development of Alzheimer's, and the other two are associated with severe COVID-19. Examining co-expression networks for microglial *Oas1*



SIGNAL TRANSDUCTION

## Signaling phage defenses in bacteria

Cyclic nucleoside monophosphates are ubiquitous signaling molecules. Indeed, cyclic adenosine monophosphate (cAMP) is an essential signaling molecule for most organisms. Tal *et al.* have been investigating the function of the more enigmatic cyclic 3′,5′-cyclic cytosine monophosphate (cCMP) and 3′,5′-cyclic uridine monophosphate (cUMP). These cyclic pyrimidines seemed to act as signaling molecules for bacterial immune effectors. Their apparent targets appear to be a transmembrane protein that may help to kill infected cells by breaking membrane integrity and a nucleosidase called NADase, which is known in other bacterial defense systems to prevent phage infection. Whether similar roles for cyclic pyrimidine nucleotides occur in eukaryotes remains an open question. —LBR

*Cell* 10.1016/j.cell.2021.09.031 (2021).

**Color-enhanced transmission electron micrograph showing T2 phage attacking *Escherichia coli* cells**

in mouse models of Alzheimer's disease, as well as in humans afflicted with dementia and COVD-19, identified an interferon-response network. —LMZ

*Brain* awab337 (2021).

SUPERCONDUCTIVITY

## Cuprates versus nickelates

The mechanism of cuprate superconductivity remains one of the most important puzzles of condensed-matter physics. The recently observed superconductivity in the nickelates may offer clues, but just how far the similarities between the two systems go is not well understood. Higashi *et al.* used dynamical mean-field theory in combination with local density approximation to simulate the results of three x-ray spectroscopic techniques applied to the parent compound of superconducting nickelates, $NdNiO_2$, and compared them with experimental results. They concluded that $NdNiO_2$ can be described by a single-band model and is proximate to, but not in, the Mott insulator state. In the superconducting nickelates, doped holes do not predominantly reside in the oxygen orbitals as they do in the cuprates. —JS

*Phys. Rev. X* **11**, 041009 (2021).

MICROBIAL METABOLISM

## Sinking fast, degrading faster

Within the water column of the ocean, organic matter from the productive surface gradually sinks to lower layers. The amount of carbon and other nutrients that reache the bottom depends on the balance of rates of sinking and degradation by attached microorganisms. Alcolombri *et al.* studied particle degradation rates in a microfluidics setup that allows for various flow rates around a stationary polysaccharide particle, mimicking the flow that would occur during sinking. They found that degradation increased with flow rate and performed follow-up experiments to show that a major factor was the removal of degradation products that would otherwise accumulate and inhibit the enzymes responsible for depolymerizing the model particles. Although real sinking particles are heterogeneous, these results illustrate the complex interplay of physical and biochemical processes occurring within the water column. —MAF

*Nat. Geosci.* **14**, 775 (2021).

IMMUNOLOGY

## DNA webs in allergic asthma

Eosinophils are important in allergic immune responses, and their infiltration of lung tissue is a key feature of asthma. Chromatin webs laced with proteins are known to be produced by neutrophils. Eosinophils can also produce extracellular chromatin webs called eosinophil extracellular traps (EETs), but their function is less well understood. Lu *et al*. found that EETs occurred in bronchoalveolar lavage fluid from individuals with asthma, and their occurrence correlated with asthma severity. EETs were produced in response to clinically relevant allergens in mice, in which they activated pulmonary neuroendocrine cells to amplify the allergic response. Neuroendocrine cells activated by EETs may offer a target for the treatment of allergic asthma. —GKA

*Nat. Cell Biol.* **23**, 1060 (2021).

SOLAR ENERGY

## Cheaper, cleaner, better

In the race to develop economically attractive alternatives to fossil fuels, no choice shines more brightly than solar photovoltaic power. Lu *et al.* demonstrate that current trends should allow China, the world's largest carbon dioxide emitter, to generate more than 40% of its energy needs with photovoltaic power by 2060—and at a price far below even that of coal. Combined solar power and storage could become a cheaper alternative to coal-fired electricity in China as soon as 2023. —HJS

*Proc. Natl. Acad. Sci. U.S.A.* **118**, e2103471118 (2021).

CREDIT: LEE D. SIMON/*SCIENCE SCIENCE*

# ALSO IN *SCIENCE* JOURNALS

*Edited by* **Michael Funk**

CORONAVIRUS

## Keeping schools open safely

There has been considerable debate about the role of schools in the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). In a Perspective, Ladhani and the sKIDs Investigation Team discuss the data, which indicate that although children and staff can acquire COVID-19 in schools, these infections largely reflect community infection rates, and schools are not a major driver of transmission when mitigation measures are in place. Those mitigations include masking, distancing, sanitization, air-quality monitoring, bubbles, and contact tracing and isolation. Children can acquire COVID-19 and experience severe disease in some cases, as well as extended symptoms called Long Covid, but children mostly benefit from being able to attend school and receive in-person teaching and other support infrastructure. —GKA

*Science*, abj2042, this issue p. 680

RADIOCARBON

## Using carbon-14

Carbon-14, or radiocarbon, a radioactive isotope produced in the upper atmosphere by cosmic rays, is rapidly incorporated into the terrestrial carbon cycle and provides a way to calculate the age of carbon-bearing materials as old as 55,000 years. Heaton *et al*. review recent progress that has allowed the construction of better radiocarbon age calibration curves and discuss the new insights into climate processes, the Sun, Earth's geodynamo, and the carbon cycle that have emerged from these efforts. —HJS

Science, abd7096, this issue p. 707

MEMBRANES

## Selective subatomic separations

Membranes are thin materials used to selectively separate gases or liquids on a range of scales from benchtop experiments to industrial processes. Challenges arise in separating materials with very similar sizes or chemical properties, particularly at the smallest scales. Kidambi *et al*. review advances in using atomically thin two-dimensional materials such as graphene or hexagonal boron nitride for the separation of subatomic species, including electrons, hydrogen isotopes, and gases. The authors explore the scope to scale up the sizes of these membranes and their potential use in applications relating to energy, microscopy, and electronics. —MSL

*Science*, abd7687, this issue p. 708

NEURODEVELOPMENT

## Synapse stabilization

Early in brain development, neurons connect to each other enthusiastically. With development, an overabundance of synapses is winnowed down to refine efficiently connected circuits. Inactive synapses are prime targets for elimination, whereas active synapses tend to be retained. Gomez-Castro *et al*. took a closer look at how these choices are made (see the Perspective by Blum and Lopes). When postsynaptic adenosine receptors are muted or do not find enough extracellular adenosine, synapses get eliminated. Neurotransmitter-dependent signaling pathways drive protein kinase A to phosphorylate the postsynaptic scaffolding molecule gephyrin. Together with a partner synaptogenic membrane protein, gephyrin is required for the stabilization of γ-aminobutyric acid receptors. Adenosine receptors thus detect synaptic activity and in turn drive the stabilization of synapses that produce such activity. —PJH

*Science*, abk2055, this issue p. 709; see also abm3902, p. 684

IMMUNOLOGY

## AMPlifying type 2 immunity

Antimicrobial proteins (AMPs) are a frontline defense against pathogenic microorganisms at mucosal surfaces. These cationic molecules inactivate their targets primarily by disrupting cell walls and membranes. Hu *et al*. found that small proline-rich protein 2A (SPRR2A) is a bactericidal protein produced in the gut that targets Gram-positive bacteria and is phylogenetically distinct from all other known AMPs (see the Perspective by Harris and Wickramasinghe). SPRR2A production is selectively enhanced by type 2 cytokines such as interleukin-4 and -13, which are elicited by helminth infection. Mice lacking SPRR2A are unable to prevent intestinal bacteria from invading the intestinal barrier after helminths damage the intestinal epithelium. SPRR2A is thus a critical component of type 2 immunity that protects against the bacterial invasion and dissemination that follow helminth infection. —STS

*Science*, abe6723, this issue p. 710; see also abm3876, p. 682

AEROSOL CHEMISTRY

## Spontaneous chemistry on aerosol surface

Interfacial redox chemistry plays an important role in the formation of gas molecules and aerosol particles. However, the characterization of such heterogeneous processes is challenging, so they are often omitted in chemical kinetics models. Using ambient-pressure x-ray photoelectron spectroscopy combined with molecular dynamics simulations, Kong *et al*. discovered spontaneous redox chemistry promoted by the surface at the first stages of the solvation process on a typical inorganic aerosol surface of ammonium sulfate (see the Perspective by Ruiz-Lopez). Several unexpected species have been identified as possible products of a sulfate-reducing ammonium oxidation reaction, and this may help to resolve some of the enduring conundrums of atmospheric chemistry. The present results could also be useful for the development of wastewater treatments and other industrial technologies. —YS

*Science*, abc5311, this issue p. 747 see also abl8914, p. 686

ORGANIC CHEMISTRY

## Asymmetric carbon coupling at boron

The Matteson reaction produces carbon–carbon bonds by coupling halocarbons such as widely available dichloromethane with an alkyl substituent on boron. Sharma *et al*. report asymmetric catalysis of this reaction. Their catalyst, derived from a chiral thiourea, a boronic ester, and an alkyl lithium base, appears to accelerate a chloride abstraction step through its lithium center. The product, still bearing a chloride, can be further modified through stereospecific displacement to generate a wide variety of trisubstituted chiral centers. —JSY

*Science*, abm0386, this issue p. 752

ORGANIC CHEMISTRY

## Careful choreography for a ketone shift

Chemists devote tremendous effort to the precise placement of oxygens in molecular frameworks. Wu *et al*. report a convenient method to shift the oxygen in a carbonyl group to an adjacent carbon center. After activation of the oxygen to an alkenyl triflate, cooperative

catalysis by palladium and norbornene adds nitrogen to the neighboring carbon while displacing the triflate with hydride. Hydrolysis then produces the desired shifted ketone. The protocol is well suited to late-stage variation of complex molecules during drug optimization. —JSY

*Science*, abl7854, this issue p. 734

AIR POLLUTION

## Little things matter

Particulate air pollution 2.5 micrometers or smaller in size (PM2.5) is a major cause of human mortality, and controlling its production is a health policy priority. Nitrogen oxides are an important precursor of PM2.5 and have been a focus of pollution control programs. However, Gu *et al*. now show that abating ammonia emissions is also an important component of PM2.5 reduction, and the societal benefits of abatement greatly outweigh the costs (see the Perspective by Erisman). Reducing ammonia emissions thus would be a cost-effective complement to nitrogen oxides and sulfur dioxide controls. —HJS

*Science*, abf8623, this issue p. 758; see also abm3492, p. 685

TUMOR IMMUNOLOGY

## A unique T cell subset in lung cancer

Although some non-small-cell lung cancers (NSCLCs) are sensitive to immune checkpoint blockade (ICB), many patients with NSCLC do not respond to ICB, and this may be related to the lack of infiltration of $CD8^+$ T cells. Horton *et al.* used mouse models of flank and lung tumors to show that $CD8^+$ T cells from lung tumors, but not flank tumors, had a dysfunctional phenotype distinct from conventional T cell exhaustion that was established in the draining lymph node and correlated to ICB resistance. Interleukin-2 and -12 treatment rescued this phenotype, leading to control of lung tumors. These data suggest that cytokine therapy might be able to rescue a specific subset of dysfunctional T cells found in lung tumors. —DAE

*Sci. Immunol.* **6**, eabi8800 (2021).

PHYSIOLOGY

## Modulating macrophages to heal muscles

Pro-inflammatory macrophages infiltrate injured skeletal muscle and become anti-inflammatory. Both macrophage phenotypes are required for skeletal muscle regeneration and are determined by metabolism, which in turn is shaped by mitochondrial calcium ions. Feno *et al.* found that the inhibitory subunit of the mitochondrial calcium uniporter MCUb promoted the acquisition of an anti-inflammatory phenotype in macrophages and the repair of damaged skeletal muscle in mice. Thus, MCUb activity is linked to a phenotypic switch in macrophages that is critical for skeletal muscle regeneration. —WW

*Sci. Signal.* **14**, eabf3838 (2021).

MARTIAN GEOLOGY

## Perseverance images of a delta on Mars

The Perseverance rover landed in Jezero crater on Mars in February 2021. Earlier orbital images showed that the crater contains an ancient river delta that was deposited by water flowing into a lake billions of years ago. Mangold *et al*. analyzed rover images taken shortly after landing that show distant cliff faces at the edge of the delta. The exposed stratigraphy and sizes of the boulders allowed them to determine the past lake level and water discharge rates. An initially steady flow transitioned into intermittent floods as the planet dried out. This history of the delta's geology provides context for the rest of the mission and improves our understanding of Mars' ancient climate. —KTS

*Science*, abl4051, this issue p. 711

REVIEW SUMMARY

RADIOCARBON

# Radiocarbon: A key tracer for studying Earth's dynamo, climate system, carbon cycle, and Sun

T. J. Heaton*, E. Bard, C. Bronk Ramsey, M. Butzin, P. Köhler, R. Muscheler, P. J. Reimer, L. Wacker

**BACKGROUND:** Radiocarbon ($^{14}C$) has long been recognized as providing an essential dating method covering the past 55,000 years. However, the further role of $^{14}C$ as a diagnostic tracer throughout the Earth and climate system is often less widely appreciated. Radiocarbon is formed by cosmic particles and then dispersed across multiple Earth system compartments. Consequently, accurate knowledge of past $^{14}C$ levels directly enables new discoveries and provides connections across broad research areas. We present recent advances in knowledge of past $^{14}C$ and the resulting insights that improve our understanding of climate processes, solar activity, geophysics, and the carbon cycle. Measurements providing improved resolution in the variations of $^{14}C$ enable us to learn more about how these system components operate and interact.

**ADVANCES:** Recent years have seen a revolution in our ability to reconstruct detailed records of $^{14}C$. Advances include an explosion in measurements of single-year samples made possible by accelerator mass spectrometry (AMS) instrumentation; new detail of pre-Holocene $^{14}C$ levels through use of speleothems, lake macrofossils, and subfossil trees; and improved modeling of marine radiocarbon reservoir ages to incorporate carbon cycle changes. Combined with advanced statistical methods, these developments have allowed the IntCal working group to estimate $^{14}C$ levels with unprecedented accuracy for the Northern and Southern hemispheres, as well as the surface ocean, back to the limit of the technique ~55,000 years ago.

Accompanying this, interdisciplinary studies have provided a new framework in which to use such comprehensive $^{14}C$ estimates. From the climate perspective, $^{14}C$ provides not only a chronology in which to place and compare diverse paleoclimate records, but also constraints on key climate forcings, such as solar variations, and changes in the carbon cycle.

Radiocarbon production is modulated by magnetic properties of the solar wind, leading to lower production during phases of high solar activity, and making atmospheric $^{14}C$ a mirror of solar activity. Conversely, brief $^{14}C$ production maxima have recently been identified and attributed to short-term solar energetic particle bursts. Comparisons of $^{14}C$ with other cosmogenic isotopes, such as $^{10}Be$ and $^{36}Cl$ in polar ice cores, allow substantial progress in documenting past behavior of the Sun, previously poorly understood by astronomers. These recent astrophysical discoveries are important beyond academia because rapid solar and space weather events could severely damage current technology.

Radiocarbon also enables insight into Earth's magnetic field, from the near-reversal of the Laschamps geomagnetic excursion through to smaller perturbations including the drop over recent centuries. Again, comparison with other cosmogenic isotopes provides key inferences concerning paleomagnetic variations, which are still difficult to simulate with geodynamo models.

Finally, $^{14}C$ offers increased understanding of the carbon cycle and its feedbacks and responses to climate change. Radiocarbon permits the identification of $CO_2$ fluxes such as the release of permafrost carbon. It also enables insight into the ocean's role in climate change via estimation of changes in the residence time of the carbon within it, and changes in the meridional overturning circulation during abrupt climate events.

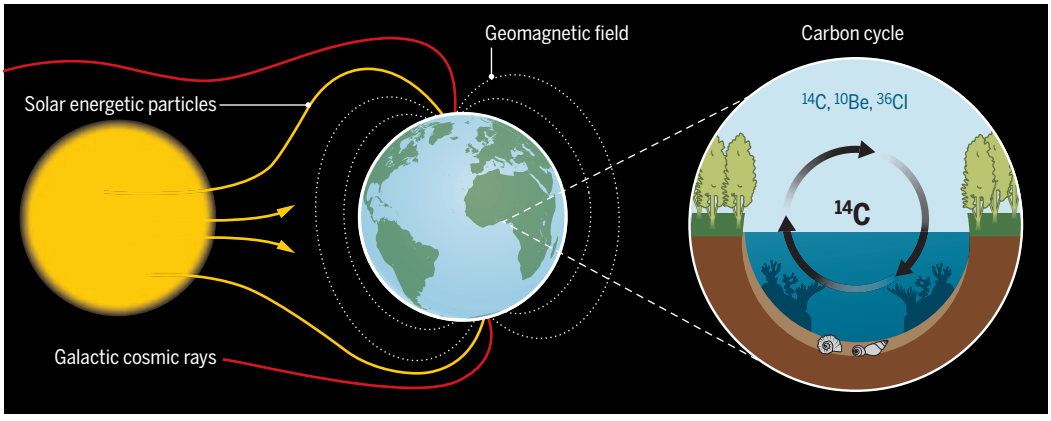
**OUTLOOK:** Harnessing $^{14}C$'s full potential involves addressing some key challenges. Current reconstructions of environmental $^{14}C$ levels beyond the most recent 14,000 years rely on a synthesis of measurements, most of which only indirectly record atmospheric levels. The search continues for a fully atmospheric reconstruction extending back 55,000 years with sufficient resolution. Such a development would greatly enhance our ability to independently test and validate climate and carbon cycle models.

Additionally, whereas many periods of the record are now annually resolved, in other periods the temporal resolution is much coarser. A continuous annual atmospheric $^{14}C$ record would enable examination of long-term trends in solar activity and provide insight into the nature and frequency of shorter-lived space weather events.

Further, accompanying improvements in our records of other cosmogenic nuclides and paleomagnetic reconstructions, in combination with modeling advances, will allow identification of key feedbacks within our Earth and climate system. This will elucidate causal chains and permit testing of key hypotheses, resulting in improved predictions of climate change. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: t.heaton@shef.ac.uk
Cite this article as T. J. Heaton *et al.*, *Science* **374**, eabd7096 (2021). DOI: 10.1126/science.abd7096

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abd7096




**Radiocarbon provides a link across diverse areas of research.** Knowledge of past $^{14}C$ levels provides insight into solar activity, the geodynamo, and the carbon cycle. It also enables the synchronized dating of key environmental records that are fundamental for the study of climate. As our estimates of past $^{14}C$ variation become more detailed and precise, we improve our understanding of the Earth and climate system.

REVIEW

RADIOCARBON

# Radiocarbon: A key tracer for studying Earth's dynamo, climate system, carbon cycle, and Sun

T. J. Heaton[1]*, E. Bard[2], C. Bronk Ramsey[3], M. Butzin[4], P. Köhler[4], R. Muscheler[5], P. J. Reimer[6], L. Wacker[7]

Radiocarbon ($^{14}$C), as a consequence of its production in the atmosphere and subsequent dispersal through the carbon cycle, is a key tracer for studying the Earth system. Knowledge of past $^{14}$C levels improves our understanding of climate processes, the Sun, the geodynamo, and the carbon cycle. Recently updated radiocarbon calibration curves (IntCal20, SHCal20, and Marine20) provide unprecedented accuracy in our estimates of $^{14}$C levels back to the limit of the $^{14}$C technique (~55,000 years ago). Such improved detail creates new opportunities to probe the Earth and climate system more reliably and at finer scale. We summarize the advances that have underpinned this revised set of radiocarbon calibration curves, survey the broad scientific landscape where additional detail on past $^{14}$C provides insight, and identify open challenges for the future.

Radiocarbon ($^{14}$C) is well known for providing chronologies and age estimates through radiocarbon dating, a technique first developed by Libby (*1*). However, its importance extends much further, allowing us to probe the fundamental relationships between multiple compartments of the Earth and climate system. It also provides a long-term perspective on past solar activity.

Excluding anthropogenic sources connected with nuclear research, industry, and weapons testing after 1940, $^{14}$C is predominantly produced in the upper atmosphere by a chain of reactions started by incoming galactic cosmic rays. The intensities of these rays, and hence $^{14}$C production rates, are spatially and temporally modulated by the Sun and Earth's magnetic field. Occasional substantial $^{14}$C production can also occur during extreme solar storms via highly energetic solar particles. After production, $^{14}$C is oxidized to $^{14}CO_2$ and redistributed through the carbon cycle (Fig. 1). Consequently, knowledge of past $^{14}$C levels in the various oceanic and terrestrial carbon reservoirs helps to quantify processes in astrophysics, geophysics, and biogeochemistry, with implications for climate science. This approach is enhanced when $^{14}$C is considered alongside other cosmogenic nuclide records and recent advances in Earth system modeling.

This review focuses on pre-1950 levels of $^{14}$C and their nonanthropogenic influences. We do not discuss the post-1950 spike in $^{14}$C levels due to nuclear weapons testing (*2*), which has provided insight into both carbon transport and biology (*3*). Similarly, we do not discuss in detail the dilution of atmospheric $^{14}CO_2$ caused by the burning of $^{14}$C-free fossil fuel and cement production since the industrial revolution (*4*), the continuation of which will introduce future ambiguity in identifying the calendar ages of Holocene samples via radiocarbon dating (*5*). Regional measurement of this "Suess effect" allows estimation of local industrial $CO_2$ emissions (*3*).

Estimates of pre-1950 atmospheric and mean surface ocean $^{14}$C levels are provided by the International Calibration (IntCal) Working Group, which regularly collates $^{14}$C measurements on samples of known, or independently estimated, calendar age from a range of archives including tree rings, lacustrine and marine sediments, speleothems, and corals. These archives are assessed for quality control and integrated to provide the IntCal set of calibration curves. The latest IntCal calibration curves, released in 2020, comprise IntCal20 for the Northern Hemisphere (NH) atmosphere (*6*), SHCal20 for the Southern Hemisphere (SH) atmosphere (*7*), and Marine20 (*8*) for the surface oceans.

Recent interdisciplinary advances—including the ability to measure $^{14}$C efficiently in very small samples such as single tree rings (*9*), the availability of many new archives including those covering the last glacial period (*10*), and improved modeling (*8*, *11*, *12*)—provide new levels of accuracy, precision, and detail in the IntCal20 set of curves when compared to the previous IntCal13 estimates (*13*, *14*). Consequently, there are new opportunities for using $^{14}$C to study the behavior of both the Sun and Earth's dynamo, undertake better absolute dating and synchronization of paleoclimate records, and obtain new insight into the carbon cycle and climate system.

When discussing $^{14}$C, calendar ages are expressed as cal BP (calendar years before present, defined as 1950 CE). $\Delta^{14}$C denotes the (age-corrected) $^{14}C/^{12}C$ ratio compared to a standard, thereby showing relative changes in the $^{14}C/^{12}C$ ratio (*15*).

## Recent improvements in our understanding of past $^{14}$C

### *Tree rings: AMS evolution and annual resolution*

Measurements of tree-ring sequences, independently dated by dendrochronology, are the gold standard for the reconstruction of past atmospheric $^{14}$C levels. The availability of new archives, as well as the higher capacity of new-generation accelerator mass spectrometry (AMS) (*16*, *17*), have enabled a rapid expansion in the availability of such data. We can now provide a precise estimate of NH atmospheric $^{14}$C from 14,000 to 0 cal BP solely from tree rings. This estimate draws on 9211 $^{14}$C measurements on tree rings of known age dated by either ring width or isotope dendrochronology extending from 12,308 to 0 cal BP, and on 1498 measurements on older tree rings for which calendar ages are estimated on the basis of $^{14}$C matching (*18*–*20*).

AMS has also enabled precise measurement of much smaller samples. Older decay counting methods require several grams of wood for a high-precision $^{14}$C measurement, typically limiting resolution to 10-year blocks. AMS is now able to provide similar precision but with only a few milligrams of material (*9*). This allows efficient and accurate measurement of $^{14}$C in single growth rings, giving insight into short-term (annual) fluctuations in $^{14}$C production (*21*).

Almost half (4952) of the tree-ring $^{14}$C measurements used to construct the IntCal20 NH estimate, covering 2731 individual calendar years, relate to determinations of single growth rings. Particular examples where annual-resolution $^{14}$C data have provided new insight include the sharp sudden spikes in $^{14}$C production (Fig. 2B) in 774–775 CE (*22*) and 993–994 CE (*23*) caused by extreme solar proton events (SPEs) (*24*) and the extension into the Younger Dryas (Fig. 2C), a critical period of late-glacial climate change (*19*, *20*).

The number of laboratories providing reliable $^{14}$C measurements has also increased, with 20 laboratories submitting data for the IntCal20 curves, reducing the previous reliance on datasets from single laboratories and associated dangers of systematic bias. The use of multiple

[1]School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, UK. [2]CEREGE, Aix-Marseille University, CNRS, IRD, INRAE, Collège de France, Technopole de l'Arbois BP 80, 13545 Aix-en-Provence Cedex 4, France. [3]Research Laboratory for Archaeology and the History of Art, University of Oxford, Oxford OX1 3TG, UK. [4]Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung (AWI), D-27515 Bremerhaven, Germany. [5]Quaternary Sciences, Department of Geology, Lund University, 223 62 Lund, Sweden. [6] $^{14}$CHRONO Centre for Climate, the Environment and Chronology, School of Natural and Built Environment, Queen's University, Belfast BT7 1NN, UK. [7]Laboratory of Ion Beam Physics, ETH Zürich, CH-8093 Zürich, Switzerland.
*Corresponding author. Email: t.heaton@shef.ac.uk

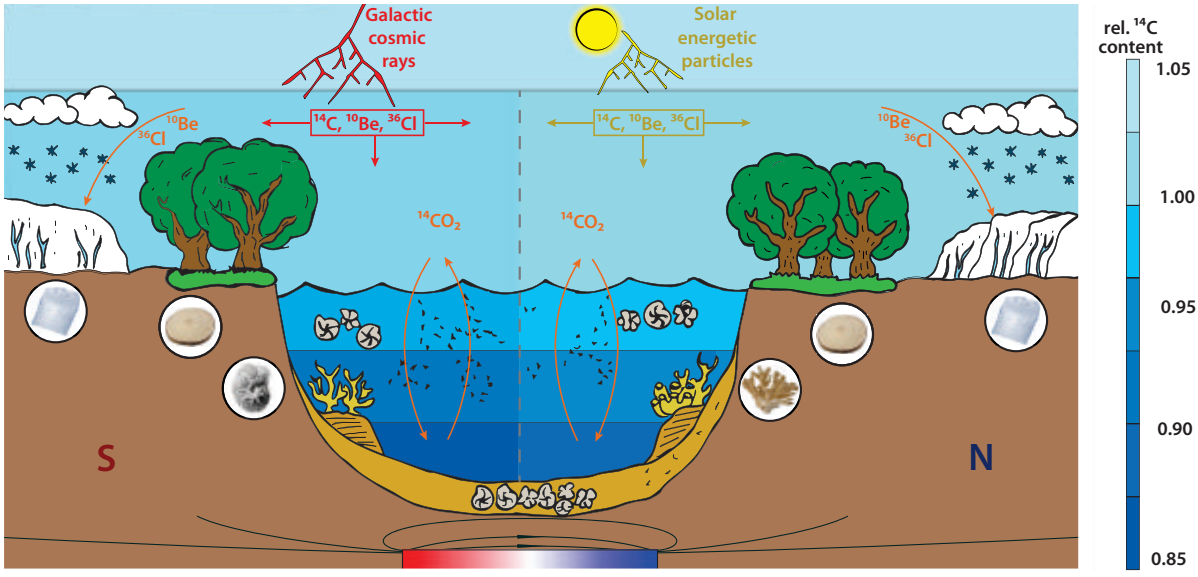



**Fig. 1. The production and subsequent distribution of $^{14}$C and other cosmogenic radionuclides throughout the Earth system.** Production of $^{14}$C and other cosmogenic nuclides (such as $^{10}$Be and $^{36}$Cl) occurs mainly in the stratosphere and upper layers of the troposphere driven by incoming galactic cosmic rays. Whereas the galactic cosmic ray flux is assumed to be constant and isotropic, nuclide production rates are spatiotemporally modulated by magnetic shielding influenced by both the Sun's activity and the strength of Earth's magnetic field. Further nuclide production can result from the release of solar energetic particles during extreme solar storms. After production, the nuclides are dispersed through the Earth system. In the case of $^{14}$C, this dispersal occurs via the carbon cycle, resulting in different ratios of $^{14}$C to stable $^{12}$C in its various oceanic and terrestrial compartments; the blue color scale shows the approximate $^{14}C/^{12}C$ ratios relative to the northern troposphere. Past levels of radionuclides are recorded in a range of archives. For $^{14}$C these archives include tree rings, stalagmites, corals, and lacustrine and marine sediments containing foraminifera; $^{10}$Be and $^{36}$Cl are recorded in ice cores and sediments. Measurements of $^{14}$C levels over time in the different carbon cycle compartments, in combination with measurements of other cosmogenic nuclides, inform us about changes to the climate system and carbon cycle processes and provide insight into the Sun and the geodynamo.

laboratories has enabled interlaboratory comparisons to improve the accuracy of $^{14}$C measurements (*25*, *26*).

### *Extending the NH atmospheric $^{14}$C estimates back to 55,000 cal BP*

Further back in time, beyond 14,000 cal BP, insufficient tree-ring measurements exist to precisely estimate the levels of $^{14}$C from these samples alone. Instead, we must rely on combining data from a wider range of archives. The backbone of the IntCal20 $^{14}$C estimate in this older time period is provided by three stalagmites from Hulu Cave, China (*10*, *27*). These stalagmites provide a continuous $^{14}$C record consisting of measurements at more than 400 unique (albeit imprecisely known) calendar ages, extending from 54,000 to 13,900 cal BP. They provide new insight into $^{14}$C levels before 30,000 cal BP, in particular around the Laschamps geomagnetic excursion (~41,000 cal BP; Fig. 2A). These data are augmented by additional $^{14}$C determinations from terrestrial macrofossils and foraminifera found in lake and ocean sediments, corals, and stalagmites from other locations (*6*).

We have also begun to find tree-ring sequences from before 14,000 cal BP for which relative internal ages are known by ring-counting but absolute ages are not (*28*–*30*). Radiocarbon measurements from these so-called "floating" sequences can be combined with the other archives to add resolution and detail to the NH estimate, as well as to enable assessment of modeling assumptions. In particular, three floating tree-ring sequences now cover the last deglaciation between 14,900 and 14,100 cal BP, revealing previously unseen $^{14}$C variations during this key warming period (*28*).

All these older data are more complicated because their calendar ages are estimated rather than precisely known. In the case of corals and stalagmites, these estimates are provided by uranium-thorium (U-Th) dating. Some of the sediment $^{14}$C archives can be provided with a relative calendar age chronology by layer-counting their annual varves (*31*–*33*). Other sediment archives (*6*) are provided with calendar age estimates by tuning abrupt climate shifts seen in their various proxies to high-resolution $\delta^{18}$O also recorded in the Hulu Cave stalagmites (*34*). Taking evidence for globally synchronous timing of the selected rapid paleoclimatic changes (*35*) into account, this allows transferral of calendar age information from Hulu Cave's U-Th time scale.

The necessity of including stalagmite and marine archives introduces further complexity because, unlike wood or terrestrial macrofossils in lake sediments, these archives do not record atmospheric $^{14}$C directly. Stalagmites contain carbon from a range of sources including a "dead carbon fraction" from ancient limestone and soil carbon, which is devoid or depleted of $^{14}$C. Marine $^{14}$C records are also depleted relative to the contemporaneous atmosphere, as a consequence of the time it takes to exchange $^{14}CO_2$ at the ocean surface and mixing/exchange with "old" carbon from the ocean interior. To use these archives to construct an atmospheric $^{14}$C record, adjustment for this depletion must be performed. The dead carbon fraction within a speleothem is estimated by comparing its $^{14}$C depletion against $^{14}$C from dendrochronologically dated tree rings, whereas marine depletion can be estimated by the same approach and computer simulations (*12*). In total, the current IntCal20 estimate of NH $^{14}$C from 55,000 to 14,000 cal BP draws upon measurements from more than 1900 unique calendar ages.

### *Combining the diverse $^{14}$C archives to estimate past NH atmospheric levels*

The varied $^{14}$C archives are carefully synthesized, recognizing their individual characteristics and potentially uncertain calendar ages, to provide a robust and reliable hemispheric-average estimate. A Bayesian spline regression approach is taken, incorporating expert prior information where available (*11*). This provides a set of complete posterior $^{14}$C realizations, each representing a plausible $^{14}$C history.

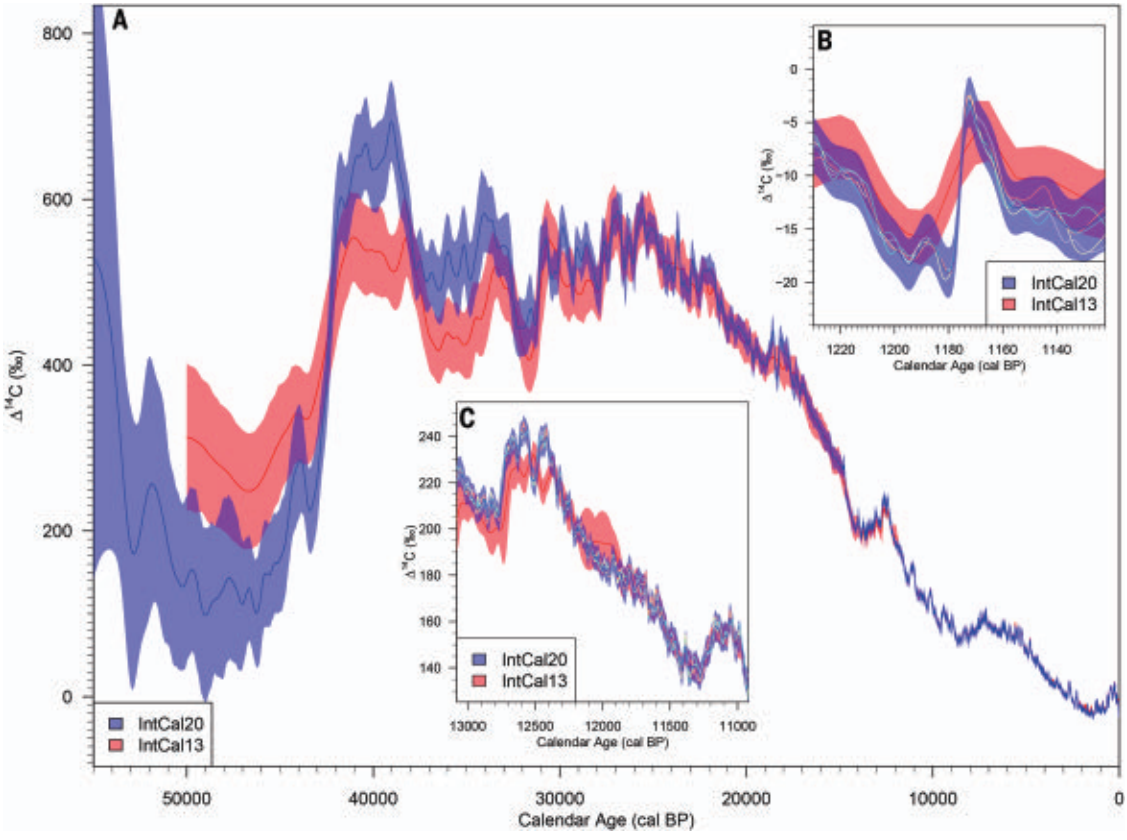



**Fig. 2. IntCal20 Northern Hemisphere estimate of atmospheric $^{14}C$ variations (expressed as $\Delta^{14}C$) over the past 55,000 years compared with the previous IntCal13 estimate.** The additional $^{14}C$ detail in the IntCal20 curve is due to the availability of new archives, combined with recent advances in measurement, statistical modeling, and computer simulation. (**A**) A plot of the posterior mean (solid line) and 2σ credible interval (shaded area) of the curves. Note in particular the large increase in $\Delta^{14}C$ identified in IntCal20 around the Laschamps geomagnetic excursion (~41,000 cal BP) observed in Hulu Cave. (**B**) The 774–775 CE event (1176–1175 cal BP) (*22*) showing the sharp increase in $\Delta^{14}C$ identified with the annual data of IntCal20. This event was not observed in the lower-resolution IntCal13. Shown are a sample of individual posterior IntCal20 curve realizations that enable a user to access covariance information (variously colored lines), the summarized posterior mean (solid blue line), and the 2σ credible interval (shaded blue). (**C**) The extension of annually resolved data into the Younger Dryas provides a much more detailed $^{14}C$ estimate during a key period of climate change than was available with the IntCal13 curve. Individual IntCal20 curve realizations are plotted (variously colored lines) together with the posterior IntCal13 and IntCal20 means and 2σ credible intervals.

These realizations allow access to covariance information on $^{14}C$ levels in adjacent years for use in subsequent modeling studies, or they can be summarized to provide pointwise estimates and probability intervals in any given year (Fig. 2).

This approach to curve construction retains the critical detail and high-frequency variations needed to unlock the full interrogative potential of $^{14}C$ from the large volume of finely resolved tree-ring data from 14,000 to 0 cal BP. However, from 55,000 to 14,000 cal BP, only a coarser-scale estimate is possible because of the sparser and less direct nature of the diverse archives used, together with their uncertain calendar ages. Radiocarbon features seen within multiple archives are identified and maintained while ensuring that individual outlying measurements, which may not provide an accurate atmospheric reconstruction, do not overly influence the estimate. Consequently, from 55,000 to 14,000 cal BP—a period containing important changes to both the carbon and climate systems—only the lower-frequency components of the $^{14}C$ signal are identifiable. Higher-frequency signals, which could be key to fully understanding the processes by which these changes occurred, cannot currently be resolved and await recovery of new archives through ongoing and future fieldwork.

### *Obtaining SH atmospheric and global surface-ocean $^{14}C$ estimates*

Also available are 2094 $^{14}C$ determinations on SH tree rings; these data enable the study of the variable $^{14}C$ gradient between the NH and SH, which is likely a result of the greater expanse of the oceans and higher air-sea $CO_2$ flux in the SH. However, fewer than 15% are determinations of single growth rings, limiting our ability to identify small interhemispheric $^{14}C$-gradient changes. Further, their calendar-age coverage is uneven, and independent estimates of atmospheric SH $^{14}C$ levels are available only in four distinct periods: 2140–0, 3520–3453, 3608–3590, and 13,140–11,375 cal BP (*20*). Outside these periods, the interhemispheric $^{14}C$ gradient is statistically modeled to construct a SH estimate extending from 55,000 to 0 cal BP (*7*).

A computer model–based estimate for $^{14}C$ levels in the global-average surface ocean (*8*) has also been constructed. This is based on a Monte Carlo ensemble of model runs driven by individual atmospheric IntCal20 $^{14}C$ realizations, $CO_2$ reconstructions from ice core records, and prior estimates for both air-sea $CO_2$ exchange and ocean circulation. It simulates large-scale changes in ocean surface $^{14}C$ due to paleoclimatic and $^{14}C$ production rate changes, and enables users to distinguish regional $^{14}C$ ocean effects from variations due to such global changes (*36*).

## Radiocarbon and the Sun

Changes in solar activity have the potential to generate direct impacts on Earth, notably on climate. Solar storms can also cause serious

harm to telecommunications, navigation systems, air and space travel, and electrical power grids. Systematic solar activity measurements started in the early 17th century when observations of sunspots with telescopes began (*37*). For the past century, the sunspot record has been complemented with other instrumental records from ground-based observatories, space probes, and satellites. However, these short-term instrumental records are insufficient for a complete understanding of the Sun, its magnetohydrodynamical behavior, and predictions on the full range of solar variability.

Cosmogenic radionuclide records, such as $^{14}C$ and $^{10}Be$, provide the best "proxy" data for extending solar activity reconstructions beyond the period of instrumental measurements. Such extensions are needed to investigate longer-term solar variations and cycles that likely have most climatic importance (*38*–*41*).

Studies of $^{14}C$ and other cosmogenic nuclides have also revealed massive past releases of solar particles in SPEs that exceed known solar storms of the instrumental era by an order of magnitude (*22*–*24*, *42*). If repeated today, such events have the potential to catastrophically damage current communications, power, and satellite systems. Understanding the size and frequency of these huge solar storms is needed to mitigate their future risk. Radionuclide records provide the greatest potential to do so.

### *Probing solar cycles*

During periods of high solar activity, a stronger solar magnetic shielding of galactic cosmic rays in the heliosphere leads to lower radionuclide production on Earth, and vice versa. This modulation of the radionuclide production rate by solar activity is well illustrated by comparison of radionuclide records and sunspot number observations over the past four centuries (Fig. 3) and can be modeled quantitatively (*43*, *44*).

The geochemical behavior of radionuclides from production to deposition does, however, complicate their interpretation. As a result of efficient mixing, atmospheric $^{14}C$ is not prone to large climatic impacts or weather noise on shorter (annual to decadal) time scales. However, atmospheric $^{14}C$ has the disadvantage that short-term variations due to the 11-year solar cycle (*21*) are damped because of the carbon cycle (*45*). In contrast, the shorter atmospheric residence times of $^{10}Be$ and $^{36}Cl$ largely preserve their cyclic 11-year variability in ice cores, but weather and climate influences on the transport and deposition add noise. Further, the relative geomagnetic modulation of cosmogenic production is maximized at the equator, whereas the solar-induced production variation is maximized at the poles. Although most production occurs in the stratosphere and is characterized by intense horizontal mixing and relatively long residence times, any deviation from complete homogeneous atmospheric mixing would affect the relative amplitudes of the geomagnetic and solar signals embedded in $^{10}Be$ and $^{36}Cl$ ice-core records.

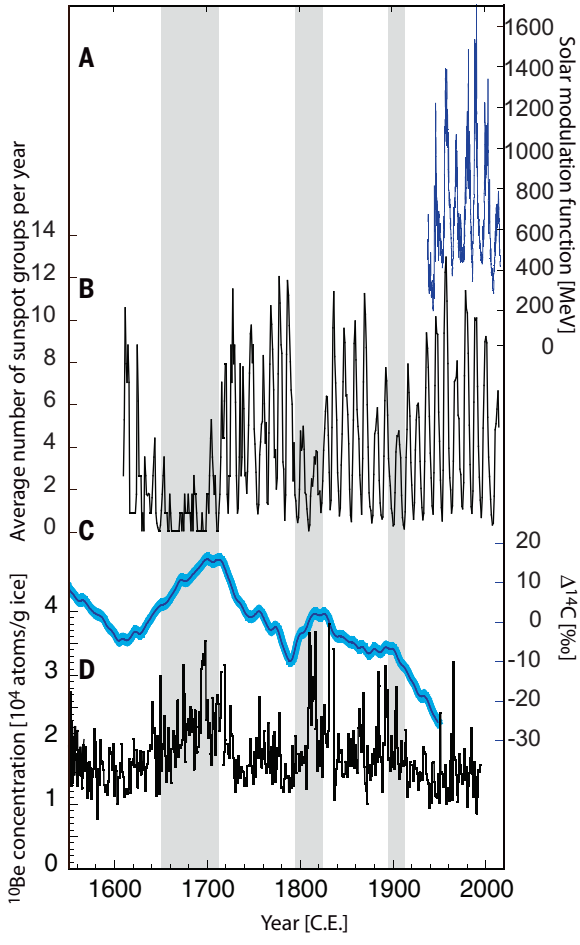



**Fig. 3. Direct and indirect data on past changes in solar activity.** (**A**) A reconstruction of the solar shielding of galactic cosmic rays based on neutron monitor measurements (*129*). (**B**) Group sunspot number reconstruction (*48*). (**C**) The IntCal20 NH $\Delta^{14}C$ estimate (mean ± 2σ) (*6*). It mirrors the ups and downs in solar activity and shows the "$^{14}C$ Suess effect" (i.e., the decrease of atmospheric $^{14}C$ in relation to the stable $^{12}C$ that began at the start of the industrial revolution in 1850 and is due to the massive release of $^{14}C$-free $CO_2$ from the burning of fossil fuel) (*4*). (**D**) Annual $^{10}Be$ data (*130*) show considerable weather noise but have a better potential to preserve the amplitude of the variations (±20%) connected to the 11-year solar cycle and show the cycle's influence on the shielding of galactic cosmic rays. The gray bands show the period of the Maunder minimum around 1700, the Dalton minimum around 1800, and the shorter period of low solar activity around 1900.

Longer-term radionuclide records over the Holocene reveal additional cyclic changes; the most prominent are the 207-year "de Vries" and 88-year "Gleissberg" cycles (*46*). Because $^{14}C$ and $^{10}Be$ records largely agree on these time scales (*47*), we can attribute common variation to production rate changes. Their solar origin is supported because there is no evidence of sufficiently large and rapid geomagnetic field changes to explain such centennial-scale cyclic changes. Further, the Maunder minimum, a 70-year period (1650 to 1715 CE) characterized by an almost complete lack of observable sunspots and coinciding with the middle of the Little Ice Age (*37*, *48*), can be regarded as one of the latest expressions of the de Vries cycle (Fig. 3).

Radiocarbon and $^{10}Be$-based solar activity records allow reconstruction of solar irradiance variations used as inputs to global climate models alongside other natural (e.g., volcanic aerosols) and anthropogenic forcings (e.g., emissions of $CO_2$ and other greenhouse gases). The first attempt (*40*) was used for the IPCC AR4 report. These solar forcing curves have been improved through the use of complex models relating the $^{10}Be$ and $^{14}C$ records to solar irradiance to provide common inputs (*49*) for the IPCC AR5 and AR6 climate models.

Radionuclides also suggest a possible bundling of periods of low solar activity with a quasi-periodicity of ~2000 to 2500 years and potentially far-reaching effects on climate (*50*). However, very little is known about changes in solar activity with such long periodicities. Tree-ring $^{14}C$ records and ice-core $^{10}Be$ records do not agree well on time scales of 1000 years and longer, even after correcting for geomagnetic field influences and the different geochemical behavior (*47*). This could be a consequence of unknown carbon cycle effects on $^{14}C$, climate effects on $^{10}Be$ transport and deposition, or uncertainties and biases within the paleomagnetic/archaeomagnetic reconstructions used to correct the $^{14}C$ and $^{10}Be$ signals. Until these effects are resolved, inferences about solar activity changes on millennial or longer time scales remain speculative.

Higher-resolution and better-quality $^{14}C$ data offer great potential to improve studies on solar

activity changes into the last ice age. This longer-term perspective might be crucial because memory effects due to the carbon cycle could affect atmospheric $^{14}C$ long into the Holocene (*51*) and bias solar activity reconstructions. In parallel, improvements in the quality and resolution of $^{10}Be$ measurements in ice cores provide the opportunity to study solar cycles several hundred thousand years in the past (*52*).

### *Massive solar proton events*

The Sun also generates solar cosmic rays—energetic particles emitted from active regions on the surface of the Sun or accelerated in so-called coronal mass ejections. Study of solar storms during the instrumental period suggests that solar particles might increase the annual radionuclide production rate by only a few percent (*53*), far too small to have an effect on the smoothed atmospheric $^{14}C$ levels or to be reliably detected in noisy $^{10}Be$ and $^{36}Cl$ records.

However, past $^{14}C$ records identify substantial interannual spikes, most clearly during an event where atmospheric $^{14}C$ increased by 1.2% over the year 774–775 CE (*22*). Considering the smoothing effect of the carbon cycle, this indicates a huge spike in $^{14}C$ production, generating almost four times the average yearly production. This cannot be explained by lower solar shielding of galactic cosmic rays. Co-occurring $^{10}Be$ and $^{36}Cl$ spikes in ice cores allow the robust attribution of this event to one or several successive massive SPEs (*24*).

Subsequent high-resolution studies of cosmogenic radionuclide records have enabled the identification of additional SPEs, such as 993–994 CE (*23*) and ~660 BCE (*42*), and a more detailed understanding of their nature (*54*). These SPEs not only provide evidence of the Sun's potential to produce extreme events; their unique global signature also offers the potential to precisely synchronize, to annual precision, climate records from tree rings and ice cores (*55*).

## Radiocarbon and the geodynamo

The magnetic field observed at Earth's surface is generated by turbulent convective flows of an electrically conducting iron-nickel fluid in Earth's outer core, a process known as the geodynamo. Since Gauss's first measurements in the mid-19th century, the intensity of the geomagnetic field has continuously decayed; it is now almost 10% weaker than in 1840 CE (*56*). This recent evolution is unexplained and must be considered in the context of longer-term geodynamics. Documenting temporal variations in the intensity of the geomagnetic field is fundamental for understanding the dynamics of Earth's deep interior and the evolution of the Earth system.

Reconstructions of the past geomagnetic field intensity can be obtained from various archives. Thermoremanent magnetization of volcanic rocks and archaeological archives (e.g., baked clays) provides absolute field intensities (*57*, *58*) but represents sporadic recording of the local and instantaneous, rather than global dipolar, field. By contrast, detrital remanent magnetization of marine and lacustrine sediments provides continuous records, which average short-term deviations. However, sediments provide only a relative estimate of the geomagnetic field. They must be normalized and calibrated to the present-day dipole intensity or absolute values for past periods based on volcanic or archaeological archives (*59*–*61*). Further, the geomagnetic field is fossilized at some depth below the sediment-water interface (lock-in depth), leading to stratigraphic and dating uncertainties. Biases can also be introduced as a result of changes in the sediment composition, notably its magnetic mineral content.

Reconstructions of geomagnetic paleointensity over the past 100,000 years (*62*) using these two sets of archives are characterized by high-amplitude variations, most notably the Laschamps geomagnetic excursion (~41,000 cal BP) (*63*). However, these reconstructions exhibit large uncertainties that impede the accurate estimation of the duration of excursions, the transitional trends into and out of excursions, and any rapid variations during the excursions themselves. These uncertainties hinder precise testing of magnetohydrodynamic models (*64*). Further progress is dependent on both advances in numerical modeling and improvements to the paleomagnetic database.

Radiocarbon and other cosmogenic nuclides also provide constraints on geomagnetic field intensity variations via integration of their production rates over the whole Earth (*65*, *66*). Periods characterized by weak magnetic shielding lead to increases in nuclide production (*67*–*70*). The largest anomaly in the atmospheric $\Delta^{14}C$ record, from 42,000 to 39,000 cal BP (Fig. 4), corresponds to the Laschamps geomagnetic excursion. This excursion also sees enhanced $^{10}Be$ and $^{36}Cl$ deposition in Antarctica and Greenland ice cores (*69*, *71*, *72*) and

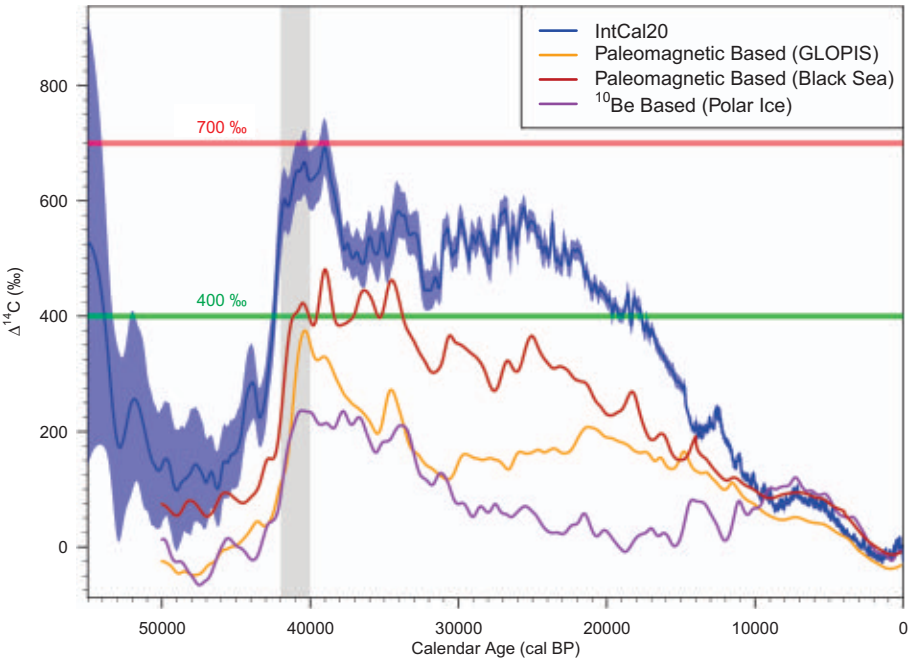



**Fig. 4. Comparison of the IntCal20 atmospheric $\Delta^{14}C$ estimate with paleomagnetic and cosmogenic radionuclide model–based reconstructions.** The measurement-based IntCal20 (blue) is shown alongside Bern3D model–based $\Delta^{14}C$ reconstructions for which $^{14}C$ production rates were calculated using geomagnetic field intensity (orange and brown) and $^{10}Be$ (purple) (*75*). Clear divergences are seen in the size of the atmospheric $\Delta^{14}C$ rise corresponding to the Laschamps geomagnetic excursion (shaded area from 42,000 to 40,000 cal BP) between the IntCal20 estimate and the paleomagnetic- and $^{10}Be$-based reconstructions. The differences extend to the beginning of the Holocene. The paleomagnetic reconstructions are based on the Global Paleointensity Stack (GLOPIS) (*76*) and a high-resolution combination of measurements from the Black Sea (*61*). The $^{10}Be$ reconstruction is based on a combination of radionuclide data from different ice cores including GRIP and GISP (*77*). All model-based reconstructions assume a constant preindustrial carbon cycle. Because of the wide spread of the various model-based reconstructions, it is not possible to conclude whether the differences from the measurement-based IntCal20 atmospheric $\Delta^{14}C$ estimate are due to limitations and uncertainties in the paleomagnetic and $^{10}Be$ reconstructions of $^{14}C$ production rate, or to a lack of knowledge in the $^{14}C$ cycle during the last glacial period.

a prominent $^{10}$Be peak in marine sediment records (*73*, *74*).

Although comparisons of $^{14}$C, $^{10}$Be, and $^{36}$Cl with paleomagnetic intensity records exhibit several correspondences, agreement remains far from optimal. Most notably, during the Laschamps geomagnetic excursion (Fig. 4), the atmospheric $\Delta^{14}$C rise between 46,000 and 41,000 cal BP is ~600‰ (*6*), twice the amplitude predicted by paleomagnetic and cosmogenic radionuclide reconstructions (*10*, *61*, *75*–*77*). This suggests serious weaknesses in existing paleorecords or in our understanding of the relationships between these geophysical and geochemical parameters.

A large discrepancy between different cosmogenic nuclides is also observed between 30,000 and 18,000 cal BP. Here, atmospheric $\Delta^{14}$C stays above 400‰, whereas a polar $^{10}$Be ice record and the mean of the geomagnetic reconstructions indicate a reduction in cosmogenic nuclide production rates that should have led to $\Delta^{14}$C values lower than 200‰ (*75*). This discrepancy could be due to major carbon cycle changes during the last glacial period (*67*, *69*, *75*, *78*). However, interpretation remains challenging because of the spread of paleomagnetic-based $\Delta^{14}$C reconstructions (*75*). Additionally, $^{10}$Be compiled from marine sediments (*73*, *79*) provides better agreement with $^{14}$C, leaving less room for a long-term impact of the carbon cycle (*80*).

### Radiocarbon and synchronization of records

Study of past environmental changes is reliant on combining information from a range of records on diverse aspects of the Earth system. These include not only climate variables such as temperature and rainfall, but also information on plant, animal, and human responses, habitat shifts, and long-term evolution. Each different record has its particular strengths, but their optimal use requires that they be placed on a time scale. However, various different time scales exist and the resultant chronological uncertainties can be considerable. This limits inference when multiple records need to be compared.

There is always the temptation to assume that changes seen in different records are synchronous, but this risks circularity if we are interested in possible leads and lags in the Earth and climate system that could inform on processes of change (*81*). We therefore need methods of comparing records that depend on variables that we have good reason to consider globally synchronous.

Particularly key for study of the past 55,000 years are the $^{14}$C-based time scale and the various ice-core chronologies. The common stratospheric production of cosmogenic isotopes provides a powerful method for their synchronization. For both $^{14}$C and $^{10}$Be, while production is greatest near the poles, intense horizontal mixing and relatively long residence times (*82*) mean that changes in production rates, including annual spikes from SPEs, will be visible in the $^{14}$C record of both hemispheres and $^{10}$Be deposited in ice cores in both polar regions.

As explained previously, although the two nuclides cannot be compared directly (Fig. 3), we can model variations in $^{14}$C levels from the production rate variations seen in $^{10}$Be. Matching common structure in the $^{10}$Be and $^{14}$C signals makes synchronization of $^{14}$C-based chronologies with the ice core time scales possible. During the Holocene, where high-resolution dendrochronologically dated $^{14}$C measurements are available, it is possible to pick up very fine structure and obtain synchronization with uncertainties of just a few years (*83*); for SPEs, annual precision is achievable (*55*). For earlier periods, synchronization is limited to larger excursions in the signal due to potentially confounding effects of carbon cycle changes, dating uncertainties, and lower data quality and resolution. Here, synchronization is also more uncertain and precision is closer to being centennial (*77*).

Both the radiocarbon and ice-core chronologies have their strengths: Ice cores have very good internal relative precision in terms of the time gaps between succeeding events; the $^{14}$C-based time scale has more precise absolute age control through both dendrochronology and U-Th dating methods. For this reason, neither time scale is currently used to correct the other, but rather a time-transfer function is used to convert one to the other, with associated uncertainties (*77*, *83*, *84*). Given the methods used for construction of the radiocarbon calibration curves (see above), the fundamental time scale underlying calibrated radiocarbon dates is dendrochronology from 14,000 to 0 cal BP, and prior to 14,000 cal BP is largely based on the highest resolution and accuracy provided by the U-Th chronology of the Hulu Cave speleothems (*10*, *27*).

Radiocarbon calibration, potentially in combination with deposition modeling (*85*–*87*), enables further environmental records to be synchronized onto this same $^{14}$C-based time scale. With new high-resolution $^{14}$C data, it is possible to get calibrated chronologies for tree-ring series approaching annual resolution, especially in those places where we have abrupt signals from SPEs. More typically, for records such as sedimentary deposits, decadal- to centennial-scale resolution is possible, except in the case of marine sediments where uncertainties in marine reservoir effects often limit our ability to synchronize records on the basis of radiocarbon alone.

Through the combined use of $^{14}$C and $^{10}$Be, it is possible to construct a synchronization framework to cover a broad range of environmental and archaeological records spanning the past 55,000 years (Fig. 5). This framework can be further enhanced with other methods. Between ice cores, the annual counting chronologies can be tied together using volcanic stratigraphic markers such as volcanic ash shards (tephra) or sulfur maxima (*55*, *88*–*94*). Analysis of trace quantities of $CH_4$, another global signal, trapped within the ice (*95*, *96*) allows further synchronization of these ice-core chronologies. On a coarser scale, geomagnetic excursions such as the Laschamps allow linkages to longer Quaternary chronologies such as U-Th–dated speleothems (*10*, *77*). However, these other synchronization techniques are limited to very specific time periods. The keystone of the framework remains the $^{14}$C-based time scale, which provides the continuous metric that can serve as the basis for studying changes over the Holocene and much of the last glacial period.

### Radiocarbon and the carbon cycle

The global carbon cycle plays a crucial role in our climate system. Changes to the carbon cycle affect the concentration of atmospheric $CO_2$, a key greenhouse gas and the main driver of current climate change. Since 1850 CE, the carbon cycle has removed nearly 60% of the anthropogenic $CO_2$ emissions from the atmosphere and stored them away in the ocean and terrestrial biosphere reservoirs (*97*). Understanding the potential for responses and feedbacks of the carbon cycle, and consequently changes to this $CO_2$ sequestration, is therefore critical for future climate projections. Past changes to the carbon cycle are, however, far from completely understood. Although ice core–based atmospheric $CO_2$ reconstructions (*98*) are available, the changes they show await detailed explanation. Because $^{14}$C is dispersed via the carbon cycle after production, it can act as a unique tracer to help investigate.

On the 55,000 to 0 cal BP time scale relevant to $^{14}$C, the dominant roles in the carbon cycle are played by the ocean [with pre-industrial stores of ~37,000 petagrams of carbon (Pg C; 1 Pg = $10^{15}$ g)], the atmosphere (with ~280 ppm of $CO_2$ or ~600 Pg C), and the land carbon cycle (~4000 Pg C) (*99*). Land carbon consists of active carbon (~2400 Pg C) and inert carbon bound in permafrost soil (~1600 Pg C). Although small relative to the ocean stores, the role of this land carbon in determining atmospheric $CO_2$ cannot be neglected. From the Last Glacial Maximum to the preindustrial era, active carbon stores increased while inert carbon stores decreased (*99*), resulting in a net land carbon increase of 450 to 1250 Pg C (*100*).

The carbon cycle's reservoirs are characterized by different $^{14}$C/$^{12}$C ratios related to their respective carbon exchange with the atmosphere. In measuring the $^{14}$C content of fossil samples characterizing these reservoirs, we can reconstruct past carbon cycle changes. The

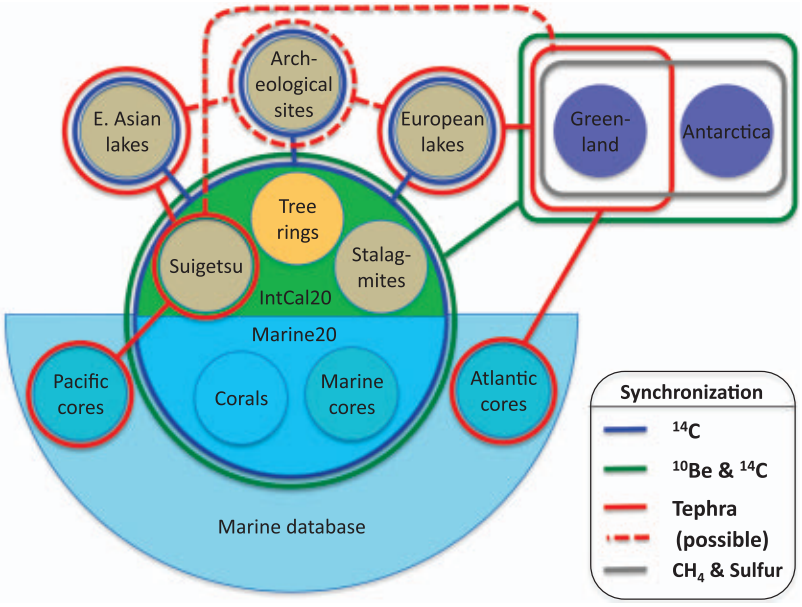



**Fig. 5. Schematic of direct synchronization methods for records spanning the past 55,000 years.** These synchronization methods can be used to enhance or replace climate-based tuning of chronologies. The central circle shows the records contained within the IntCal20 dataset and directly synchronized through the statistical processes involved in the compilation using $^{14}C$. Radiocarbon is also the primary method for relating terrestrial environmental and archaeological records to the primary IntCal $^{14}C$-based time scale. Variation in marine offsets means that other methods such as volcanic ash shards (tephra) and other event-based methods are normally used to synchronize marine records. The link to ice-core chronologies is provided by looking at the relationship between $^{10}Be$ and $^{14}C$. Between ice cores, we can use a combination of local events and global signals, in particular $CH_4$, sulfur maxima, and $^{10}Be$. Volcanic tephra layers provide a check and give precise synchronization at specific points in the chronology; these can be found in sedimentary records, ice cores, and some archaeological sites.

interhemispheric $^{14}C$ gradient can be studied by comparing $^{14}C$ in subfossil trees collected in both hemispheres. Further, planktonic and benthic foraminifera, surface corals, and deep-water corals allow reconstructions of $^{14}C$ gradients at different depths and locations of the world ocean. Probing $^{14}C$ in terrestrial organic matter is more challenging because of the lack of large-scale exchanges between reservoirs and the intense mixing of carbon of different ages at local scale (e.g., in soils) (*101*).

Benthic-planktonic and benthic-atmospheric $^{14}C$ differences have been used to constrain changes in mean ocean circulation. Recent research estimated an increase of ~700 $^{14}C$ years in the average residence time of deep ocean carbon at the Last Glacial Maximum compared to the preindustrial era, which suggests that much of the reduction in atmospheric $CO_2$ levels seen in this glacial period may have been due to greater carbon storage in the mid-depth Pacific (*102*). However, some glacial marine $^{14}C$ levels are so low that they suggest local influx of $^{14}C$-depleted hydrothermal $CO_2$ at particular sites (*103*, *104*). Our understanding of highly depleted deep-ocean $^{14}C$ data is therefore far from complete, requiring more data and the combination of climate- and process-based solid Earth models (*105*).

Most abrupt carbon cycle changes are identifiable by jumps in atmospheric $CO_2$. Across the last deglaciation, three events (~16,500, ~14,600, and ~11,500 cal BP) have been detected where $CO_2$ rose by more than 10 ppm in just 100 to 200 years (*106*). $CO_2$ behavior also changed at the onset of the Younger Dryas (~12,900 cal BP) when levels began to consistently increase after more than 1500 years of relative stability.

Marine data and box-model simulations attribute most centennial- and millennial-scale $CO_2$ changes in the past 70,000 years to ocean circulation, specifically invigoration or weakening of the Atlantic meridional ocean circulation and associated northward and southward shifts of the intertropical convergence zone (*107*, *108*). Radiocarbon can shed further light on this. Recent $^{14}C$ data suggest changes in the $^{14}C$ interhemispheric gradient at the onset of the Younger Dryas (*20*) and reduced North Atlantic Deep Water formation rates (*36*). These temporal interhemispheric $^{14}C$-gradient changes and proposed ocean circulation changes need further understanding with high-resolution Earth system modeling.

The NH permafrost area was reduced by half across the last deglaciation (*109*), implying a massive permafrost thawing and related carbon release detectable by permafrost-specific biomarkers in marine sediment cores bathed in runoff waters (*110*–*115*). Using $^{14}C$, the pre-depositional age of the permafrost carbon can be determined, helping to constrain site-specific climate conditions in the catchment areas. Arctic permafrost thaw due to sea-level rise has been suggested as potentially responsible for the abrupt $CO_2$ jumps around 14,600 cal BP, at the onset of the NH Bølling warming, and at 11,500 cal BP. As such thawing would have released large amounts of old ($^{14}C$-depleted) carbon, the contribution of permafrost can potentially be constrained by changes in atmospheric $^{14}C$ levels (*111*, *116*). These two $CO_2$ jumps are accompanied by synchronous rapid jumps in methane. Because $^{14}C$ measured in methane extracted from ice cores indicates no substantial methane releases from old carbon sources (*117*), it is possible that permafrost thaw may not be the only contributory process.

More detailed inference on climate-based changes to the ocean reservoir of the carbon cycle is available through ocean general circulation models. This is possible because the deep-sea ocean circulation time scale coincides with the $^{14}C$ half-life. Typically, these models prescribe atmospheric $^{14}C$ levels and focus on understanding its ocean dispersal, specifically the levels of oceanic $^{14}C$ depletion. This depletion is location and time specific, and not only influenced by ocean circulation (Fig. 6). Variations in wind stress and sea-ice cover cause spatiotemporal differences in air-sea $^{14}CO_2$ exchange (*118*, *119*). Glacial-interglacial variation in atmospheric partial pressure of $CO_2$ further affects this exchange (*119*, *120*), while turbulent mixing within the ocean may also introduce $^{14}C$ variations in the deep sea unrelated to ocean circulation changes (*121*).

Modern-day $^{14}C$ measurements on seawater have been the primary means used to assess these ocean circulation models (*122*). However, incorporation of further constraints provided by marine $^{14}C$ paleorecords holds promise to improve our understanding of the ocean's role in climate change and its response to rising atmospheric $CO_2$ (*118*, *123*). Unfortunately, this is complicated by the lack of resolution, among those global models that can run long-term simulations, near the ocean margins that provide the majority of $^{14}C$ paleorecords (*124*, *125*). Global multiresolution models may provide a solution (*126*), ideally coupled with models of all other carbon cycle components.

## Future work

Despite the developments in our understanding and use of $^{14}C$, considerable work still remains if we wish to harness its full potential as a tracer. This requires improvements to cosmogenic isotope and paleomagnetic records, as well as improved modeling techniques.

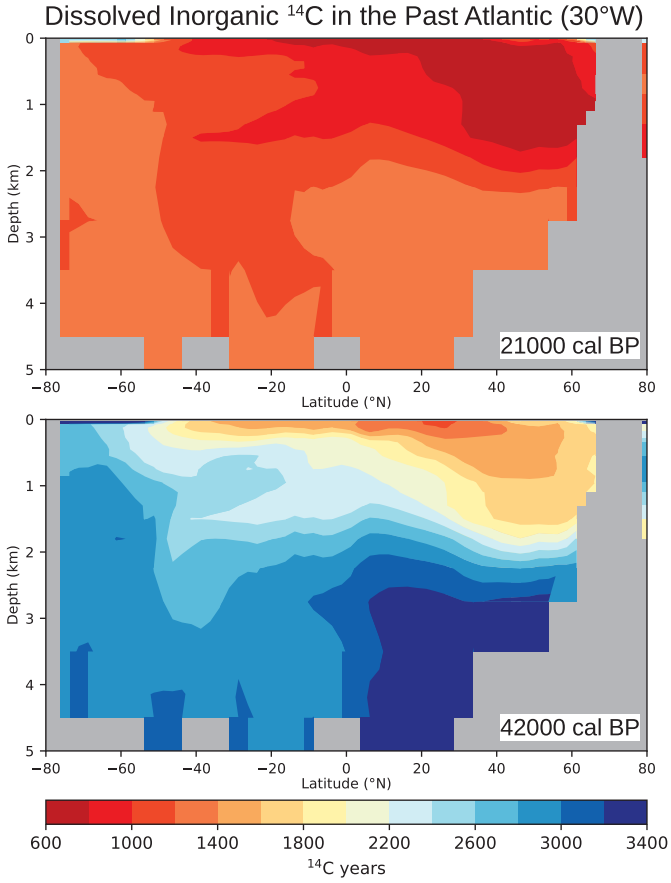



**Fig. 6. Radiocarbon distributions in the past Atlantic.** Depletion of dissolved $^{14}C$ with respect to the contemporaneous atmosphere for meridional sections along 30°W at 21,000 cal BP (around the Last Glacial Maximum, top) and 42,000 cal BP (the onset of the Laschamps geomagnetic excursion, bottom) according to an ocean general circulation model (*12*, *124*). This model was run from 55,000 cal BP to 0 cal BP, forced with IntCal20 atmospheric $\Delta^{14}C$ and ice core $CO_2$ (*6*). Each panel displays the median of nine simulations. Depletion is expressed in terms of $^{14}C$ age. Low $^{14}C$ concentrations translate to high $^{14}C$ ages and vice versa. In both panels, the average ocean circulation is the same. The differences between 21,000 cal BP and 42,000 cal BP are instead due to changes in atmospheric $\Delta^{14}C$ levels in combination with different leads and lags between these atmospheric $\Delta^{14}C$ changes and their oceanic uptake and dispersal.

### *Improving our $^{14}C$ estimates*

Annual-resolution NH $^{14}C$ measurements are currently available on less than 20% of the time scale back to 14,000 cal BP, and in some periods the underlying data remain sparse. Extending annual-resolution tree-ring $^{14}C$ data would allow much longer-term research into past solar activity. Further, this would enable inference on the frequency and size of SPEs and potentially would provide a wider set of markers on which to precisely synchronize $^{14}C$ and ice core records. Annual-resolution extensions to the SH $^{14}C$ data sets would also help us to better understand the interhemispheric $^{14}C$ gradient and constrain past air-sea $CO_2$ fluxes.

Additionally, open questions remain on potential intrahemispheric offsets in $^{14}C$ due to differences in growing season, species, altitude, latitude, or location of the recording tree rings. Currently, deconvolution of these potential intrahemispheric influences is confounded by interlaboratory differences, which are typically of similar scale to any potential offsets (*18*). The current NH $^{14}C$ record is predominantly reliant on trees from European oak and pine. An extension to other tree-ring records would be aided by the development of a broader set of dendrochronologically dated tree rings.

Most critically, we require a truly atmospheric $^{14}C$ record extending back to 55,000 cal BP that also provides sufficient detail to reliably and precisely reconstruct the high-frequency component of the $^{14}C$ signal. Extending annual-resolution $^{14}C$ records back before 14,000 cal BP would provide new information on Earth system processes in a very different climate regime. Further, it would enable $^{14}C$ records that arise from different compartments of the carbon cycle, in particular the marine environment, to be used independently. This would improve our predictions and understanding of climate and carbon cycle changes through enhanced validation of relevant models.

Potential sources of such terrestrial $^{14}C$ data include further macrofossil archives to add to those from Lake Suigetsu (*33*) and, ideally, the discovery of further floating tree-ring sequences. Both of these sources would, however, likely be complicated by the need to estimate their calendar ages. The most immediate hope for an extension perhaps lies in subfossil New Zealand kauri (*Agathis australis*) logs, which have been found to cover much of the $^{14}C$ range, including the Laschamps geomagnetic excursion (*127*). Such logs, even though they would only provide information on SH $^{14}C$ levels, would still offer unprecedented precision and resolution.

### *Improvements to other records and models*

Extensions of $^{10}Be$ and $^{36}Cl$ records to a wider range of archives would be valuable—in particular new polar ice cores, but also less used archives such as marine and lacustrine sediments, low-latitude glacier ice cores, and packrat middens (*128*). Special emphasis should be made to better understand the relationships among production signal, atmospheric transport and deposition, and the $^{10}Be$ and $^{36}Cl$ concentrations in these archives. Radionuclide records obtained at different latitudes, combined with numerical modeling of their production and spread in the atmosphere, would enable checks that $^{10}Be$ and $^{36}Cl$ records provide an accurate picture of the global solar and geomagnetic modulation. Work is required to resolve remaining differences and uncertainties in the relationships between the geomagnetic and heliomagnetic fields and absolute production rates of $^{14}C$, $^{10}Be$, and $^{36}Cl$. These may be due to uncertainties in the galactic cosmic ray flux outside the heliosphere and the production cross sections for cosmogenic radionuclides.

Turning to the geodynamo, research is needed to improve databases on volcanic rocks and sediments, and to correct for the effects of non-dipolar components and various sources of blurring of individual paleomagnetic records (bioturbation, variable lock-in depth). Further, many paleomagnetic records based on marine sediments are not precisely dated, which could introduce biases when they are combined (stacked) to estimate global geodynamo changes. Improvements in carbon cycle modeling are also needed to better understand the differences that currently exist between measurement-based atmospheric $\Delta^{14}C$ estimates and model-based reconstructions using $^{14}C$ production rates obtained from $^{10}Be$ and paleomagnetic records.

Combined, such advances would provide a step-change in our ability to separate the geomagnetic, solar, and carbon cycle components

of the $^{14}C$ record. This would potentially permit detection of short-term geomagnetic events and the resolution of hypothesized long-term trends in solar activity and changes in the carbon cycle.

## REFERENCES AND NOTES


1. W. F. Libby, E. C. Anderson, J. R. Arnold, Age Determination by Radiocarbon Content: World-Wide Assay of Natural Radiocarbon. *Science* **109**, 227–228 (1949). doi: 10.1126/science.109.2827.227; pmid: 17818054
2. Q. Hua, M. Barbetti, A. Z. Rakowski, Atmospheric Radiocarbon for the Period 1950–2010. *Radiocarbon* **55**, 2059–2072 (2013). doi: 10.2458/azu_js_rc.v55i2.16177
3. H. Graven, R. F. Keeling, J. Rogelj, Changes to Carbon Isotopes in Atmospheric $CO_2$ Over the Industrial Era and Into the Future. *Global Biogeochem. Cycles* **34**, GB006170 (2020). doi: 10.1029/2019GB006170; pmid: 33380771
4. H. E. Suess, Radiocarbon Concentration in Modern Wood. *Science* **122**, 415–417 (1955). doi: 10.1126/science.122.3166.415-a
5. P. Köhler, Using the Suess effect on the stable carbon isotope to distinguish the future from the past in radiocarbon. *Environ. Res. Lett.* **11**, 124016 (2016). doi: 10.1088/1748-9326/11/12/124016
6. P. J. Reimer *et al*., The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020). doi: 10.1017/RDC.2020.41
7. A. G. Hogg *et al*., SHCal20 Southern Hemisphere Calibration, 0–55,000 Years cal BP. *Radiocarbon* **62**, 759–778 (2020). doi: 10.1017/RDC.2020.59
8. T. J. Heaton *et al*., Marine20—The Marine Radiocarbon Age Calibration Curve (0–55,000 cal BP). *Radiocarbon* **62**, 779–820 (2020). doi: 10.1017/RDC.2020.68
9. L. Wacker, M. Němec, J. Bourquin, A revolutionary graphitisation system: Fully automated, compact and simple. *Nucl. Instrum. Methods Phys. Res. B* **268**, 931–934 (2010). doi: 10.1016/j.nimb.2009.10.067
10. H. Cheng *et al*., Atmospheric $^{14}C/^{12}C$ changes during the last glacial period from Hulu Cave. *Science* **362**, 1293–1297 (2018). doi: 10.1126/science.aau0747; pmid: 30545886
11. T. J. Heaton *et al*., The IntCal20 Approach to Radiocarbon Calibration Curve Construction: A New Methodology Using Bayesian Splines and Errors-in-Variables. *Radiocarbon* **62**, 821–863 (2020). doi: 10.1017/RDC.2020.46
12. M. Butzin, T. J. Heaton, P. Köhler, G. Lohmann, A Short Note on Marine Reservoir Age Simulations Used in IntCal20. *Radiocarbon* **62**, 865–871 (2020). doi: 10.1017/RDC.2020.9
13. P. J. Reimer *et al*., IntCal13 and Marine13 radiocarbon age calibration curves 0-50,000 years cal BP. *Radiocarbon* **55**, 1869–1887 (2013). doi: 10.2458/azu_js_rc.55.16947
14. A. G. Hogg *et al*., SHCal13 Southern Hemisphere calibration, 0-50,000 years cal BP. *Radiocarbon* **55**, 1889–1903 (2013). doi: 10.2458/azu_js_rc.55.16783
15. M. Stuiver, H. A. Polach, Discussion Reporting of $^{14}C$ Data. *Radiocarbon* **19**, 355–363 (1977). doi: 10.1017/S0033822200003672
16. M. Suter, R. Huber, S. A. W. Jacob, H.-A. Synal, J. B. Schroeder, A new small accelerator for radiocarbon dating. *AIP Conf. Proc.* **475**, 665–667 (1999). doi: 10.1063/1.59210
17. H.-A. Synal, M. Stocker, M. Suter, MICADAS: A new compact radiocarbon AMS system. *Nucl. Instrum. Methods Phys. Res. B* **259**, 7–13 (2007).
18. A. Bayliss *et al*., IntCal20 Tree Rings: An Archaeological Swot Analysis. *Radiocarbon* **62**, 1045–1078 (2020). doi: 10.1017/RDC.2020.77
19. F. Reinig *et al*., Illuminating IntCal During the Younger Dryas. *Radiocarbon* **62**, 883–889 (2020). doi: 10.1017/RDC.2020.15
20. M. Capano *et al*., Onset of the Younger Dryas Recorded with $^{14}C$ at Annual Resolution in French Subfossil Trees. *Radiocarbon* **62**, 901–918 (2020). doi: 10.1017/RDC.2019.116
21. N. Brehm *et al*., Eleven-year solar cycles over the last millennium revealed by radiocarbon in tree rings. *Nat. Geosci.* **14**, 10–15 (2021). doi: 10.1038/s41561-020-00674-0
22. F. Miyake, K. Nagaya, K. Masuda, T. Nakamura, A signature of cosmic-ray increase in AD 774-775 from tree rings in Japan. *Nature* **486**, 240–242 (2012). doi: 10.1038/nature11123; pmid: 22699615
23. F. Miyake, K. Masuda, T. Nakamura, Another rapid event in the carbon-14 content of tree rings. *Nat. Commun.* **4**, 1748 (2013). doi: 10.1038/ncomms2783; pmid: 23612289
24. F. Mekhaldi *et al*., Multiradionuclide evidence for the solar origin of the cosmic-ray events of AD 774/5 and 993/4. *Nat. Commun.* **6**, 8611 (2015). doi: 10.1038/ncomms9611; pmid: 26497389
25. L. Wacker *et al*., Findings from an in-Depth Annual Tree-Ring Radiocarbon Intercomparison. *Radiocarbon* **62**, 873–882 (2020). doi: 10.1017/RDC.2020.49
26. E. M. Scott, P. Naysmith, G. T. Cook, Should Archaeologists Care about $^{14}C$ Intercomparisons? Why? A Summary Report on SIRI. *Radiocarbon* **59**, 1589–1596 (2017). doi: 10.1017/RDC.2017.12
27. J. Southon, A. L. Noronha, H. Cheng, R. L. Edwards, Y. Wang, A high-resolution record of atmospheric $^{14}C$ based on Hulu Cave speleothem H82. *Quat. Sci. Rev.* **33**, 32–41 (2012). doi: 10.1016/j.quascirev.2011.11.022
28. F. Adolphi *et al*., Radiocarbon calibration uncertainties during the last deglaciation: Insights from new floating tree-ring chronologies. *Quat. Sci. Rev.* **170**, 98–108 (2017). doi: 10.1016/j.quascirev.2017.06.026
29. C. S. M. Turney *et al*., The potential of New Zealand kauri (Agathis australis) for testing the synchronicity of abrupt climate change during the Last Glacial Interval (60,000–11,700 years ago). *Quat. Sci. Rev.* **29**, 3677–3682 (2010). doi: 10.1016/j.quascirev.2010.08.017
30. C. S. M. Turney *et al*., High-precision dating and correlation of ice, marine and terrestrial sequences spanning Heinrich Event 3: Testing mechanisms of interhemispheric change using New Zealand ancient kauri (Agathis australis). *Quat. Sci. Rev.* **137**, 126–134 (2016). doi: 10.1016/j.quascirev.2016.02.005
31. K. A. Hughen, J. R. Southon, C. J. H. Bertrand, B. Frantz, P. Zermeño, Cariaco Basin Calibration Update: Revisions to Calendar and $^{14}C$ Chronologies for Core PI07-58Pc. *Radiocarbon* **46**, 1161–1187 (2004). doi: 10.1017/S0033822200033075
32. C. Bronk Ramsey *et al*., Reanalysis of the Atmospheric Radiocarbon Calibration Record from Lake Suigetsu, Japan. *Radiocarbon* **62**, 989–999 (2020). doi: 10.1017/RDC.2020.18
33. C. Bronk Ramsey *et al*., A complete terrestrial radiocarbon record for 11.2 to 52.8 kyr B.P. *Science* **338**, 370–374 (2012). doi: 10.1126/science.1226660; pmid: 23087245
34. H. Cheng *et al*., The Asian monsoon over the past 640,000 years and ice age terminations. *Nature* **534**, 640–646 (2016). doi: 10.1038/nature18591; pmid: 27357793
35. E. C. Corrick *et al*., Synchronous timing of abrupt climate changes during the last glacial period. *Science* **369**, 963–969 (2020). doi: 10.1126/science.aay5538; pmid: 32820122
36. F. Muschitiello *et al*., Deep-water circulation changes lead North Atlantic climate during deglaciation. *Nat. Commun.* **10**, 1272 (2019). doi: 10.1038/s41467-019-09237-3; pmid: 30894523
37. J. A. Eddy, The Maunder minimum. *Science* **192**, 1189–1202 (1976). doi: 10.1126/science.192.4245.1189; pmid: 17771739
38. F. Steinhilber *et al*., 9,400 years of cosmic radiation and solar activity from ice cores and tree rings. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5967–5971 (2012). doi: 10.1073/pnas.1118965109; pmid: 22474348
39. R. Roth, F. Joos, A reconstruction of radiocarbon production and total solar irradiance from the Holocene $^{14}C$ and $CO_2$ records: Implications of data and model uncertainties. *Clim. Past* **9**, 1879–1909 (2013). doi: 10.5194/cp-9-1879-2013
40. E. Bard, G. Raisbeck, F. Yiou, J. Jouzel, Solar irradiance during the last 1200 years based on cosmogenic nuclides. *Tellus B* **52**, 985–992 (2000). doi: 10.3402/tellusb.v52i3.17080
41. F. Adolphi *et al*., Persistent link between solar activity and Greenland climate during the Last Glacial Maximum. *Nat. Geosci.* **7**, 662–666 (2014). doi: 10.1038/ngeo2225
42. P. O'Hare *et al*., Multiradionuclide evidence for an extreme solar proton event around 2,610 B.P. (660 BC). *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5961–5966 (2019). doi: 10.1073/pnas.1815725116; pmid: 30858311
43. G. A. Kovaltsov, A. Mishev, I. G. Usoskin, A new model of cosmogenic production of radiocarbon $^{14}C$ in the atmosphere. *Earth Planet. Sci. Lett.* **337–338**, 114–120 (2012). doi: 10.1016/j.epsl.2012.05.036
44. J. Masarik, J. Beer, Simulation of particle fluxes and cosmogenic nuclide production in the Earth's atmosphere. *J. Geophys. Res.* **104**, 12099–12111 (1999). doi: 10.1029/1998JD200091
45. U. Siegenthaler, M. Heimann, H. Oeschger, $^{14}C$ Variations Caused by Changes in the Global Carbon Cycle. *Radiocarbon* **22**, 177–191 (1980). doi: 10.1017/S0033822200009449
46. P. E. Damon, C. P. Sonett, in *The Sun in Time*, C. P. Sonett, M. S. Giampapa, M. S. Matthews, Eds. (Univ. of Arizona Press, 1991), p. 360.
47. M. Vonmoos, J. Beer, R. Muscheler, Large variations in Holocene solar activity: Constraints from $^{10}Be$ in the Greenland Ice Core Project ice core. *J. Geophys. Res.* **111**, A10105 (2006). doi: 10.1029/2005JA011500
48. L. Svalgaard, K. H. Schatten, Reconstruction of the Sunspot Group Number: The Backbone Method. *Sol. Phys.* **291**, 2653–2684 (2016). doi: 10.1007/s11207-015-0815-8
49. J. H. Jungclaus *et al*., The PMIP4 contribution to CMIP6 – Part 3: The last millennium, scientific objective, and experimental design for the PMIP4 *past1000* simulations. *Geosci. Model Dev.* **10**, 4005–4033 (2017). doi: 10.5194/gmd-10-4005-2017
50. G. Bond *et al*., Persistent solar influence on North Atlantic climate during the Holocene. *Science* **294**, 2130–2136 (2001). doi: 10.1126/science.1065680; pmid: 11739949
51. R. Muscheler *et al*., Changes in the carbon cycle during the last deglaciation as indicated by the comparison of $^{10}Be$ and $^{14}C$ records. *Earth Planet. Sci. Lett.* **219**, 325–340 (2004). doi: 10.1016/S0012-821X(03)00722-2
52. A. Cauquoin, G. M. Raisbeck, J. Jouzel, E. Bard, No evidence for planetary influence on solar activity 330 000 years ago. *Astron. Astrophys.* **561**, A132 (2014). doi: 10.1051/0004-6361/201322879
53. W. R. Webber, P. R. Higbie, K. G. McCracken, Production of the cosmogenic isotopes $^{3}H$, $^{7}Be$, $^{10}Be$, and $^{36}Cl$ in the Earth's atmosphere by solar and galactic cosmic rays. *J. Geophys. Res.* **112**, A10106 (2007). doi: 10.1029/2007JA012499
54. U. Büntgen *et al*., Tree rings reveal globally coherent signature of cosmogenic radiocarbon events in 774 and 993 CE. *Nat. Commun.* **9**, 3605 (2018). doi: 10.1038/s41467-018-06036-0; pmid: 30190505
55. M. Sigl *et al*., Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature* **523**, 543–549 (2015). doi: 10.1038/nature14565; pmid: 26153860
56. D. Gubbins, A. L. Jones, C. C. Finlay, Fall in Earth's magnetic field is erratic. *Science* **312**, 900–902 (2006). doi: 10.1126/science.1124855; pmid: 16690863
57. E. Thellier, O. Thellier, Sur l'intensité du champ magnétique terrestre, en France, trois siècles avant les premières mesures directes. Application, au problème de la désaimantation du globe. *C. R. Acad. Sci. Paris* **214**, 382–384 (1942).
58. M. C. Brown *et al*., GEOMAGIA50.v3: 1. general structure and modifications to the archeological and volcanic database. *Earth Planets Space* **67**, 83 (2015). doi: 10.1186/s40623-015-0232-0
59. C. Laj, C. Kissel, A. Mazaud, J. E. T. Channell, J. Beer, North Atlantic palaeointensity stack since 75ka (NAPIS–75) and the duration of the Laschamp event. *Philos. Trans. R. Soc. London. Ser. A* **358**, 1009–1025 (2000). doi: 10.1098/rsta.2000.0571
60. N. Thouveny, J. Carcaillet, E. Moreno, G. Leduc, D. Nérini, Geomagnetic moment variation and paleomagnetic excursions since 400 kyr BP: A stacked record from sedimentary sequences of the Portuguese margin. *Earth Planet. Sci. Lett.* **219**, 377–396 (2004). doi: 10.1016/S0012-821X(03)00701-5
61. N. R. Nowaczyk, U. Frank, J. Kind, H. W. Arz, A high-resolution paleointensity stack of the past 14 to 68 ka from Black Sea sediments. *Earth Planet. Sci. Lett.* **384**, 1–16 (2013). doi: 10.1016/j.epsl.2013.09.028
62. S. Panovska, M. Korte, C. G. Constable, One Hundred Thousand Years of Geomagnetic Field Evolution. *Rev. Geophys.* **57**, 1289–1337 (2019). doi: 10.1029/2019RG000656
63. I. Lascu, J. M. Feinberg, J. A. Dorale, H. Cheng, R. L. Edwards, Age of the Laschamp excursion determined by U-Th dating of a speleothem geomagnetic record from North America. *Geology* **44**, 139–142 (2016). doi: 10.1130/G37490.1
64. G. A. Glatzmaiers, P. H. Roberts, A three-dimensional self-consistent computer simulation of a geomagnetic field reversal. *Nature* **377**, 203–209 (1995). doi: 10.1038/377203a0
65. W. Elsasser, E. P. Ney, J. R. Winckler, Cosmic-Ray Intensity and Geomagnetism. *Nature* **178**, 1226–1227 (1956). doi: 10.1038/1781226a0
66. S. V. Poluianov, G. A. Kovaltsov, A. L. Mishev, I. G. Usoskin, Production of cosmogenic isotopes $^{7}Be$, $^{10}Be$, $^{14}C$, $^{22}Na$, and $^{36}Cl$ in the atmosphere: Altitudinal profiles of yield functions.

*J. Geophys. Res.* **121**, 8125–8136 (2016). doi: 10.1002/2016JD025034

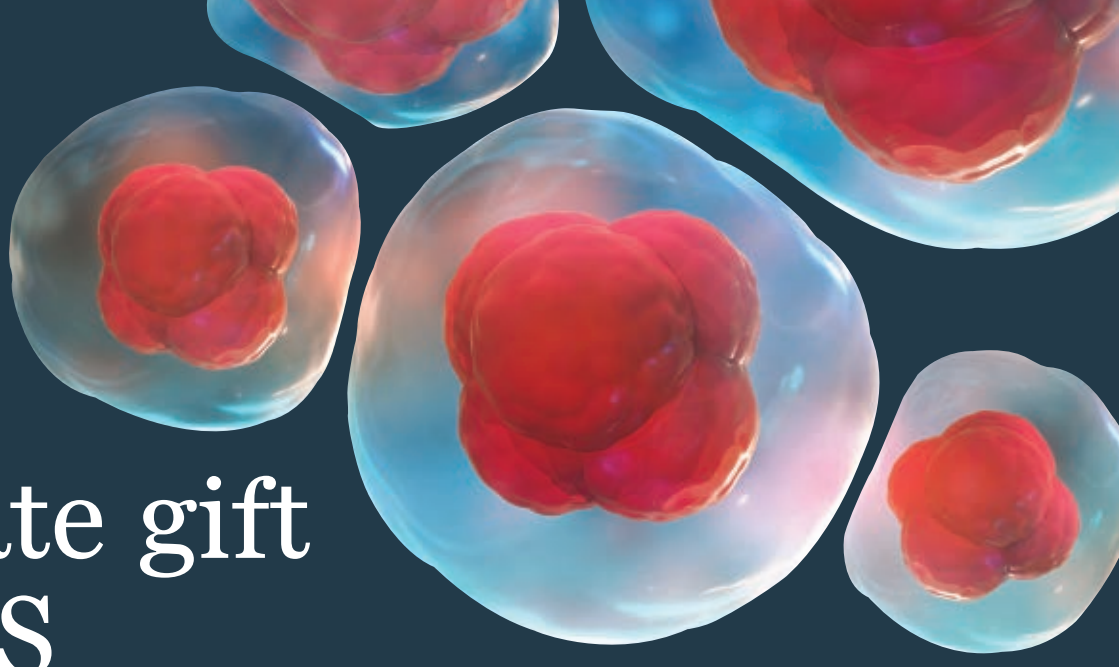
67. E. Bard, B. Hamelin, R. G. Fairbanks, A. Zindler, Calibration of the $^{14}C$ timescale over the past 30,000 years using mass spectrometric U–Th ages from Barbados corals. *Nature* **345**, 405–410 (1990). doi: 10.1038/345405a0
68. K. Hughen *et al*., $^{14}C$ activity and global carbon cycle changes over the past 50,000 years. *Science* **303**, 202–207 (2004). doi: 10.1126/science.1090300; pmid: 14716006
69. R. Muscheler, J. Beer, P. W. Kubik, H.-A. Synal, Geomagnetic field intensity during the last 60,000 years based on $^{10}Be$ and $^{36}Cl$ from the Summit ice cores and $^{14}C$. *Quat. Sci. Rev.* **24**, 1849–1860 (2005). doi: 10.1016/j.quascirev.2005.01.012
70. J. E. T. Channell, D. A. Hodell, S. J. Crowhurst, L. C. Skinner, R. Muscheler, Relative paleointensity (RPI) in the latest Pleistocene (10–45 ka) and implications for deglacial atmospheric radiocarbon. *Quat. Sci. Rev.* **191**, 57–72 (2018). doi: 10.1016/j.quascirev.2018.05.007
71. G. M. Raisbeck *et al*., Evidence for two intervals of enhanced $^{10}Be$ deposition in Antarctic ice during the last glacial period. *Nature* **326**, 273–277 (1987). doi: 10.1038/326273a0
72. G. M. Raisbeck *et al*., An improved north–south synchronization of ice core records around the 41 kyr $^{10}Be$ peak. *Clim. Past* **13**, 217–229 (2017). doi: 10.5194/cp-13-217-2017
73. M. Frank *et al*., A 200 kyr record of cosmogenic radionuclide production rate and geomagnetic field intensity from $^{10}Be$ in globally stacked deep-sea sediments. *Earth Planet. Sci. Lett.* **149**, 121–129 (1997). doi: 10.1016/S0012-821X(97)00070-8
74. Q. Simon, N. Thouveny, D. L. Bourlès, J.-P. Valet, F. Bassinot, Cosmogenic $^{10}Be$ production records reveal dynamics of geomagnetic dipole moment (GDM) over the Laschamp excursion (20–60 ka). *Earth Planet. Sci. Lett.* **550**, 116547 (2020). doi: 10.1016/j.epsl.2020.116547
75. A. Dinauer, F. Adolphi, F. Joos, Mysteriously high $\Delta^{14}C$ of the glacial atmosphere: Influence of $^{14}C$ production and carbon cycle changes. *Clim. Past* **16**, 1159–1185 (2020). doi: 10.5194/cp-16-1159-2020
76. C. Laj, C. Kissel, J. Beer, in *Timescales of the Paleomagnetic Field*, J. E. T. Channell, D. V. Kent, W. Lowrie, J. G. Meert, Eds. (American Geophysical Union, 1994), pp. 255–265. doi: 10.1029/145GM19
77. F. Adolphi *et al*., Connecting the Greenland ice-core and U/Th timescales via cosmogenic radionuclides: Testing the synchroneity of Dansgaard–Oeschger events. *Clim. Past* **14**, 1755–1781 (2018). doi: 10.5194/cp-14-1755-2018
78. P. Köhler, R. Muscheler, H. Fischer, A model-based interpretation of low-frequency changes in the carbon cycle during the last 120,000 years and its implications for the reconstruction of atmospheric $\Delta^{14}C$. *Geochem. Geophys. Geosyst.* **7**, Q11N06 (2006). doi: 10.1029/2005GC001228
79. Y. Lao *et al*., Increased production of cosmogenic $^{10}Be$ during the Last Glacial Maximum. *Nature* **357**, 576–578 (1992). doi: 10.1038/357576a0
80. E. Bard, Geochemical and geophysical implications of the radiocarbon calibration. *Geochim. Cosmochim. Acta* **62**, 2025–2038 (1998). doi: 10.1016/S0016-7037(98)00130-6
81. S. P. E. Blockley *et al*., Synchronisation of palaeoenvironmental records over the last 60,000 years, and an extended INTIMATE event stratigraphy to 48,000 b2k. *Quat. Sci. Rev.* **36**, 2–10 (2012). doi: 10.1016/j.quascirev.2011.09.017
82. U. Heikkilä, J. Beer, J. Feichter, Meridional transport and deposition of atmospheric $^{10}Be$. *Atmos. Chem. Phys.* **9**, 515–527 (2009). doi: 10.5194/acp-9-515-2009
83. F. Adolphi, R. Muscheler, Synchronizing the Greenland ice core and radiocarbon timescales over the Holocene-Bayesian wiggle-matching of cosmogenic radionuclide records. *Clim. Past* **12**, 15–30 (2016). doi: 10.5194/cp-12-15-2016
84. C. Bronk Ramsey *et al*., Integrating timescales with time-transfer functions: A practical approach for an INTIMATE database. *Quat. Sci. Rev.* **106**, 67–80 (2014). doi: 10.1016/j.quascirev.2014.05.028
85. M. Blaauw, J. A. Christen, Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Anal.* **6**, 457–474 (2011). doi: 10.1214/ba/1339616472
86. C. Bronk Ramsey, Deposition models for chronological records. *Quat. Sci. Rev.* **27**, 42–60 (2008). doi: 10.1016/j.quascirev.2007.01.019
87. J. Haslett, A. Parnell, A Simple Monotone Process with Application to Radiocarbon-Dated Depth Chronologies. *J. R. Stat. Soc. Ser. C* **57**, 399–418 (2008). doi: 10.1111/j.1467-9876.2008.00623.x
88. L. Bazin *et al*., An optimized multi-proxy, multi-site Antarctic ice and gas orbital chronology (AICC2012): 120–800 ka. *Clim. Past* **9**, 1715–1731 (2013). doi: 10.5194/cp-9-1715-2013
89. C. Bronk Ramsey *et al*., Improved age estimates for key Late Quaternary European tephra horizons in the RESET lattice. *Quat. Sci. Rev.* **118**, 18–32 (2015). doi: 10.1016/j.quascirev.2014.11.007
90. C. Buizert *et al*., Abrupt ice-age shifts in southern westerly winds and Antarctic climate forced from the north. *Nature* **563**, 681–685 (2018). doi: 10.1038/s41586-018-0727-5; pmid: 30487614
91. S. O. Rasmussen *et al*., A new Greenland ice core chronology for the last glacial termination. *J. Geophys. Res.* **111**, D06102 (2006). doi: 10.1029/2005JD006079
92. M. Sigl *et al*., The WAIS Divide deep ice core WD2014 chronology – Part 2: Annual-layer counting (0–31 ka BP). *Clim. Past* **12**, 769–786 (2016). doi: 10.5194/cp-12-769-2016
93. A. Svensson *et al*., Bipolar volcanic synchronization of abrupt climate change in Greenland and Antarctic ice cores during the last glacial period. *Clim. Past* **16**, 1565–1580 (2020). doi: 10.5194/cp-16-1565-2020
94. D. Veres *et al*., The Antarctic ice core chronology (AICC2012): An optimized multi-parameter and multi-site dating approach for the last 120 thousand years. *Clim. Past* **9**, 1733–1748 (2013). doi: 10.5194/cp-9-1733-2013
95. EPICA Community Members, One-to-one coupling of glacial climate variability in Greenland and Antarctica. *Nature* **444**, 195–198 (2006). doi: 10.1038/nature05301; pmid: 17099953
96. C. Buizert *et al*., The WAIS Divide deep ice core WD2014 chronology – Part 1: Methane synchronization (68–31 ka BP) and the gas age–ice age difference. *Clim. Past* **11**, 153–173 (2015). doi: 10.5194/cp-11-153-2015
97. P. Friedlingstein *et al*., Global Carbon Budget 2019. *Earth Syst. Sci. Data* **11**, 1783–1838 (2019). doi: 10.5194/essd-11-1783-2019
98. B. Bereiter *et al*., Revision of the EPICA Dome C $CO_2$ record from 800 to 600 kyr before present. *Geophys. Res. Lett.* **42**, 542–549 (2015). doi: 10.1002/2014GL061957
99. P. Ciais *et al*., Large inert carbon pool in the terrestrial biosphere during the Last Glacial Maximum. *Nat. Geosci.* **5**, 74–79 (2012). doi: 10.1038/ngeo1324
100. A. Jeltsch-Thömmes, G. Battaglia, O. Cartapanis, S. L. Jaccard, F. Joos, Low terrestrial carbon storage at the Last Glacial Maximum: Constraints from multi-proxy data. *Clim. Past* **15**, 849–879 (2019). doi: 10.5194/cp-15-849-2019
101. Z. Shi *et al*., The age distribution of global soil carbon inferred from radiocarbon measurements. *Nat. Geosci.* **13**, 555–559 (2020). doi: 10.1038/s41561-020-0596-z
102. L. C. Skinner *et al*., Radiocarbon constraints on the glacial ocean circulation and its impact on atmospheric $CO_2$. *Nat. Commun.* **8**, 16010 (2017). doi: 10.1038/ncomms16010; pmid: 28703126
103. L. Stott, J. Southon, A. Timmermann, A. Koutavas, Radiocarbon age anomaly at intermediate water depth in the Pacific Ocean during the last deglaciation. *Paleoceanography* **24**, PA2223 (2009). doi: 10.1029/2008PA001690
104. T. A. Ronge *et al*., Radiocarbon constraints on the extent and evolution of the South Pacific glacial carbon pool. *Nat. Commun.* **7**, 11487 (2016). doi: 10.1038/ncomms11487; pmid: 27157845
105. J. Hasenclever *et al*., Sea level fall during glaciation stabilized atmospheric $CO_2$ by enhanced volcanic degassing. *Nat. Commun.* **8**, 15867 (2017). doi: 10.1038/ncomms15867; pmid: 28681844
106. S. A. Marcott *et al*., Centennial-scale changes in the global carbon cycle during the last deglaciation. *Nature* **514**, 616–619 (2014). doi: 10.1038/nature13799; pmid: 25355363
107. T. K. Bauska *et al*., Carbon isotopes characterize rapid changes in atmospheric carbon dioxide during the last deglaciation. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3465–3470 (2016). doi: 10.1073/pnas.1513868113; pmid: 26976561
108. T. K. Bauska *et al*., Controls on Millennial-Scale Atmospheric $CO_2$ Variability During the Last Glacial Period. *Geophys. Res. Lett.* **45**, 7731–7740 (2018). doi: 10.1029/2018GL077881
109. J. Vandenberghe *et al*., The Last Permafrost Maximum (LPM) map of the Northern Hemisphere: Permafrost extent and mean annual air temperatures, 25–17 ka BP. *Boreas* **43**, 652–666 (2014). doi: 10.1111/bor.12070
110. F. Rostek, E. Bard, Hydrological changes in eastern europe during the last 40,000 yr inferred from biomarkers in Black Sea Sediments. *Quat. Res.* **80**, 502–509 (2013). doi: 10.1016/j.yqres.2013.07.003
111. M. Winterfeld *et al*., Deglacial mobilization of pre-aged terrestrial carbon from degrading permafrost. *Nat. Commun.* **9**, 3666 (2018). doi: 10.1038/s41467-018-06080-w; pmid: 30201999
112. T. Tesi *et al*., Massive remobilization of permafrost carbon during post-glacial warming. *Nat. Commun.* **7**, 13653 (2016). doi: 10.1038/ncomms13653; pmid: 27897191
113. J. Martens *et al*., Remobilization of Old Permafrost Carbon to Chukchi Sea Sediments During the End of the Last Deglaciation. *Global Biogeochem. Cycles* **33**, 2–14 (2019). doi: 10.1029/2018GB005969; pmid: 31007381
114. V. D. Meyer *et al*., Permafrost-carbon mobilization in Beringia caused by deglacial meltwater runoff, sea-level rise and warming. *Environ. Res. Lett.* **14**, 085003 (2019). doi: 10.1088/1748-9326/ab2653
115. J. Martens *et al*., Remobilization of dormant carbon from Siberian-Arctic permafrost during three past warming events. *Sci. Adv.* **6**, eabb6546 (2020). doi: 10.1126/sciadv.abb6546; pmid: 33067229
116. P. Köhler, G. Knorr, E. Bard, Permafrost thawing as a possible source of abrupt carbon release at the onset of the Bølling/Allerød. *Nat. Commun.* **5**, 5520 (2014). doi: 10.1038/ncomms6520; pmid: 25409739
117. M. N. Dyonisius *et al*., Old carbon reservoirs were not important in the deglacial methane budget. *Science* **367**, 907–910 (2020). doi: 10.1126/science.aax0504; pmid: 32079770
118. M. Butzin, M. Prange, G. Lohmann, Radiocarbon simulations for the glacial ocean: The effects of wind stress, Southern Ocean sea ice and Heinrich events. *Earth Planet. Sci. Lett.* **235**, 45–61 (2005). doi: 10.1016/j.epsl.2005.03.003
119. E. Bard, Correction of accelerator mass spectrometry $^{14}C$ ages measured in planktonic foraminifera: Paleoceanographic implications. *Paleoceanography* **3**, 635–645 (1988). doi: 10.1029/PA003i006p00635
120. E. D. Galbraith, E. Y. Kwon, D. Bianchi, M. P. Hain, J. L. Sarmiento, The impact of atmospheric $pCO_2$ on carbon isotope ratios of the atmosphere and ocean. *Global Biogeochem. Cycles* **29**, 307–324 (2015). doi: 10.1002/2014GB004929
121. T. DeVries, F. Primeau, An improved method for estimating water-mass ventilation age from radiocarbon data. *Earth Planet. Sci. Lett.* **295**, 367–378 (2010). doi: 10.1016/j.epsl.2010.04.011
122. K. Matsumoto *et al*., Evaluation of ocean carbon cycle models with data-based metrics. *Geophys. Res. Lett.* **31**, L07303 (2004). doi: 10.1029/2003GL018970
123. T. Tschumi, F. Joos, M. Gehlen, C. Heinze, Deep ocean ventilation, carbon isotopes, marine sedimentation and the deglacial $CO_2$ rise. *Clim. Past* **7**, 771–800 (2011). doi: 10.5194/cp-7-771-2011
124. M. Butzin, P. Köhler, G. Lohmann, Marine radiocarbon reservoir age simulations for the past 50,000 years. *Geophys. Res. Lett.* **44**, 8473–8480 (2017). doi: 10.1002/2017GL074688
125. N. Zhao, O. Marchal, L. Keigwin, D. Amrhein, G. Gebbie, A Synthesis of Deglacial Deep-Sea Radiocarbon Records and Their (In)Consistency With Modern Ocean Ventilation. *Paleoceanogr. Paleoclimatol.* **33**, 128–151 (2018). doi: 10.1002/2017PA003174
126. G. Lohmann, M. Butzin, N. Eissner, X. Shi, C. Stepanek, Abrupt climate and weather changes across time scales. Paleoceanogr. *Paleoclimatol.* **35**, e2019PA003782 (2020). doi: 10.1029/2019PA003782
127. A. Cooper *et al*., A global environmental crisis 42,000 years ago. *Science* **371**, 811–818 (2021). doi: 10.1126/science.abb8677; pmid: 33602851
128. M. A. Plummer *et al*., Chlorine-36 in fossil rat urine: An archive of cosmogenic nuclide deposition during the past 40,000 years. *Science* **277**, 538–541 (1997). doi: 10.1126/science.277.5325.538; pmid: 9227999
129. K. G. McCracken, J. Beer, Long-term changes in the cosmic ray intensity at Earth, 1428–2005. *J. Geophys. Res.* **112**, A10101 (2007). doi: 10.1029/2006JA012117
130. A.-M. Berggren *et al*., A 600-year annual $^{10}Be$ record from the NGRIP ice core, Greenland. *Geophys. Res. Lett.* **36**, L11801 (2009). doi: 10.1029/2009GL038004

**ACKNOWLEDGMENTS**

We thank E. Rhodes and S. Quegan for insightful comments that improved our proposal; F. Adolphi for discussion on $^{14}C$ production rate reconstructions; L. Johnstone and S. Renwick for help with the illustrations; and C. Pratt, J. Ely, and S. Whyte for feedback on earlier drafts. **Funding:** Supported by Leverhulme Trust Fellowship RF-2019-140\9 (T.J.H.); EQUIPEX ASTER-CEREGE and ANR CARBOTRYDH (E.B.); the German Federal Ministry of Education and Research (BMBF) as a Research for Sustainability initiative (FONA) through the PalMod project (grant 01LP1919A) (M.B., P.K.); Swedish Research Council grants DNR2013-8421 and DNR2018-05469 (R.M.); the Natural Environment Research Council (NERC) as part of the National Environmental Isotope Facility (NEIF) (C.B.R.); and UKRI Natural Environment Research Council grant NE/M004619/1 (P.J.R.). **Author contributions:** All authors contributed to the development of ideas for the review, the writing of initial drafts, and subsequent revision. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data for Fig. 6 are accessible at https://doi.org/10.1594/PANGAEA.928497. All other data are available in the accompanying references.

10.1126/science.abd7096

1848 SOCIETY
AAAS



# An estate gift to AAAS

**Going all the way back to 1848,** our founding year, the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

Today, we are dedicated to advocating for science and scientific evidence to be fully and positively integrated into public policy and for the community to speak with one voice to advance science and engineering in the United States and around the world.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you will become a member of our 1848 Society and will help fuel our work on behalf of science and society – including publishing the world's most promising, innovative research in the *Science* family of journals and engaging in the issues that matter locally, nationally and around the world.

*"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."*

**—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member**

---

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

**Legal Name:** American Association for the Advancement of Science
**Federal Tax ID Number:** 53-0196568
**Address:** 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

cut here

## Yes, I would like more information about joining the AAAS 1848 Society.

1848 SOCIETY
AAAS

**PLEASE CONTACT ME AT:**

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

**RETURN THIS FORM TO:**
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

REVIEW SUMMARY

MEMBRANES

# Subatomic species transport through atomically thin membranes: Present and future applications

Piran R. Kidambi*, Pavan Chaturvedi, Nicole K. Moehring

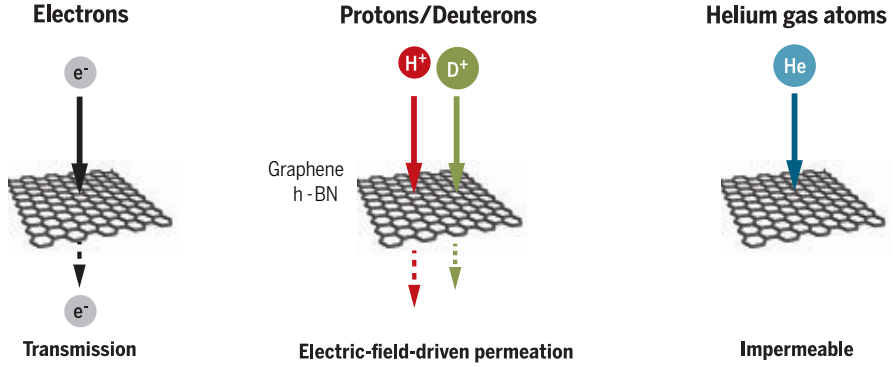
**BACKGROUND:** Membranes are thin physical barriers that allow for transport of certain species while limiting others. Atomically thin two-dimensional (2D) materials are crystalline solids with constituent atoms bonded in a planar 2D sheet and present opportunities to realize ultrathin membranes that allow for selective transport of subatomic species. The pristine lattice of monolayer graphene (honeycomb mesh of carbon atoms) and hexagonal boron nitride (h-BN, honeycomb mesh of alternating B and N atoms) is impermeable to small atoms such as helium (at room temperature) but allows for energy-dependent transmission of electrons and electric-field-driven transport of protons and deuterons.

Energy-dependent electron transmission through graphene, coupled with its atomic thinness, impermeability to gases, high strength, and electrical conductivity, provides opportunities to overcome material challenges limiting the advance of electron microscopy and spectroscopy to sample environments beyond high vacuum. Selective proton transport through h-BN, graphene, and other 2D materials with impermeability to even small atoms can help advance energy storage and conversion processes by providing avenues to address persistent problems of crossover and poor cation selectivity in conventional proton exchange membranes that typically result in long-term efficiency degradation. The differences in the rates of selective proton and deuteron transport through h-BN and graphene offer transformative opportunities for advancing isotope separations over current processes that are extremely energy intensive.

**ADVANCES:** The impermeability of graphene and h-BN to helium and other gases was initially demonstrated by using monolayer flakes to seal micrometer-sized gas-filled cavities, forming atomically thin balloons. However, subatomic species such as electrons tunnel through the lattice of monolayer graphene and h-BN and show promise for applications in vertical transistors and tunnel barriers for spintronic devices. The transmission of energetic electrons through graphene varies with their kinetic energy and offers transformative opportunities for advancing electron microscopy (EM). Graphene's high electron transparency enables its use as an ultrathin, crystalline, conductive sample substrate for transmission electron microscopy (up to ~80 kV), minimizing background noise (compared with ~3- to 20-nm-thick holey amorphous carbon films) and also allowing the formation of uniform ice thickness for cryogenic electron microscopy. The use of graphene as an electron transparent molecularly impermeable barrier to isolate detectors from the sample environment (while still allowing for transmission of electrons to and from the sample) allows for advancing conventional in situ metrology techniques requiring high-vacuum environments (EM and spectroscopy) to new frontiers such as ambient pressures, liquid systems, and biology.

Permeation of protons through h-BN and graphene with impermeability to atoms provides avenues to mitigate crossover of reactants and undesired species, and enhance selectivity of proton exchange membranes for improved efficiencies over a wide spectrum of energy generation and conversion processes, including (i) fuels cells for clean/green transportation and distributed or mobile auxiliary power generation, (ii) electrochemical hydrogen pumps for distributed hydrogen production and purification, and (iii) redox flow batteries for grid storage. The different rates of selective proton and deuteron permeation through h-BN and graphene (due to differences in vibrational zero-point energy) results in separation factors for $H^+/D^+$ of ~8 to 14 and offers transformational advances for isotope separations.

**OUTLOOK:** Selective transport of subatomic species through graphene and h-BN presents the potential for breakthrough advances in several fields. However, detailed insights into the mechanisms of transport are still emerging. Small-scale applications will be the first to emerge, e.g., transmission electron microscopy grids with graphene as the sample substrate are available commercially, and the use of graphene as an electron transparent barrier for imaging and spectroscopy is steadily rising. Scalable synthesis of high-quality 2D materials and facile cost-effective processes to integrate them into devices are imperative to enable large-area applications. However, understanding the influence of defects inevitably introduced during scalable synthesis and the resulting differences in transport characteristics compared to pristine materials will be essential. Studies are also needed to assess long-term 2D material durability under realistic application conditions. With the potential for substantial reduction in energy consumption compared to existing technologies, $H^+/D^+$ isotope separations are most likely to be explored in the nuclear industry. Other energy-related applications are expected to become viable, with scaled-up 2D material production offering economies of scale and by considering energy savings over the application life cycle. ■




**Subatomic species transport through atomically thin membranes.** (Left) Electrons tunnel through the lattice of monolayer graphene and h-BN. The transmission of energetic electrons through graphene depends on their kinetic energy. (Middle) Transport of thermal protons and deuterons through the graphene and h-BN lattice occurs via pores in the electron cloud. (Right) The graphene and h-BN lattice is impermeable to helium atoms and other gases.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: piran.kidambi@vanderbilt.edu
Cite this article as P. R. Kidambi *et al.*, *Science* 374, eabd7687 (2021). DOI: 10.1126/science.abd7687

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abd7687

REVIEW

MEMBRANES

# Subatomic species transport through atomically thin membranes: Present and future applications

Piran R. Kidambi[1,2,3,4]*, Pavan Chaturvedi[1], Nicole K. Moehring[2,4]

**Atomically thin two-dimensional materials present opportunities for selective transport of subatomic species. The pristine lattice of monolayer graphene and hexagonal boron nitride, although impermeable to helium atoms, allows for transmission of electrons and permits transport of thermal protons and its isotopes. We discuss advances in selective subatomic species transport through atomically thin membranes and their potential for transformative advances in energy storage and conversion, isotope separations, in situ electron microscopy and spectroscopy, and future electronic applications. We outline technological challenges and opportunities for these applications and discuss early adoption in imaging and spectroscopy that are starting to become commercially available, as well as emerging applications in the nuclear industry and future application potential in grid storage, clean/green transportation, environmental remediation, and others.**

Atomically thin two-dimensional (2D) materials are crystalline solids with constituent atoms bonded in a planar 2D sheet and exhibit distinctly different properties compared to bulk materials (*1*). Graphene, a monolayer mesh of carbon atoms arranged in a hexagonal lattice (~3.4 Å thick), was initially isolated as a model 2D material and can be considered as a building block for other carbon materials such as graphite (stack of graphene sheets held together via van der Waals forces), carbon nanotubes (seamless cylinders of graphene), and buckyballs (seamless spheres of graphene) (*1*). Interest in the properties of 2D materials have since resulted in successful isolation of monolayers from other layered materials, e.g., monolayer hexagonal boron nitride (h-BN, isomorph of graphene with alternating B and N atoms), graphene oxide, 2D chalcogenides (about three atoms thick), 2D oxides, 2D mica (~10 Å thick), 2D metal-organic frameworks, 2D covalent-organic frameworks, and others (*2*–*6*), as well as combinations in lateral (in-plane bonding of different 2D materials) or vertical (stacking different 2D materials) heterostructures (*2*, *5*).

Membranes are typically thin physical barriers that allow for transport of certain species while hindering others, and their ratio is commonly defined as selectivity (*7*, *8*). Monolayer graphene and h-BN represent the thinnest possible physical barrier and allow for selective permeation of subatomic species (Fig. 1)—i.e., the pristine lattice of graphene and h-BN is impermeable to small atoms such as helium at room temperature (*9*) but allows for transport of protons (*10*) and deuterons (*11*) and shows energy-dependent transparency to electrons (*12*). Graphene exhibits excellent thermal conductivity [$(4.84 \pm 0.44) \times 10^3$ to $(5.30 \pm 0.48) \times 10^3$ W/mK] (*13*) and electron mobility (~3000 to 230,000 $cm^2 V^{-1} s^{-1}$) (*1*, *14*, *15*), whereas h-BN is an insulator (bandgap ~5.9 to 6 eV) (*16*) with high thermal conductivity (~751 W/mK at room temperature) (*17*). The exceptional mechanical strength of monolayer graphene (Young's modulus ~1 TPa, breaking strength ~42 N $m^{-1}$) (*18*) and h-BN (Young's modulus ~0.86 TPa, fracture strength of ~70 GPa) (*19*) when suspended over micrometer-scale apertures, coupled with the high adhesion energy (graphene ~0.45 $Jm^{-2}$) (*20*), enables the fabrication of functional atomically thin membranes. Monolayer graphene in particular can withstand ~100 bar of applied pressure (*21*) and several orders of magnitude pressure differential (*22*). The pressure tolerance, however, depends on the area over which graphene is suspended (*21*), and calculations indicate that aperture diameters <1 μm are ideal for extreme pressures (up to ~570 bar) (*23*). This review will focus on selective transport of electrons, protons, and deuterons through atomically thin membranes and discuss the most promising present and future applications.

## Impermeability to atoms, gases, and molecules

Graphene and h-BN were initially shown to be impermeable to helium and other gases (Fig. 1C, inset) (*9*, *20*). Atomically thin balloons formed by sealing micrometer-sized gas-filled cavities with mechanically exfoliated monolayer graphene or h-BN flakes exhibited negligible leakage (transport of gas through the 2D lattice) rates, indicating that their lattice is impermeable to helium and larger atoms or molecules at room temperature (*9*, *24*, *25*). Monolayer molybdenum disulfide ($MoS_2$) is also impermeable to $H_2$ at ~50°C (1 bar and over 3 days) (*26*). By contrast, high gas leakage rates were observed for even nanoscale defects (*24*, *25*, *27*). Theoretical calculations were in agreement with experimental observations of gas impermeability and predicted large energy barriers ($E_B$) for the permeation of atoms through the hexagonal rings (radius ~1.42 to 1.45Å) in the pristine lattice of graphene (H atoms, $E_B$ ~2.86 to 4.61 eV; He atoms, $E_B$ ~3.5 to 18.77 eV; O atoms, $E_B$ ~5.5 eV; and N atoms, $E_B$ ~3.2 eV) and h-BN (H atoms, $E_B$ ~6.38 eV) (*9*, *28*–*31*).

The impermeability of graphene and h-BN to even small gas atoms at room temperature offers routes to creating atomically thin barriers or atomically sharp interfaces that separate two distinct reservoirs of atoms or molecules. However, observations of anomalous $H_2$ permeation through monolayer graphene, while maintaining impermeability to the smaller He atoms as well as the absence of molecular deuterium ($D_2$) permeation (*26*), taken together with the stochastic switching behavior of leakage rates observed in some atomically thin balloon experiments (*24*, *27*, *32*), raise fundamentally interesting questions on the absolute limits of gas impermeability, as well as the transport mechanisms (and associated $E_B$) for selective $H_2$ permeation (*26*).

By contrast, the transmission of energetic He ions or alpha particles and other larger ions through graphene, h-BN, and other 2D materials depends on the ion energy, size, and incident angle wherein collisions with the 2D lattice can also result in the formation of defects with varying yields (*33*, *34*). Irradiation with energetic ions and protons, as well as with alpha, beta, and gamma particles, is typically used to assess the reliability of electronic devices incorporating 2D materials for space applications (*35*, *36*). Finally, selective transport of atoms, molecules, and ions through defects in 2D materials presents the potential for advancing separation processes (*7*, *8*).

## Transmission of electrons

Electron transmission through 2D materials entails combinations of electron-electron and electron-phonon interactions in the lattice, via elastic and inelastic collisions with the nuclei of the atoms in the lattice (*37*). Electrons tunnel through the lattice of monolayer graphene (*38*–*40*) and h-BN (tunnel-barrier height ~3.07 eV, dielectric strength ~7.94 MV/cm) (*41*, *42*), and the in-plane electron conductivity of graphene can be used to modulate tunneling via changes to the potential or charge distribution, which the tunneling electrons encounter during transport in the normal direction

[1]Department of Chemical and Bimolecular Engineering, Vanderbilt University, Nashville, TN, USA. [2]Vanderbilt Institute of Nanoscale Sciences and Engineering, Vanderbilt University, Nashville, TN, USA. [3]Department of Mechanical Engineering, Vanderbilt University, Nashville, TN, USA. [4]Interdisciplinary Graduate Program in Material Science, Vanderbilt University, Nashville, TN, USA.
*Corresponding author. Email: piran.kidambi@vanderbilt.edu

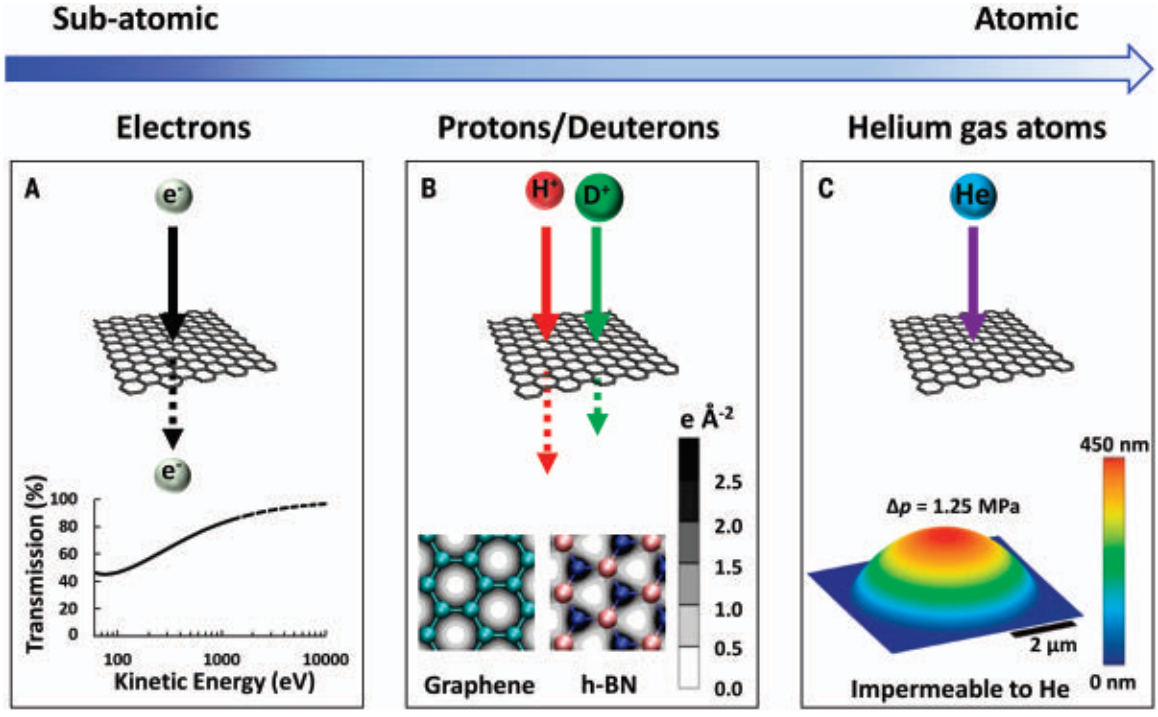



**Fig. 1. Transport through the atomically thin lattice of graphene and h-BN.** (**A**) Electrons tunnel through the lattice of monolayer graphene and h-BN. The transmission of energetic electrons, however, depends on their kinetic energy. Solid line up to ~1600 eV indicates regions where measured values are available in the literature. [Redrawn with data from (*12*, *49*)] (**B**) Transport of electric-field–driven thermal protons and deuterons through the graphene and h-BN lattice occurs via pores in the electron cloud. Inset shows integrated charge density (electrons/Å$^2$) for graphene and h-BN. [Adapted with permission from (*10*)] (**C**) The graphene lattice is impermeable to helium atoms and other gases. Inset shows a 3D rendering of an atomic force microscopy image of the graphene sealing gas molecules under a pressure difference of 1.25 MPa. [Adapted with permission from (*20*)]

(*39*). In this context, emerging electronics applications such as the use of graphene and h-BN as tunnel barriers for spintronic devices, (*43*) as well as vertical transistors exploiting electron tunneling (*39*, *42*, *44*), show promise.

The transmission of energic electrons through graphene varies with their kinetic energy (Fig. 1A). For low-energy electrons (~40 to 200 eV), simulations predict transmission >80% through monolayer graphene (*45*) that appear to be in agreement with experimentally measured transmission coefficients [ratio of transmitted electron current ($I_0$) to incident electron current($I$)] of ~0.60 to 0.74 for electrons with ~2- to 205-eV kinetic energy (*46*–*48*). For electrons with kinetic energy of ~200 to 1600 eV, the measured transmission coefficients for monolayer graphene increases with electron kinetic energy (Fig. 1A) (*12*, *49*). The electron attenuation lengths [$\lambda_{EAL} = d_G/\ln(I_0/I)\cos\theta$] computed from the measured transmission coefficients $I_0/I$ (assuming graphene thickness $d_G$ ~3.35 Å and incident angle $\theta = 0$), show reasonable agreement with values computed using the inelastic mean free path predictive formula (TPP-2M model) for graphite (*50*), and deviations at lower energy were ascribed to elastic scattering (*12*). These theoretical and experimental observations suggest that the inelastic mean free path of electrons provides a reasonable measure of the electron transparency of graphene, indicating its potential as an electron transparent gas-impermeable barrier for electron microscopy and spectroscopy (*12*, *51*). Electron transmission decreases with increasing number of graphene layers (*12*), and therefore bulk graphite is expected to show substantially lower electron transmission.

Graphene and h-BN show reasonable tolerance to irradiation-induced damage from electron beams with energies of ~60 to 80 keV, even for high beam doses under vacuum (*37*, *52*). At electron energies >80 keV, knock-on damage with ejection of atoms from the lattice is observed for monolayer graphene with formation of structural and vacancy defects, as well as defect clusters in the lattice upon prolonged irradiation (*52*). Graphene synthesized using $C^{13}$ isotope shows a slightly higher threshold of >95 keV for knock-on damage. (*52*) For monolayer h-BN, irradiation with an electron beam ~80 keV results in preferential ejection of B atoms, and the metastable nitrogen-terminated zig-zag edges preserve a triangular vacancy defect shape (*53*, *54*). However, the electron irradiation tolerance at higher pressures for graphene (*55*) and h-BN (*56*) depends on the gas composition of the environment with the possibility of beam-induced reactive chemical degradation and etching (*57*).

### Permeation of protons and deuterons

The impermeability to helium atoms and electron transparency of the pristine graphene and h-BN lattice raised fundamentally important scientific questions regarding the transport of protons. Protons represent an interesting intermediate case (*10*), and theoretical and experimental understanding of transport is still emerging.

Initially, permeation of protons through the pristine graphene lattice was assessed to be improbable on the basis of the calculated energy barriers ($E_B$) of ~1.41 to 2.21 eV, and transport was hypothesized to only occur in the presence of lattice defects, which reduce the barrier height (*29*). However, areal conductivity of ~3 mS cm$^{-2}$ for graphene ($E_B$ ~0.78 ± 0.03 eV) and ~100 mS cm$^{-2}$ for h-BN ($E_B$ ~0.3 ± 0.02 eV) were measured at room temperature during electric-field–driven transport of thermal protons through micrometer-scale mechanically exfoliated flakes sandwiched between Nafion, an ionomer that conducts protons in the hydrated state (Fig. 2, A and B) (*10*). Because Nafion shows negligible electron conductivity, the current obtained was a measure of proton transport through the graphene and h-BN lattice (*10*). Nafion sandwich devices with monolayer $MoS_2$ did not show measurable proton current under similar

experimental conditions (*10*). The measured increase in proton conductivity with temperature (Arrhenius dependence) indicated that values >1 S $cm^{-2}$, required for practical energy relevant applications, could be attained at relatively mild temperatures for h-BN (>80°C) and graphene (>110°C) (*10*).

The large difference in proton conductivity for graphene and h-BN, despite only ~1.8% lattice mismatch (*40*), is attributed to the polar nature of bonds in h-BN, which results in valence electrons concentrating around the N atom, leading to larger openings or pores in the electron cloud for the h-BN lattice compared to graphene (Fig. 1B, inset) (*10*). The proton conductivity decreases with an increasing number of layers because of an increase in integrated electron density, i.e., bilayer h-BN (~5 mS $cm^{-2}$) and trilayer h-BN (~0.1 mS $cm^{-2}$), whereas bilayer graphene shows negligible transport (*10*). Differences between bilayer graphene and h-BN are attributed to different stacking order, i.e., the AA' stacking in h-BN aligns the hexagonal rings between different layers preserving the central pore, whereas the AB stacking in graphene positions a carbon atom in one layer in the center of the hexagonal ring in the next layer, effectively blocking the pores in the electron cloud (*10*). Bulk graphite and h-BN are hence expected to be impermeable to protons.

Proton conductivity values similar to Nafion sandwich devices were also measured for

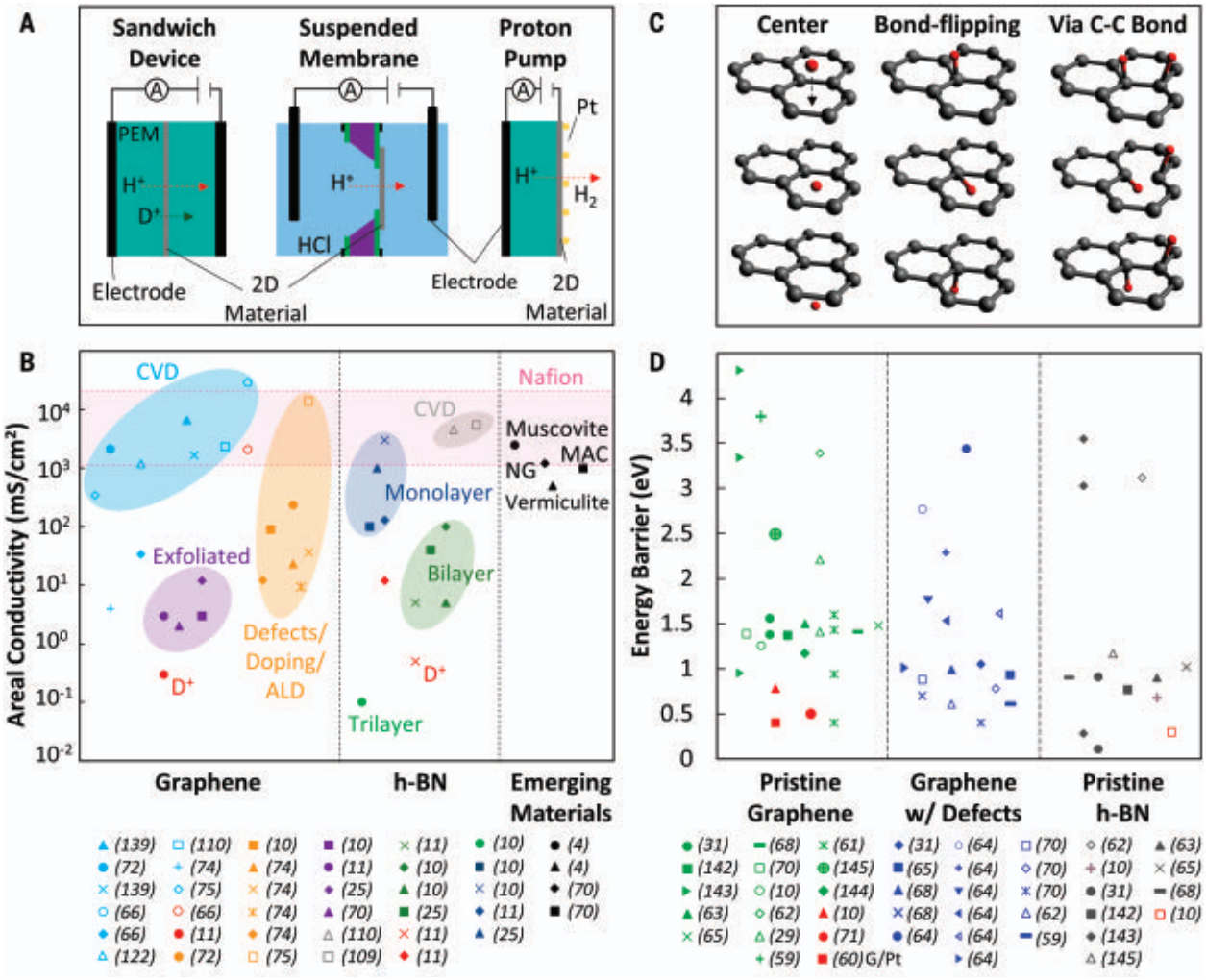



**Fig. 2. Experimental and theoretical aspects of proton permeation through 2D materials.** (**A**) Schematics of experimental device configurations used to measure proton transport (*10*, *60*). Nafion–2D material–Nafion sandwich devices (left), suspended 2D-material membranes separating liquid electrolytes (middle), Nafion 2D–material–Pt proton pump devices (right). (**B**) Experimentally measured areal proton conductivity values for graphene, h-BN, other emerging materials, and the industry standard ionomer Nafion reported in the literature. Red symbols represent deuteron areal conductivities. Most of the reported conductivity values in the literature are measured with 0.1 M HCl electrolyte or $H_2$ gas/methanol feed to the devices or fuel cells, except (*25*), (*75*) with 1 M HCl, and (*122*) with 0.5 M $H_2SO_4$ in vanadium flow batteries. Values for Nafion are extracted from (*66*, *139*–*141*). Different symbols with the same reference represent distinct material structure or defects. (**C**) Schematics of mechanisms of proton transport through graphene proposed by theoretical studies calculating energy barriers ($E_B$) (*29*, *61*–*65*). Proton transport via a straight perpendicular path through the center of the hexagonal ring in the lattice (left). Proton transport via chemisorption and subsequent bond-flipping through the hexagonal ring (middle). Proton transport via a cooperative mechanism involving multiprotonation (shown here for two protons), where neighboring chemisorbed protons facilitate subsequent protons to first chemisorb on carbon atoms in the hexagonal ring, followed by bond-flipping via the C-C bond (right). (**D**) Calculated and experimentally measured (red symbols) energy barriers ($E_B$) for proton transport through pristine graphene, graphene with defects, and pristine h-BN reported in the literature. Different symbols with the same reference represent distinct material structure, including extent of hydrogenation, defects, and others. Initial calculations predicted high $E_B$, but recent studies are exploring transport pathways and mechanisms that allow for lower $E_B$. References (*142*–*145*) are used in the plot and are not discussed in the paper.

micrometer-scale mechanically exfoliated graphene (~3 mS $cm^{-2}$) and h-BN (~100 mS $cm^{-2}$) separating liquid electrolytes (0.1 M HCl, Fig. 2A) (*10*, *25*). The areal conductivity showed a linear dependence on the electrolyte concentration, i.e., at 1 M HCl, ~1 S $cm^{-2}$ for monolayer h-BN and ~12 mS $cm^{-2}$ for monolayer graphene were measured (Fig. 2B), indicating that the concentration of protons interfacing the 2D lattice could influence transport rates (*25*). A Nernst analysis of membrane potentials showed that protons account for nearly all the observed currents with no detectable flow of counter ions ($Cl^-$), indicating near perfect proton selectivity (*25*). Although these studies implemented rigorous controls and measured negligible gas leakage rates through the membranes compared to noticeable gas fluxes for graphene synthesized via chemical vapor deposition (CVD) with intrinsic defects (*10*, *25*), some studies attribute proton transport through CVD graphene in the liquid phase to atomic-scale defects in the lattice (*58*, *59*). Further research is expected to improve fundamental understanding and shed light on the origins of these differences, e.g., the influence of bond rotations, single vacancies, etc. that could allow for enhanced proton transport while hindering He permeation.

The deposition of a discontinuous layer of Pt on the 2D lattice to form Nafion–2D material–Pt devices (where protons permeating the 2D lattice recombine on Pt to evolve $H_2$ gas, Fig. 2A) further increased proton conductivity of graphene to ~90 mS $cm^{-2}$ (reducing $E_B$ by ~0.5 eV), while ~3 S$cm^{-2}$ was measured as a lower bound for h-BN (Nafion resistance limited proton current) (*10*, *60*). Although the attraction of transient protons to Pt has been suggested to play a role (*10*, *60*), further research may provide mechanistic insights. Illuminating the Nafion-graphene-Pt devices with visible light (100 mW $cm^{-2}$) resulted in a further increase in proton conductivity to ~20 S $cm^{-2}$ (at ~2.8 V) and up to 10 times higher proton fluxes than for devices in the dark (*60*). The measured gain of ~$10^4$ protons per photon (photo-responsivity ~$10^4$ $AW^{-1}$) and response times in the microsecond range can possibly enable photodetector applications (*60*). Similar effects are also seen for Pd and Ni nanoparticles (albeit not as effective as Pt) that also strongly interact and n-dope graphene, and the observed effects are attributed to photovoltages created from hot electrons generated in graphene (upon illumination of in-plane electric fields and built-in junctions formed in areas surrounding the nanoparticle due to n doping) that funnel protons and electrons toward the metal nanoparticle, resulting in an enhanced rate of electron–proton conversion to atomic hydrogen on the nanoparticle (*60*).

Theoretical studies have proposed different mechanisms (via hydrogenation or protonation) and transport pathways (straight perpendicular path through the center of the hexagonal ring, $E_B$ ~1.41 eV or via chemisorption, $E_B$ ~2.21 eV) to explain the measured proton conductivity for graphene and h-BN (Fig. 2C) (*29*). Although the computational methods vary, most studies computed $E_B$ >1.4 eV for graphene, (*28*, *29*, *59*, *61*–*65*) and $E_B$ >0.9 eV for h-BN (*31*), and the inclusion of quantum effects such as tunneling and zero-point energy can further reduce $E_B$ by ~0.5 eV (*61*), bringing the theoretical values (Fig. 2D) closer to the experimentally measured values $E_B$ ~0.8 eV (graphene) and $E_B$ ~0.3 eV (h-BN) (*10*). Theoretical studies of proton transport in the presence of water molecules, i.e. proton transfer from $H_3O^+$ on one side to $H_2O$ on the other (comparable to experimental approaches interfacing 2D crystals with hydrated Nafion or aqueous electrolytes), suggest that hydrogenation of graphene reduces $E_B$ from >3 eV to <1 eV (*62*). $E_B$ ~1 eV for proton transport through multiprotonated graphene was also computed for a cooperative mechanism in which nearby chemisorbed protons facilitate chemisorption of subsequent protons onto a carbon atom in the hexagonal ring, followed by bond flipping via the C-C bond to allow for proton transfer to the other side (Fig. 2C) (*64*).

In addition to proton transport, transport of the heavier isotope deuteron was also measured through pristine monolayer flakes of graphene and h-BN using Nafion–2D material–Nafion devices (Fig. 2A) (*11*). The rate of transport for deuterons was an order of magnitude lower than for protons, allowing for a separation factor or selectivity ~10 ($10H^+$:$1D^+$) arising from the difference in vibrational zero-point energies (~60 meV) of transient protons and deuterons bound to the $SO_3^-$ groups in the Nafion before being incident on the graphene or h-BN lattice (*11*).

Although the measured areal proton conductivity for pristine graphene (~3 mS $cm^{-2}$) and h-BN (~100 mS $cm^{-2}$) at ambient temperature is lower than that for state-of-the-art industry-standard ionomers such as Nafion (thickness-dependent proton conductivity ranging from ~1 to 20 S $cm^{-2}$, Fig. 2B) (*10*, *66*), atomic-scale defects in the lattice, as well as domain boundaries in CVD graphene, have been shown to increase proton conductivity to ~29 S $cm^{-2}$ (after subtracting Nafion resistance) while maintaining negligible transport of $K^+$ in Nafion-graphene-Nafion devices (*66*). Despite the presence of defects in CVD graphene, a $H^+/D^+$ isotope selectivity ~14 ($14H^+$:$1D^+$) was maintained (*66*), indicating that the mere presence of defects alone did not change the dominant mode of rate-based $H^+/D^+$ isotope separation in these devices (*66*, *67*).

Theoretical calculations for $H^+$ and $D^+$ transport through topological Stone–Wales (SW) defects ( SW 5757) predict an energy barrier <1 eV and $H^+/D^+$ selectivity ~ 7 at ambient conditions (*68*). These calculations offer an alternative interpretation, with proton transport occurring primarily through defects in the 2D lattice rather than permeation through the pristine lattice. Such an interpretation would be consistent with proton conductivity along with negligible gas leakage rates measured for mechanically exfoliated graphene and h-BN (*10*, *25*) and very high proton conductivity along with high $H^+/D^+$ selectivity for CVD graphene with grain boundaries (*66*, *67*) containing pentagon-heptagon rings (*69*), as well as the high proton conductivity ~1 S $cm^{-2}$ for nanocrystalline graphene (NG) and monolayer amorphous carbon (MAC, with eight carbon atom rings) along with negligible gas leakage rates (*4*, *70*).

Practical applications will require large-area 2D materials synthesized via scalable approaches such as CVD that inevitably introduce intrinsic defects, grain boundaries, wrinkles, and other defects (Fig. 3), and the proton transport properties may be quite different than those of pristine 2D materials (Fig. 2B) (*58*, *59*, *66*, *67*, *71*–*74*). Hence, understanding the selective proton transport behavior of CVD-grown 2D materials is imperative for advancing applications. Notably, nanoscale defects introduced in CVD graphene via ion-beam bombardment (*74*), plasma treatments, and incorporation of dopants can increase ionic conductance (*74*, *75*), with pH dependence confirming protons as the main contributors (*74*). However, nonselective defects can also increase ionic conductivity with loss in proton selectivity. Finally, the emergence of new materials such as 2D mica (~10 Å in thickness and ~5-Å-wide tubular channels) that exhibit 1 to 2 orders of magnitude higher proton conductivity (~100 S $cm^{-2}$ at 500 °C) than graphene or h-BN have ignited research interest in ultrathin membranes that could operate under high temperatures, increasing the efficiency of fuel cells, as well as under dry or unhydrated conditions. (*4*) Other 2D materials, e.g., phosphorene ($E_B$ ~0.48) and silicene ($E_B$ ~0.12), have been explored theoretically (*31*) for enhanced proton transport, but their limited stability under ambient conditions presents challenges.

### Advances in synthesis and processing of atomically thin membranes

Mechanical exfoliation was initially used to isolate monolayers of 2D materials and suspend them over apertures to form atomically thin membranes (Fig. 3) (*1*, *9*, *20*). Although it produces the highest quality of pristine flakes ideally suited for probing fundamental material and transport properties, the flake sizes remain limited to a few micrometers (*9*, *10*, *24*–*27*). Practical membrane applications

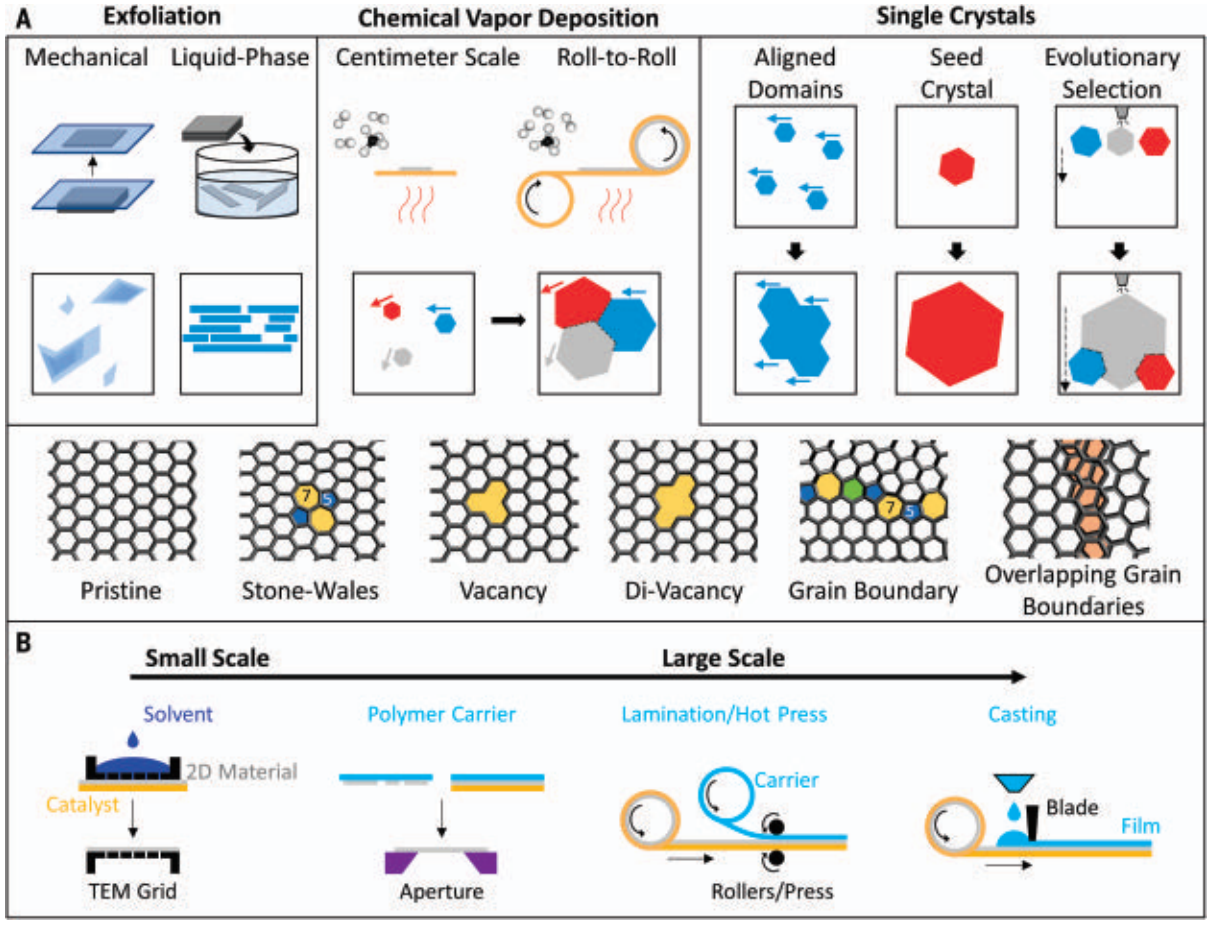



**Fig. 3. Synthesis and processing of large-area atomically thin 2D materials.** (**A**) Mechanical exfoliation, liquid phase exfoliation, and chemical vapor deposition (CVD) represent the main synthesis methods for 2D materials. Although mechanical exfoliation produces high-quality flakes, it is not scalable to produce continuous layers. Liquid phase exfoliation allows for scalable synthesis, but realizing atomically thin membranes from flakes produced is challenging. CVD (and its variants) allows for scalable continuous monolayer synthesis including roll-to-roll processes, but the films produced are typically polycrystalline with intrinsic defects such as Stone-Wales defects, mono- and multivacancy, and grain boundaries. Efforts to synthesize single-crystal 2D materials via CVD have focused on (i) alignment of domains on a single crystalline catalyst or substrate, (ii) growing a single nucleus larger, and (iii) evolutionary selection where controlled feeding of the fastest-growing domain results in it outgrowing others. (**B**) Solvent-assisted 2D material transfer and the use of polymer carrier layers are some of the most commonly used approaches for fabrication of small-area (few micrometers to centimeter scale) membranes for applications, e.g., TEM grids. Large-area (centimeter to meter scale) energy and separation applications, e.g., fuel cell isotope separations, respectively, require the development of scalable approaches such as lamination/hot press and polymer casting.

will require scalability over much larger areas and will need higher levels of homogeneity in terms of layer numbers, film coverage, and quality (*8*). Although liquid-phase exfoliation of 2D materials allows for scalability, fabricating atomically thin membranes from a randomly oriented collage of flakes of multiple sizes, shapes, and thicknesses is inherently challenging (*8*, *76*). Bottom-up synthesis via CVD (using a catalytic substrate to dissociate a precursor vapor followed by self-assembly via nucleation and subsequent growth at elevated temperatures) and its variants allow for continuous monolayer films of graphene (*77*), h-BN (*78*), nanocrystalline graphene (*79*), monolayer amorphous carbon (*80*), and other 2D materials (*81*), where the scalability is in principle only limited by the catalyst or substrate surface area and reactor design (*82*), e.g., synthesis of ~10 cm by 10 cm monolayer h-BN (*83*) and ~100 m of monolayer graphene via roll-to-roll processes (*84*) has been demonstrated. However, CVD-grown monolayers are typically polycrystalline with domain boundaries (*69*) as well as intrinsic defects within the domains that could allow for nonselective leakage (Fig. 3) (*85*).

Efforts toward improving the quality of CVD-grown 2D materials were largely driven by requirements for electronic applications, i.e., minimizing electron scattering at grain boundaries to achieve performance comparable to that of mechanically exfoliated flakes (*86*). Hence, increasing domain sizes within the polycrystalline film has been the primary focus, with the ultimate goal of producing single crystalline 2D materials without grain boundaries via three main approaches (Fig. 3). In the first approach, a single crystalline substrate is used to nucleate the domains of a 2D material that align with respect to each other and orient in the most energetically favorable crystallographic direction with respect to the substrate (*83*, *87*, *88*), or electrostatic interaction between domains is used to achieve alignment on a liquid metal (*89*). Such alignment results in seamless merging of the individual domains into a continuous film without domain boundaries (*83*, *87*–*89*). In practice, single-crystal substrates are expensive to produce, and ensuring perfect alignment of all domains with respect to each other is nontrivial because a small fraction of unaligned domains can result in polycrystalline films (*87*). In the

second approach, a single nucleus of the 2D material is grown larger in size without the nucleation of additional domains by carefully controlling the supply of precursor (*90*), but the processing is typically long, i.e., a few hours compared to processes that require a few minutes or seconds (*91*). The third approach involves evolutionary selection in which multiple domains nucleate but careful control of the gas supply through a nozzle on a moving substate results in one nuclei outgrowing the others (*92*). This method allows for the use of inexpensive polycrystalline catalyst foils to produce single crystalline monolayer 2D materials via kinetic control (*92*) and is adaptable to roll-to-roll CVD, offering process scalability (*82*, *84*).

However, single crystalline 2D materials are still not completely devoid of defects including vacancy defects (*8*, *93*), and the quality requirement for membrane applications tends to be somewhat different than electronic applications (*7*, *8*, *94*). For example, intrinsic defects (e.g., SW 5757) and grain boundaries can allow for enhanced proton transport (*66*, *68*), but larger defect sizes could result in leakage of gases, ions and larger molecules, compromising selectivity (*7*, *8*, *24*, *25*, *27*). Additionally, nanoscale defects show a propensity to cluster along wrinkles originating from differences in thermal expansion coefficients between the 2D material and catalyst, and the detrimental effects of leakage (loss in selectivity) through even a very small number of nanoscale defects is greatly exacerbated for atomically thin membranes applications compared to most electronic applications (*94*). In this context, recent synthesis of nanocrystalline graphene (*79*) and monolayer amorphous carbon (*80*) shows promise, but attaining atomic-scale control and complete absence of larger defects over large areas remains nontrivial, necessitating leakage sealing approaches (*95*, *96*), and large-area synthesis of 2D mica remains to be demonstrated (*4*).

In addition to synthesis, the interfacing of the 2D material by transferring it from the growth substrate to an appropriate support is crucial to enable membrane applications (Fig. 3B) (*7*, *8*). Transfer procedures developed for 2D materials device fabrication, e.g., sacrificial polymer scaffolds (*9*, *24*–*27*, *83*, *87*, *88*) or the use of an evaporating solvent to adhere 2D materials to transmission electron microscopy (TEM) grids (*97*), can aid small-area membrane (few micrometers to centimeter scale) fabrication, whereas scalable approaches such as hot-pressing (*66*), polymer support casting (*98*), and roll-to-roll lamination (*99*), along with approaches for effective reuse of the catalyst (*100*), may enable scalable, cost-effective synthesis of atomically thin membranes for large-area (centimeter to meter scale) applications (*7*, *8*). Leveraging synergistic opportunities such as roll-to-roll processes for 2D material synthesis, as well as membrane fabrication, may provide rapid advances in this area (*8*, *82*, *99*).

Commercial production of large-area 2D materials is slowly maturing. Although polycrystalline centimeter scale CVD graphene has been available commercially for almost a decade, innovations in processing and economies of scale have allowed for the cost of bulk orders of CVD graphene on Cu foil to steadily drop from ~ €1000/cm$^2$ in 2010 to ~ €2/cm$^2$, indicating that applications using small areas of graphene can already be cost competitive while delivering improved functionality (*101*, *102*). On the basis of current trends, a further reduction in price for graphene may be expected, and translating lessons from developments with graphene can perhaps aid the compressed time scale for commercial production of other 2D materials.

## Applications of selective proton transport through atomically thin membranes

Selective proton transport through atomically thin membranes, along with impermeability to atoms, gases, and hydrated ions, presents opportunities to improve efficiencies across a wide spectrum of energy generation and conversion processes in the hydrogen economy, as well as enabling additional separation processes (Fig. 4).

### *Fuel cells and hydrogen pumps*

Fuel cells powered by renewably generated hydrogen or methanol and ethanol are expected to play an important role in environmentally sustainable advances toward clean/green transportation, distributed, and mobile auxiliary power generation (Fig. 4) (*103*). Nafion and sulfonated polyether ether ketone (S-PEEK) currently represent the most widely used fuel-cell membranes but suffer from crossover of reactants (leakage of undesired species through the membrane, thereby reducing selectivity), along with swelling and softening at high relative humidity, and require hydrated environments for proton conductance (*104*, *105*). Hydration requirements also limit the maximum operating temperatures, hindering efficiency gains (*103*–*107*).

Among the many approaches to improving the operability of proton-conducting polymers (*108*), coating continuous layers of graphene or h-BN has been shown to limit cross-over of hydrogen (*109*) or methanol in fuel cells (*110*, *111*). In hydrogen fuel cells, the effect was more pronounced in accelerated stress tests for ~100 hours, i.e., $H_2$ permeation currents (measured at 0.4 V, 30% relative humidity, and 90°C) for one or three layers (AA' stacked) of h-BN-coated-Nafion showed almost no change compared to a >100× increase for bare Nafion, indicating long-term benefits (*109*). In the case of methanol fuel cells, a decrease in crossover of up to ~68% is observed for monolayer graphene-coated Nafion compared with bare Nafion, allowing higher concentrations of methanol (up to ~10 M compared with typical concentrations of 1 to 5 M) to obtain enhanced power density (*111*).

The reduction in crossover can, however, come at the expense of low proton conductivity of 2D materials that can decrease the overall efficiency of the fuel cell by increasing the ohmic losses (*109*–*111*). Higher operating temperatures (*110*, *111*) or the incorporation of selective defects in the 2D lattice can enhance proton transport (*66*, *74*). However, achieving atomic precision in defect sizes over large-area membranes using scalable processes remains nontrivial, and leakage-sealing techniques (*7*, *8*, *96*), as well as appropriate support selection (*112*), could play a crucial role in increasing functionality by minimizing nonselective leakage. h-BN and graphene membranes supported on porous polymers or even ceramics can allow for membranes with high proton selectivity that can operate in anhydrous environments and medium temperatures of 110° to 160°C [h-BN is stable up to ~500° to 700°C in air (*113*)], aiding increased fuel-cell efficiencies and advancing future fuel-cell designs. (*103*, *104*, *106*, *107*).

Although Li-ion battery (specific energy >250 Wh kg$^{-1}$) vehicles currently have a higher market penetration for passenger vehicles and will continue to in the near future owing to established electricity infrastructure (despite relatively long charging times), the extra weight required to enhance the range of fuel-cell vehicles is negligible compared with the drastic weight compounding for battery-powered vehicles (*106*, *107*, *114*). These differences are exacerbated in the context of clean/green cargo and commercial transport, both on land (via heavy vehicles, trucks, and buses) and on shipping routes (*106*, *107*, *114*). Hydrogen fuel-cell–powered vehicles have also been under development for underwater transport systems and military applications, where the longer ranges enabled by the high energy density of $H_2$ (~120 to 142 MJ/kg), as well as low thermal and acoustic signatures, allow for extended periods of quiet submergence and undetected transport (*107*). Finally, the use of hydrogen fuel cells for aviation is also under consideration, particularly for unmanned aerial vehicles (*107*). The already high cost of Nafion membranes ($75 to 250/kg ionomer cost for high-volume manufacturing) (*115*) accounts for ~9 to 17% of the fuel-cell cost (*107*), and the addition of 2D materials is likely to increase costs further. However, advantages in increased efficiency over the targeted membrane life cycle of operation, i.e., ~40,000 hours (stationary fuel cells) and ~5000 hours (transportation fuel cells) (*116*), need to be considered during

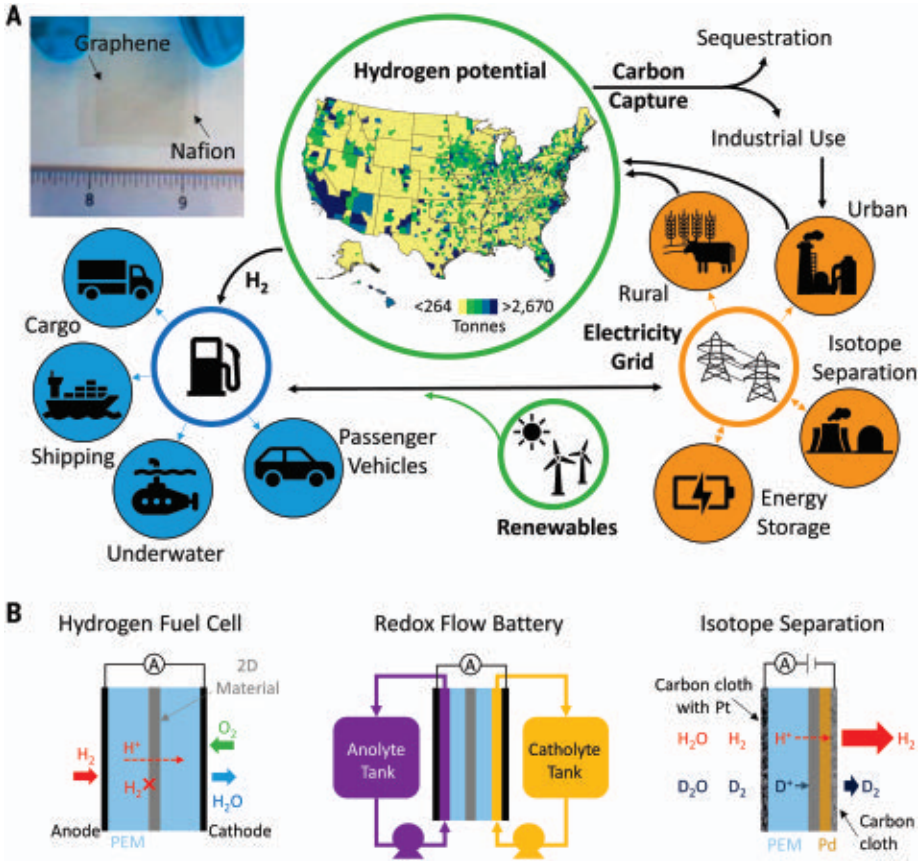



**Fig. 4. Applications of proton transport through atomically thin membranes.** (**A**) Atomically thin membranes present opportunities to improve efficiency across a wide spectrum of energy generation and conversion processes. Coating a layer of graphene or h-BN onto conventional proton-conducting polymers [adapted with permission from (*67*)] can allow for reduced crossover and increased efficiency of fuel cells for transportation and auxiliary power generation, as well as redox flow batteries for grid storage. Electrochemical hydrogen/proton pumps incorporating 2D materials can allow for compact separators for hydrogen purification and pumping, aiding distributed hydrogen production from rural and urban waste streams. US map showing the potential for hydrogen production from biogas generated from waste streams adapted from (*117*). Atomically thin membranes can allow for hydrogen isotope ($H^+/D^+$) separation (*67*) and environmental remediation efforts. (**B**) Schematic showing the integration of atomically thin membranes in hydrogen fuel cells to mitigate crossover of hydrogen as well as reactants in a redox flow battery while allowing for selective $H^+$ transport. In hydrogen fuel cells, $H_2$ and $O_2$ are supplied at opposite electrodes. The $H_2$ dissociates to form $H^+$ that is transported through the membrane and form $H_2O$ as a by-product upon combining with dissociated $O_2$, and the electrons move via the external circuit. In a redox flow battery, redox ions change oxidation state during charge and discharge cycles via the exchange of $H^+$ through a membrane and electrons through the external circuit. The use of monolayer graphene and h-BN with Nafion in electrochemical proton pumps can enable new approaches for $H^+/D^+$ separation with separation factor or selectivity ≥8.

feasibility analysis. Studies are also needed to evaluate long-term durability and chemical stability of 2D materials under real-world fuel-cell operating conditions.

The lack of hydrogen infrastructure perhaps presents the most severe challenge to the hydrogen economy (*103*, *106*, *107*, *114*). Leveraging the potential of agricultural waste in rural areas, and waste water and industrial waste in urban areas, to generate biogas or methane (and $H_2$ via reforming; predicted total capacity of US ~2.8 million tonnes and enough to power ~11 million fuel-cell vehicles per year), as well as the use of existing natural gas infrastructure, are some viable approaches for possible distributed $H_2$ production (Fig. 4A) (*117*). Separating $H_2$ from reformate mixtures can allow for carbon capture (which can further be converted into economically valuable organic molecules such as methanol, formic acid, olefins, etc.) or carbon storage for negative carbon emissions technologies, particularly for $H_2$ sourced from biogas generated from waste streams (*117*). Here, compact separators using electrochemical hydrogen pumps can aid distributed $H_2$ production (*107*). Electrochemical hydrogen pumps incorporating proton-selective atomically thin membranes can allow for facile, compact, single-step $H_2$ separation and purification (>99.97% purity needed for fuel cells), alleviating the need for multistage conventional separation processes (*116*). In addition to separation from reformate mixtures, such proton pumps can also be used to pump against a pressure gradient to allow for a single-step purification and pressurization of $H_2$ and facilitate dispensing for fuel cells.

### *Redox flow batteries and grid storage*

Grid-scale energy storage is emerging as a critical requirement for future electricity grids powered by periodic and/or intermittent renewable energy such as solar and wind (Fig. 4). Redox flow batteries, where ions in electrolyte solutions undergo a change in the redox states via the exchange of protons through a cation-conducting polymer separating two reservoirs, are emerging as promising technology platforms (*118*). Vanadium redox flow batteries, in particular, offer distinct advantages with scalability in storage capacity from stand-alone units of a few kilowatts to 200 MW (under construction) (*119*), thousands of deep-discharge cycles (>15,000), and >20 years of life-time; they are also nonflammable and allow for independent control of power and energy storage (*118*, *120*, *121*). Nafion, S-PEEK, and other proton-conducting polymers are the current industry standard for vanadium flow batteries but do not offer high selectivity between protons and vanadium ions and thereby suffer from crossover of vanadium ions as well as other undesired redox species, leading to long-term efficiency losses (*118*, *120*, *121*). Sandwiching monolayer graphene or h-BN between layers of Nafion or S-PEEK has been shown to increase proton selectivity (*120*–*122*). In particular for CVD graphene sandwiched between Nafion, proton transport ($0.02 \pm 0.005$ ohm cm$^2$) four orders of magnitude faster than vanadium ion transport ($223 \pm 4$ ohm cm$^2$) and near complete elimination of crossover were observed under lab-scale test conditions (*121*). The chemical robustness of graphene or h-BN to highly acidic electrolyte environments, coupled with selective proton transport, presents the potential for advancing new kinds of proton exchange membranes by using inert polymers, e.g., Teflon, to support the 2D material, although long-term studies under realistic operating conditions are required to evaluate durability and stability.

### *Isotope separation and nuclear technologies*

Isotopes of hydrogen (hydrogen H, deuterium D, and tritium T) are widely used for nuclear, military, medical, and research applications (*67*). Conventional $H^+/D^+$ separation processes for heavy-water production ($D_2O$, used as a moderator for neutrons as well as a coolant in nuclear reactors) are extremely energy intensive, requiring multistage separations owing to poor selectivity, e.g., Girdler-sulfide process ($H^+/D^+$ ~1.3) and monothermal-$NH_3/H_2$ process ($H^+/D^+$ ~6) (*67*, *123*, *124*). The naturally occurring low concentration of D ~0.015% in water, coupled with the low selectivity of conventional processes, necessitates processing of enormous quantities to produce ~1 kg $D_2O$ with an energy consumption of ~10 MWh, thereby increasing capital and operating costs (*67*, *123*, *124*).

Here, selective $H^+$ and $D^+$ transport through graphene and h-BN with rate-based separations factors $H^+/D^+$ ≥10 offers transformational advances (*11*, *66*). Large-area electrochemical proton pumps with monolayer CVD graphene show a $H^+/D^+$ selectivity ~8 despite the presence of cracks and defects, as well as ~1 to 2 orders of magnitude reduction in energy consumption compared with current technologies (*67*). Extrapolating the measured flux ~0.8 mmol $hour^{-1}$ $cm^{-2}$ at 0.5 V in the electrochemical pumps indicates that a graphene membrane ~30 $m^2$ would produce ~40 tons of heavy water per year, comparable to that of a modern plant (*67*), and h-BN is expected to offer even higher performance. Similar advantages for removal of radioactive tritium (T) from contaminated water in commercial and research nuclear reactors are predicted with higher $H^+/T^+$ selectivity ~37 (*11*, *65*–*67*). Finally, such technologies are expected to be leveraged to supply experimental fusion reactors that require T as a fuel (*125*), as well as separating mixtures of T-$^3$He to recover $^3$He for applications in radiation monitors for border security at ports of entry to detect illicit transport of radiological or nuclear materials (*126*).

### Applications for the electron transparency of atomically thin membranes

The energy-dependent electron transparency of graphene, along with its atomic thinness, electrical conductivity, excellent mechanical strength, and impermeability to gases, offers transformative opportunities for advancing electron microscopy and spectroscopy (*49*, *57*, *97*, *127*–*129*), as well as extending conventional in situ metrology techniques requiring high-vacuum environments to new frontiers (Fig. 5) such as ambient pressures (*6*, *49*), liquid systems (*12*, *129*, *130*), and biology (*127*, *131*).

### Atomically thin substrates for TEM

Graphene represents the ideal substrate material to support samples for TEM imaging and offers distinct advantages over the typically used ~3- to 20-nm-thick holey amorphous carbon films (*57*). The electron transparency of atomically thin graphene (~0.34 nm) at acceleration voltages typically used in TEM minimizes background noise, and contributions at high resolution can be effectively filtered out using the periodicity of the crystalline lattice (*37*, *57*, *132*). The high mechanical strength and chemical inertness of graphene (compared to amorphous carbon), along with its ability to maintain integrity under ~80-keV electron beams (even at high beam dose) and effective charge dissipation, have enabled advancements in imaging of individual low–atomic-number atoms and adsorbates, 1 D materials such as carbon nanotubes and nanowires, 2D materials and heterostructures,

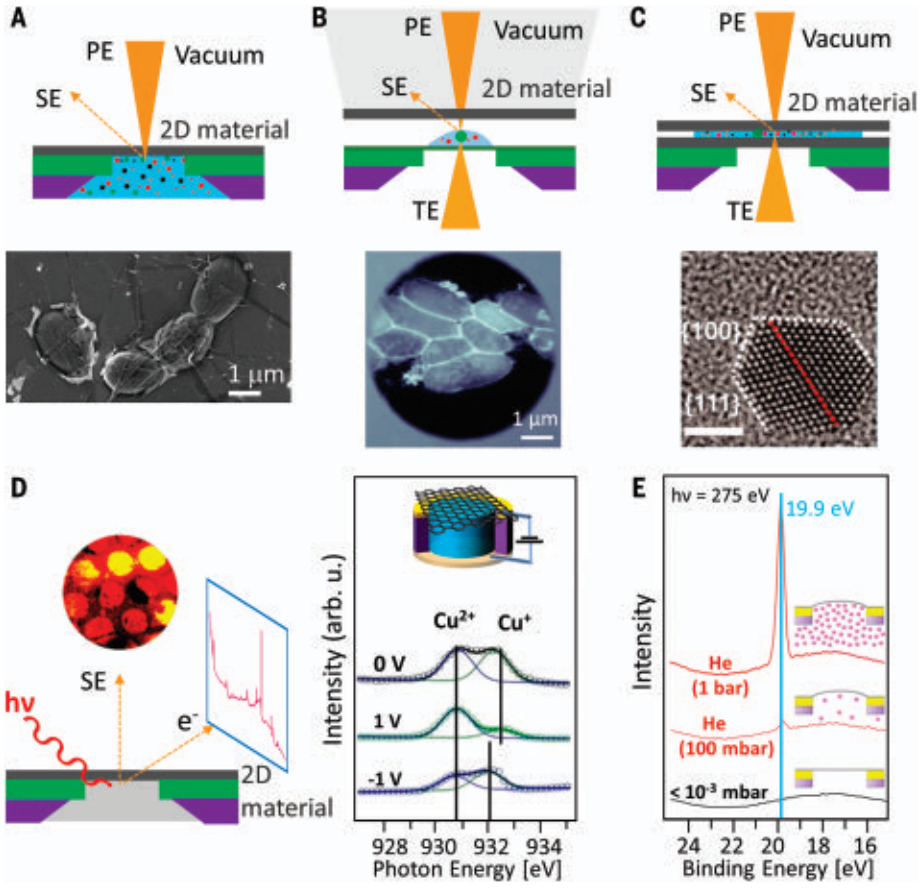



**Fig. 5. Applications for the electron transparency of atomically thin membranes.** (**A**) Schematic of graphene covering a wet sample and isolating it from the vacuum environment for scanning electron microscopy (SEM). SEM image of *Lactococcus lactis* bacteria using graphene as a veil or covering. [Adapted with permission from (*138*); scale bar superimposed from the figure provided in the main text.] (**B**) Schematic of graphene membrane sealing the optics in a SEM, eliminating the need for a specimen vacuum chamber (airSEM). Dark-field scanning transmission electron microscope (STEM) image of uranium-stained *Escherichia* coli bacteria. [Adapted with permission from (*131*)] (**C**) Schematic of encapsulated samples between two graphene layers. TEM image of Pt nanoparticle in liquid solution (acquired at 64.22 s during growth in solution; scale bar, 2 nm). [Adapted with permission from (*97*)] (**D**) Schematic of graphene sealing a gaseous or wet sample for SEM, as well as photoelectron spectroscopy and microscopy. Inset shows photoemission electron microscopy (PEEM) image of a graphene capped multichannel array at the O K-edge energy. X-ray absorption spectroscopy (averaged Cu $L_3$-edge spectra and their Voigt fits) collected on graphene-capped 0.1 M $CuSO_4$ solution showing changes in concentration of the mono- and bivalent copper ions just below the graphene membrane as a function of its potential. Inset shows a single channel (from the multichannel array used for PEEM) for clarity. [Adapted with permission from (*137*)] (**E**) X-ray photoelectron spectra of He (He 1s spectra) collected through a single-layer graphene membrane capping a reaction cell filled with He gas at different pressures. [Adapted with permission from (*22*)]. PE, SE, and TE stand for primary, secondary, and transmitted electrons, respectively.

organic compounds and molecules, biological materials and biomolecules, nanoparticles, and clusters supported on graphene, including their dynamic interactions with graphene and among themselves (*57*, *132*, *133*). The uniform thickness of monolayer graphene and opportunities for chemical modification of its surface (for enhanced interaction with biological specimens) present advantages for sample preparation for cryogenic electron microscopy of biological samples by enabling the formation of uniform ice thicknesses along with minimal signal from graphene used as a sample support (*134*, *135*). The use of graphene for TEM grid supports [aided by the small areas required (a few square millimeters) (*134*) and high transfer yields (*135*)] was one of the first application to progress successfully into the commercial arena (*102*), with several companies, e.g., Ted Pella Inc., ACS materials, etc., offering graphene-coated TEM grids in their product line. However, electron beam–induced knock-on damage >80 keV and limited chemical and thermal stability compared with silicon nitride ($SiN_x$) present some limitations, but these may be mitigated to some extent by advances in aberration correctors that allow for high-resolution imaging <80 keV (*37*, *57*).

### Electron transparent windows for in situ electron microscopy and spectroscopy

The use of graphene as an atomically thin electron transparent barrier [replacing the typically used ~15- to 50-nm-thick $SiN_x$ membranes with high atomic number Z and electron scattering (*130*, *131*)] to isolate the sample environment from the vacuum environment of analyzers and detectors (Fig. 5) presents transformative opportunities for advancing in situ electron microscopy and spectroscopy (*12*, *49*, *129*, *136*). The simplest configuration involves covering the sample with a layer or two (where defects in the first layer are sealed by the second layer) of graphene (Fig. 5, A, D, and E). Because the characteristic mean free path of the generated secondary or photoelectrons is typically larger than the graphene thickness, such an approach allows for probing the top few nanometers of the sample surface via photoelectron spectroscopy (PES) and microscopy (*12*, *49*, *137*) and scanning electron microscopy (SEM) (*136*, *138*). Such approaches have enabled PES of wet and gaseous samples (Fig. 5D) (*12*, *136*, *137*), up to ~1- to 1.5-bar pressure (Fig. 5E) (*22*), as well as SEM imaging of live bacterial cells without labels (Fig. 5A) (*138*). In the latter, graphene's impermeability protected the cells from the vacuum environment, and the scavenging of radicals (generated due to radiolysis of water by the e-beam) by graphene along with its electrical conductivity allowed for a 100-fold increase in electron dose, after which the cells still appear to maintain structure and function (*138*). An alternative approach in airSEM (*131*) seals the electron optics in an SEM with graphene and uses the mean free path of electrons in air ~10 to 100 μm to enable high-contrast imaging and improved spatial resolution (at 7-kV acceleration voltage, simulated electron scattering from bilayer graphene is ~3% compared with ~75% for 10-nm $SiN_x$ membrane) under ambient conditions, eliminating the need for a specimen vacuum chamber (Fig. 5B) (*131*). A third configuration isolates the sample (usually a liquid solution or biological samples) between two graphene layers and allows for atomic-resolution TEM of nanoparticles in liquids (and their growth and dynamics) (*97*, *129*) as well as nanometer resolution electron energy loss spectroscopy of biomolecules (*127*). Although early approaches used static solutions in hermetically sealed graphene pouches (Fig. 5C), limiting studies to e-beam–induced or time-lapsed studies (*97*), systematic advances have resulted in the development of multichannel arrays capped with graphene (Fig. 5D) and complete flow cells with graphene windows that allow for in situ and in operando studies on a range of materials, including biological systems (*129*). Similar advances are also seen for PES and SEM, with combinatorial studies being made possible (Fig. 5D) (*128*). Taken together, graphene windows present the potential for transformational insights into reaction mechanisms at the solid-liquid interface; heterogeneous catalysis under realistic pressures; crystal nucleation, growth, and dissolution; material performance in batteries; and biological and other processes that have previously remained inaccessible (*12*, *49*, *127*, *129*, *130*, *136*–*138*). Finally, the electrical conductivity of graphene enables its use as a model electrode, facilitating fundamental understanding of electrochemical processes in operando (Fig. 5D) (*137*). Despite these advantages, issues regarding bubble formation due to radiolysis of water by the e-beam (which can alter solution pH and concentrations of electrolytes), mechanical integrity of graphene membranes (damage to the graphene from bubble collapse or attack from free radicals), perturbations or damage to the interface being probed, and any influence from the graphene surface or residual contaminants on the graphene surface are issues that still need to be addressed (*12*, *129*, *130*, *136*–*138*).

### Outlook

Electron, proton, and deuteron permeation through atomically thin graphene and h-BN presents the potential for breakthrough advances in several fields. However, fundamental understanding of transport mechanisms is still emerging, and advances in measurement techniques and resolution are furthering insights, e.g., anomalous $H_2$ transport through graphene (proposed to occur via $H_2$ dissociation on catalytic ripples or wrinkles and subsequent flipping of adsorbed atoms to the other side of the 2D lattice) (*26*) has renewed focus on the limits of gas impermeability of graphene.

Small-scale applications, such as graphene-coated TEM grids, are already available, and electron transparent windows for imaging and spectroscopy are being increasingly used. Large-area energy-related applications will require advances in scalable, cost-effective 2D material synthesis and membrane fabrication. Isotope separation for $H^+/D^+$ is most likely to see rapid development and commercialization owing to the potential for substantial reduction in energy consumption compared with existing technologies, and these approaches are also likely to be explored for tritium decontamination efforts. Considering that processes to commercially produce large-area CVD graphene and h-BN are starting to mature and facile membrane fabrication using hot pressing and lamination and polymer casting have already been demonstrated, these disruptive innovations are likely to be deployed in the nuclear industry in the near future. Incorporation of 2D materials into proton exchange membranes for flow batteries, fuel cells, and proton pumps is expected to become viable in the next 5 to 10 years, with scaled-up production offering economies of scale and considering energy saving over the application life cycle. Further research is needed to inform and guide technological advances toward each of these applications, including long-term durability studies under realistic conditions to assess material performance, device integration approaches, and membrane manufacturing processes for practical applications.

### REFERENCES AND NOTES


1. A. K. Geim, K. S. Novoselov, The rise of graphene. *Nat. Mater.* **6**, 183–191 (2007). doi: 10.1038/nmat1849; pmid: 17330084
2. A. K. Geim, I. V. Grigorieva, Van der Waals heterostructures. *Nature* **499**, 419–425 (2013). doi: 10.1038/nature12385; pmid: 23887427
3. Y. Peng *et al*., Membranes. Metal-organic framework nanosheets as building blocks for molecular sieving membranes. *Science* **346**, 1356–1359 (2014). doi: 10.1126/science.1254227; pmid: 25504718
4. L. Mogg *et al*., Atomically thin micas as proton-conducting membranes. *Nat. Nanotechnol.* **14**, 962–966 (2019). doi: 10.1038/s41565-019-0536-5; pmid: 31477802
5. Y. Zhong *et al*., Wafer-scale synthesis of monolayer two-dimensional porphyrin polymers for hybrid superlattices. *Science* **366**, 1379–1384 (2019). doi: 10.1126/science.aax9385; pmid: 31699884
6. A. Kolmakov *et al*., Graphene oxide windows for in situ environmental cell photoelectron spectroscopy. *Nat. Nanotechnol.* **6**, 651–657 (2011). doi: 10.1038/nnano.2011.130; pmid: 21873990
7. L. Wang *et al*., Fundamental transport mechanisms, fabrication and potential applications of nanoporous atomically thin membranes. *Nat. Nanotechnol.* **12**, 509–522 (2017). doi: 10.1038/nnano.2017.72; pmid: 28584292
8. L. Prozorovska, P. R. Kidambi, State-of-the-Art and Future Prospects for Atomically Thin Membranes from 2D Materials. *Adv. Mater.* **30**, e1801179 (2018). doi: 10.1002/adma.201801179; pmid: 30085371
9. J. S. Bunch *et al*., Impermeable atomic membranes from graphene sheets. *Nano Lett.* **8**, 2458–2462 (2008). doi: 10.1021/nl801457b; pmid: 18630972

10. S. Hu *et al*., Proton transport through one-atom-thick crystals. *Nature* **516**, 227–230 (2014). doi: 10.1038/nature14015; pmid: 25470058
11. M. Lozada-Hidalgo *et al*., Sieving hydrogen isotopes through two-dimensional crystals. *Science* **351**, 68–70 (2016). doi: 10.1126/science.aac9726; pmid: 26721995
12. J. Kraus *et al*., Photoelectron spectroscopy of wet and gaseous samples through graphene membranes. *Nanoscale* **6**, 14394–14403 (2014). doi: 10.1039/C4NR03561E; pmid: 25333337
13. A. A. Balandin *et al*., Superior thermal conductivity of single-layer graphene. *Nano Lett.* **8**, 902–907 (2008). doi: 10.1021/nl0731872; pmid: 18284217
14. C. R. Dean *et al*., Boron nitride substrates for high-quality graphene electronics. *Nat. Nanotechnol.* **5**, 722–726 (2010). doi: 10.1038/nnano.2010.172; pmid: 20729834
15. K. I. I. Bolotin *et al*., Ultrahigh electron mobility in suspended graphene. *Solid State Commun.* **146**, 351–355 (2008). doi: 10.1016/j.ssc.2008.02.024
16. G. Cassabois, P. Valvin, B. Gil, Hexagonal boron nitride is an indirect bandgap semiconductor. *Nat. Photonics* **10**, 262–266 (2016). doi: 10.1038/nphoton.2015.277
17. Q. Cai *et al*., High thermal conductivity of high-quality monolayer boron nitride and its thermal expansion. *Sci. Adv.* **5**, eaav0129 (2019). doi: 10.1126/sciadv.aav0129; pmid: 31187056
18. C. Lee, X. Wei, J. W. Kysar, J. Hone, Measurement of the elastic properties and intrinsic strength of monolayer graphene. *Science* **321**, 385–388 (2008). doi: 10.1126/science.1157996; pmid: 18635798
19. A. Falin *et al*., Mechanical properties of atomically thin boron nitride and the role of interlayer interactions. *Nat. Commun.* **8**, 15815 (2017). doi: 10.1038/ncomms15815; pmid: 28639613
20. S. P. Koenig, N. G. Boddeti, M. L. Dunn, J. S. Bunch, Ultrastrong adhesion of graphene membranes. *Nat. Nanotechnol.* **6**, 543–546 (2011). doi: 10.1038/nnano.2011.123; pmid: 21841794
21. L. Wang, C. M. C. M. Williams, M. S. H. Boutilier, P. R. Kidambi, R. Karnik, Single-Layer Graphene Membranes Withstand Ultrahigh Applied Pressure. *Nano Lett.* **17**, 3081–3088 (2017). doi: 10.1021/acs.nanolett.7b00442; pmid: 28434230
22. R. S. Weatherup, B. Eren, Y. Hao, H. Bluhm, M. B. Salmeron, Graphene Membranes for Atmospheric Pressure Photoelectron Spectroscopy. *J. Phys. Chem. Lett.* **7**, 1622–1627 (2016). doi: 10.1021/acs.jpclett.6b00640; pmid: 27082434
23. D. Cohen-Tanugi, J. C. Grossman, Mechanical strength of nanoporous graphene as a desalination membrane. *Nano Lett.* **14**, 6171–6178 (2014). doi: 10.1021/nl502399y; pmid: 25357231
24. S. P. Koenig, L. Wang, J. Pellegrino, J. S. Bunch, Selective molecular sieving through porous graphene. *Nat. Nanotechnol.* **7**, 728–732 (2012). doi: 10.1038/nnano.2012.162; pmid: 23042491
25. L. Mogg *et al*., Perfect proton selectivity in ion transport through two-dimensional crystals. *Nat. Commun.* **10**, 4243 (2019). doi: 10.1038/s41467-019-12314-2; pmid: 31534140
26. P. Z. Sun *et al*., Limits on gas impermeability of graphene. *Nature* **579**, 229–232 (2020). doi: 10.1038/s41586-020-2070-x; pmid: 32161387
27. L. Wang *et al*., Molecular valves for controlling gas phase transport made from discrete ångström-sized pores in graphene. *Nat. Nanotechnol.* **10**, 785–790 (2015). doi: 10.1038/nnano.2015.158; pmid: 26237344
28. L. Tsetseris, S. T. Pantelides, Graphene: An impermeable or selectively permeable membrane for atomic species? *Carbon* **67**, 58–63 (2014). doi: 10.1016/j.carbon.2013.09.055
29. M. Miao, M. B. Nardelli, Q. Wang, Y. Liu, First principles study of the permeability of graphene to hydrogen atoms. *Phys. Chem. Chem. Phys.* **15**, 16132–16137 (2013). doi: 10.1039/c3cp52318g; pmid: 23986179
30. O. Leenaerts, B. Partoens, F. M. Peeters, Graphene: A perfect nanoballoon. *Appl. Phys. Lett.* **93**, 193107 (2008). doi: 10.1063/1.3021413
31. M. Seel, R. Pandey, Proton and hydrogen transport through two-dimensional monolayers. *2D Mater.* **3**, 025004 (2016). doi: 10.1088/2053-1583/3/2/025004
32. L. W. Drahushuk, L. Wang, S. P. Koenig, J. S. Bunch, M. S. Strano, Analysis of time-varying, stochastic gas transport through graphene membranes. *ACS Nano* **10**, 786–795 (2016). doi: 10.1021/acsnano.5b05870; pmid: 26720748
33. O. Lehtinen *et al*., Effects of ion bombardment on a two-dimensional target: Atomistic simulations of graphene irradiation. *Phys. Rev. B Condens. Matter Mater. Phys.* **81**, 153401 (2010). doi: 10.1103/PhysRevB.81.153401
34. K. Yoon *et al*., Atomistic-Scale Simulations of Defect Formation in Graphene under Noble Gas Ion Irradiation. *ACS Nano* **10**, 8376–8384 (2016). doi: 10.1021/acsnano.6b03036; pmid: 27532882
35. K. Alexandrou, Ionizing Radiation Effects on Graphene Based Field Effects Transistors. *ProQuest Diss. Theses*, 148 (2016).
36. A. V. Krasheninnikov, Are two-dimensional materials radiation tolerant? *Nanoscale Horiz.* **5**, 1447–1452 (2020). doi: 10.1039/D0NH00465K; pmid: 32969454
37. T. Susi, J. C. Meyer, J. Kotakoski, Quantifying transmission electron microscopy irradiation effects using two-dimensional materials. *Nat. Rev. Phys.* **1**, 397–405 (2019). doi: 10.1038/s42254-019-0058-y
38. E. Cobas, A. L. Friedman, O. M. J. Van't Erve, J. T. Robinson, B. T. Jonker, Graphene as a tunnel barrier: Graphene-based magnetic tunnel junctions. *Nano Lett.* **12**, 3000–3004 (2012). doi: 10.1021/nl3007616; pmid: 22577860
39. X. Zhu *et al*., A Study of Vertical Transport through Graphene toward Control of Quantum Tunneling. *Nano Lett.* **18**, 682–688 (2018). doi: 10.1021/acs.nanolett.7b03221; pmid: 29300487
40. J. Xue *et al*., Scanning tunnelling microscopy and spectroscopy of ultra-flat graphene on hexagonal boron nitride. *Nat. Mater.* **10**, 282–285 (2011). doi: 10.1038/nmat2968; pmid: 21317900
41. G. H. Lee *et al*., Electron tunneling through atomically flat and ultrathin hexagonal boron nitride. *Appl. Phys. Lett.* **99**, 243114 (2011). doi: 10.1063/1.3662043
42. L. Britnell *et al*., Electron tunneling through ultrathin boron nitride crystalline barriers. *Nano Lett.* **12**, 1707–1710 (2012). doi: 10.1021/nl3002205; pmid: 22380756
43. E. C. Ahn, 2D materials for spintronic devices. *npj 2D Mater. Appl.* **4**, 17 (2020). doi: 10.1038/s41699-020-0152-0
44. F. Giannazzo, G. Greco, F. Roccaforte, S. Sonde, Vertical Transistors Based on 2D Materials: Status and Prospects. *Crystals* **8**, 70 (2018). doi: 10.3390/cryst8020070
45. J. A. Yan, J. A. Driscoll, B. K. Wyatt, K. Varga, S. T. Pantelides, Time-domain simulation of electron diffraction in crystals. *Phys. Rev. B Condens. Matter Mater. Phys.* **84**, 224117 (2011). doi: 10.1103/PhysRevB.84.224117
46. J.-N. N. Longchamp, T. Latychevskaia, C. Escher, H.-W. W. Fink, Low-energy electron transmission imaging of clusters on free-standing graphene. *Appl. Phys. Lett.* **101**, 113117 (2012). doi: 10.1063/1.4752717
47. J. Y. Mutus *et al*., Low-energy electron point projection microscopy of suspended graphene, the ultimate "microscope slide.". *New J. Phys.* **13**, 063011 (2011). doi: 10.1088/1367-2630/13/6/063011
48. G. Hassink *et al*., Transparency of graphene for low-energy electrons measured in a vacuum-triode setup. *APL Mater.* **3**, 076106 (2015). doi: 10.1063/1.4927406
49. R. S. Weatherup; R. S. Weatherup, 2D Material Membranes for Operando Atmospheric Pressure Photoelectron Spectroscopy. *Top. Catal.* **61**, 2085–2102 (2018). doi: 10.1007/s11244-018-1075-2
50. S. Tanuma, C. J. Powell, D. R. Penn, Calculations of electron inelastic mean free paths. IX. Data for 41 elemental solids over the 50 eV to 30 keV range. *Surf. Interface Anal.* **43**, 689–713 (2011). doi: 10.1002/sia.3522
51. M. Xu, D. Fujita, J. Gao, N. Hanagata, Auger electron spectroscopy: A rational method for determining thickness of graphene films. *ACS Nano* **4**, 2937–2945 (2010). doi: 10.1021/nn100276w; pmid: 20373812
52. J. C. Meyer *et al*., Accurate measurement of electron beam induced displacement cross sections for single-layer graphene. *Phys. Rev. Lett.* **108**, 196102 (2012). doi: 10.1103/PhysRevLett.108.196102; pmid: 23003063
53. J. C. Meyer, A. Chuvilin, G. Algara-Siller, J. Biskupek, U. Kaiser, Selective sputtering and atomic resolution imaging of atomically thin boron nitride membranes. *Nano Lett.* **9**, 2683–2689 (2009). doi: 10.1021/nl9011497; pmid: 19480400
54. S. M. Gilbert *et al*., Fabrication of Subnanometer-Precision Nanopores in Hexagonal Boron Nitride. *Sci. Rep.* **7**, 15096 (2017). doi: 10.1038/s41598-017-12684-x; pmid: 29118413
55. C. Thiele *et al*., Electron-beam-induced direct etching of graphene. *Carbon* **64**, 84–91 (2013). doi: 10.1016/j.carbon.2013.07.038
56. C. Elbadawi *et al*., Electron beam directed etching of hexagonal boron nitride. *Nanoscale* **8**, 16182–16186 (2016). doi: 10.1039/C6NR04959A; pmid: 27603125
57. S. Sinha, J. H. Warner, Recent Progress in Using Graphene as an Ultrathin Transparent Support for Transmission Electron Microscopy. *Small Struct.* **2**, 2000049 (2021). doi: 10.1002/sstr.202000049
58. M. I. Walker, P. Braeuninger-Weimer, R. S. Weatherup, S. Hofmann, U. F. Keyser, Measuring the proton selectivity of graphene membranes. *Appl. Phys. Lett.* **107**, 213104 (2015). doi: 10.1063/1.4936335
59. J. L. Achtyl *et al*., Aqueous proton transfer across single-layer graphene. *Nat. Commun.* **6**, 6539 (2015). doi: 10.1038/ncomms7539; pmid: 25781149
60. M. Lozada-Hidalgo *et al*., Giant photoeffect in proton transport through graphene membranes. *Nat. Nanotechnol.* **13**, 300–303 (2018). doi: 10.1038/s41565-017-0051-5; pmid: 29358638
61. I. Poltavsky, L. Zheng, M. Mortazavi, A. Tkatchenko, Quantum tunneling of thermal protons through pristine graphene. *J. Chem. Phys.* **148**, 204707 (2018). doi: 10.1063/1.5024317; pmid: 29865849
62. Y. Feng *et al*., Hydrogenation Facilitates Proton Transfer through Two-Dimensional Honeycomb Crystals. *J. Phys. Chem. Lett.* **8**, 6009–6014 (2017). doi: 10.1021/acs.jpclett.7b02820; pmid: 29185752
63. J. M. H. Kroes, A. Fasolino, M. I. Katsnelson, Density functional based simulations of proton permeation of graphene and hexagonal boron nitride. *Phys. Chem. Chem. Phys.* **19**, 5813–5817 (2017). doi: 10.1039/C6CP08923B; pmid: 28177003
64. M. Bartolomei, M. I. Hernández, J. Campos-Martínez, R. Hernández-Lamoneda, Graphene multi-protonation: A cooperative mechanism for proton permeation. *Carbon* **144**, 724–730 (2019). doi: 10.1016/j.carbon.2018.12.086
65. Q. Zhang, M. Ju, L. Chen, X. C. Zeng, Differential Permeability of Proton Isotopes through Graphene and Graphene Analogue Monolayer. *J. Phys. Chem. Lett.* **7**, 3395–3400 (2016). doi: 10.1021/acs.jpclett.6b01507; pmid: 27522866
66. S. Bukola, Y. Liang, C. Korzeniewski, J. Harris, S. Creager, Selective Proton/Deuteron Transport through Nafion|Graphene|Nafion Sandwich Structures at High Current Density. *J. Am. Chem. Soc.* **140**, 1743–1752 (2018). doi: 10.1021/jacs.7b10853; pmid: 29350035
67. M. Lozada-Hidalgo *et al*., Scalable and efficient separation of hydrogen isotopes using graphene-based electrochemical pumping. *Nat. Commun.* **8**, 15215 (2017). doi: 10.1038/ncomms15215; pmid: 28485380
68. Y. An, A. F. Oliveira, T. Brumme, A. Kuc, T. Heine, Stone-Wales Defects Cause High Proton Permeability and Isotope Selectivity of Single-Layer Graphene. *Adv. Mater.* **32**, e2002442 (2020). doi: 10.1002/adma.202002442; pmid: 32743870
69. P. Y. Huang *et al*., Grains and grain boundaries in single-layer graphene atomic patchwork quilts. *Nature* **469**, 389–392 (2011). doi: 10.1038/nature09718; pmid: 21209615
70. E. Griffin *et al*., Proton and Li-Ion Permeation through Graphene with Eight-Atom-Ring Defects. *ACS Nano* **14**, 7280–7286 (2020). doi: 10.1021/acsnano.0c02496; pmid: 32427466
71. S. Bukola, S. E. Creager, A charge-transfer resistance model and Arrhenius activation analysis for hydrogen ion transmission across single-layer graphene. *Electrochim. Acta* **296**, 1–7 (2019). doi: 10.1016/j.electacta.2018.11.005
72. M. I. Walker, R. S. Weatherup, N. A. W. Bell, S. Hofmann, U. F. Keyser, Free-standing graphene membranes on glass nanopores for ionic current measurements. *Appl. Phys. Lett.* **106**, 023119 (2015). doi: 10.1063/1.4906236
73. M. I. Walker *et al*., Extrinsic Cation Selectivity of 2D Membranes. *ACS Nano* **11**, 1340–1346 (2017). doi: 10.1021/acsnano.6b06034; pmid: 28157333
74. P. Chaturvedi *et al*., Ionic conductance through graphene: Assessing its applicability as a proton selective membrane. *ACS Nano* **13**, 12109–12119 (2019). doi: 10.1021/acsnano.9b06505; pmid: 31592639
75. Z. Zeng *et al*., Biomimetic N-Doped Graphene Membrane for Proton Exchange Membranes. *Nano Lett.* **21**, 4314–4319 (2021). doi: 10.1021/acs.nanolett.1c00813; pmid: 33848172
76. Y. Hernandez *et al*., High-yield production of graphene by liquid-phase exfoliation of graphite. *Nat. Nanotechnol.* **3**, 563–568 (2008). doi: 10.1038/nnano.2008.215; pmid: 18772919
77. X. Li *et al*., Large-area synthesis of high-quality and uniform graphene films on copper foils. *Science* **324**, 1312–1314 (2009). doi: 10.1126/science.1171245; pmid: 19423775

78. K. K. Kim *et al*., Synthesis of monolayer hexagonal boron nitride on Cu foil using chemical vapor deposition. *Nano Lett.* **12**, 161–166 (2012). doi: 10.1021/nl203249a; pmid: 22111957
79. T. Zhao *et al*., Ultrafast growth of nanocrystalline graphene films by quenching and grain-size-dependent strength and bandgap opening. *Nat. Commun.* **10**, 4854 (2019). doi: 10.1038/s41467-019-12662-z; pmid: 31649240
80. C. T. Toh *et al*., Synthesis and properties of free-standing monolayer amorphous carbon. *Nature* **577**, 199–203 (2020). doi: 10.1038/s41586-019-1871-2; pmid: 31915396
81. Z. Cai, B. Liu, X. Zou, H. M. Cheng, Chemical Vapor Deposition Growth and Applications of Two-Dimensional Materials and Their Heterostructures. *Chem. Rev.* **118**, 6091–6133 (2018). doi: 10.1021/acs.chemrev.7b00536; pmid: 29384374
82. P. R. Kidambi *et al*., A Scalable Route to Nanoporous Large-Area Atomically Thin Graphene Membranes by Roll-to-Roll Chemical Vapor Deposition and Polymer Support Casting. *ACS Appl. Mater. Interfaces* **10**, 10369–10378 (2018). doi: 10.1021/acsami.8b00846; pmid: 29553242
83. L. Wang *et al*., Epitaxial growth of a 100-square-centimetre single-crystal hexagonal boron nitride monolayer on copper. *Nature* **570**, 91–95 (2019). doi: 10.1038/s41586-019-1226-z; pmid: 31118514
84. T. Kobayashi *et al*., Production of a 100-m-long high-quality graphene transparent conductive film by roll-to-roll chemical vapor deposition and transfer process. *Appl. Phys. Lett.* **102**, 023112 (2013). doi: 10.1063/1.4776707
85. S. C. O'Hern *et al*., Selective molecular transport through intrinsic defects in a single layer of CVD graphene. *ACS Nano* **6**, 10130–10138 (2012). doi: 10.1021/nn303869m; pmid: 23030691
86. N. Petrone *et al*., Chemical vapor deposition-derived graphene with electrical performance of exfoliated graphene. *Nano Lett.* **12**, 2751–2756 (2012). doi: 10.1021/nl204481s; pmid: 22582828
87. X. Xu *et al*., Ultrafast epitaxial growth of metre-sized single-crystal graphene on industrial Cu foil. *Sci. Bull. (Beijing)* **62**, 1074–1080 (2017). doi: 10.1016/j.scib.2017.07.005
88. J. H. Lee *et al*., Wafer-scale growth of single-crystal monolayer graphene on reusable hydrogen-terminated germanium. *Science* **344**, 286–289 (2014). doi: 10.1126/science.1252268; pmid: 24700471
89. J. S. Lee *et al*., Wafer-scale single-crystal hexagonal boron nitride film via self-collimated grain formation. *Science* **362**, 817–821 (2018). doi: 10.1126/science.aau2132; pmid: 30442807
90. T. Wu *et al*., Fast growth of inch-sized single-crystalline graphene from a controlled single nucleus on Cu-Ni alloys. *Nat. Mater.* **15**, 43–47 (2016). doi: 10.1038/nmat4477; pmid: 26595118
91. X. Xu *et al*., Ultrafast growth of single-crystal graphene assisted by a continuous oxygen supply. *Nat. Nanotechnol.* **11**, 930–935 (2016). doi: 10.1038/nnano.2016.132; pmid: 27501317
92. I. V. Vlassiouk *et al*., Evolutionary selection growth of two-dimensional materials on polycrystalline substrates. *Nat. Mater.* **17**, 318–322 (2018). doi: 10.1038/s41563-018-0019-3; pmid: 29531368
93. P. R. Kidambi *et al*., Selective Nanoscale Mass Transport across Atomically Thin Single Crystalline Graphene Membranes. *Adv. Mater.* **29**, 1605896 (2017). doi: 10.1002/adma.201605896; pmid: 28306180
94. P. R. Kidambi *et al*., Assessment and control of the impermeability of graphene for atomically thin membranes and barriers. *Nanoscale* **9**, 8496–8507 (2017). doi: 10.1039/C7NR01921A; pmid: 28604878
95. S. C. O'Hern *et al*., Nanofiltration across defect-sealed nanoporous monolayer graphene. *Nano Lett.* **15**, 3254–3260 (2015). doi: 10.1021/acs.nanolett.5b00456; pmid: 25915708
96. P. Cheng *et al*., Facile Size-Selective Defect Sealing in Large-Area Atomically Thin Graphene Membranes for Sub-Nanometer Scale Separations. *Nano Lett.* **20**, 5951–5959 (2020). doi: 10.1021/acs.nanolett.0c01934; pmid: 32628858
97. J. M. Yuk *et al*., High-resolution EM of colloidal nanocrystal growth using graphene liquid cells. *Science* **336**, 61–64 (2012). doi: 10.1126/science.1217654; pmid: 22491849
98. P. R. Kidambi *et al*., Facile Fabrication of Large-Area Atomically Thin Membranes by Direct Synthesis of Graphene with Nanoscale Porosity. *Adv. Mater.* **30**, e1804977 (2018). doi: 10.1002/adma.201804977; pmid: 30368941
99. P. Cheng, N. K. Moehring, J. C. Idrobo, I. N. Ivanov, P. R. Kidambi, Scalable synthesis of nanoporous atomically thin graphene membranes for dialysis and molecular separations *via* facile isopropanol-assisted hot lamination. *Nanoscale* **13**, 2825–2837 (2021). doi: 10.1039/D0NR07384A; pmid: 33508042
100. Y. Wang *et al*., Electrochemical delamination of CVD-grown graphene film: Toward the recyclable use of copper catalyst. *ACS Nano* **5**, 9927–9933 (2011). doi: 10.1021/nn203700w; pmid: 22034835
101. The Price Of Graphene – Graphenea, (available at https://www.graphenea.com/pages/graphene-price#.X-uoxdhKhPY).
102. A. Zurutuza, C. Marinelli, Challenges and opportunities in graphene commercialization. *Nat. Nanotechnol.* **9**, 730–734 (2014). doi: 10.1038/nnano.2014.225; pmid: 25286257
103. S. Satyapal, U. Department of Energy Hydrogen, F. C. Program, "DOE Hydrogen and Fuel Cell Perspectives and Overview of the International Partnership for Hydrogen and Fuel Cells in the Economy (IPHE)" (2020), (available at https://www.energy.gov/sites/prod/files/2020/07/f77/hfto-satyapal-gabi-workshop-jul20.pdf).
104. G. Alberti, M. Casciola, L. Massinelli, B. Bauer, Polymeric proton conducting membranes for medium temperature fuel cells (110-160°C). *J. Membr. Sci.* **185**, 73–81 (2001). doi: 10.1016/S0376-7388(00)00635-9
105. K. A. Mauritz, R. B. Moore, State of understanding of nafion. *Chem. Rev.* **104**, 4535–4586 (2004). doi: 10.1021/cr0207123; pmid: 15669162
106. C. E. Thomas, Fuel cell and battery electric vehicles compared. *Int. J. Hydrogen Energy* **34**, 6005–6020 (2009). doi: 10.1016/j.ijhydene.2009.06.003
107. Y. Wang, D. F. Ruiz Diaz, K. S. Chen, Z. Wang, X. C. Adroher, Materials, technological status, and fundamentals of PEM fuel cells – A review. *Mater. Today* **32**, 178–203 (2020). doi: 10.1016/j.mattod.2019.06.005
108. M. B. Karimi, F. Mohammadi, K. Hooshyari, Recent approaches to improve Nafion performance for fuel cell applications: A review. *Int. J. Hydrogen Energy* **44**, 28919–28938 (2019). doi: 10.1016/j.ijhydene.2019.09.096
109. S. I. Yoon *et al*., AA'-Stacked Trilayer Hexagonal Boron Nitride Membrane for Proton Exchange Membrane Fuel Cells. *ACS Nano* **12**, 10764–10771 (2018). doi: 10.1021/acsnano.8b06268; pmid: 30335961
110. S. M. Holmes *et al*., 2D Crystals Significantly Enhance the Performance of a Working Fuel Cell. *Adv. Energy Mater.* **7**, 1601216 (2017). doi: 10.1002/aenm.201601216
111. X. H. H. Yan, R. Wu, J. B. B. Xu, Z. Luo, T. S. S. Zhao, A monolayer graphene – Nafion sandwich membrane for direct methanol fuel cells. *J. Power Sources* **311**, 188–194 (2016). doi: 10.1016/j.jpowsour.2016.02.030
112. M. S. H. Boutilier *et al*., Molecular Sieving Across Centimeter-Scale Single-Layer Nanoporous Graphene Membranes. *ACS Nano* **11**, 5726–5736 (2017). doi: 10.1021/acsnano.7b01231; pmid: 28609103
113. L. H. Li, J. Cervenka, K. Watanabe, T. Taniguchi, Y. Chen, Strong oxidation resistance of atomically thin boron nitride nanosheets. *ACS Nano* **8**, 1457–1462 (2014). doi: 10.1021/nn500059s; pmid: 24400990
114. Z. P. Cano *et al*., Batteries and fuel cells for emerging electric vehicle markets. *Nat. Energy* **3**, 279–289 (2018). doi: 10.1038/s41560-018-0108-1
115. B. D. James, J. A. Kalinoski, K. N. Baum, *Mass Production Cost Estimation for Direct H 2 PEM Fuel Cell Systems for Automotive Applications : 2009 Update, US department of energy report* (2010); www.energy.gov/sites/prod/files/2017/06/f34/fcto_sa_2016_pemfc_transportation_cost_analysis.pdf.
116. Fuel Cells (Hydrogen and Fuel Cell Technologies Office, US Department of Energy); www.energy.gov/eere/fuelcells.
117. G. Saur, A. Milbrandt, Renewable Hydrogen Potential from Biogas in the United States. *Natl. Renew. Energy Lab.*, 1–44 (2014).
118. G. L. Soloveichik, Flow Batteries: Current Status and Trends. *Chem. Rev.* **115**, 11533–11558 (2015). doi: 10.1021/cr500720t; pmid: 26389560
119. UniEnergy Technologies–Rongke Power, EcoPartnerships; https://electrek.co/2017/12/21/worlds-largest-battery-200mw-800mwh-vanadium-flow-battery-rongke-power/.
120. J. Liu *et al*., Sandwiching h-BN monolayer films between sulfonated poly(ether ether ketone) and nafion for proton exchange membranes with improved ion selectivity. *ACS Nano* **13**, 2094–2102 (2019). doi: 10.1021/acsnano.8b08680; pmid: 30768234
121. S. Bukola *et al*., Single-layer graphene as a highly selective barrier for vanadium crossover with high proton selectivity. *J. Energy Chem.* **59**, 419–430 (2021). doi: 10.1016/j.jechem.2020.11.025
122. Q. Chen *et al*., Graphene enhances the proton selectivity of porous membrane in vanadium flow batteries. *Mater. Des.* **113**, 149–156 (2017). doi: 10.1016/j.matdes.2016.10.019
123. A. I. Miller, Heavy Water: A Manufacturers' Guide for the Hydrogen Century. *Can. Nucl. Soc. Bull.* **22**, 1–14 (2001).
124. H. K. Rae, Selecting Heavy Water Processes. *ACS Symp. Ser.* **68**, 1–26 (1978).
125. S. Tosti, A. Pozio, Membrane Processes for the Nuclear Fusion Fuel Cycle. *Membranes* **8**, 96 (2018). doi: 10.3390/membranes8040096; pmid: 30322084
126. "Report for the Second of Two 2008 Charges to the Nuclear Science Advisory Committee on the Isotope Development and Production for Research and Applications Program Isotopes for the Nation's Future A Long Range Plan NSAC Isotopes Subcommittee" (2009), (available at https://science.osti.gov/-/media/np/nsac/pdf/docs/nsaci_ii_report.pdf).
127. C. Wang, Q. Qiao, T. Shokuhfar, R. F. Klie, High-resolution electron microscopy and spectroscopy of ferritin in biocompatible graphene liquid cells and graphene sandwiches. *Adv. Mater.* **26**, 3410–3414 (2014). doi: 10.1002/adma.201306069; pmid: 24497051
128. A. Yulaev *et al*., Graphene Microcapsule Arrays for Combinatorial Electron Microscopy and Spectroscopy in Liquids. *ACS Appl. Mater. Interfaces* **9**, 26492–26502 (2017). doi: 10.1021/acsami.7b02824; pmid: 28447785
129. J. Park *et al*., Graphene Liquid Cell Electron Microscopy: Progress, Applications, and Perspectives. *ACS Nano* **15**, 288–308 (2021). doi: 10.1021/acsnano.0c10229; pmid: 33395264
130. F. M. Ross, Opportunities and challenges in liquid cell electron microscopy. *Science* **350**, aaa9886–aaa9886 (2015). doi: 10.1126/science.aaa9886; pmid: 26680204
131. Y. Han, K. X. Nguyen, Y. Ogawa, J. Park, D. A. Muller, Atomically Thin Graphene Windows That Enable High Contrast Electron Microscopy without a Specimen Vacuum Chamber. *Nano Lett.* **16**, 7427–7432 (2016). doi: 10.1021/acs.nanolett.6b03016; pmid: 27960512
132. J. C. Meyer, C. O. Girit, M. F. Crommie, A. Zettl, Imaging and dynamics of light atoms and molecules on graphene. *Nature* **454**, 319–322 (2008). doi: 10.1038/nature07094; pmid: 18633414
133. R. R. Nair *et al*., Graphene as a transparent conductive support for studying biological molecules by transmission electron microscopy. *Appl. Phys. Lett.* **97**, 153102 (2010). doi: 10.1063/1.3492845
134. K. Naydenova, M. J. Peet, C. J. Russo, Multifunctional graphene supports for electron cryomicroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11718–11724 (2019). pmid: 31127045
135. Y. Han *et al*., High-yield monolayer graphene grids for near-atomic resolution cryoelectron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1009–1014 (2020). doi: 10.1073/pnas.1919114117; pmid: 31879346
136. J. D. Stoll, A. Kolmakov, Electron transparent graphene windows for environmental scanning electron microscopy in liquids and dense gases. *Nanotechnology* **23**, 505704 (2012). doi: 10.1088/0957-4484/23/50/505704; pmid: 23165114
137. S. Nemšák *et al*., Interfacial Electrochemistry in Liquids Probed with Photoemission Electron Microscopy. *J. Am. Chem. Soc.* **139**, 18138–18141 (2017). doi: 10.1021/jacs.7b07365; pmid: 29148738
138. K. Koo, K. S. Dae, Y. K. Hahn, J. M. Yuk, Live Cell Electron Microscopy Using Graphene Veils. *Nano Lett.* **20**, 4708–4713 (2020). doi: 10.1021/acs.nanolett.0c00715; pmid: 32369367
139. S. Bukola, K. Beard, C. Korzeniewski, J. M. Harris, S. E. Creager, Single-Layer Graphene Sandwiched between Proton-Exchange Membranes for Selective Proton Transmission. *ACS Appl. Nano Mater.* **2**, 964–974 (2019). doi: 10.1021/acsanm.8b02270
140. B. Jiang, L. Wu, L. Yu, X. Qiu, J. Xi, A comparative study of Nafion series membranes for vanadium redox flow batteries. *J. Membr. Sci.* **510**, 18–26 (2016). doi: 10.1016/j.memsci.2016.03.007
141. S. Slade, S. A. Campbell, T. R. Ralph, F. C. Walsh, Ionic Conductivity of an Extruded Nafion 1100 EW Series of Membranes. *J. Electrochem. Soc.* **149**, A1556 (2002). doi: 10.1149/1.1517281
142. N. T. Ekanayake, J. Huang, J. Jakowski, B. G. Sumpter, S. Garashchuk, Relevance of the Nuclear Quantum Effects on the Proton/Deuteron Transmission through Hexagonal Boron

Nitride and Graphene Monolayers. *J. Phys. Chem. C* **121**, 24335–24344 (2017). doi: 10.1021/acs.jpcc.7b08152
143. L. Shi, A. Xu, G. Chen, T. Zhao, Theoretical Understanding of Mechanisms of Proton Exchange Membranes Made of 2D Crystals with Ultrahigh Selectivity. *J. Phys. Chem. Lett.* **8**, 4354–4361 (2017). doi: 10.1021/acs.jpclett.7b01999; pmid: 28846433
144. W. L. Wang, E. Kaxiras, Graphene Hydrate: Theoretical Prediction of a New Insulating Form of Graphene. *New J. Phys.* **12**, 125012 (2010). doi: 10.1088/1367-2630/12/12/125012
145. T. Hao, Y. Xu, T. Hao, Conductivity Equations of Protons Transporting through 2D Crystals Obtained with the Rate Process Theory and Free Volume Concept. *Chem. Phys. Lett.* **698**, 67–71 (2018). doi: 10.1016/j.cplett.2018.02.059

## ACKNOWLEDGMENTS

We acknowledge helpful input from A. Kolmakov and R. Schrimpf and stimulating discussions with S. Pantelides, S. Creager, R. Karnik, P. N. Pintauro, J. A. Hart, C. Payne, E. Balkin, and D. Matuszak. **Funding:** P.R.K. acknowledges funding from NSF CAREER award 1944134 and 2020-2021 ECS-Toyota Young Investigator Fellowship. **Author contributions:** P.R.K. wrote the paper with input from P.C. and N.K.M. N.K.M. and P.C. made the figures, contributed to discussions and writing. P.C. and N.K.M. contributed equally. **Competing interests:** P.R.K. acknowledges a stake in a company aimed at commercializing 2D materials with patents awarded and pending. **Data and materials availability:** References to all data are provided in the manuscript.

10.1126/science.abd7687

RESEARCH ARTICLE SUMMARY

NEURODEVELOPMENT

# Convergence of adenosine and GABA signaling for synapse stabilization during development

**Ferran Gomez-Castro†, Stefania Zappettini†, Jessica C. Pressey†, Carla G. Silva, Marion Russeau, Nicolas Gervasi, Marta Figueiredo, Claire Montmasson, Marianne Renner, Paula M. Canas, Francisco Q. Gonçalves, Sofia Alçada-Morais, Eszter Szabó, Ricardo J. Rodrigues, Paula Agostinho, Angelo R. Tomé, Ghislaine Caillol, Olivier Thoumine, Xavier Nicol, Christophe Leterrier, Rafael Lujan, Shiva K. Tyagarajan, Rodrigo A. Cunha, Monique Esclapez, Christophe Bernard*, Sabine Lévi***

**INTRODUCTION:** During development, brain circuits go through phases of synapse formation, stabilization, or elimination. γ-aminobutyric acid–mediated (GABAergic) synapse formation depends mainly on cell adhesion molecules, such as neuroligins and leucine-rich repeat transmembrane proteins, that interact with presynaptic neurexins and Slit- and Trk-like family proteins that bind to presynaptic protein tyrosine phosphatases. GABA and GABA type A ($GABA_A$) receptors are involved in an activity-dependent manner in the maturation and pruning of synapses. Adenosine triphosphate (ATP) and adenosine can be coreleased with GABA at synapses to be perceived by adenosine $A_{2A}$ receptors. We tested the role of adenosine signaling in the stabilization and elimination of GABAergic synapses.

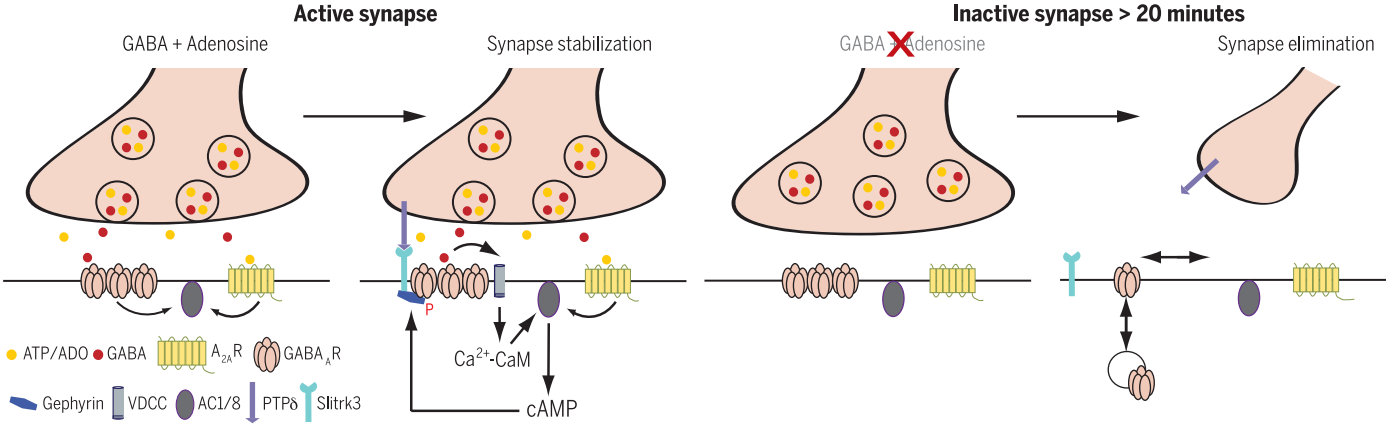
**RATIONALE:** $A_{2A}$ receptors control migration speed, axonal elongation, and dendrite branching. Whether $A_{2A}$ receptors control synapse formation, stabilization, or elimination in the brain is not known. In the adult brain, $A_{2A}$ receptors are mostly expressed on presynaptic terminals, where they control the probability of synaptic vesicle release. The amount, location, and function of $A_{2A}$ receptors at neural synapses during early brain development has been unclear.

**RESULTS:** During synaptogenesis in the developing mouse hippocampus, between postnatal days P5 and P16, the density of $A_{2A}$ receptors increases transiently around the postsynaptic density. Activity-dependent release of its endogenous ligand, adenosine, increases as well. $A_{2A}$ receptors control the fate of GABAergic synapses. Suppression of $A_{2A}$ receptors, their pharmacological blockade, or the removal of adenosine results in the destabilization of pre- and postsynaptic sites in vivo, ex vivo, and in vitro. If $A_{2A}$ receptors remain inactive for >20 min, synapse destabilization is irreversible. We found that $A_{2A}$ receptor activation is necessary and sufficient for GABAergic synapse stabilization, whereas $GABA_A$ receptor activation is not necessary as long as $A_{2A}$ receptors remain activated. We studied the molecular mechanism at play. $A_{2A}$ receptor and $GABA_A$ receptor signaling pathways converge onto calcium-calmodulin–sensitive adenylyl cyclases to produce adenosine 3′,5′-monophosphate (cAMP). The resulting activation of protein kinase A leads to phosphorylation of the postsynaptic scaffolding molecule gephyrin at the protein kinase A–sensitive serine residue 303 site. Expression of the gephyrin mutant mimicking this phosphorylation state prevents synapse loss upon the removal of extracellular adenosine. Phosphorylated gephyrin can be coimmunoprecipitated with the postsynaptic transmembrane Slit- and Trk-like family protein 3 that binds in the synaptic cleft to presynaptic protein tyrosine phosphatase σ to organize inhibitory synapses. The contribution of Slit- and Trk-like family protein 3 in stabilizing GABAergic synapses through adenosine signaling is demonstrated with a short hairpin RNA (shRNA) approach or after the expression of a mutant. Finally, antagonizing $A_{2A}$ receptors during synaptogenesis in vivo results in the loss of GABAergic synapses during development and cognitive deficits when animals reach adulthood.

**CONCLUSION:** $A_{2A}$ receptors regulate the elimination of certain GABAergic synapses when they become inactive. $A_{2A}$ receptors are poised to detect active presynaptic terminals and trigger synapse removal after a defined period of synaptic inactivity. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: sabine.levi@inserm.fr (S.L.); christophe.bernard@univ-amu.fr (C.B.)
†These authors contributed equally to this work.
Cite this article as F. Gomez-Castro *et al.*, *Science* **374**, eabk2055 (2021). DOI: 10.1126/science.abk2055

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abk2055




**Adenosine signaling stabilizes nascent GABAergic synapses.** (Left) Active synapse: Corelease of adenosine, ATP, and GABA activates $A_{2A}$ receptors and $GABA_A$ receptors, whose signaling pathways converge on $Ca^{2+}$-calmodulin–dependent adenylyl cyclases and cAMP production, which in turn may stabilize the nascent synapse through recruitment of the Slitrk3-PTPσ transsynaptic organizers by gephyrin phosphorylated at a protein kinase A (PKA) site. (Right) Inactive synapse: In the absence of adenosine, ATP, and GABA release at inactive synapses, this pathway is not activated, and the synapse is eliminated. ADO, adenosine; $A_{2A}R$, adenosine type 2A receptor; $GABA_AR$, $GABA_A$ receptor; VDCC, voltage-dependent calcium channel; AC1 or AC8, adenylyl cyclase 1 or 8; PTPδ, protein tyrosine phosphatase δ; Slitrk3, Slit- and Trk-like family member 3; CaM, calmodulin.

RESEARCH ARTICLE

NEURODEVELOPMENT

# Convergence of adenosine and GABA signaling for synapse stabilization during development

Ferran Gomez-Castro[1]†, Stefania Zappettini[2]†, Jessica C. Pressey[1,3]†, Carla G. Silva[2,4], Marion Russeau[1], Nicolas Gervasi[1,5], Marta Figueiredo[6], Claire Montmasson[1], Marianne Renner[1], Paula M. Canas[4], Francisco Q. Gonçalves[4], Sofia Alçada-Morais[4], Eszter Szabó[4], Ricardo J. Rodrigues[4], Paula Agostinho[4,7], Angelo R. Tomé[4,8], Ghislaine Caillol[9], Olivier Thoumine[10], Xavier Nicol[11], Christophe Leterrier[9], Rafael Lujan[12], Shiva K. Tyagarajan[6], Rodrigo A. Cunha[4,7], Monique Esclapez[2], Christophe Bernard[2]*, Sabine Lévi[1]*

During development, neural circuit formation requires the stabilization of active γ-aminobutyric acid–mediated (GABAergic) synapses and the elimination of inactive ones. Here, we demonstrate that, although the activation of postsynaptic GABA type A receptors ($GABA_AR$s) and adenosine $A_{2A}$ receptors ($A_{2A}$Rs) stabilizes GABAergic synapses, only $A_{2A}$R activation is sufficient. Both $GABA_AR$- and $A_{2A}$R-dependent signaling pathways act synergistically to produce adenosine 3′,5′-monophosphate through the recruitment of the calcium–calmodulin–adenylyl cyclase pathway. Protein kinase A, thus activated, phosphorylates gephyrin on serine residue 303, which is required for $GABA_AR$ stabilization. Finally, the stabilization of pre- and postsynaptic GABAergic elements involves the interaction between gephyrin and the synaptogenic membrane protein Slitrk3. We propose that $A_{2A}$Rs act as detectors of active GABAergic synapses releasing GABA, adenosine triphosphate, and adenosine to regulate their fate toward stabilization or elimination.

During development, brain circuits go through different phases of synapse formation, stabilization, and elimination, which involve numerous molecular mechanisms, in particular at γ-aminobutyric acid–mediated (GABAergic) synapses. Synaptic cell adhesion molecules are essential for synapse formation and maturation, including neuroligins and leucine-rich repeat transmembrane proteins (LRRTMs) that interact with presynaptic neurexins, Slit- and Trk-like family proteins (Slitrks), which bind to presynaptic protein tyrosine phosphatases (PTPs), immunoglobulin superfamily proteins (IgSFs), cadherin family proteins, and transmembrane tyrosine kinase receptors (*1*–*4*). The γ-aminobutyric acid (GABA) neurotransmitter itself regulates the maturation and innervation patterns of GABAergic synapses as well as their elimination and pruning (*4*–*8*). GABA operates through the activation of GABA type A receptors ($GABA_AR$s) and the elevation in intraneuronal calcium $Ca^{2+}$ levels after the activation of voltage-dependent $Ca^{2+}$ channels (*9*, *10*). Additionally, the adenosine signaling pathway is also involved during development: Extracellular adenosine builds up with synaptic activity (*11*, *12*), and adenosine $A_{2A}$ receptors ($A_{2A}$Rs) control the migration speed of GABAergic neurons (*13*), axonal elongation and dendritic branching (*14*), and synapse stabilization or elimination at the neuromuscular junction (*15*). Because GABA could be coreleased with adenosine triphosphate (ATP) and adenosine in a synaptic activity–dependent manner (*16*, *17*), we tested the role of adenosine signaling on GABAergic synaptogenesis in the brain.

[1]INSERM UMR-S 1270, Sorbonne Université, Institut du Fer à Moulin, Paris, France. [2]Aix Marseille Université, INSERM, INS, Institut de Neurosciences des Systèmes, Marseille, France. [3]Department of Cell and Systems Biology, University of Toronto, Toronto, ON M5S 3G5, Canada. [4]CNC-Center for Neuroscience and Cell Biology, University of Coimbra, 3004-504 Coimbra, Portugal. [5]Center for Interdisciplinary Research in Biology, College de France, INSERM U1050, CNRS UMR7241, Labex Memolife, Paris, France. [6]Institute of Pharmacology and Toxicology, University of Zürich, 8057 Zürich, Switzerland. [7]Faculty of Medicine, University of Coimbra, 3004-504 Coimbra, Portugal. [8]Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal. [9]Aix Marseille Université, CNRS, INP UMR7051, NeuroCyto, Marseille, France. [10]Université Bordeaux, CNRS, Interdisciplinary Institute for Neuroscience, IINS, UMR 5297, Bordeaux, France. [11]Sorbonne Université, Inserm, CNRS, Institut de la Vision, Paris, France. [12]Synaptic Structure Laboratory, Instituto de Investigación en Discapacidades Neurológicas (IDINE), Departamento Ciencias Médicas, Facultad de Medicina, Universidad Castilla-La Mancha, Campus Biosanitario, 02008 Albacete, Spain.
***Corresponding author. Email: sabine.levi@inserm.fr (S.L.); christophe.bernard@univ-amu.fr (C.B.)**
†These authors contributed equally to this work.

## Results

### $A_{2A}$Rs are transiently expressed at developing GABAergic synapses

In the adult hippocampus, there is a low density of $A_{2A}$Rs, which are essentially presynaptic (*18*). During the peak of synaptogenesis, between postnatal days P5 and P16, we found a transient increased density of $A_{2A}$Rs in mouse hippocampi, in particular in purified synaptic contacts (fig. S1). Electron microscopy showed their post- and perisynaptic localizations on symmetric, presumably GABAergic, synapses (fig. S1). We also confirmed the presence of $A_{2A}$Rs in primary cultures of hippocampal neurons. $A_{2A}$Rs were clustered at synapses containing glutamic acid decarboxylase 67 kD [(GAD67) a GABA synthesizing enzyme], and clustering increased during synaptogenesis between 7 and 14 days in vitro (DIV) (fig. S1). $A_{2A}$Rs were not present at all GABAergic synapses, but they accumulated at a subset (39.1 ± 3.7%, $n$ = 34 cells, three independent experiments) of synapses. DNA points accumulation for imaging in nanoscale topography (DNA-PAINT) further confirmed that $A_{2A}$Rs were located near the GABAergic postsynapse identified by the presence of the scaffolding molecule gephyrin at DIV 10 (fig. S1). We then asked whether such enrichment of post- and perisynaptic $A_{2A}$Rs at GABAergic synapses was accompanied by an increase in activity-dependent release of their ligand adenosine.

### ATP and adenosine release is increased during synaptogenesis

Adenosine can originate from its direct activity-dependent release by presynaptic terminals or neighboring glial cells and/or from the conversion of ATP released by neurons or glial cells through the ecto-5′-nucleotidase CD73 (*11*, *12*). The evoked release of adenosine was larger from P7 than P60 (adult) hippocampal synaptosomes (fig. S2), and blocking CD73 with adenosine 5′-(α,β-methylene)diphosphate (AMPCP) (100 μM) decreased extracellular adenosine by 25% (fig. S2). Accordingly, we found a large density of CD73 in synapses during the peak of synaptogenesis (fig. S2), in keeping with the tight association of the enzyme with $A_{2A}$Rs (*19*). Therefore, one fraction of extracellular adenosine comes from local extracellular ATP metabolism, and most of the adenosine is likely released as such through nonconcentrative nucleoside transporters upon its intracellular formation as a by-product of metabolic activity sustaining synaptic activity (*11*, *12*, *20*). The activity-dependent secretion of ATP and adenosine was also larger at P7 than at P60 (fig. S2). Thus, the activation of synaptic terminals at the early stage of synaptogenesis bolsters the release of adenosine (via yet-unclear mechanisms) and of vesicular ATP, which is efficiently converted into extracellular adenosine by CD73. Given the presence of $A_{2A}$Rs at inhibitory synapses and the activity-dependent production of its ligand adenosine during the period of synaptogenesis, we hypothesized a role for this pathway in GABAergic synapse formation and elimination. Because GABA controls the number of GABAergic synapses during development (*4*–*8*), we tested the relative contribution of $GABA_AR$ and $A_{2A}$R pathways.

### $A_{2A}$Rs and $GABA_AR$s control GABAergic synapse fate

Incubating neurons with tetanus toxin (TeNT) (1 to 40 nM) in vitro, which abolishes vesicular release of neurotransmitters (*21*), resulted in

the disappearance of 43% of GABAergic synaptic terminals labeled for the vesicular GABA transporter (VGAT) and 62% of postsynaptic $GABA_AR$ γ2 subunit clusters, which indicates synapse dismantling (fig. S3). Synaptic loss was fully prevented by the activation of $A_{2A}Rs$ with CGS21680 (30 nM) or $GABA_ARs$ with muscimol (10 μM) for 30 min (fig. S3). The rescue of GABAergic synapses occurred within 30 min of $A_{2A}R$ or $GABA_AR$ activation, ruling out de novo synapse formation, which requires hours (*22*). Together, these results show that, in the absence of synaptic transmission, the activation of either $A_{2A}Rs$ or $GABA_ARs$ by their respective ligands is sufficient to maintain GABAergic synapses that would have otherwise disappeared. This led us to hypothesize that both adenosine and GABA signaling pathways play a role in activity-dependent synapse stabilization and that when presynaptic sites are inactive, such synapses are eliminated. Because the activation of either receptor could rescue GABAergic synapses to an extent similar to that observed when both receptors were simultaneously activated (fig. S3), we further hypothesized a convergence of both $A_{2A}R$ and $GABA_AR$ signaling pathways.

### *$A_{2A}Rs$ are sufficient to stabilize GABAergic synapses*

To test the interdependence of the two $A_{2A}R$ and $GABA_AR$ pathways, we blocked one while activating the other, and we investigated the fate of GABAergic synapses in vitro. With specific short hairpin RNAs (shRNAs), we knocked down the $GABA_AR$ γ2 subunit—which leads to synaptic loss (*23*)—or $A_{2A}Rs$ (*24*), and we activated $A_{2A}Rs$ or $GABA_ARs$ with their selective agonists CGS21680 (30 nM) or muscimol (10 μM), respectively. Neurons transfected at DIV 6 to DIV 7 with shRNAs against $GABA_ARγ2$ (shγ2) or $A_{2A}Rs$ ($shA_{2A}R$) showed, respectively, a 51 and 43% reduction of GABAergic synapses (Fig. 1, A and B). Because <10 out of 100,000 cells were transfected and because we did not find any evidence of synapse destabilization in the nontransfected cells surrounding the transfected ones, we conclude that the destabilization of GABAergic synapses occurs as a result of the loss of postsynaptic $GABA_ARs$ or $A_{2A}Rs$. The activation of $A_{2A}Rs$ with CGS21680 rescued GABAergic synapses in neurons lacking $GABA_ARγ2$ (Fig. 1A). By contrast, activation of $GABA_ARs$ by muscimol failed to rescue GABAergic synapses in $shA_{2A}R$-expressing neurons (Fig. 1B). Thus, $A_{2A}R$ activation appears to be necessary and sufficient for GABAergic synapse stabilization, whereas $GABA_AR$ activation is not necessary as long as $A_{2A}Rs$ remain activated. According to this scheme, blocking $A_{2A}Rs$ should result in synapse loss.

### *$A_{2A}R$ blockade triggers GABAergic synapse loss*

A 30-min application of the selective $A_{2A}R$ antagonist SCH58261 (100 nM) in vitro decreased the number of clusters of the presynaptic protein VGAT as well as postsynaptic $GABA_ARγ2$ and scaffolding protein gephyrin (fig. S4), demonstrating a destabilization of GABAergic synapses at both pre- and postsynaptic sites. Acute treatment with SCH58261 induced a loss of GABAergic synapses expressing the $GABA_AR$ γ2 subunit but also of synapses containing the α1 or α2 subunits (fig. S4), which suggests that several types of GABAergic synapses are controlled by $A_{2A}Rs$. Not all synapses disappeared, as ~30% of GABAergic synapses were destabilized (fig. S4), in keeping with the finding that not all GABAergic synapses are equipped with $A_{2A}Rs$ at any given developmental time (fig. S1). The fact that the loss of GABAergic synapses induced by $A_{2A}R$ blockade was similar to that found in $shA_{2A}R$-expressing neurons (Fig. 1B) suggests that synaptic loss was not a result of an indirect effect on network activity. In keeping with the reduced synaptic clustering of $GABA_ARγ2$, quantum dot–based single-particle tracking of $GABA_ARγ2$ in DIV 8 hippocampal

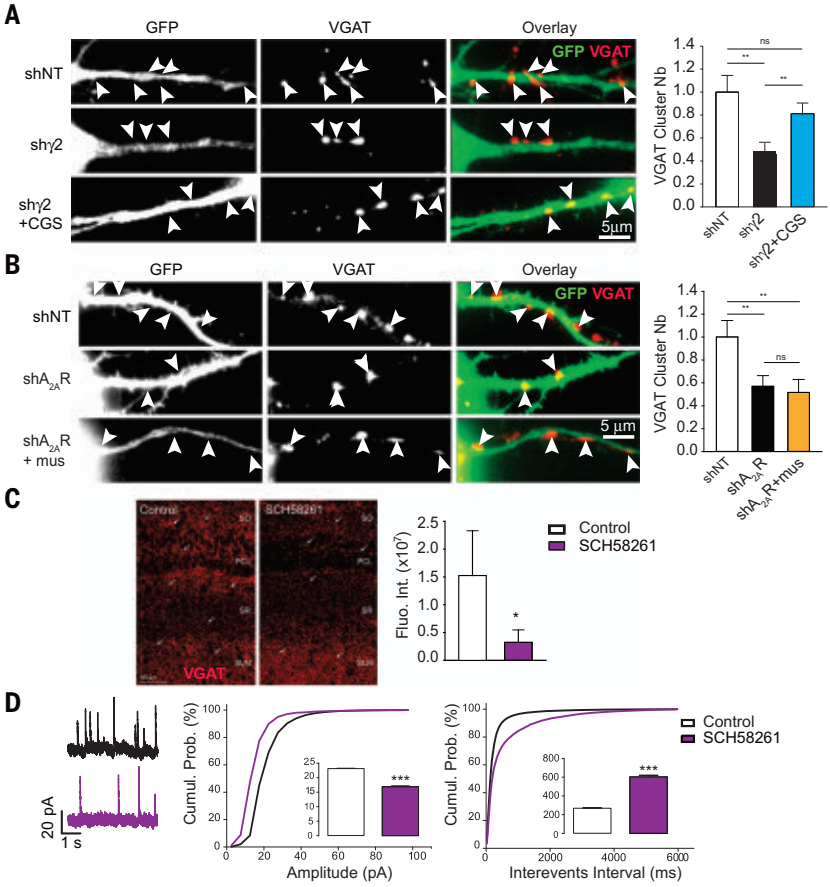



**Fig. 1. $A_{2A}R$ activation stabilizes GABAergic synapses.** (**A**) VGAT staining (left) and quantification (right) in DIV 10 to 11 neurons transfected with nontarget (shNT) or on-target $GABA_ARγ2$ (shγ2) shRNAs exposed or not exposed to CGS21680 (CGS) (30 nM) for 30 min. shNT, *n* = 60; shγ2, *n* = 52; shγ2 and CGS21680, *n* = 54; four cultures. The loss of GABAergic synapses upon suppression of $GABA_ARγ2$ was rescued by activation of $A_{2A}Rs$. Cluster Nb, cluster number. (**B**) VGAT staining (left) and quantification (right) in DIV 10 to 11 neurons transfected with nontarget (shNT) or on-target $A_{2A}R$ ($shA_{2A}R$) shRNAs exposed or not exposed to muscimol (mus) (10 μM) for 30 min. shNT, *n* = 44; $shA_{2A}R$, *n* = 47; $shA_{2A}R$ and muscimol, *n* = 36; four cultures. The loss of GABAergic synapses upon suppression of $A_{2A}R$ was not rescued by $GABA_AR$ activation. Scale bar, 5 μm. Arrowheads in (A) and (B) show examples of inhibitory synapses labeled for VGAT. (**C**) Acute application of SCH58261 (100 nM) for 30 min in hippocampal slices decreased the density of VGAT-immunoreactive terminals by 47% (eight slices, three P6 pups) compared with controls (six slices, three P6 pups). Fluo. Int., fluorescence intensity. (**D**) Application of SCH58261 (100 nM) for 30 min reduced the amplitude and frequency of $GABA_AR$-mediated mIPSCs in CA1 pyramidal cells (eight cells, eight slices, six P6 pups). (Left) Examples of mIPSCs recorded before and after application of SCH58261. (Insets) IPSC amplitude or interevent interval. Cumul. Prob., cumulative probability. Values were normalized to the control values [(A) and (B)]; histograms represent means and SEMs. Statistics were calculated using the Mann-Whitney test [(A), (B), and (C)] and the Kolmogorov-Smirnov test (D). ns, not significant; $*P \le 0.05$; $**P < 0.01$; $***P < 0.001$.

neurons revealed that the membrane diffusion of $GABA_A$Rs increased within 10 min after bath application of SCH58261 (100 nM) (fig. S5). Finally, this adenosine-dependent mechanism seems specific to $A_{2A}$Rs because blocking adenosine $A_1$ receptors had no effect on the density of GABAergic synapses (fig. S6). We then set out to confirm these results ex vivo.

Immunohistochemistry revealed that blocking $A_{2A}$Rs for 30 min with SCH58261 (100 nM) in P6 hippocampal slices triggered a 47% loss of VGAT-containing synaptic terminals (Fig. 1C). Single-cell recordings of miniature inhibitory postsynaptic currents (mIPSCs) in hippocampal CA1 pyramidal cells at P6 showed that a 30-min application of SCH58261 (100 nM) decreased the frequency of mIPSCs by ~40% (Fig. 1D). The amplitude of mIPSCs was decreased by ~30% (Fig. 1D). We found the same effect on mIPSCs recorded in CA1 interneurons (fig. S7) and in layer V pyramidal cells of the V1 visual cortex (fig. S7). Because the effects of SCH58261 on GABAergic currents could be an indirect consequence of the modulation of glutamatergic neurotransmission by $A_{2A}$Rs, we verified that the effect was still present in the presence of glutamate receptor blockers (fig. S8). Together, these results show that $A_{2A}$Rs control the stability of GABAergic synapses in different cell types and brain regions at early stages of development. We then assessed whether this role of $A_{2A}$Rs was developmentally regulated and studied the minimum time of $A_{2A}$R inactivity necessary to trigger synapse removal.

### *$A_{2A}$R-mediated stabilization of GABAergic synapses is time dependent*

In keeping with the transient increased density of $A_{2A}$Rs during the first 2 postnatal weeks (fig. S1), we found that the control of GABAergic synapses by $A_{2A}$Rs ex vivo was restricted to the period of synaptogenesis (between P4 and P12), with a maximum effect at P6 (fig. S9). A morphological analysis performed in cultures confirmed that the action of $A_{2A}$Rs was also developmentally regulated (fig. S10).

We then assessed whether there is a specific time window beyond which $A_{2A}$R inactivity triggers an irreversible synapse destabilization. Varying the time of application of SCH58261, we found that the effect on mIPSCs was reversible if $A_{2A}$R blockade did not exceed 10 min (fig. S9). Beyond 20 min, the effect was irreversible (fig. S9). Similarly, there was no loss of GABAergic synapses 10 min after $A_{2A}$R blockade in cultures, whereas it occurred after 30 min of treatment (fig. S11). Thus, the $A_{2A}$R-dependent control of GABAergic synapse density is developmentally regulated, and a period of at least 20 min of inactivity of $A_{2A}$Rs is sufficient to trigger the removal of GABAergic synapses.

Together, our results show that $A_{2A}$Rs satisfy the three requirements of a system able to control the fate of synapses toward stabilization or elimination during development. It is time dependent, its activation maintains synapses, and the absence of activation results in synapse elimination. We then investigated the molecular mechanisms upstream and downstream of $A_{2A}$R activation.

### *Adenosine is required for synapse stabilization*

Application of the CD73 inhibitor AMPCP (100 μM) together with adenosine deaminase (ADA) (4 to 20 U/mL), which hydrolyses adenosine into inosine, decreased mIPSC amplitude (30%) and frequency (74%) ex vivo (Fig. 2A). The same treatment reduced the time $GABA_A$Rγ2 spent at inhibitory synapses in vitro—an effect that was prevented by the direct activation of $A_{2A}$Rs with their agonist CGS21680 (fig. S5). This increased escape of $GABA_A$Rγ2 from inhibitory synapses upon the removal of ambient adenosine was accompanied by a rapid disappearance of active inhibitory synaptic boutons (labeled by preloading of a VGAT-oyster[550] antibody into synaptic vesicles during multiple vesicular exocytosis-endocytosis cycles) and increased trafficking of vesicle packets (*22*) into axons (fig. S11), probably resulting from the destabilization of presynaptic active zones. This rapid dismantling of GABAergic pre- and postsynaptic compartments induced by the withdrawal of ambient adenosine with ADA and AMPCP resulted in a net loss of the number of GABAergic synapses that could be prevented by $A_{2A}$R activation in vitro (Fig. 2B).

To test whether glial cells are necessary to produce the adenosine required to activate $A_{2A}$Rs, we pretreated cultures with the gliotoxin cytosine arabinoside (araC), which resulted in the complete loss of glial fibrillary acidic protein (GFAP)–stained glial cells (fig. S12). The application of ADA and AMPCP still induced synapse loss, which was prevented by CGS21680 (fig. S12). This indicates that glial

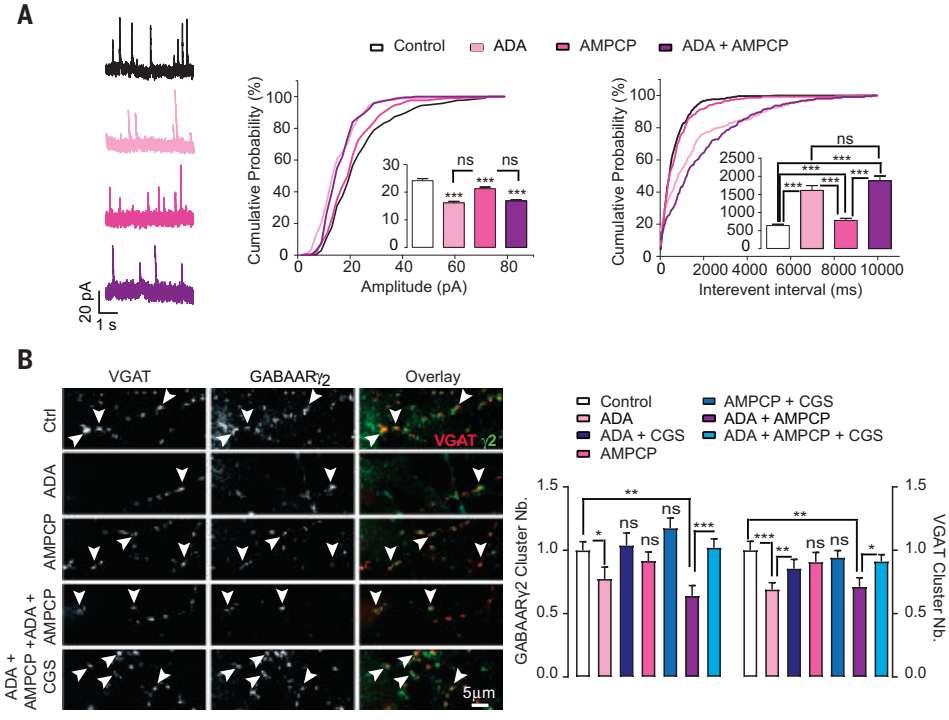



**Fig. 2. Synapse stabilization requires extracellular adenosine production.**
(**A**) Decreasing extracellular adenosine levels with a 30-min treatment of ADA (4 to 20 U/mL), AMPCP (100 μM), or ADA and AMPCP led to a decrease of mIPSCs' amplitude and frequency in CA1 pyramidal cells. *N* = 12 cells, 12 slices, 6 pups. (**B**) Immunostaining and quantification of VGAT and $GABA_A$Rγ2 in DIV 10 neurons in the absence or presence of the indicated drugs for 30 min. Arrowheads show examples of inhibitory synapses labeled for VGAT and $GABA_A$Rγ2. Scale bar, 5 μm. Decreasing extracellular adenosine levels lead to synapse destabilization, an effect prevented by the direct activation of $A_{2A}$Rs with 30 nM CGS21680. *N* = 26 to 41 cells, three cultures. Ctrl, control. Histograms represent means and SEMs. Statistics were calculated using the Kolmogorov-Smirnov test (A) and the Mann-Whitney test (B). ns, not significant; $*P \leq 0.05$; $**P < 0.01$; $***P < 0.001$.

cells are not the main source of adenosine and that neurons can produce adenosine (directly or indirectly as a by-product of its phosphorylated forms) in the absence of glial cells in a sufficient quantity to activate $A_{2A}$Rs.

Altogether, these results show that direct adenosine activation of $A_{2A}$Rs is necessary and sufficient to maintain synapse integrity. We then looked at the downstream effects of $A_{2A}$R activation, starting with the possible cooperativity between GABAergic and $A_{2A}$R signaling pathways, as suggested above (fig. S3).

### *$A_{2A}$Rs and GABA$_A$Rs converge to stabilize synapses*

#### *$A_{2A}$Rs stabilize GABAergic synapses via the AC-cAMP pathway*

$A_{2A}$Rs are G protein–coupled receptors, and their prime transducing system is the activation of adenylyl cyclases (ACs), which generate adenosine 3′,5′-monophosphate (cAMP). If synapse stabilization is G protein–dependent, blocking G protein activity should destabilize synapses. Recording CA1 pyramidal cells with an intracellular solution containing guanosine diphosphate-β-S to block G protein activity directly led to a decrease of mIPSC frequency and amplitude ex vivo (~20 and 15%, respectively; fig. S13), reproducing the effect of $A_{2A}$R blockade.

We then tested the involvement of the AC-cAMP signaling cascade, which is downstream of $A_{2A}$R activation. Bath application of 3-isobutyl-1-methylxanthine (IBMX) (100 μM, a nonselective inhibitor of phosphodiesterases that metabolizes cAMP) and forskolin (10 μM, an activator of ACs) prevented the SCH58261-induced decrease of mIPSC frequency and amplitude in CA1 pyramidal cells ex vivo (fig. S13). Similarly, IBMX and forskolin prevented the loss of GABAergic synapses induced by SCH58261 in hippocampal neuron cultures (fig. S13). Therefore, the $A_{2A}$R-dependent control of GABAergic synapses occurs through the activation of the AC-cAMP signaling cascade.

#### *$A_{2A}$Rs and GABA$_A$Rs elevate intracellular cAMP levels*

At what level do the $A_{2A}$R and GABA$_A$R signaling pathways converge to regulate synaptogenesis? $A_{2A}$Rs and GABA$_A$Rs are linked to AC-cAMP and $Ca^{2+}$ signaling, respectively. During development, the activation of GABA$_A$Rs leads to a rise in intracellular $Ca^{2+}$ concentration ($[Ca^{2+}]_i$) (*9*, *10*). Because hippocampal neurons express type 1 and type 8 $Ca^{2+}$–calmodulin (CaM)–activated ACs (*25*), we hypothesized that GABA signaling converges on $A_{2A}$R signaling by tuning CaM-stimulated ACs and their downstream effectors.

Acute activation of GABA$_A$Rs with 10 μM muscimol significantly increased $[Ca^{2+}]_i$ in DIV 8 hippocampal neurons, whereas it was decreased in DIV 14 neurons (fig. S14), in keeping with the developmentally regulated effect of GABA$_A$Rs on $[Ca^{2+}]_i$. By contrast, an acute application of CGS21680 did not change $[Ca^{2+}]_i$ in DIV 8 or DIV 14 neurons (fig. S14). Buffering intracellular $Ca^{2+}$ with BAPTA-AM (20 μM) lead to the destabilization of inhibitory synapses—an effect that could be prevented by $A_{2A}$R activation with CGS21680 (Fig. 3A). Thus, activation of $A_{2A}$Rs is sufficient to maintain synapse integrity in the absence of intracellular $Ca^{2+}$ signaling.

Because GABA but not adenosine signaling can elevate $[Ca^{2+}]_i$ and because $A_{2A}$R-mediated activation of ACs can be boosted by CaM (*10*), we tested the convergence of the GABA$_A$R and

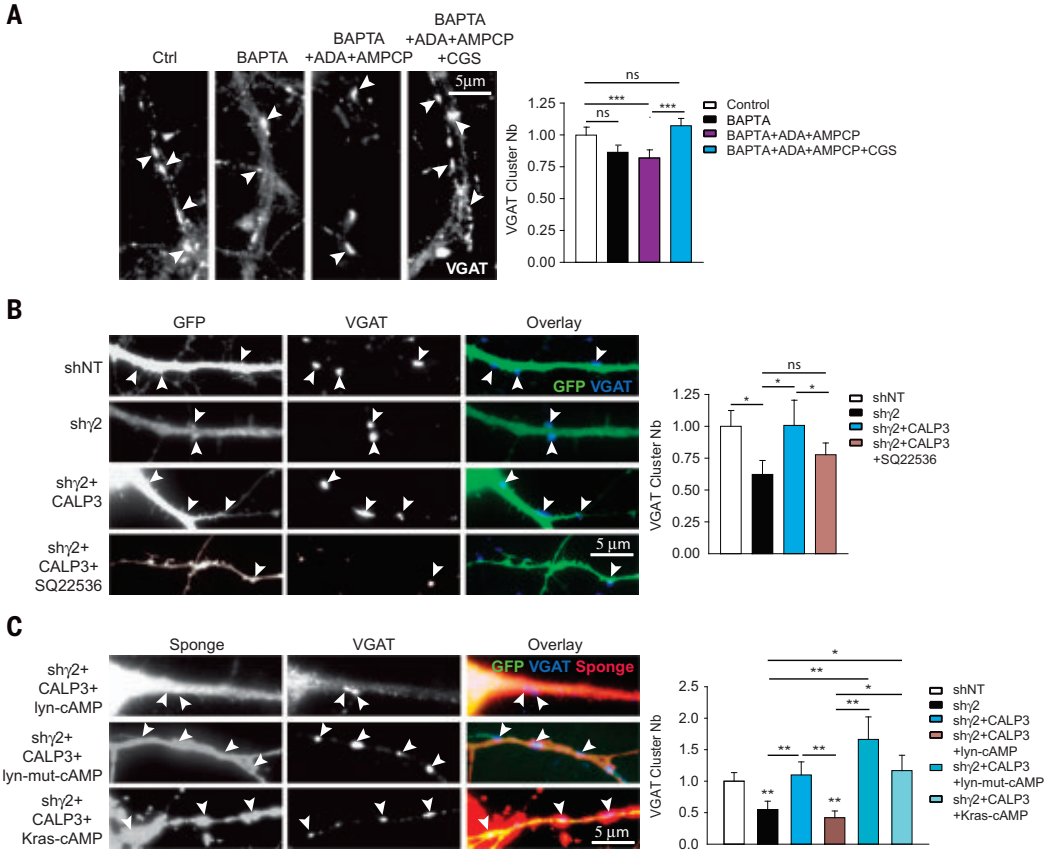



**Fig. 3. GABAergic synapse stabilization by calmodulin-activated, calcium-sensitive ACs.** (**A**) Immunostaining and quantification of VGAT in DIV 10 neurons in the absence or presence of the indicated drugs for 30 min. Scale bar, 5 μm. Decreasing intracellular calcium levels with BAPTA-AM (20 μM) in the absence or presence of ADA (4 to 20 U/mL) and AMPCP (100 μM) leads to synapse destabilization, an effect prevented by the direct activation of $A_{2A}$Rs with CGS21680 (30 nM). *N* = 35 to 38 cells, three cultures. (**B**) Calp3 (100 μM) activation of calmodulin for 30 min rescued the loss of GABAergic synapses in DIV 10 to 11 neurons expressing GABA$_A$Rγ2 shRNA (shγ2). The effect of calmodulin was blocked by an adenylyl cyclase inhibitor SQ22536 (20 μM). Scale bar, 5 μm. shNT, *n* = 38; shγ2, *n* = 42; shγ2 and CALP3, *n* = 43; shγ2 and CALP3 and SQ22536, *n* = 30; three cultures. (**C**) VGAT staining and quantification of DIV 10 to 11 neurons transfected with nontarget (shNT) or on-target GABA$_A$Rγ2 (shγ2) and/or the different sponges constructs (lyn-cAMP, lyn-mut-cAMP, and Kras-cAMP) exposed to CALP3 for 30 min. Scale bar, 5 μm. shNT, *n* = 40; shγ2, *n* = 37; shγ2 and CALP3, *n* = 40; shγ2 and CALP3 and lyn-cAMP sponge, *n* = 23; shγ2 and CALP3 and lyn-mut-cAMP sponge, *n* = 27; shγ2 and CALP3 and Kras-cAMP sponge, *n* = 31; four cultures. Arrowheads show examples of inhibitory synapses labeled for VGAT. In all graphs, histograms represent means and SEMs; values normalized to their respective controls. Statistics were calculated using the Mann-Whitney test. ns, not significant; *$P \leq 0.05$; **$P < 0.01$; ***$P < 0.001$.

$A_{2A}R$ pathways on cAMP. To address this question, we performed cAMP imaging in DIV 8 hippocampal neurons infected with recombinant Sindbis virus encoding EPAC-SH150, a sensor with increased sensitivity for cAMP (*26*). As expected, $A_{2A}R$ activation with CGS21680 increased intracellular cAMP levels (fig. S15). $GABA_AR$ activation with muscimol also increased cAMP levels within a similar range to that observed with CGS21680 (fig. S15). Application of both CGS21680 and muscimol strongly elevated intracellular cAMP levels (fig. S15), demonstrating that the activation of $GABA_ARs$ and $A_{2A}Rs$ converge with additive effects on cAMP production.

### *Stabilization of GABAergic synapses requires cAMP production*

We then tested the implication of intracellular $Ca^{2+}$ and CaM in GABAergic synapse stabilization. The loss of GABAergic synapses induced by the expression of shRNA against the $GABA_AR$ γ2 subunit (shγ2) in DIV 10 neurons was prevented by the activation of CaM with the cell-permeable activator CALP3 (100 μM) for 30 min (Fig. 3B). Thus, CaM is involved in the stabilization of GABAergic synapses. Application of the AC inhibitor SQ22536 (20 μM) prevented the CALP3-mediated rescue of GABAergic synapses (Fig. 3B), which demonstrates that CaM stabilizes synapses through the activation of ACs.

Hippocampal neurons express $Ca^{2+}$-dependent and -independent ACs targeted, respectively, to lipid raft and nonraft plasma membranes (*27*). To investigate the involvement of $Ca^{2+}$-dependent and -independent ACs, we used genetically encoded cAMP sponges targeted to (lyn-cAMP sponge) or outside (Kras-cAMP sponge) lipid rafts, which enabled the local perturbation of cAMP downstream signaling (*28*). A variant of lyn-cAMP sponge (lyn-mut-cAMP sponge) unable to bind and buffer cAMP was used as a control. In basal conditions, none of the cAMP sponges altered the density of GABAergic synapses (fig. S16). However, buffering cAMP near lipid rafts blocked the CALP3-mediated rescue of GABAergic synapses in neurons expressing shγ2 (Fig. 3C). This effect was specific to sponges targeted to lipid rafts because the density of synapses in neurons expressing lyn-mut-cAMP or Kras-cAMP sponges were not distinguishable from CALP3-treated neurons lacking $GABA_ARγ2$ (Fig. 3C). Thus, the convergence of $GABA_AR$ and $A_{2A}R$ pathways on the stabilization of GABAergic synapses relies on cAMP production by $Ca^{2+}$-dependent, CaM-activated ACs. We then investigated the mechanism downstream of cAMP production.

### *Adenosine stabilizes synapses through PKA-phosphoregulation of gephyrin*

Protein kinase A (PKA) is a downstream effector of cAMP after $A_{2A}R$ activation. As expected, intracellular inhibition of PKA with the protein kinase inhibitor peptide (PKI) decreased mIPSC amplitude and frequency ex vivo (26 and 50%, respectively; fig. S17) to the same extent as that found with SCH58261 (100 nM). Because PKA-mediated phosphorylation of gephyrin is required for its postsynaptic stabilization and the anchoring of $GABA_ARs$ at synapses (*29*), we hypothesized that PKA activation via the $A_{2A}R$–AC-cAMP cascade ensures the phosphorylation of gephyrin required to maintain $GABA_ARs$ at the synapse.

Gephyrin is a direct substrate of PKA, so we tested whether gephyrin phosphorylation at the $Ser^{303}$ PKA-sensitive site was increased after $A_{2A}R$ activation by CGS21680. For this, we used HEK293 cells and measured gephyrin phosphorylation using $Ser^{303}$ site–specific gephyrin phospho-antibody (*30*). Removal of ambient adenosine with ADA and AMPCP and simultaneous activation of $A_{2A}Rs$ with CGS21680 significantly increased gephyrin phosphorylation at the $Ser^{303}$ site (Fig. 4A). These effects of ADA and AMPCP or ADA, AMPCP, and CGS21680 were not observed in cells that expressed a gephyrin phospho-null mutant (gephyrin-S303A) (Fig. 4A). This demonstrates that gephyrin can be phosphorylated on its specific PKA site upon $A_{2A}R$ activation.

We then directly tested the contribution of PKA-dependent phosphorylation of gephyrin in synapse stabilization in vitro by expressing

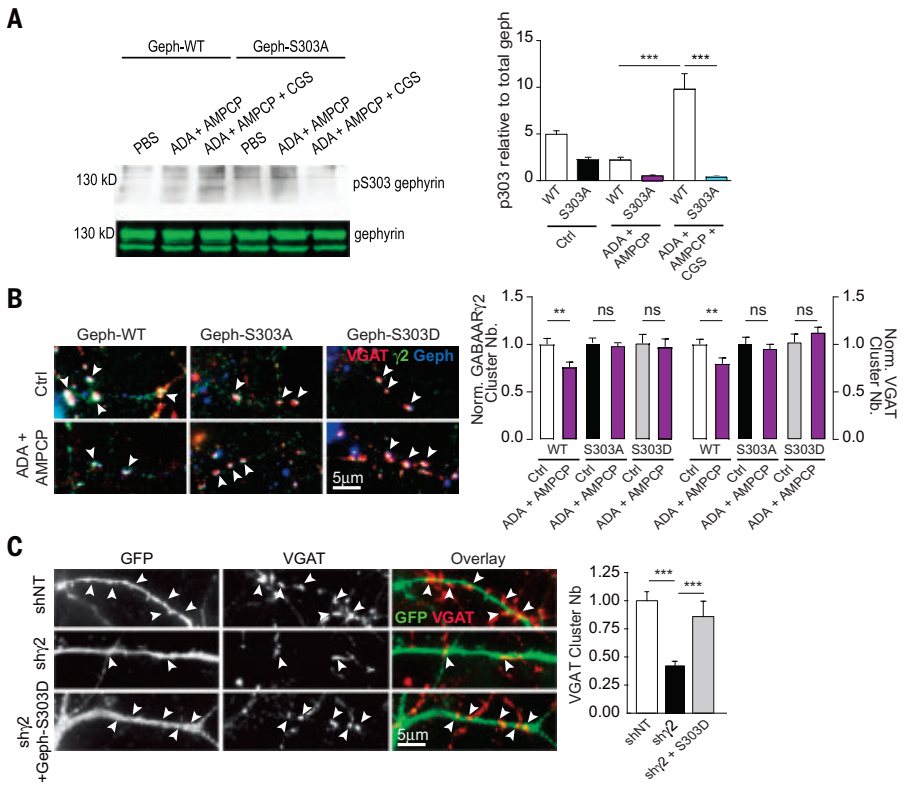



**Fig. 4. GABAergic synapse stabilization requires phosphorylation of gephyrin on $Ser^{303}$.** (**A**) Western blot and quantification of the gephyrin $Ser^{303}$ phosphorylation in control and after 30-min exposure to ADA (4 to 20 U/mL) and AMPCP (100 μM) without or with CGS21680 (30 nM) in HEK293 cells transfected with gephyrin-WT or gephyrin-S303A constructs. There was significant increased phosphorylation of P-S303 of gephyrin-WT but not gephyrin-S303A upon acute $A_{2A}R$ activation with CGS21680. PBS, phosphate-buffered saline. (**B**) Immunostaining (left) and quantification (right) of $GABA_ARγ2$ and VGAT in the absence or presence of ADA and AMPCP in DIV 10 neurons transfected with 3′ untranslated region (3′UTR) gephyrin shRNA and eGFP-gephyrin-WT, S303A, or S303D constructs. Scale bar, 5 μm. Arrowheads show examples of inhibitory synapses triple-stained for VGAT, $GABA_ARγ2$, and gephyrin. WT, $n$ = 53; WT and ADA and AMPCP, $n$ = 49; S303A, $n$ = 48; S303A and ADA and AMPCP, $n$ = 53; S303D, $n$ = 45; S303D and ADA and AMPCP, $n$ = 47; four cultures. Gephyrin-S303A or S303D block the ADA and AMPCP effect. (**C**) VGAT staining (left) and quantification (right) of DIV 10 to 11 neurons transfected with nontarget (shNT) or on-target $GABA_ARγ2$ (shγ2) shRNAs with eGFP-gephyrin-WT or eGFP-gephyrin-S303D. Scale bar, 5 μm. Arrowheads show examples of inhibitory synapses labeled for VGAT. shNT, $n$ = 35; shγ2, $n$ = 49; shγ2 and geph-S303D, $n$ = 60; three cultures. The loss of GABAergic synapses upon suppression of $GABA_ARγ2$ can be rescued upon overexpression of gephyrin-S303D. In all graphs, histograms represent means and SEMs; values were normalized to corresponding controls [(B) and (C)]. Statistics were calculated using the analysis of variance (ANOVA) test (A) and the Mann-Whitney test [(B) and (C)]. ns, not significant; $^{**}P < 0.01$; $^{***}P < 0.001$.

constructs harboring mutations of the PKA $Ser^{303}$ site to aspartate (S303D) or alanine (S303A) to mimic the gephyrin phosphorylated and dephosphorylated states (*30*). Decreasing extracellular adenosine with AMPCP and ADA led to synapse destabilization in neurons expressing wild-type (WT) gephyrin (Fig. 4B), whereas the expression of the gephyrin-S303D or gephyrin-S303A mutants prevented synapse loss upon the removal of extracellular adenosine (Fig. 4B). Thus, the PKA $Ser^{303}$ phosphosite of gephyrin is required for the $A_{2A}R$-mediated synapse stabilization. This effect was specific to this phospho site because decreasing extracellular adenosine led to synapse destabilization in neurons expressing the glycogen synthase kinase 3β (GSK3β) phospho-null gephyrin mutant (gephyrin-S270A) (fig. S18). We conclude that activation of $A_{2A}Rs$ leads to GABAergic synapse stabilization through CaM-AC-cAMP–PKA–mediated phosphorylation of gephyrin at position $Ser^{303}$. This mechanism accounts for the disappearance of the postsynaptic site when $A_{2A}Rs$ are not activated by adenosine for >10 min (fig. S11).

### *Gephyrin phosphorylated on $Ser^{303}$ interacts with Slitrk3*

How can we explain the concurrent loss of the presynaptic element? Overexpression of the PKA-phosphorylation site gephyrin-S303D mutant was sufficient to rescue the loss of VGAT-containing terminals on neurons transfected with shγ2 (Fig. 4C). This demonstrates that gephyrin is sufficient to stabilize synapses. We therefore hypothesized that gephyrin phosphorylated at $Ser^{303}$ interacts with the postsynaptic neuroligin-2 (*31*, *32*) or Slitrk3 (*33*) transmembrane proteins that bind in the synaptic cleft to presynaptic neurexins or PTPδ to organize inhibitory synapses. The knockdown of postsynaptic neuroligin-2 with a specific shRNA or the antibody-induced blockade of presynaptic neurexin led, as expected (*34*–*36*), to a loss of GABAergic synapses, which could be rescued upon $A_{2A}R$ activation with CGS21680 (fig. S19). Thus, $A_{2A}R$-dependent synapse stabilization does not depend upon the neurexin–neuroligin-2 transsynaptic complex. By contrast, $GABA_AR$ activation did not rescue the synapse loss induced by the expression of a shRNA against neuroligin-2 (fig. S19).

$A_{2A}R$ or $GABA_AR$ activation did not prevent synaptic loss induced by a shRNA against postsynaptic Slitrk3 (*33*) (Fig. 5A), which supports the involvement of Slitrk3 in $A_{2A}R$- and $GABA_AR$-dependent synapse stabilization. We then overexpressed the Slitrk3-Y969A mutant because this site is involved in the formation of GABAergic synapses (*37*) (Fig. 5B). $A_{2A}R$ activation was not able to restore synapses in cells overexpressing Slitrk3-Y969A (Fig. 5B). Therefore, deletion of Slitrk3 or a single amino acid mutation of Slitrk3 was enough to prevent $A_{2A}R$-dependent stabilization of GABAergic synapses. Finally, using coimmunoprecipitation experiments, we found that gephyrin-WT and gephyrin-S303D interact directly with Slitrk3 (Fig. 5C). These results suggest that $A_{2A}R$ signaling stabilizes the pre- and postsynaptic GABAergic elements through PKA-dependent gephyrin phosphorylation and Slitrk3 recruitment.

### *Transient blockade of $A_{2A}Rs$ induces cognitive deficits*

All previous experiments were performed in vitro and ex vivo to unravel the molecular mechanism. We next tested whether the role of $A_{2A}Rs$ could also be observed in vivo during the postnatal synaptogenesis period. We could not use $A_{2A}R$ knockout animals because a delayed migration of GABAergic neurons occurs in these animals (*13*), which would result in a decreased number of GABAergic synapses. This would prevent us from distinguishing between effects resulting from a delayed migration and those resulting from a direct effect on the stability of GABAergic synapses. We therefore injected shRNAs against $A_{2A}Rs$ in vivo in the hippocampus at P3. This led to the loss of GABAergic synapses evaluated at P16 (fig. S20), which confirmed the physiological role of $A_{2A}Rs$ in the stabilization of GABAergic synapses. Additionally, we treated pups during the peak of synaptogenesis between P3

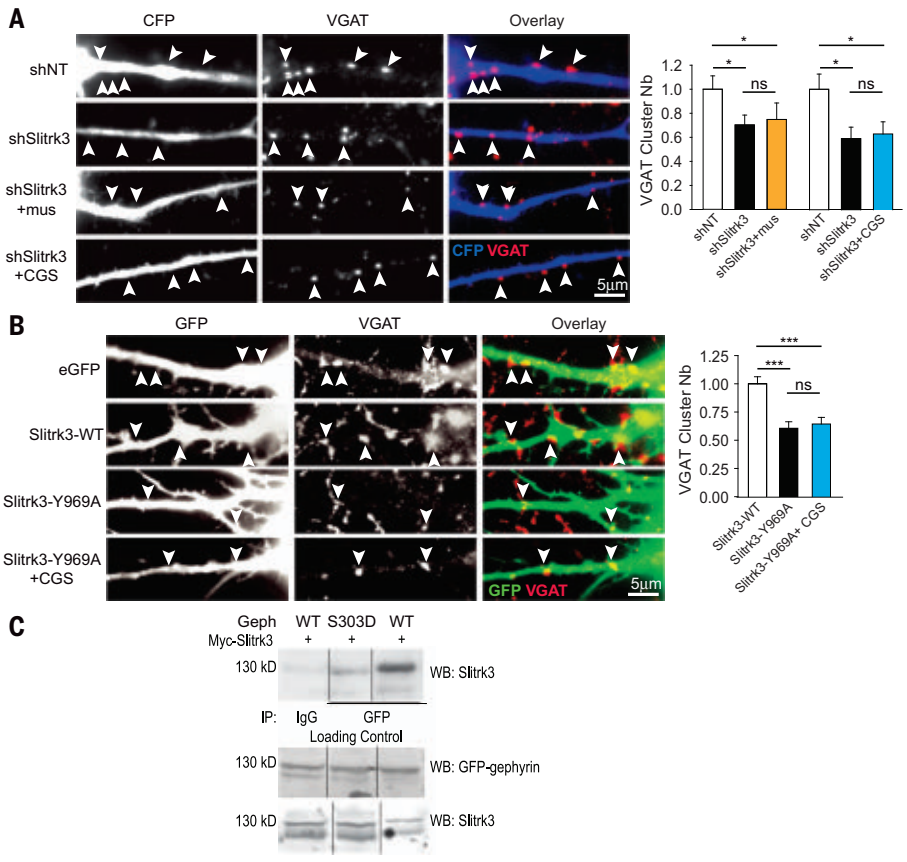



**Fig. 5. $A_{2A}Rs$ stabilize GABAergic synapses through the synaptogenic organizer Slitrk3.** (**A**) The loss of GABAergic synapses after Slitrk3 suppression cannot be rescued by $A_{2A}R$ or $GABA_AR$ activation. VGAT staining (left) and quantification (right) of DIV 10 to 11 neurons transfected with shMock or shSlitrk3 exposed or not exposed to muscimol (Mus) (10 µM) or CGS21680 (CGS) (30 nM) for 30 min. CFP, cyan fluorescent protein. Scale bar, 5 µm. Muscimol: shMock, *n* = 42; shSlitrk3, *n* = 34; shSlitrk3 and muscimol, *n* = 41; three cultures. CGS: shNT, *n* = 41; shSlitrk3, *n* = 37; shSlitrk3 and CGS, *n* = 36; three cultures. (**B**) The loss of GABAergic synapses after overexpression of the Slitrk3-Y969A mutant cannot be rescued by the activation of $A_{2A}Rs$. VGAT staining (left) and quantification (right) of DIV 10 to 11 neurons transfected with Slitrk3-WT or Slitrk3-Y969A exposed or not exposed to CGS21680 for 30 min. Scale bar, 5 µm. Muscimol: eGFP, *n* = 35; Slitrk3-WT, *n* = 38; Slitrk3-Y969A, *n* = 39; Slitrk3-Y969A and CGS, *n* = 41; two cultures. (**C**) Pull down of gephyrin-WT using the GFP tag followed by a blot for Slitrk3 showing gephyrin-Slitrk3 interaction in HEK293 cells. IP, immunoprecipitation; WB, Western blot; IgG, immunoglobulin G. In all images, arrowheads show examples of inhibitory synapses labeled for VGAT. In all graphs, histograms represent means and SEMs; values were normalized to corresponding controls. Statistics were calculated using the Mann-Whitney test [(A) and (B)]. ns, not significant; *$P < 0.05$; ***$P < 0.001$.

and P16 with SCH58261 (0.1 mg/kg). Although we cannot rule out nonspecific effects, the treatment produced a loss of GABAergic synapses at P16 in the hippocampus (fig. S20) similar to that found in vitro (fig. S4) and ex vivo (Fig. 1C). Given such synaptic loss, we predicted detrimental functional consequences, in particular for hippocampus-dependent spatial memory. We therefore tested P70 animals, which had been treated between P3 and P16 with SCH58261. As predicted, we found deficits in the novel object location task (fig. S20). By contrast, open field, anxiety, and novel object recognition tests were not modified (fig. S20). These results indicate that interfering with $A_{2A}$Rs in vivo during synaptogenesis has long-term deleterious effects on cognitive function.

## Discussion

This $A_{2A}$R-dependent mechanism adds to other molecules and signaling pathways known to control hippocampal GABAergic synapses during development (*1*, *4*, *38*), including GABA (*4*–*8*) and $GABA_A$R-induced elevation of intraneuronal $Ca^{2+}$ levels after the activation of voltage-dependent $Ca^{2+}$ channels (*9*, *10*). We propose that this in turn activates $Ca^{2+}$-CaM–sensitive ACs, leading to a rise in intracellular cAMP in neurons. Thus, one outcome of $GABA_A$R activation in immature neurons is the elevation of intracellular cAMP levels. A possible explanation is that adenosine and ATP are coreleased with GABA (*16*, *17*) during development, which would provide $A_{2A}$Rs with direct information that the presynaptic terminal is active. Once activated, G protein–coupled $A_{2A}$Rs will trigger a rise in intracellular cAMP levels, leading in turn to PKA activation and stabilization of GABAergic synapses (fig. S21). Such $A_{2A}$R-mediated control of synapse selection seems to be autonomous because it does not require brain-derived neurotrophic factor (BDNF)–TrkB receptor signaling (fig. S22).

We found additive effects of GABA and adenosine pathways on cAMP levels. ACs can act as coincidence detectors, promoting cellular responses only when convergent regulatory signals occur close in time and space (*39*). Our results show that the $GABA_A$R and $A_{2A}$R systems are spatially close and operate within a similar time frame to control the fate of some GABAergic synapses. This effect of GABA-adenosine cosignaling at GABAergic synapses would occur only during a specific period of development because of the transient expression of $A_{2A}$Rs and because of the transient depolarizing action of $GABA_A$Rs and their ability to activate CaM-sensitive ACs.

The $A_{2A}$R mechanism we have described therefore provides a conceptual framework to understand how some synapses are stabilized or removed during development. During neural development, inactive synapses are eliminated. This requires the existence of a machinery that includes a detector of activity from the presynaptic terminal and a mechanism that removes the synapse when the detector is not activated. The adenosine-operated $A_{2A}$R can subserve all of these functions, in addition to the involvement of GABA itself. The post- and perisynaptic localizations of $A_{2A}$Rs can detect the activity-dependent release of adenosine and ATP (*20*, *40*, *41*), which mostly involves a direct release of adenosine as a signal proportional to the metabolic support of synaptic activity as well as a CD73-mediated extracellular formation of vesicular ATP–derived adenosine (*11*, *12*, *20*). The nonactivation of $A_{2A}$Rs will prevent the continuous phosphorylation of gephyrin, which is essential for synapse stabilization. Synapse removal requires a certain period of $A_{2A}$R inactivity (20 min), which permits the persistence of somewhat quiet but not silent synapses.

Caffeine, the most commonly consumed psychoactive drug in the world, including during pregnancy and lactation, is a natural antagonist of $A_{2A}$Rs. Exposure to caffeine during the perinatal period of synaptogenesis could trigger suppression of some synapses, with deleterious effects.

## Materials and methods summary

### *Electrophysiology*

mIPSCs were recorded in coronal CA1 hippocampal slices (350 μm) from male GIN-mice, as in (*42*).

### *Immunohistochemistry*

For acute SCH58261 treatment, P6 slices were fixed with 4% paraformaldehyde (PFA), cryoprotected in 20% sucrose, frozen on dry ice, and incubated in rabbit anti-VGAT antibody (1:1000; SYSY). Images were acquired with a Zeiss LSM 510 microscope. For chronic SCH58261 treatment or $shA_{2A}$R expression, P16 brains were postfixed in 4% PFA, cryoprotected (30% sucrose), and cut in parasagittal free-floating sections (30 μm) (*43*) that were incubated in rabbit anti-VGAT (1:1000; from B. Gasnier). Z-stacks were acquired with a Leica SP5 confocal microscope and quantified using Imaris. For $A_{2A}$R detection, P3 to P12 coronal free-floating sections (100 μm) were incubated in goat anti-$A_{2A}$R antibody (1:200; Santa Cruz). Images were acquired with a Zeiss Z2 microscope. Ultrastructural analysis of $A_{2A}$R was performed using the pre-embedding immunogold method (*44*) with guinea pig anti-$A_{2A}$R (Frontier Institute, Japan). Images were acquired with a Jeol-1010 microscope.

### *Western blotting*

$A_{2A}$R density (*18*, *24*) was studied with goat (Santa Cruz) or mouse (Millipore) anti-$A_{2A}$R antibodies. Validation of $shA_{2A}$Rs was done with rabbit anti-$A_{2A}$R (1:500; Alomone) antibodies in N2a cells infected with AAV2.1-U6-shNT-GFP or AAV2.1-U6-shA2AR-GFP. Gephyrin-Slitrk3 interaction was assessed with rabbit anti-Slitrk3 (1:1000; Sigma) antibodies in HEK293T cells transfected with pCMV3-Myc-Slitrk3, eGFP-gephyrin-WT (eGFP, enhanced GFP) (*45*), or eGFP-gephyrin-S303D (*30*). Gephyrin phosphorylation was analyzed with rabbit custom made anti–phosphogephyrin and mouse anti-gephyrin (3B11, 1:1000; SYSY) antibodies in HEK293T cells transfected with eGFP-gephyrin-WT or eGFP-gephyrin-S303A. $A_{2A}$R binding (*18*) was done with 2 nM3HSCH58261 using 140 to 210 mg of protein. Release of ATP and adenosine (*20*) from hippocampal synaptosomes was assessed using the luciferin-luciferase assay (ATP) and highperformance liquid chromatography (HPLC) (adenosine) upon K+- induced depolarization (*20*).

### *Behavioral testing*

Open-field, novel object location (NOL) tasks, novel object recognition (NOR) tasks, and elevated plus maze were done as in (*46*) at P70 on WT males treated with saline or SCH58261 (0.1 mg/kg) between P3 and P16.

### *Neuronal culture*

Cultures of hippocampal neurons were prepared as described in (*47*).

Immunocytochemistry was done using rabbit anti-VGAT (1:500; from B. Gasnier), mouse anti-gephyrin (mAb7a, 1:500; SYSY) and guinea pig anti-$GABA_A$Rγ2 (1:2000; from J. M. Fristchy), or mouse anti-VGAT (1:500; SYSY) and rabbit anti-$GABA_A$Rα1 or $GABA_A$Rα2 (1:500; SYSY), or mouse anti-GAD 67 (1:500; Chemicon) and rabbit anti-$A_{2A}$R (1:100; Alomone). Image acquisition and cluster analysis was performed as in (*47*).

Single-particle tracking of $GABA_A$Rγ2 was performed as described in (*47*, *48*).

DNA-PAINT was performed with rabbit anti-$A_{2A}$R (1:100; Alomone) and mouse anti-gephyrin (mAb7a, 1:500; SYSY) on an inverted Nikon Eclipse Ti microscope using 561- and 647-nm lasers.

Calcium imaging of AAV-GCaMP6-ruby–infected neurons was done on a Leica DMI4000 microscope (Yokogawa CS20 spinning Nipkow disk) with a 491-nm laser. Time-lapse images (0.33 Hz for 600 s) of stacks (~21 sections; step, 0.3 μm) were acquired.

cAMP imaging of EPAC-sh150 (*26*) was done on a two-photon Leica TCS/MP5 microscope with a Ti:sapphire laser (Coherent). Z-stacks (8 to 10 sections; step, 1 to 2 μm) were acquired every 15 s.

Video-microscopy of presynaptic terminals (*22*) stained with rabbit anti–VGAT-oyster$^{550}$ (1:200; SYSY) was done using an Olympus IX71 microscope. Time-lapse images (one image every 5 min) were acquired.

A detailed materials and methods section can be found in the supplementary materials.

## REFERENCES AND NOTES


1. D. Krueger-Burg, T. Papadopoulos, N. Brose, Organizers of inhibitory synapses come of age. *Curr. Opin. Neurobiol.* **45**, 66–77 (2017). doi: 10.1016/j.conb.2017.04.003; pmid: 28460365
2. A. M. Craig, Y. Kang, Neurexin-neuroligin signaling in synapse development. *Curr. Opin. Neurobiol.* **17**, 43–52 (2007). doi: 10.1016/j.conb.2007.01.011; pmid: 17275284
3. T. C. Südhof, Towards an understanding of synapse formation. *Neuron* **100**, 276–293 (2018). doi: 10.1016/j.neuron.2018.09.040; pmid: 30359597
4. Z. J. Huang, P. Scheiffele, GABA and neuroligin signaling: Linking synaptic activity and adhesion in inhibitory synapse development. *Curr. Opin. Neurobiol.* **18**, 77–83 (2008). doi: 10.1016/j.conb.2008.05.008; pmid: 18513949
5. Z. J. Huang, G. Di Cristo, F. Ango, Development of GABA innervation in the cerebral and cerebellar cortices. *Nat. Rev. Neurosci.* **8**, 673–686 (2007). doi: 10.1038/nrn2188; pmid: 17704810
6. W. C. Oh, S. Lutzu, P. E. Castillo, H. B. Kwon, De novo synaptogenesis induced by GABA in the developing mouse cortex. *Science* **353**, 1037–1040 (2016). doi: 10.1126/science.aaf5206; pmid: 27516412
7. B. Chattopadhyaya *et al*., GAD67-mediated GABA synthesis and signaling regulate inhibitory synaptic innervation in the visual cortex. *Neuron* **54**, 889–903 (2007). doi: 10.1016/j.neuron.2007.05.015; pmid: 17582330
8. X. Wu *et al*., GABA signaling promotes synapse elimination and axon pruning in developing cortical inhibitory interneurons. *J. Neurosci.* **32**, 331–343 (2012). doi: 10.1523/JNEUROSCI.3189-11.2012; pmid: 22219294
9. X. Leinekugel, V. Tseeb, Y. Ben-Ari, P. Bregestovski, Synaptic $GABA_A$ activation induces $Ca^{2+}$ rise in pyramidal cells and interneurons from rat neonatal hippocampal slices. *J. Physiol.* **487**, 319–329 (1995). doi: 10.1113/jphysiol.1995.sp020882; pmid: 8558466
10. T. S. Perrot-Sinal, A. P. Auger, M. M. McCarthy, Excitatory actions of GABA in developing brain are mediated by l-type $Ca^{2+}$ channels and dependent on age, sex, and brain region. *Neuroscience* **116**, 995–1003 (2003). doi: 10.1016/S0306-4522(02)00794-7; pmid: 12617940
11. B. P. Klyuch, N. Dale, M. J. Wall, Deletion of ecto-5′-nucleotidase (CD73) reveals direct action potential-dependent adenosine release. *J. Neurosci.* **32**, 3842–3847 (2012). doi: 10.1523/JNEUROSCI.6052-11.2012; pmid: 22423104
12. M. J. Wall, N. Dale, Neuronal transporter and astrocytic ATP exocytosis underlie activity-dependent adenosine release in the hippocampus. *J. Physiol.* **591**, 3853–3871 (2013). doi: 10.1113/jphysiol.2013.253450; pmid: 23713028
13. C. G. Silva *et al*., Adenosine receptor antagonists including caffeine alter fetal brain development in mice. *Sci. Transl. Med.* **5**, 197ra104 (2013). doi: 10.1126/scitranslmed.3006258; pmid: 23926202
14. S. Alçada-Morais *et al*., Adenosine $A_{2A}$ receptors contribute to the radial migration of cortical projection neurons through the regulation of neuronal polarization and axon formation. *Cereb. Cortex*, 10.1093/cercor/bhab188 (2021). doi: 10.1093/cercor/bhab188; pmid: 34184030
15. L. Nadal *et al*., Presynaptic muscarinic acetylcholine autoreceptors (M1, M2 and M4 subtypes), adenosine receptors ($A_1$ and $A_{2A}$) and tropomyosin-related kinase B receptor (TrkB) modulate the developmental synapse elimination process at the neuromuscular junction. *Mol. Brain* **9**, 67 (2016). doi: 10.1186/s13041-016-0248-9; pmid: 27339059
16. Y. H. Jo, R. Schlichter, Synaptic corelease of ATP and GABA in cultured spinal neurons. *Nat. Neurosci.* **2**, 241–245 (1999). doi: 10.1038/6344; pmid: 10195216
17. Y. H. Jo, L. W. Role, Coordinate release of ATP and GABA at in vitro synapses of lateral hypothalamic neurons. *J. Neurosci.* **22**, 4794–4804 (2002). doi: 10.1523/JNEUROSCI.22-12-04794.2002; pmid: 12077176
18. N. Rebola, P. M. Canas, C. R. Oliveira, R. A. Cunha, Different synaptic and subsynaptic localization of adenosine $A_{2A}$ receptors in the hippocampus and striatum of the rat. *Neuroscience* **132**, 893–903 (2005). doi: 10.1016/j.neuroscience.2005.01.014; pmid: 15857695
19. E. Augusto *et al*., Ecto-5′-nucleotidase (CD73)-mediated formation of adenosine is critical for the striatal adenosine $A_{2A}$ receptor functions. *J. Neurosci.* **33**, 11390–11399 (2013). doi: 10.1523/JNEUROSCI.5817-12.2013; pmid: 23843511
20. R. A. Cunha, E. S. Vizi, J. A. Ribeiro, A. M. Sebastião, Preferential release of ATP and its extracellular catabolism as a source of adenosine upon high- but not low-frequency stimulation of rat hippocampal slices. *J. Neurochem.* **67**, 2180–2187 (1996). doi: 10.1046/j.1471-4159.1996.67052180.x; pmid: 8863529
21. E. Link *et al*., Tetanus toxin action: Inhibition of neurotransmitter release linked to synaptobrevin proteolysis. *Biochem. Biophys. Res. Commun.* **189**, 1017–1023 (1992). doi: 10.1016/0006-291X(92)92305-H; pmid: 1361727
22. F. A. Dobie, A. M. Craig, Inhibitory synapse dynamics: Coordinated presynaptic and postsynaptic mobility and the major contribution of recycled vesicles to new synapse formation. *J. Neurosci.* **31**, 10481–10493 (2011). doi: 10.1523/JNEUROSCI.6023-10.2011; pmid: 21775594
23. R. W. Li *et al*., Disruption of postsynaptic $GABA_A$ receptor clusters leads to decreased GABAergic innervation of pyramidal neurons. *J. Neurochem.* **95**, 756–770 (2005). doi: 10.1111/j.1471-4159.2005.03426.x; pmid: 16248887
24. A. P. Simões *et al*., Adenosine $A_{2A}$ receptors in the amygdala control synaptic plasticity and contextual fear memory. *Neuropsychopharmacology* **41**, 2862–2871 (2016). doi: 10.1038/npp.2016.98; pmid: 27312408
25. A. C. Conti *et al*., Distinct regional and subcellular localization of adenylyl cyclases type 1 and 8 in mouse brain. *Neuroscience* **146**, 713–729 (2007). doi: 10.1016/j.neuroscience.2007.01.045; pmid: 17335981
26. M. Polito *et al*., The NO/cGMP pathway inhibits transient cAMP signals through the activation of PDE2 in striatal neurons. *Front. Cell. Neurosci.* **7**, 211 (2013). doi: 10.3389/fncel.2013.00211; pmid: 24302895
27. D. M. F. Cooper, V. G. Tabbasum, Adenylate cyclase-centred microdomains. *Biochem. J.* **462**, 199–213 (2014). doi: 10.1042/BJ20140560; pmid: 25102028
28. S. Averaimo *et al*., A plasma membrane microdomain compartmentalizes ephrin-generated cAMP signals to prune developing retinal axon arbors. *Nat. Commun.* **7**, 12896 (2016). doi: 10.1038/ncomms12896; pmid: 27694812
29. P. Zacchi, R. Antonelli, E. Cherubini, Gephyrin phosphorylation in the functional organization and plasticity of GABAergic synapses. *Front. Cell. Neurosci.* **8**, 103 (2014). doi: 10.3389/fncel.2014.00103; pmid: 24782709
30. C. E. Flores *et al*., Activity-dependent inhibitory synapse remodeling through gephyrin phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E65–E72 (2015). doi: 10.1073/pnas.1411170112; pmid: 25535349
31. S. K. Tyagarajan, J.-M. Fritschy, Gephyrin: A master regulator of neuronal function? *Nat. Rev. Neurosci.* **15**, 141–156 (2014). doi: 10.1038/nrn3670; pmid: 24552784
32. V. Tretter *et al*., Gephyrin, the enigmatic organizer at GABAergic synapses. *Front. Cell. Neurosci.* **6**, 23 (2012). doi: 10.3389/fncel.2012.00023; pmid: 22615685
33. H. Takahashi *et al*., Selective control of inhibitory synapse development by Slitrk3-PTPδ trans-synaptic interaction. *Nat. Neurosci.* **15**, 389–398 (2012). doi: 10.1038/nn.3040; pmid: 22286174
34. B. Chih, H. Engelman, P. Scheiffele, Control of excitatory and inhibitory synapse formation by neuroligins. *Science* **307**, 1324–1328 (2005). doi: 10.1126/science.1107470; pmid: 15681343
35. M. Mondin, B. Tessier, O. Thoumine, Assembly of synapses: Biomimetic assays to control neurexin/neuroligin interactions at the neuronal surface. *Curr. Protoc. Neurosci.* **64**, 2.19.1–2.19.30 (2013). doi: 10.1002/0471142301.ns0219s64; pmid: 23853109
36. A. A. Chubykin *et al*., Activity-dependent validation of excitatory versus inhibitory synapses by neuroligin-1 versus neuroligin-2. *Neuron* **54**, 919–931 (2007). doi: 10.1016/j.neuron.2007.05.029; pmid: 17582332
37. J. Li *et al*., A conserved tyrosine residue in Slitrk3 carboxyl-terminus is critical for GABAergic synapse development. *Front. Mol. Neurosci.* **12**, 213 (2019). doi: 10.3389/fnmol.2019.00213; pmid: 31551708
38. J. Ko, G. Choii, J. W. Um, The balancing act of GABAergic synapse organizers. *Trends Mol. Med.* **21**, 256–268 (2015). doi: 10.1016/j.molmed.2015.01.004; pmid: 25824541
39. A. G. Nair, O. Gutierrez-Arenas, O. Eriksson, P. Vincent, J. Hellgren Kotaleski, Sensing positive versus negative reward signals through adenylyl cyclase-coupled GPCRs in direct and indirect pathway striatal medium spiny neurons. *J. Neurosci.* **35**, 14017–14030 (2015). doi: 10.1523/JNEUROSCI.0730-15.2015; pmid: 26468202
40. J. B. Mitchell, C. R. Lupica, T. V. Dunwiddie, Activity-dependent release of endogenous adenosine modulates synaptic responses in the rat hippocampus. *J. Neurosci.* **13**, 3439–3447 (1993). doi: 10.1523/JNEUROSCI.13-08-03439.1993; pmid: 8393482
41. A. Wieraszko, G. Goldsmith, T. N. Seyfried, Stimulation-dependent release of adenosine triphosphate from hippocampal slices. *Brain Res.* **485**, 244–250 (1989). doi: 10.1016/0006-8993(89)90567-2; pmid: 2566360
42. R. Cossart *et al*., Dendritic but not somatic GABAergic inhibition is decreased in experimental epilepsy. *Nat. Neurosci.* **4**, 52–62 (2001). doi: 10.1038/82900; pmid: 11135645
43. T. Notter, P. Panzanelli, S. Pfister, D. Mircsof, J. M. Fritschy, A protocol for concurrent high-quality immunohistochemical and biochemical analyses in adult mouse central nervous system. *Eur. J. Neurosci.* **39**, 165–175 (2014). doi: 10.1111/ejn.12447; pmid: 24325300
44. R. Luján, Z. Nusser, J. D. B. Roberts, R. Shigemoto, P. Somogyi, Perisynaptic location of metabotropic glutamate receptors mGluR1 and mGluR5 on dendrites and dendritic spines in the rat hippocampus. *Eur. J. Neurosci.* **8**, 1488–1500 (1996). doi: 10.1111/j.1460-9568.1996.tb01611.x; pmid: 8758956
45. S. K. Tyagarajan *et al*., Regulation of GABAergic synapse formation and plasticity by GSK3beta-dependent phosphorylation of gephyrin. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 379–384 (2011). doi: 10.1073/pnas.1011824108; pmid: 21173228
46. M. P. Kaster *et al*., Caffeine acts through neuronal adenosine $A_{2A}$ receptors to prevent mood and memory dysfunction triggered by chronic stress. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7833–7838 (2015). doi: 10.1073/pnas.1423088112; pmid: 26056314
47. S. Battaglia *et al*., Activity-dependent inhibitory synapse scaling is determined by gephyrin phosphorylation and subsequent regulation of $GABA_A$ receptor diffusion. *eNeuro* **5**, ENEURO.0203-17.2017 (2018). doi: 10.1523/ENEURO.0203-17.2017; pmid: 29379879
48. H. Bannai, S. Lévi, C. Schweizer, M. Dahan, A. Triller, Imaging the lateral diffusion of membrane molecules with quantum dots. *Nat. Protoc.* **1**, 2628–2634 (2006). doi: 10.1038/nprot.2006.429; pmid: 17406518


## ACKNOWLEDGMENTS

We thank B. Tessier and O. Thoumine for providing recombinant neurexin1β-Fc fragments, A. Triller for providing gephyrin-mRFP construct, A. M. Craig for shRNA against Slitrk3 coupled to CFP, Q. Tian and W. Lu for the Slitrk3-Y969A mutant, and P. Scheiffele for shRNA against neuroligin-2. We are also grateful to the Animal Facility and Cell and Tissue Imaging Facility of Institut du Fer à Moulin (IFM). **Funding:** This study was supported by Inserm (S.L. and C.B.), Sorbonne Université-UPMC (S.L.), Agence Nationale de la Recherche ADONIS ANR-14-CE13-0032 (S.L. and C.B.), DIM NeRF from Région Ile-de-France (S.L.), AXA Research Fund (S.Z.), Fondation pour la Recherche sur le Cerveau FRC/Rotary Espoir en tête (S.L.), La Caixa Foundation LCF/PR/HP17/52190001 (R.A.C.), Centro 2020 CENTRO-01-0145-FEDER-000008: BrainHealth 2020 and CENTRO-01-0246-FEDER-000010 (R.A.C.), FCT POCI-01-0145-FEDER-03127 and UIDB/04539/2020 (R.A.C.), Spanish Ministerio de Economía y Competitividad (RTI2018-095812-B-I00) (R.L.), Junta de Comunidades de Castilla-La Mancha (SBPLY/17/180501/000229) (R.L.), and CNRS ATIP AO2016 (C.L.). **Author contributions:** Conceptualization was by S.L., C.B., and R.A.C. F.G.-C., M.Ru., and C.M. performed most (F.G.-C.) and some (M.Ru. and C.M.) immunofluorescence experiments in hippocampal cultures and analyzed the data. M.Ru. performed hippocampal cultures and molecular biology. S.Z., C.G.S., and C.B. performed electrophysiological experiments, and M.E. performed post hoc morphology. J.C.P. performed calcium imaging, characterized the $shA_{2A}R$ by Western blot, performed the stereotaxic injection of $AAVshA_{2A}R$ in the hippocampus in vivo, chronically treated animals with SCH58261, performed immunohistochemistry in chronically treated and $shA_{2A}R$-expressing animals, and quantified the data. S.Z. and M.E. performed and analyzed the immunofluorescence in acute hippocampal slices treated with $A_{2A}R$ antagonists. R.L. performed the electron microscopy of the $A_{2A}R$. S.L. and M.Ru. performed the single-particle tracking experiments and analyzed the data. G.C. and C.L. performed DNA-PAINT experiments, and M.Re. analyzed the data. N.G. performed cAMP imaging and analyzed the data. P.M.C., F.Q.G., S.A.-M., E.S., R.J.R., P.A., and A.R.T. performed and analyzed the biochemical experiments exploring ATP and adenosine release as well as $A_{2A}R$ density during the synaptogenesis period of development in vivo. S.K.T. designed the biochemical experiments testing gephyrin phosphorylation and gephyrin-Slitrk3 interaction, and M.F. performed the experiments and analyzed the data. X.N. produced the cAMP sponges. F.G.-C., S.Z., C.B., and S.L. prepared the figures. Funding acquisition was done by S.L., C.B., and R.A.C. Project administration was by S.L., C.B., and R.A.C. S.L., C.B., and R.A.C. supervised the work. S.L., C.B., and R.A.C. wrote the original draft. S.L. and C.B. wrote the revised manuscript. Reviewing and editing was done by S.L., C.B., R.A.C., F.G.-C., C.L., X.N., O.T., and M.Re. **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.abk2055
Materials and Methods
Figs. S1 to S22
MDAR Reproducibility Checklist

29 June 2021; accepted 9 September 2021
10.1126/science.abk2055

RESEARCH ARTICLE SUMMARY

IMMUNOLOGY

# Small proline-rich protein 2A is a gut bactericidal protein deployed during helminth infection

**Zehan Hu†*, Chenlu Zhang†, Luis Sifuentes-Dominguez, Christina M. Zarek, Daniel C. Propheter, Zheng Kuang, Yuhao Wang, Mihir Pendse, Kelly A. Ruhn, Brian Hassell, Cassie L. Behrendt, Bo Zhang, Prithvi Raj, Tamia A. Harris-Tryon, Tiffany A. Reese*, Lora V. Hooper***

**INTRODUCTION:** The mammalian intestinal tract is colonized with bacteria, fungi, viruses, and parasites, which present a diverse array of immunological challenges. A number of antimicrobial mechanisms have evolved to cope with these challenges. A key defense mechanism is the production of antimicrobial proteins (AMPs), which rapidly inactivate microorganisms by targeting their cell wall or cell membrane structures. However, we still have a limited understanding of the repertoire of gut AMPs and how these AMPs work cooperatively to regulate intestinal microbial communities and protect against infection in varied immunological settings.

**RATIONALE:** Prior studies have shown that the gene encoding small proline-rich protein 2A (SPRR2A) is highly up-regulated in the mouse small intestine after colonization with commensal bacteria. To elucidate the function of SPRR2A in the intestine, we studied its regulation, mechanism of action, impact on the microbiota, and ability to limit pathogenic bacterial infection. When mice are infected with intestinal helminths (parasitic worms), *Sprr2a* expression is further elevated above the levels induced by the commensal microbiota. Helminth infections drastically alter intestinal immunological responses and cause epithelial damage and increased tissue invasion by intestinal bacteria. We therefore also addressed the molecular basis for SPRR2A up-regulation and its physiological role during helminth infection in our study.

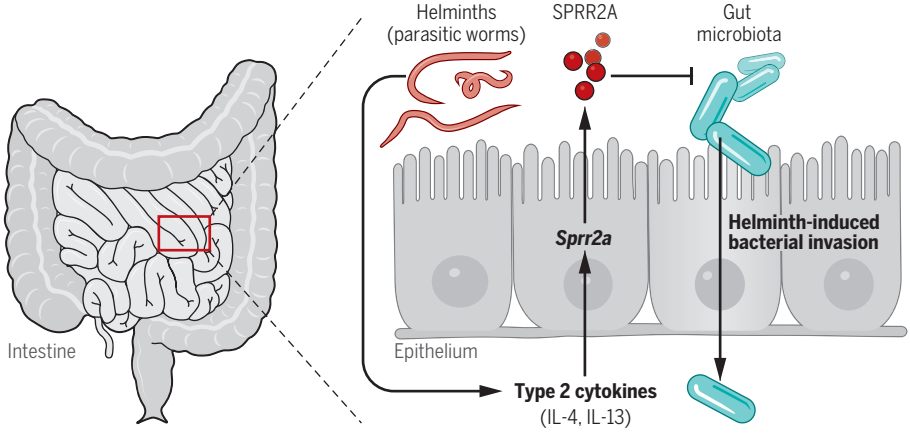
**RESULTS:** We report that SPRR2A is an intestinal bactericidal protein that is not phylogenetically related to any previously discovered mammalian AMP. Experiments in mice revealed that SPRR2A was induced by the intestinal microbiota in specialized secretory cell lineages (Paneth cells and goblet cells) of the intestinal epithelium. SPRR2A was also expressed in the human intestine, where it selectively localized to goblet cells. Transcripts encoding SPRR2A were detected in large intestinal biopsies and stool from inflammatory bowel disease patients but not from those of healthy controls.

We conducted in vitro biochemical studies to identify the biological activity of SPRR2A. Purified recombinant SPRR2A selectively killed Gram-positive bacteria in vitro, whereas Gram-negative bacteria were resistant to killing by SPRR2A. The SPRR2A bactericidal mechanism involved binding to negatively charged lipids and disrupting bacterial membranes.

We generated *Sprr2a*-deficient mice and found that overall bacterial loads remained the same. However, the absence of SPRR2A led to the expansion of Gram-positive bacteria in the small intestinal lumen and mucus layer, consistent with the selective bactericidal activity of SPRR2A against Gram-positive bacteria in vitro. *Sprr2a*-deficient mice also showed an increased susceptibility to infection by *Listeria monocytogenes*, a Gram-positive intestinal pathogen.

Intestinal helminth infections can cause damage to the intestinal epithelium, which promotes bacterial breach of the intestinal barrier and highlights the need for AMP expression during infection with these parasitic worms. Although other AMPs [such as regenerating family member 3 gamma (REG3G) and lysozyme] showed reduced expression during infection of mice with the intestinal helminth *Heligmosomoides polygyrus*, SPRR2A expression was induced above the levels elicited by bacterial colonization. Helminth infection elicits the production of type 2 cytokines, such as interleukin (IL)–4 and IL-13. SPRR2A differed from other intestinal AMPs in that it was further induced by type 2 cytokines during infection. Accordingly, *Sprr2a* induction during helminth infection required STAT6, a transcription factor that acts downstream of IL-4 and IL-13 to induce type 2 immunity. Finally, SPRR2A was necessary to prevent intestinal bacteria, mainly the Gram-positive subset, from invading the intestinal barrier during *H. polygyrus* infection.

**CONCLUSION:** SPRR2A is an intestinal antibacterial protein that is unrelated to any previously discovered mammalian AMP. SPRR2A contributes to intestinal innate immunity by shaping the composition of the gut microbiota and by limiting pathogenic infection by Gram-positive bacteria. A distinctive aspect of SPRR2A function is its essential role during helminth infection of the intestine. Although such infection lowers the expression of a number of intestinal AMPs, SPRR2A expression is selectively increased by the type 2 cytokines that are produced during helminth infection. SPRR2A defends the intestinal barrier against bacterial invasion and dissemination that is provoked by helminth infection. This study highlights how different AMPs may be produced in distinctive immunological settings and deepens our understanding of why a diverse repertoire of epithelial AMPs has evolved for host immune defense. ■




**SPRR2A protects the intestinal barrier during helminth infection.** SPRR2A is an antimicrobial protein produced by intestinal epithelial cells. During infection with the parasitic helminth *H. polygyrus*, SPRR2A expression is boosted by the type 2 cytokines IL-4 and IL-13. SPRR2A kills bacteria by damaging their membranes and thus limits helminth-induced invasion of intestinal tissue by the gut microbiota.

The list of author affiliations is available in the full article online.
*Corresponding author: Email: zehan.hu@sjtu.edu.cn (Z.H.); tiffany.reese@utsouthwestern.edu (T.A.R.); lora.hooper@utsouthwestern.edu (L.V.H.)
†These authors contributed equally to this work.
Cite this article as Z. Hu *et al.*, *Science* **374**, eabe6723 (2021). DOI: 10.1126/science.abe6723

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abe6723

RESEARCH ARTICLE

IMMUNOLOGY

# Small proline-rich protein 2A is a gut bactericidal protein deployed during helminth infection

Zehan Hu[1,2]†‡*, Chenlu Zhang[3]‡, Luis Sifuentes-Dominguez[4], Christina M. Zarek[1], Daniel C. Propheter[1], Zheng Kuang[1], Yuhao Wang[1], Mihir Pendse[1], Kelly A. Ruhn[1], Brian Hassell[1], Cassie L. Behrendt[1], Bo Zhang[1], Prithvi Raj[1], Tamia A. Harris-Tryon[3], Tiffany A. Reese[1]*, Lora V. Hooper[1,2]*

**A diverse group of antimicrobial proteins (AMPs) helps protect the mammalian intestine from varied microbial challenges. We show that small proline-rich protein 2A (SPRR2A) is an intestinal antibacterial protein that is phylogenetically unrelated to previously discovered mammalian AMPs. In this study, SPRR2A was expressed in Paneth cells and goblet cells and selectively killed Gram-positive bacteria by disrupting their membranes. SPRR2A shaped intestinal microbiota composition, restricted bacterial association with the intestinal surface, and protected against *Listeria monocytogenes* infection. SPRR2A differed from other intestinal AMPs in that it was induced by type 2 cytokines produced during helminth infection. Moreover, SPRR2A protected against helminth-induced bacterial invasion of intestinal tissue. Thus, SPRR2A is a distinctive AMP triggered by type 2 immunity that protects the intestinal barrier during helminth infection.**

The mammalian intestine contains a complex community of microorganisms that present a diverse array of immunological challenges (*1*). This community is mostly composed of commensal bacteria that are essential for digestion but can also include opportunistic pathobionts and overtly pathogenic bacteria and fungi. In addition, the intestine can become infected with parasites such as helminths, a class of parasitic worms. Helminths can cause epithelial damage and provoke increased tissue invasion by intestinal bacteria (*2*, *3*).

To cope with these varied microbiological challenges, the intestinal epithelium produces a diverse repertoire of antimicrobial proteins (AMPs), which have widely divergent primary sequences and rapidly kill or inactivate microorganisms (*4*, *5*). Intestinal AMPs are secreted into the mucus layer that overlays the intestinal epithelium, where they restrict bacterial–epithelial contact and thus limit bacterial invasion of host tissues. One reason for the evolution of a diverse array of epithelial AMPs is that individual antimicrobial proteins target distinct populations of bacteria (e.g., Gram-positive versus Gram-negative bacteria), thus necessitating multiple AMPs to defend against challenges from complex bacterial populations.

A less well explored idea is that a diverse AMP repertoire may also have evolved to provide antimicrobial protection in varied immunological settings. For example, helminth infections alter the immunological landscape of the intestine by triggering expression of type 2 cytokines such as interleukin (IL)–4 and IL-13 (*6*, *7*). These cytokines direct numerous changes to epithelial cell function that promote worm expulsion, such as increased mucus secretion and turnover of epithelial cells (*8*, *9*). However, little is known about how helminth infection and the ensuing production of type 2 cytokines shape the repertoire of epithelial AMPs, or which AMPs protect the host in the face of the pronounced helminth-induced alterations of the gut microbiota. Here, we identify small proline-rich protein A (SPRR2A) as an epithelial antibacterial protein, phylogenetically distinct from previously discovered mammalian AMPs, that plays a distinctive role in defending the intestinal barrier during helminth infection.

[1]Department of Immunology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.
[2]Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.
[3]Department of Dermatology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.
[4]Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.
***Corresponding author: Email: zehan.hu@sjtu.edu.cn (Z.H.); tiffany.reese@utsouthwestern.edu (T.A.R.); lora.hooper@utsouthwestern.edu (L.V.H.)**
†Present address: State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China.
‡These authors contributed equally to this work.

## Results

### *SPRR2A is expressed in goblet cells and Paneth cells in the mouse intestine*

In a previous study, colonization of the germ-free mouse intestine with the Gram-negative commensal *Bacteroides thetaiotaomicron* elicited markedly increased expression of the gene encoding SPRR2A (*10*). SPRR2A belongs to the small proline-rich protein family of genes, which are encoded tandemly within a 170-kb region of the epidermal differentiation complex in both humans and mice (*11*). SPRR family genes are also induced during bacterial infection of the stomach, lung, and skin, suggesting a role in responding to infection (*12*–*16*). Although SPRRs have proposed functions in the development of the cornified envelope of squamous epithelia (*17*), virtually nothing is known about the function of SPRR2A in the nonsquamous intestinal epithelium and its role during bacterial colonization and infection.

In mice, *Sprr2a* transcripts were most abundant in tissues that interface with the external environment. These included the gastrointestinal tract, bladder, and skin (Fig. 1A and fig. S1A). In the intestine, *Sprr2a* transcripts were selectively expressed in intestinal epithelial cells (IEC) (Fig. 1B), with expression of SPRR2A transcripts and protein restricted to secretory epithelial cell lineages including goblet cells and Paneth cells (Fig. 1, C to E, and figs. S1, B and C, and S2, A to E). This accords with prior findings from single-cell RNA sequencing of the mouse small intestinal epithelium (fig. S1D) (*18*). SPRR2A was present in secretory granule-like structures within goblet cells and Paneth cells (Fig. 1F and fig. S1E), suggesting that SPRR2A might be secreted into the intestinal lumen. We detected SPRR2A in the stool and colonic mucus layer (Fig. 1G and fig. S1F), which supports this idea. Together, these observations indicate that mouse SPRR2A is selectively expressed in intestinal secretory epithelial cells and is secreted into the intestinal lumen.

SPRR2A was also expressed in the human intestine. Several human colonic cell lines had abundant *SPRR2A* transcripts (table S1 and fig. S3, A and B) (*19*), and *SPRR2A* transcripts were detected in large intestinal biopsies from inflammatory bowel disease (IBD) patients but not in those from healthy controls (table S2 and fig. S3C). As in mice, human SPRR2A was selectively localized to goblet cells (fig. S3D) and was detected in the stool of the IBD patients (fig. S3E).

We next examined the impact of the gut microbiota on SPRR2A expression in the small intestine. The prior study showing induction of *Sprr2a* expression upon *B. thetaiotaomicron* colonization was conducted in germ-free NMRI mice, an outbred strain, and reported a ~200-fold increase in *Sprr2a* transcript abundance (*10*). By contrast, when we conventionalized germ-free C57BL/6 mice, we detected only a twofold increase in *Sprr2a* transcript abundance (fig. S4A). However, conventionalization of germ-free Swiss Webster mice resulted in a 20-fold increase (fig. S4A), and treatment of conventional BALB/c mice with broad-spectrum antibiotics reduced *Sprr2a* transcript abundance by ~50-fold (fig. S4A)

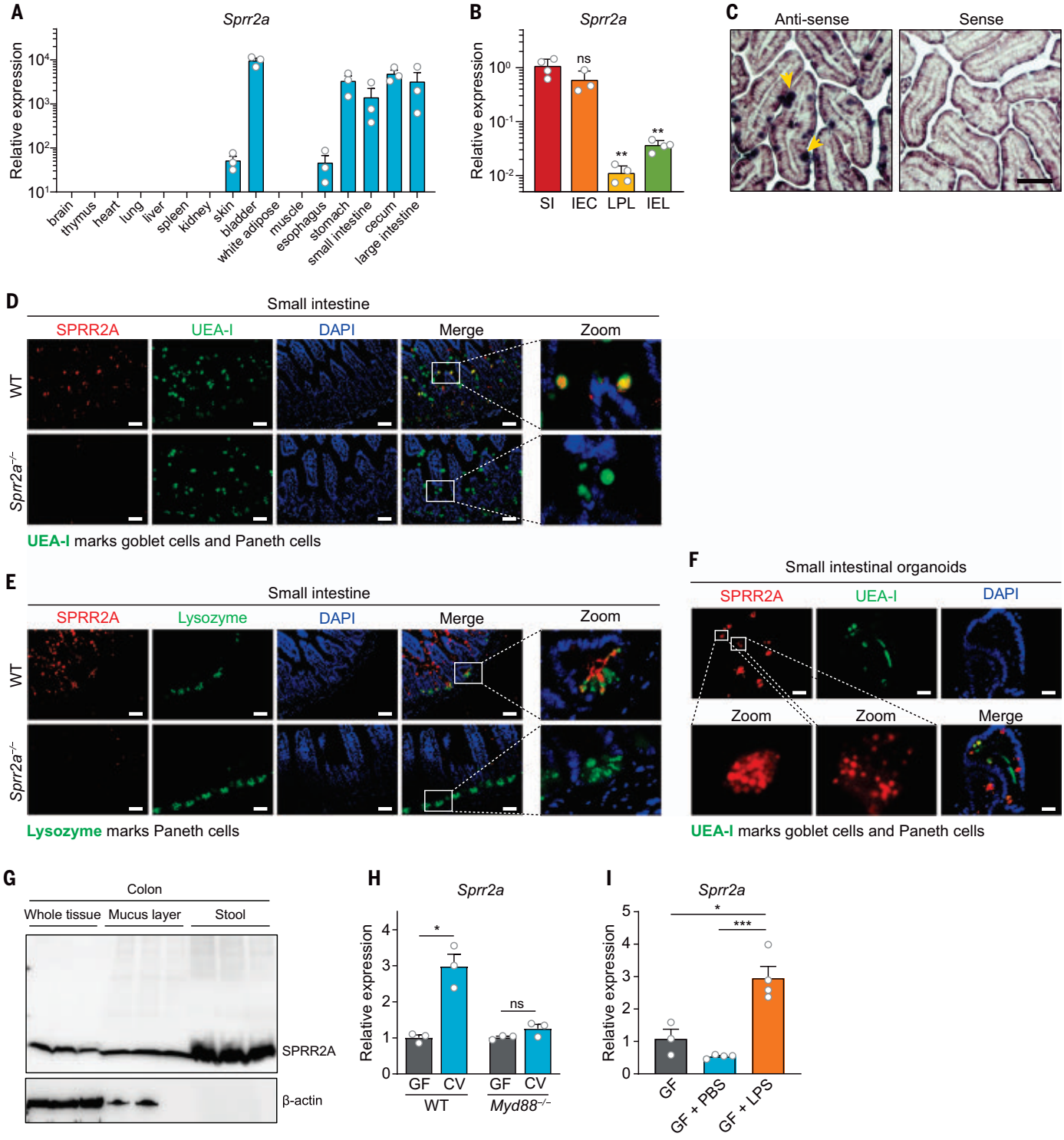



**Fig. 1. SPRR2A is expressed in goblet cells and Paneth cells in the mouse intestine.** (**A**) Quantitative PCR (qPCR) analysis of *Sprr2a* expression in various organs of wild-type C57BL/6 mice. Values were normalized to 18*S* rRNA expression. *n* = 3 mice per group. (**B**) qPCR analysis of *Sprr2a* expression in mouse small intestine (SI), small intestinal epithelial cells (IEC) acquired by laser capture microdissection, lamina propria lymphocytes (LPL), and intraepithelial lymphocytes (IEL). Values were normalized to 18*S* rRNA expression. (**C**) In situ hybridization of small intestine sections with antisense and sense *Sprr2a* RNA probes. Examples of positive signals are indicated by yellow arrowheads. Scale bar, 50 μm. (**D** and **E**) Immunofluorescence detection of SPRR2A (red) in sections of paraffin-embedded mouse small intestine. *Ulex europaeus* agglutinin-I (UEA-I) (D) and anti-lysozyme antibody (E) were used to identify goblet cells (UEA-I) and Paneth cells (UEA-I and anti-lysozyme). Nuclei were detected with 4′,6-diamidino-2-phenylindole (DAPI). Sections from *Sprr2a*$^{-/-}$ mice (generated as shown in fig. S2) were used as negative controls. Scale bars, 50 μm. (**F**) Immunofluorescence detection of SPRR2A (red) in OCT-embedded frozen sections of mouse small-intestinal organoids. Scale bars, 25 μm. (**G**) Immunoblot detection of SPRR2A protein in mouse colon, the colon mucus layer, and stool samples. (**H**) qPCR analysis of epithelial cell *Sprr2a* expression in germ-free (GF) wild-type (WT) and *Myd88*-deficient (*Myd88*$^{-/-}$) C57BL/6 mice with or

without conventionalization (CV). Epithelial cells were harvested by laser capture microdissection. Values were normalized to *Gapdh* expression. $n = 3$ mice per group. (**I**) qPCR analysis of *Sprr2a* expression in the small intestines of GF Swiss-Webster mice treated with lipopolysaccharide (LPS). PBS was administered as a vehicle control. Values were normalized to *Gapdh* expression. $n = 3$ or 4 mice per group. Means ± SEM (error bars) are plotted. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; ns, not significant per two-tailed *t* test.

with an accompanying decrease in protein levels (fig. S4, B and C). Notably, the *Sprr2a* transcript numbers in germ-free C57BL/6 mice were two orders of magnitude higher than those in germ-free Swiss Webster mice or antibiotic-treated BALB/c mice (fig. S4D), which could explain why the induction of SPRR2A expression by the microbiota was less pronounced in C57BL/6 mice than in the other two strains. Additionally, monocolonization of germ-free C57BL/6 and Swiss Webster mice with *Listeria monocytogenes* increased *Sprr2a* transcript abundance (fig. S4E). The microbiota therefore induces intestinal expression of *Sprr2a*, but the magnitude of the increase varies according to mouse strain.

The microbiota-dependent increase in *Sprr2a* transcript abundance in C57BL/6 mice required the Toll-like receptor (TLR) signaling adaptor myeloid differentiation primary response 88 (MyD88) (Fig. 1H). Accordingly, *Sprr2a* expression in germ-free mouse small intestine was triggered by lipopolysaccharide (LPS), a major component of the outer membrane of Gram-negative bacteria that activates TLR4–MyD88 signaling (Fig. 1I and fig. S4F). Thus, *Sprr2a* expression is induced by the intestinal microbiota through TLR–MyD88 signaling.

### *SPRR2A is a bactericidal protein that targets Gram-positive bacteria by membrane permeabilization*

SPRR2A has several characteristics that are shared with intestinal AMPs such as angiogenin-4 (ANG4), regenerating family member 3 gamma (REG3G), and resistin-like molecule beta (RELMβ) (*20*–*23*). These shared characteristics include expression in secretory epithelial cells and induction by bacterial colonization. Additionally, SPRR2A is basic, with a predicted isoelectric point of 8.4 (human) and 7.7 (mouse). This is also a shared characteristic of AMPs that target bacterial membranes, which are acidic owing to the presence of negatively charged lipid headgroups (*24*). We therefore hypothesized that SPRR2A is an AMP.

To test for SPRR2A bactericidal activity, we produced recombinant human SPRR2A in insect cells and purified it by size exclusion chromatography (fig. S5, A and B). We added the purified SPRR2A to a panel of enteric commensal and pathogenic bacteria that included both Gram-positive and Gram-negative species (Fig. 2A and fig. S5C). We observed a dose-dependent reduction in the viability of the Gram-positive species *Lactobacillus reuteri*, *Enterococcus faecalis*, and *L. monocytogenes* when exposed to low micromolar concentrations of SPRR2A. By contrast, the Gram-negative species *B. thetaiotaomicron*, *Escherichia coli*, and *Citrobacter rodentium* were resistant to SPRR2A (Fig. 2A and fig. S5C). SPRR2A bactericidal activity against Gram-positive bacteria was inhibited by an anti-SPRR2A polyclonal antibody, indicating that the bactericidal activity was specific to the SPRR2A protein (Fig. 2B). Thus, we conclude that SPRR2A is a bactericidal protein that selectively kills Gram-positive bacteria.

The bactericidal activity of SPRR2A, like that of other intestinal AMPs, required low salt concentrations and an acidic pH (fig. S5, D and E) (*23*, *25*). This accords with the fact that intestinal AMPs such as SPRR2A must function in the acidic, low-salt environment of the mucus layer and intestinal lumen (*26*). However, unlike other AMPs such as RELMβ (*23*), the bactericidal activity of SPRR2A was not dependent on the growth phase of the target bacteria, as SPRR2A killed logarithmic-phase and stationary-phase bacteria with similar efficiency (fig. S5F).

To acquire initial insight into how SPRR2A kills bacteria, we used transmission electron microscopy to visualize morphological changes in Gram-positive bacteria after exposure to SPRR2A. The images showed evidence of bacterial cell wall damage and cytoplasmic leakage (Fig. 2C and fig. S5G), suggesting that SPRR2A kills bacteria by permeabilizing their membranes. To test this idea, we measured bacterial membrane permeabilization by assaying for bacterial uptake of propidium iodide (PI), a membrane-impermeant dye. SPRR2A promoted the dose-dependent uptake of PI by *L. monocytogenes*, *L. reuteri*, and *E. faecalis* but not by *B. thetaiotaomicron* (Fig. 2D and fig. S6A). Membrane permeabilization by SPRR2A required low salt concentrations and an acidic pH (fig. S6, B and C), consistent with the requirements for bactericidal activity. These data reveal that SPRR2A kills bacteria by disrupting their membranes.

The requirement for low salt concentrations and an acidic pH suggested that electrostatic interactions might be driving SPRR2A interactions with bacterial membranes, which tend to be acidic (*24*). We found, in support of this idea, that SPRR2A bound selectively to immobilized lipids bearing negatively charged lipid headgroups, including phosphatidic acid, phosphatidylserine (PS), cardiolipin (CL), and phosphatidylinositol phosphates, but not to zwitterionic or neutral lipids (Fig. 2E). Moreover, SPRR2A was precipitated by liposomes containing negatively charged PS or CL but not by liposomes composed of only the neutral lipid phosphatidylcholine (PC) (Fig. 2F). These interactions were also salt- and pH-dependent (fig. S6, D and E), supporting the concept that electrostatic interactions drive SPRR2A–lipid interactions. Thus, SPRR2A binds to lipids bearing negatively charged headgroups that reflect the acidic characteristic of most bacterial membranes.

To test whether SPRR2A disrupts membranes, we added purified SPRR2A to liposomes. Liposomes that were exposed to SPRR2A and visualized by negative-stain electron microscopy were completely disrupted, with only membrane fragments remaining (Fig. 2G). We next conducted dye efflux assays on liposomes encapsulating fluorescent dyes of different sizes. SPRR2A induced rapid dye efflux from PC/PS and PC/CL liposomes loaded with either carboxyfluorescein (~10-Å Stokes diameter) or fluorescein isothiocyanate–dextran 10 (~44-Å Stokes diameter) (Fig. 2H and fig. S6, F to H), indicating a lack of size selectivity in SPRR2A-induced dye efflux. There was no dye efflux from liposomes composed only of PC (Fig. 2H). By contrast, the microbiota-inducible intestinal AMPs REG3A and RELMβ form structured pores that are visible by negative-stain electron microscopy and that induce size-selective dye efflux from liposomes (*23*, *27*). These results indicate that SPRR2A disrupts membranes but uses a mechanism that is distinct from that of other known microbiota-inducible AMPs.

The SPRR2A-mediated killing of *L. monocytogenes* was inhibited by LPS (Fig. 2I), suggesting that the resistance of Gram-negative bacteria may be due to the LPS in their outer membranes. LPS did not inhibit SPRR2A binding to either Gram-positive or Gram-negative bacteria or to PC/PS liposomes (fig. S7, A to C), arguing against binding interference as the mechanism of inhibition. Rather, LPS inhibited SPRR2A disruption of liposomes (Fig. 2, J and K). Thus, LPS appears to inhibit SPRR2A bactericidal activity by interfering with membrane permeabilization.

Many proline-rich proteins are intrinsically unstructured owing to proline's disruptive effect on secondary structures such as α helices and β sheets (*28*). However, in addition to being proline-rich, SPRR2A is also cysteine-rich (fig. S8), with 11 cysteines that form five pairs of intrachain disulfide bonds (fig. S9, A and B). Because disulfide bonds can impart higher-order structure to proteins, we assessed whether the disulfide bonds were essential for SPRR2A bactericidal function. The bactericidal activity

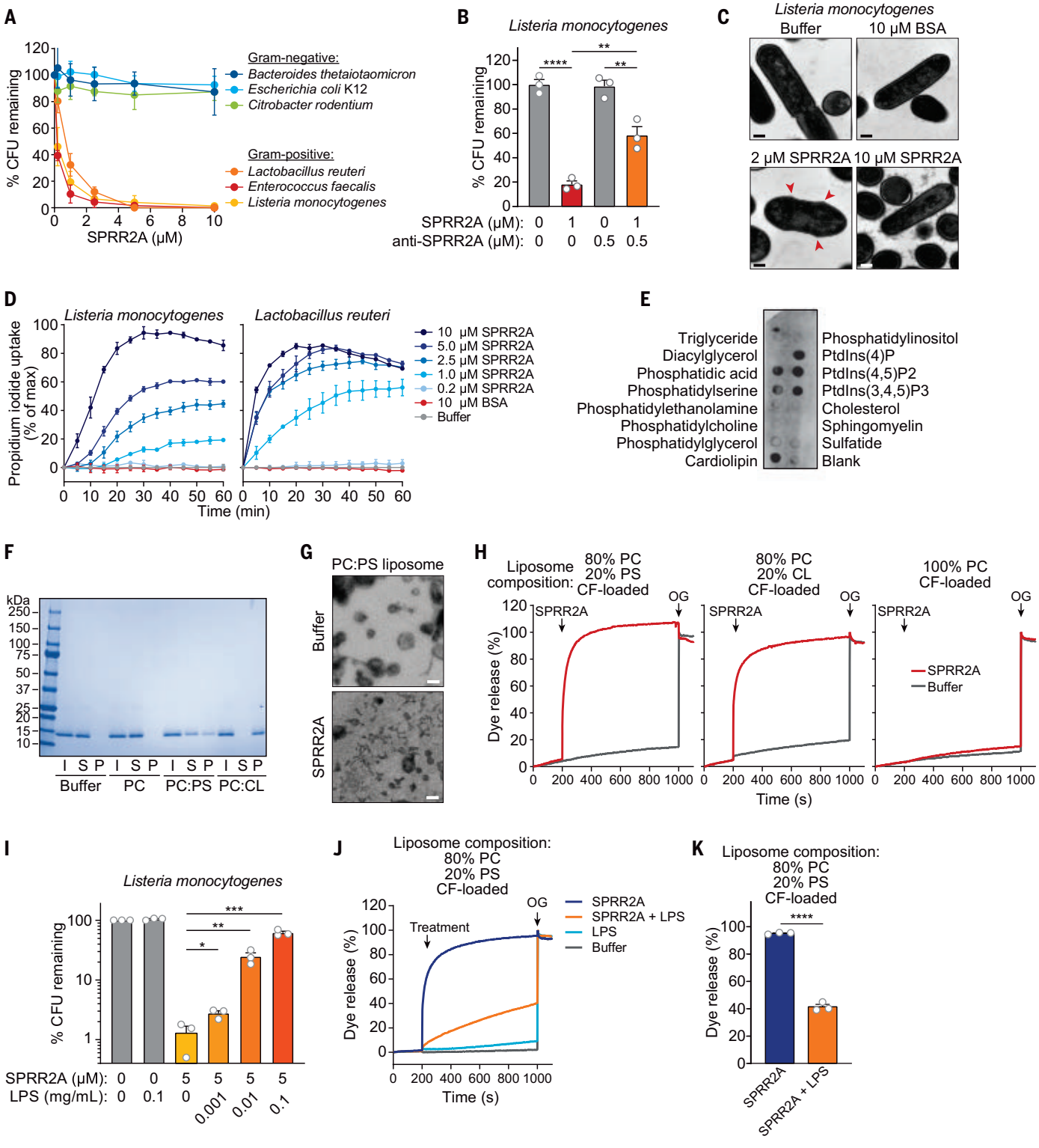



**Fig. 2. SPRR2A is a bactericidal protein that targets Gram-positive bacteria by membrane permeabilization.** (**A**) Recombinant SPRR2A was expressed in a baculovirus expression system and purified by size exclusion chromatography. SPRR2A was added to ~$10^4$ CFU of log-phase bacteria for 2 hours, and surviving bacteria were enumerated by dilution plating. (**B**) SPRR2A was added to log-phase *L. monocytogenes* in the presence of anti-SPRR2A antibody, and surviving bacteria were enumerated by dilution plating. (**C**) Transmission electron microscopy of *L. monocytogenes* after incubation with SPRR2A. Bovine serum albumin (BSA) was used as a negative control. Examples of cell surface damage and cytoplasmic leakage are indicated with red arrowheads. Scale bars, 200 nm. (**D**) *L. monocytogenes* (left) and *L. reuteri* (right) were treated with SPRR2A or BSA as control, and propidium iodide (PI) uptake was measured over 1 hour. (**E**) Membranes displaying various lipids were incubated with 1 μg/ml of SPRR2A and detected with anti-SPRR2A antibody. (**F**) SPRR2A was incubated with liposomes having the indicated lipid compositions. After ultracentrifugation, the liposome-free supernatant (S)

and the liposome pellet (P) were analyzed by SDS-PAGE and Coomassie blue staining. I, input; PC, phosphatidylcholine; PS, phosphatidylserine; CL, cardiolipin. (**G**) Transmission electron microscopy of PC:PS liposomes after treatment with 5 μM SPRR2A. Scale bars, 100 nm. (**H**) Carboxyfluorescein (CF)–loaded liposomes with the indicated lipid compositions were treated with 10 μM SPRR2A. Dye efflux was monitored over time and is expressed as a percentage of maximal efflux in the presence of the detergent octyl glucoside (OG). (**I**) 5 μM SPRR2A was added to log-phase *L. monocytogenes* in the presence of increasing lipopolysaccharide (LPS) concentrations. Surviving bacteria were then enumerated by dilution plating. (**J**) CF-loaded PC:PS liposomes were treated with 10 μM SPRR2A in the presence or absence of 0.2 mg/ml of LPS, and dye efflux was monitored over time and is expressed as a percentage of maximal efflux in the presence of OG. (**K**) Quantification of the results in (J); $n = 3$ experimental replicates. All assays were performed in triplicate, and results are representative of at least two independent experiments. Means ± SEM (error bars) are plotted. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; ns, not significant per two-tailed *t* test.

of SPRR2A was inhibited by the reducing agent 1,4-dithiothreitol (DTT) (fig. S9C), indicating that the disulfide bonds are required for SPRR2A bactericidal activity. DTT also inhibited lipid binding by SPRR2A (fig. S9, D to F), which is consistent with this finding. Finally, SPRR2A purified from *E. coli* (fig. S9D), which does not support the formation of disulfide bonds, did not bind negatively charged lipids, disrupt liposomes, or kill Gram-positive bacteria (fig. S9, E, G, and H). Therefore, we conclude that disulfide bonds are necessary for lipid binding and bacterial killing by SPRR2A.

### *Mice lacking SPRR2A have an altered intestinal microbiota and are more susceptible to* L. monocytogenes *infection*

The bactericidal activity of SPRR2A suggested that it might regulate intestinal microbiota composition, contact between bacteria and the intestinal surface, and/or pathogen colonization. To test for these effects in vivo, we created mice lacking SPRR2A. Mice have three copies of the *Sprr2a* gene (*Sprr2a1*, *Sprr2a2*, and *Sprr2a3*), all of which encode the same protein (figs. S2A and S8). We therefore used CRISPR-Cas9–mediated gene targeting to delete the entire mouse *Sprr2a* locus (fig. S2, A and B) and verified that *Sprr2a*$^{-/-}$ mice lacked SPRR2A transcript and protein expression (fig. S2, C to E). *Sprr2a*$^{-/-}$ mice were born in normal Mendelian ratios, were healthy when reared in a specific pathogen–free (SPF) facility, and showed normal intestinal morphology with no signs of inflammation (fig. S10A) and no increased paracellular permeability (fig. S10B).

Although overall small intestinal bacterial loads in wild-type and *Sprr2a*$^{-/-}$ littermates were similar (fig. S10C), 16*S* rRNA gene sequencing analysis of male wild-type and *Sprr2a*$^{-/-}$ littermates revealed an increased abundance of Gram-positive bacteria in the small intestinal lumen of *Sprr2a*$^{-/-}$ mice, with a marked increase in the relative abundance of *Lactobacillus*, *Turicibacter*, and *Candidatus Arthromitus* (segmented filamentous bacteria, or SFB, belonging to the class *Clostridia*). At the same time, there was a reduction in the abundance of *Bacteroidetes*, a class of Gram-negative bacteria (Fig. 3, A to C, and fig. S10D). We observed similar microbiota alterations in both male and female wild-type and *Sprr2a*$^{-/-}$ littermates under different caging conditions (fig. S11, A to F), indicating that the alterations were not caging or sex dependent. The microbiota composition in the colonic lumen was similar between *Sprr2a*$^{-/-}$ mice and their wild-type littermates (Fig. 3, A to C). This similarity may reflect the increased density of the bacterial communities in the colon, or the increased rigidity and thickness of the colonic mucus layer compared with the small intestinal mucus layer (*29*), which could limit the effect of a single AMP deficiency on the whole microbial community. These small intestinal phenotypes are consistent with our finding that SPRR2A selectively kills Gram-positive bacteria (Fig. 2A), and they indicate that SPRR2A shapes the composition of bacterial communities in the small intestinal lumen.

A key function of other microbiota-inducible intestinal AMPs is to limit contact between bacteria and the intestinal epithelial surface by restricting bacterial colonization of the mucus layer that overlays the intestinal epithelium (*22*, *23*). When analyzing mucus-associated bacteria in the small intestine, we detected an increase in the abundance of the Gram-positive SFB and a decrease in the abundance of the Gram-negative *Bacteroidetes* (Fig. 3, A and C). The differences in mucus-associated bacteria were confined to the small intestine, as we did not detect marked differences in the composition of mucus-associated bacterial communities in the colon (Fig. 3, A and C). Finally, visualization of bacteria in the small intestinal mucus layer showed more bacteria in close association with the epithelial surface of *Sprr2a*$^{-/-}$ mice than wild-type mice (Fig. 3D). The filamentous morphology of the surface-associated bacteria in *Sprr2a*$^{-/-}$ mice (Fig. 3D) suggested that they were mostly SFB, consistent with the compositional analysis of mucus-associated bacteria (Fig. 3, A and C). These observations indicate that SPRR2A regulates Gram-positive colonization of the small intestinal mucus layer and limits interactions between bacteria and the small intestinal surface.

We next assessed the susceptibility of *Sprr2a*$^{-/-}$ mice to infection with a bacterial pathogen. Oral infection of *Sprr2a*$^{-/-}$ mice with the Gram-positive pathogen *L. monocytogenes* resulted in higher bacterial burdens in mesenteric lymph nodes, liver, and spleen as compared with wild-type mice (Fig. 3E). *Sprr2a*$^{-/-}$ mice also showed higher mortality after a high-dose oral *L. monocytogenes* infection (Fig. 3F). These findings show that SPRR2A protects against infection by a Gram-positive bacterial pathogen in vivo.

### *SPRR2A expression is induced by type 2 cytokines during helminth infection of the intestine*

Helminth infections can cause numerous pathologies, including damage to the intestinal epithelium, which promotes bacterial breach of the intestinal barrier (*3*) and underscores the need for AMP expression. Prior studies have shown that SPRR2A expression increases markedly during intestinal infection by parasitic helminths such as *Trichinella spiralis*, *Nippostrongylus brasiliensis*, and *Heligmosomoides polygyrus* (*30*–*32*). Accordingly, the abundance of *Sprr2a* transcripts increased in the duodenum and proximal jejunum when we infected C57BL/6 and BALB/c mice with *H. polygyrus* (Fig. 4A and fig. S12). The magnitude of the change was greater in BALB/c than in C57BL/6 mice (Fig. 4A and fig. S12), likely because of the higher baseline levels of SPRR2A expression in C57BL/6 mice (fig. S4D). The increase in SPRR2A expression was less pronounced in the ileum and colon (fig. S12), consistent with the fact that *H. polygyrus* resides predominantly in the duodenum and proximal jejunum of mice. After helminth infection, there were more SPRR2A-producing goblet cells in the jejunum, and SPRR2A expression in individual cells was also increased (Fig. 4B). Additionally, helminth-induced SPRR2A expression was maintained in antibiotic-treated mice and was therefore independent of the intestinal microbiota (fig. S13, A and B). These data show that helminth infection induces SPRR2A expression above the levels elicited by bacterial colonization.

Given that helminth infection induces strong type 2 immunity in the intestinal tract, we asked whether SPRR2A expression is induced by type 2 cytokines such as IL-4 or IL-13. We first addressed this question using small intestinal organoids, which are established from Lgr5$^{+}$ stem cells and develop all of the epithelial cell lineages of the in vivo mouse epithelium (*33*). SPRR2A expression increased when organoids from wild-type C57BL/6 small intestines were exposed to IL-4 and IL-13, but not when exposed to other cytokines, including interferon-γ, IL-1β, and IL-22 (Fig. 4C). Intraperitoneal injection of IL-13, but not IL-22, into wild-type BALB/c mice also increased expression of SPRR2A in the small intestine (Fig. 4D), demonstrating

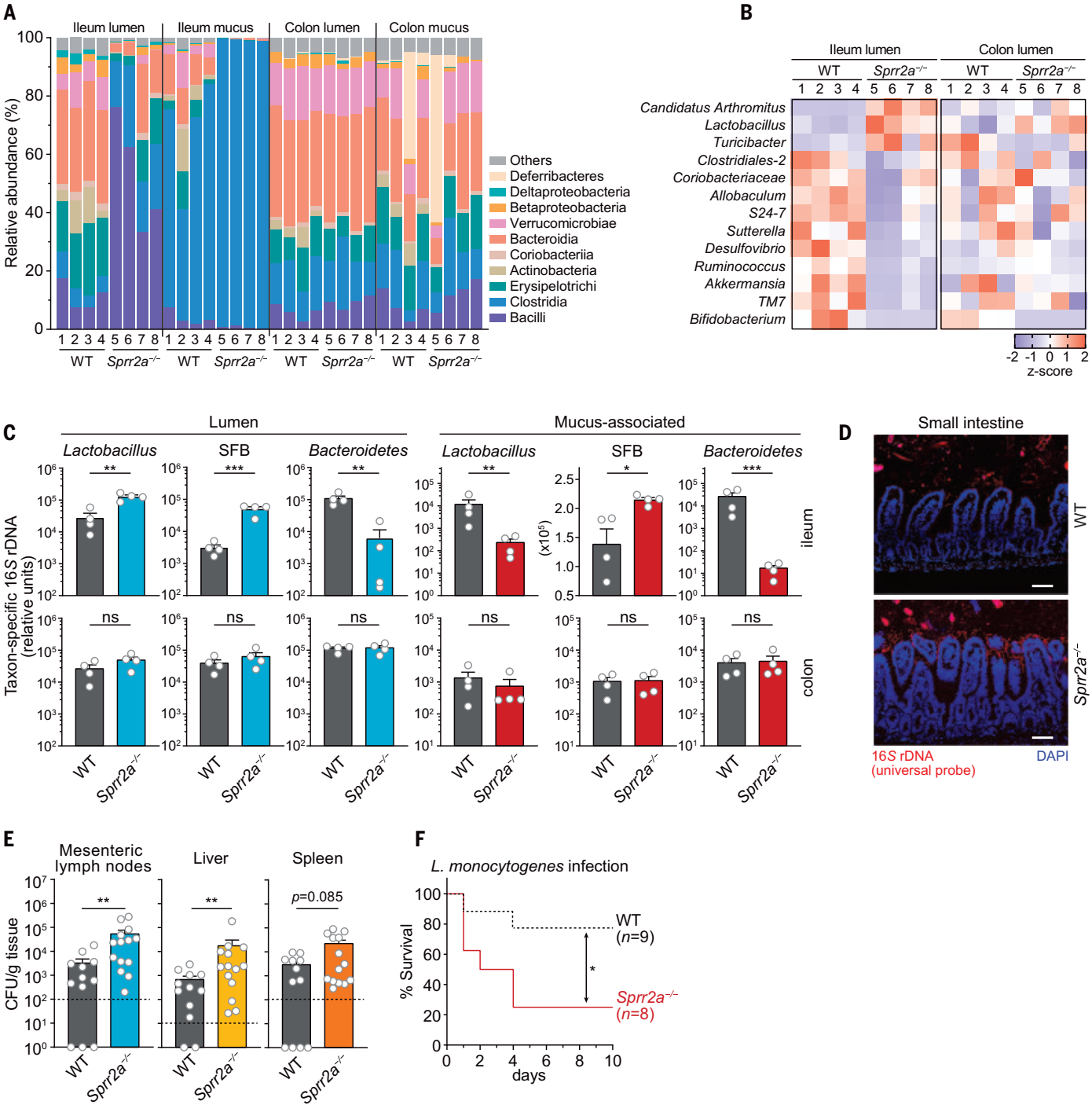



**Fig. 3. Mice lacking SPRR2A have an altered intestinal microbiota and are more susceptible to *L. monocytogenes* infection.** (**A**) Male WT and *Sprr2a*$^{-/-}$ littermates were separated at weaning and caged by genotype for 6 weeks. Intestinal mucus–associated and luminal microbial communities were characterized by 16*S* rRNA sequencing. Relative abundances of bacterial classes in WT and *Sprr2a*$^{-/-}$ mice are shown. $n = 4$ mice per group. (**B**) Relative abundances of bacterial genera in WT and *Sprr2a*$^{-/-}$ mice. (**C**) qPCR analysis of specific bacterial groups in the intestinal mucus–associated and luminal microbial communities of wild-type and *Sprr2a*$^{-/-}$ littermates. Values for each bacterial group are expressed relative to total 16*S* rDNA levels. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; ns, not significant by two-tailed *t* test. (**D**) Fluorescence in situ hybridization (FISH) detection of bacteria in the small intestines of WT and *Sprr2a*$^{-/-}$ mice. Scale bars, 50 μm. (**E**) 8- to 10-week-old WT ($n = 12$) and *Sprr2a*$^{-/-}$ mice ($n = 14$) were treated with antibiotics and orally infected with $1 \times 10^9$ CFU of log-phase *L. monocytogenes*. Liver, spleen, and mesenteric lymph nodes (MLN) were collected after 24 hours, and bacterial counts were determined by dilution plating. Results were pooled from three independent experiments. $^{**}P < 0.01$ by Mann-Whitney *U* test. Dotted line indicates limit of detection. (**F**) 8- to 10-week-old WT ($n = 9$) and *Sprr2a*$^{-/-}$ mice ($n = 8$) were orally infected with $2.5 \times 10^9$ CFU of log-phase *L. monocytogenes*. Survival rates were monitored over 10 days. Results were pooled from two independent experiments. $^{*}P < 0.05$ by the log-rank test. SFB, segmented filamentous bacteria; CFU, colony-forming units.

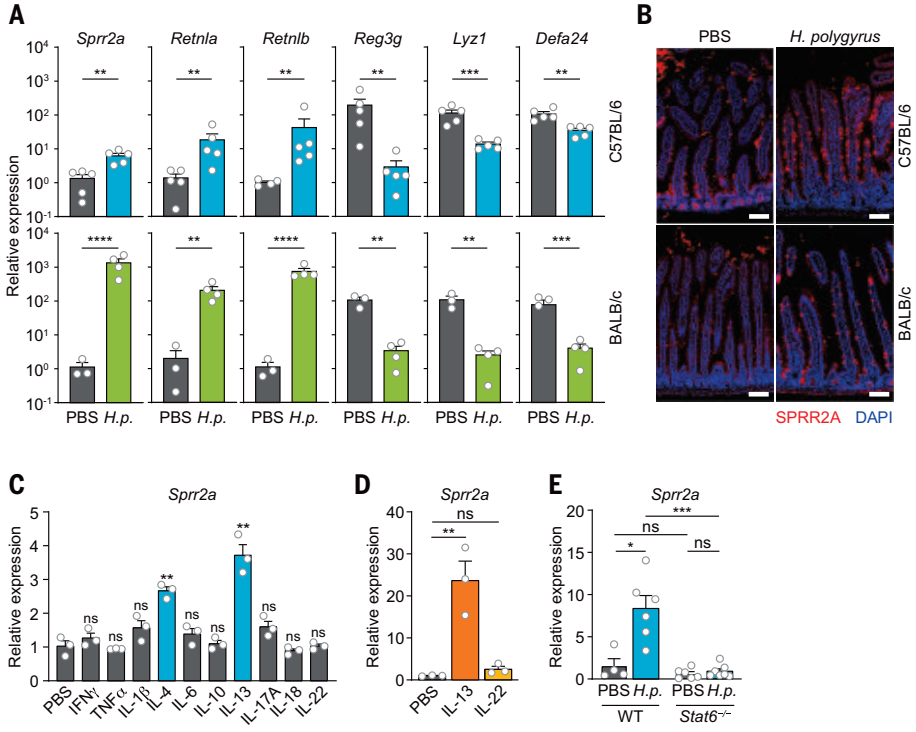



**Fig. 4. SPRR2A expression is induced by type 2 cytokines during helminth infection of the intestine.** (**A**) WT C57BL/6 (top) and BALB/c (bottom) mice were orally infected with 200 infective L3 *H. polygyrus* larvae for 2 weeks, and gene expression in the proximal jejunum was analyzed by qPCR. All values were normalized to *Gapdh* expression. (**B**) WT C57BL/6 or BALB/c mice were orally infected with 200 infective L3 *H. polygyrus* larvae for 2 weeks. Control littermates were treated with PBS. SPRR2A was detected by immunofluorescence in sections of proximal jejunum. Nuclei were detected with DAPI. Scale bars, 100 μm. Representative images are shown for 3 or 4 mice per group. (**C**) qPCR analysis of *Sprr2a* expression in small intestinal organoids derived from WT C57BL/6 mice treated with the indicated cytokines. Values were normalized to *Gapdh* expression. (**D**) qPCR analysis of intestinal epithelial *Sprr2a* expression in WT BALB/c mice intraperitoneally injected with recombinant IL-13, IL-22, or vehicle. Intestinal epithelial cells were harvested by laser-capture microdissection. Values were normalized to *Gapdh* expression. $n = 3$ mice per group. (**E**) WT and *Stat6*$^{-/-}$ mice were orally infected with 200 infective L3 *H. polygyrus* larvae, and *Sprr2a* expression in the proximal jejunum was analyzed by qPCR after 2 weeks. All values were normalized to *Gapdh* expression. Means ± SEM (error bars) are plotted. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$; ns, not significant per two-tailed *t* test.

that a type 2 cytokine can induce SPRR2A expression in vivo. Thus, we conclude that type 2 cytokines increase SPRR2A expression in the small intestine.

IL-4 and IL-13 induce type 2 immunity during helminth infection through the transcription factor STAT6. We therefore studied *Stat6*$^{-/-}$ mice to further assess the involvement of type 2 cytokines in boosting *Sprr2a* expression during helminth infection. Although *H. polygyrus* infection elicited a ninefold increase in *Sprr2a* transcript abundance in the proximal jejunum of wild-type C57BL/6 mice, there was no increase in *Sprr2a* transcript abundance in infected *Stat6*$^{-/-}$ mice (Fig. 4E). However, microbiota-dependent baseline expression of *Sprr2a* was maintained in *Stat6*$^{-/-}$ mice (fig. S13C). These results indicate that type 2 cytokines and their downstream signaling pathways induce *Sprr2a* expression during helminth infection but are not required for baseline induction of *Sprr2a* expression by the microbiota.

### *SPRR2A protects against helminth-induced bacterial invasion of intestinal tissue*

We next assessed the impact of SPRR2A on intestinal helminth infection and the ensuing type 2 immune response. Intestinal worm burden and egg counts were similar in wild-type and *Sprr2a*$^{-/-}$ mice after *H. polygyrus* infection (Fig. 5A). Additionally, markers of intestinal type 2 immunity, including *Il13* expression, goblet cell numbers, and tuft cell numbers, were increased in both wild-type and *Sprr2a*$^{-/-}$ mice (fig. S14, A to F). Therefore, SPRR2A does not directly affect the ability of *H. polygyrus* to establish an infection nor does it alter the type 2 immune response. However, it remains possible that SPRR2A could affect worm fitness or expulsion at later time points.

Although *H. polygyrus* infection increased *Sprr2a* expression in the small intestine, there was decreased expression of transcripts encoding other AMPs, including REG3B, REG3G, lysozyme, intelectin-1, and several α-defensins (Fig. 4A and fig. S15), consistent with prior findings (*32*). This suggested that SPRR2A may be essential for defense of the intestinal barrier during *H. polygyrus* infection. To test this idea, we infected wild-type and *Sprr2a*$^{-/-}$ mice with *H. polygyrus* for 2 weeks and then assessed bacterial burdens in host tissues. Immunofluorescence and 16*S* quantitative polymerase chain reaction (PCR) analysis of *H. polygyrus*–infected *Sprr2a*$^{-/-}$ mice revealed more bacteria in intestinal tissues than in infected wild-type mice (Fig. 5, B and C) and higher bacterial burdens in *Sprr2a*$^{-/-}$ mesenteric lymph nodes (Fig. 5C). The tissue-associated bacteria in *Sprr2a*$^{-/-}$ mice were characterized by an increased abundance of Gram-positive *Bacilli* and a decreased abundance of Gram-negative *Bacteroidia* (fig. S16, A to D), consistent with the selective killing of Gram-positive bacteria by SPRR2A in vitro (Fig. 2A). These data reveal that SPRR2A is essential for resistance to bacterial invasion of intestinal tissues during *H. polygyrus* infection (fig. S17).

## Discussion

We have identified SPRR2A as an intestinal antibacterial protein that is unrelated to any previously discovered mammalian AMP. Other SPRR proteins are expressed in the lung and skin of mice and humans and are induced by infection, thus suggesting that SPRR2A may be representative of a previously unknown family of AMPs that defend mammalian body surfaces.

Although this is the first report of bactericidal activity in a member of the SPRR family, other proline-rich AMPs, unrelated to the SPRR family, have been identified in flies, frogs, cows, and sheep, but not in mice or humans (*34*). These proteins selectively target Gram-negative bacteria and use a different mechanism of bacterial killing that involves translocation across the bacterial membrane and inhibition of specific cellular processes in bacterial cells. Thus, SPRR2A differs from previously characterized proline-rich AMPs by disrupting bacterial membranes and selectively targeting Gram-positive bacteria.

SPRR2A contributes to intestinal innate immunity by shaping microbiota composition, restricting bacterial access to host tissues, and limiting infection by the Gram-positive bacterial pathogen *L. monocytogenes*. A distinctive aspect of SPRR2A function is its essential role in defending the intestinal barrier against bacterial invasion during helminth infection. Although helminth infection lowers the expression of a number of intestinal AMPs,

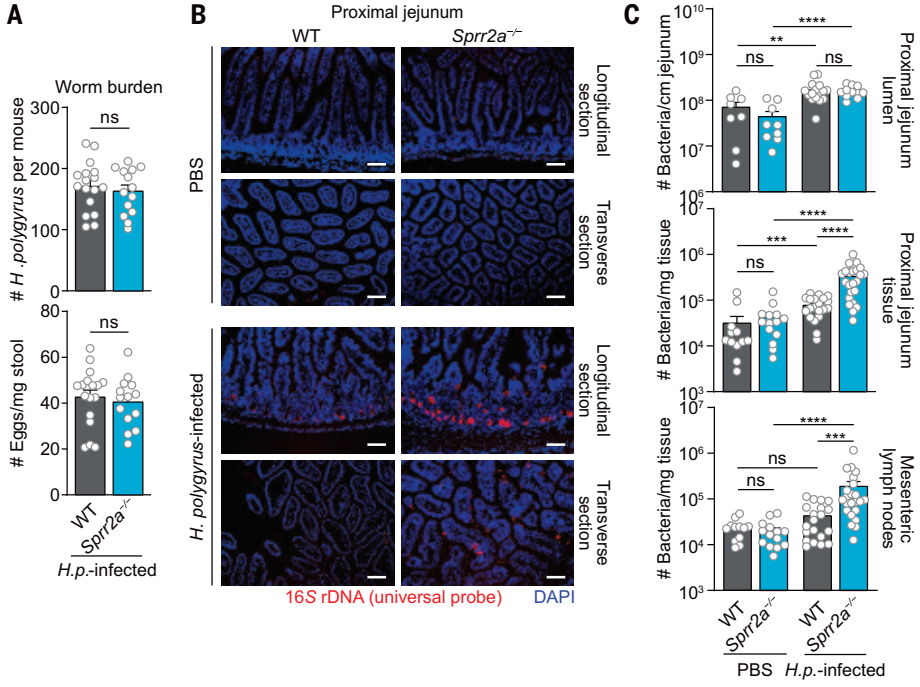



**Fig. 5. SPRR2A protects against helminth-induced bacterial invasion of intestinal tissue.** (**A**) WT and *Sprr2a*$^{-/-}$ mice were orally infected with 200 infective L3 *H. polygyrus* larvae for 2 weeks, and helminth burden in the intestine (upper panel) and egg numbers in the stool (lower panel) were counted. Results were pooled from two independent experiments. (**B**) Detection of bacteria by fluorescence in situ hybridization in the proximal jejunum of WT and *Sprr2a*$^{-/-}$ mice after *H. polygyrus* infection for 2 weeks. Nuclei were detected with DAPI. Results are representative of three independent experiments. (**C**) qPCR determination of total bacterial numbers in the proximal jejunal tissue, proximal jejunal lumen, and mesenteric lymph nodes of WT and *Sprr2a*$^{-/-}$ mice after *H. polygyrus* infection for 2 weeks. Values were normalized to tissue weight or length. Results are from three independent experiments. Means ± SEM (error bars) are plotted. $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; ns, not significant by two-tailed *t* test. *H.p.*, *Heligmosomoides polygyrus*.

SPRR2A expression is selectively increased by the type 2 cytokines that are abundantly produced during helminth infection. At the same time, RELMβ, which preferentially kills Gram-negative bacteria, is also up-regulated after *H. polygyrus* infection (Fig. 4A and fig. S15) (*23*, *32*, *35*). This suggests that SPRR2A may work in concert with RELMβ to kill both Gram-positive and Gram-negative bacteria and thus maintain the intestinal barrier during helminth infection. The elevated production of these two antibacterial proteins during helminth infection could help explain how helminth-infected mice resist colonization by bacteria that promote inflammation (*36*). These findings also illuminate how different AMPs may be deployed in distinctive immunological settings and provide insight into why a diverse repertoire of AMPs has evolved to defend the intestine.

## Materials and methods

### Mice

Conventionally raised C57BL/6 wild-type, *Sprr2a*$^{-/-}$, *Myd88*$^{-/-}$ (*37*), and *Stat6*$^{-/-}$ mice (*38*) and conventionally raised BALB/c wild-type mice were bred and maintained in the SPF barrier facility at the University of Texas (UT) Southwestern Medical Center at Dallas. The generation of *Sprr2a*$^{-/-}$ mice is described below. *Stat6*$^{-/-}$ mice (*38*) were acquired from Jackson Laboratory (stock number 005977, B6.129S2(C)-*Stat6*$^{tm1Gru}$/J; and stock number 002828, C.129S2-Stat6$^{tm1Gru}$/J). Germ-free C57BL/6 wild-type and *Myd88*$^{-/-}$ mice and Swiss Webster wild-type mice were bred and maintained in the gnotobiotic mouse facility at the UT Southwestern Medical Center. All mice were housed under a 12-hour-light:12-hour-dark cycle. Mice were fed ad libitum with free access to drinking water according to protocols approved by the Institutional Animal Care and Use Committees (IACUC) of UT Southwestern Medical Center.

*Sprr2a*$^{-/-}$ mice lacking all three mouse *Sprr2a* genes (*Sprr2a1*, *Sprr2a2*, and *Sprr2a3*) were generated using CRISPR-Cas9 genome editing with guide RNAs that targeted regions upstream and downstream of the *Sprr2a* locus (fig. S2, A to E). Guide RNAs were injected into fertilized C57BL/6J embryos along with in vitro transcribed Cas9 mRNA by the UT Southwestern Transgenic Core facility. Healthy blastocysts were implanted into pseudo-pregnant mice. The resulting litters were screened by genomic sequencing to detect the deletion of the *Sprr2a* locus, and mice harboring the deleted allele were bred to homozygosity. *Sprr2a*$^{-/-}$ mice were then backcrossed with wild-type C57BL/6 mice for five generations.

### Cell lines and bacterial strains

HT-29 and LS513 cells were obtained from ATCC and were cultured in Dulbecco's modified Eagle's medium (DMEM, Thermo Fisher) and Roswell Park Memorial Institute 1640 medium (RPMI 1640, ATCC), respectively. Both media were supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 10 U/ml Penicillin-Streptomycin (Thermo Fisher). Cells were cultured in the $CO_2$ incubator at 37°C.

Bacterial strains, including *E. faecalis*, *L. monocytogenes* strain EGDe, *E. coli* strain K-12, *B. thetaiotaomicron* strain VPI-5482, and *C. rodentium* strain DBS100, were obtained from ATCC. *L. reuteri* strain 100-23 was a gift from J. Walter (University of Alberta, Canada) (*39*). *E. faecalis* and *L. monocytogenes* were grown in brain-heart infusion (BHI) broth, *E. coli* and *C. rodentium* were grown in Luria broth (LB), *L. reuteri* was grown in Lactobacilli MRS broth in an anaerobic chamber, and *B. thetaiotaomicron* was grown in tryptone yeast extract glucose (TYG) medium anaerobically.

### Antibodies and chemicals

Polyclonal anti-mouse SPRR2A was produced in rabbits by Pacific Immunology (Ramona, CA). The synthesized peptide (CLPSVWPGP) was conjugated to a carrier protein and the conjugate was used to generate the antibody. Other antibodies were purchased from various vendors: anti-human SPRR2A antibody (Abcam, ab125385), anti-MUC2 antibody (Thermo Fisher, MA5-12345), anti-Lysozyme antibody (Santa Cruz, sc-27958), anti-DCLK1 antibody (Abcam, ab202754), anti-GAPDH antibody (Thermo Fisher, PA1-988), anti-β-actin antibody (Cell Signaling, 8457S), anti-β-tubulin (Cell Signaling, 2128S). Fluorescein isothiocyanate (FITC)-conjugated UEA-1 lectin was purchased from GeneTex (GTX01512).

### Isolation of intraepithelial lymphocytes (IELs) and lamina propria lymphocytes (LPLs)

Small intestines were dissected from mice, flushed with the ice-cold PBS, and cut open longitudinally. The intestines were cut into small sections that were thoroughly washed with ice-cold PBS. Tissues were incubated at 37°C in Hank's buffered salt solution (HBSS) supplemented with 3% FBS, 1 mM EDTA, and 1 mM DTT. After 30 min, tissues were vortexed for 2 min, and the cell suspensions containing intestinal epithelial cells and the IELs were filtered through a 100-μm cell strainer followed by passage through a glass wool column (Ohio Valley). The cell suspensions were washed and

suspended in ice-cold RPMI 1640 and then applied to a 40%:80% Percoll gradient (GE Healthcare). After centrifugation at 800*g* for 20 min, the IELs were collected from the Percoll gradient interface and prepared for RNA isolation and purification.

For LPL isolation, the residual tissue fragments were digested for 1 hour at 37°C in RPMI 1640 containing 0.025 mg/ml of Collagenase IV (Sigma-Aldrich), 0.05 mg/ml of DNase I (Sigma-Aldrich), and 0.25 units/ml of Dispase (BD Biosciences). Cells were filtered through 40-μm cell strainers and applied to a 40%:80% Percoll gradient (GE Healthcare). After centrifugation, the LPL were collected from the gradient interface and prepared for RNA isolation and purification.

### *Sample collection from human patients*

Procedures involving human subjects were approved by the UT Southwestern Medical Center's Institutional Review Board (STU 112010-130). Written informed consent was obtained from all human participants or legal guardians of participating minors.

Briefly, stool was collected at the time of colonoscopy through direct aspiration of colonic contents. A total of 10 ml was collected per individual. Samples were aliquoted into 1-ml cryovials and immediately frozen in liquid nitrogen. Colonic mucosal tissue samples were collected at time of colonoscopy by standard endoscopic tissue sampling technique. A total of two standard forceps biopsies were obtained per individual. Immediately after collection, samples were placed in RNA*later* stabilization solution (Invitrogen) and stored at −80°C. RNA extraction was performed using the Qiagen RNeasy mini kit on thawed RNA*later* stabilized tissues following the manufacturer's protocol.

### *Laser capture microdissection and RNA purification*

Epithelial cells were captured from mouse small intestines by laser capture microdissection using an Arcturus PixCell IIe system as previously described (*40*). Total RNA was isolated from the captured cells using the PicoPure RNA Isolation Kit (Thermo Fisher, KIT0204) following the manufacturer's protocol. RNA from intact tissue or cells was isolated with the RNeasy Plus Universal Mini Kit (Qiagen, 73404) following the manufacturer's protocol.

### *Quantitative real-time PCR (qPCR) analysis of gene expression*

cDNA was synthesized from extracted RNA using M-MLV Reverse Transcriptase (Thermo Fisher, 28025021) following the manufacturer's protocol. qPCR was performed on a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems) as previously described (*41*). Transcript abundances were normalized to *Gapdh* or 18*S* ribosomal RNA transcript abundance, and relative expression values were calculated by the comparative Ct (ΔΔCt) method. Primer sequences are provided in table S3.

### *In situ hybridization (ISH)*

Plasmids containing the entire open reading frame of the mouse *Sprr2a1* gene (NCBI accession: NM_011468.4) with 252 nucleotides were used as templates of RNA probes. Digoxigenin (DIG)–labeled antisense or sense probes were prepared with the HiScribe T7 Quick High Yield RNA Synthesis Kit (NEB, E2050S) following the manufacturer's protocol. The concentrations of riboprobe were determined on a Nanodrop spectrophotometer, and the DIG labeling was verified by dot blot using peroxidase-conjugated anti-DIG antibody (Roche, 11207733910).

Paraformaldehyde-fixed paraffin-embedded ileum tissue sections were washed twice in xylene followed by rehydration in decreasing concentrations of ethanol. Tissue sections were permeabilized with 20 μg/ml of proteinase K for 10 min and treated with ice-cold 20% acetic acid for 20 s. The sections were then prehybridized with 50% formamide, 5X saline sodium citrate, 1 mg/ml of yeast tRNA (Roche, 10109517001), 0.1 mg/ml of heparin (Sigma-Aldrich, H3149), 1X Denhardt's solution (Sigma-Aldrich, D2532), and 0.1% Tween-20 at 60°C for 2 hours and hybridized with 0.5 μg/ml riboprobe at 60°C for 16 hours. After washing, sections were incubated with peroxidase-conjugated anti-DIG antibody (Roche, 11207733910), followed by biotin-labeled tyramide (Thermo Fisher, B40951) for signal amplification. Hybridized probes were detected by ABC-Alkaline phosphatase (Vector Laboratories, AK-5000) and NBT/BCIP (Roche, 11681451001). Images were captured using a Zeiss Axio Imager M1 microscope with Zeiss AxioCam MRm Rev3 and Zeiss EC Plan-Neofluar 10X/0.30 M27 objective lens.

### *Intestinal organoid culture*

Organoids were cultured from small intestinal crypts recovered from 6- to 8-week-old mice, using a protocol based on previously described methods (*33*). The protocol was carried out as previously described (*42*), except that organoid cultures were passaged every 4 to 7 days.

### *Immunofluorescence microscopy*

For paraffin-embedded sections, mouse intestinal tissues were washed with PBS, fixed in freshly made 4% paraformaldehyde (PFA) at room temperature for 6 hours, and then embedded in paraffin by the UT Southwestern histology core. Sections were washed twice in xylene followed by rehydration in decreasing concentrations of ethanol. Antigen retrieval was performed by boiling in 10 mM sodium citrate for 15 min followed by washing in PBS.

For frozen sections, mouse intestinal tissues or organoids were snap-frozen in optimum cutting temperature (OCT) compound (Fisher, 23-730.571). Ten-micron frozen sections were cut, fixed in 4% PFA at room temperature for 15 min, and washed in PBS. After washing, slides were blocked with 10% FBS, 1% bovine serum albumin (BSA), 1% Trixon X-100 in PBS for 1 hour and incubated with the following primary antibodies at 4°C overnight: anti-mouse SPRR2A at 1:100 dilution, anti-Lysozyme at 1:200 dilution, and FITC-conjugated UEA-1 at 20 μg/ml. Secondary antibodies AlexaFluor 488/647 (Thermo Fisher) were diluted 1:400 and applied to slides for 1 hour at room temperature in the dark in a humidified chamber. Slides were washed and mounted with DAPI Fluoromount-G (Southern Biotechnology, 0100-20). Images were captured using a Zeiss Axio Imager M1 microscope with Zeiss AxioCam MRm Rev3 and Zeiss EC Plan-Neofluar 10X/0.30 M27 objective lens.

### *Fluorescence in situ hybridization (FISH)*

Mouse intestinal tissues were cut, directly fixed in Methacarn fixative solution at room temperature for 4 hours, and then embedded in paraffin by the UT Southwestern histology core. Sections were washed twice in xylene followed by rehydration in decreasing concentrations of ethanol. The universal bacterial probe (/5Alex594N/GCTGCCTCCCGTAGGAGT/3AlexF594N/) or the control nonspecific probe (/5Alex594N/ACTCCTACGGGAGGCAGC/3AlexF594N/) was diluted to 100 nM in FISH hybridization buffer (20 mM Tris pH 7.2, 0.9 M NaCl, 0.1% SDS) and applied to slides overnight at 56°C in a humidified chamber. Slides were washed and mounted with DAPI Fluoromount-G (Southern Biotechnology, 0100-20). Images were captured using a Discover Echo Revolve microscope.

### *Immunoblot*

For protein extraction from intestinal tissue and stool samples, samples were added to Lysine Matrix E tubes (MP Biomedicals, 116914050), and then 1 ml of 1X SDS loading buffer (2% SDS, 10% glycerol, 100 mM DTT, 0.1% bromophenol blue, and 50 mM Tris-HCl pH 6.8) was added. Tubes were then added to a FastPrep-24 5G Homogenizer. After homogenization, the lysates were boiled at 100°C for 10 min and then centrifuged for 10 min at 4°C at 16,100*g*. The supernatants were then transferred to clean tubes.

For protein extraction from cultured cells and intestinal organoids, cells were resuspended in 100 μl RIPA buffer (Thermo Fisher, 89900), and then 100 μl of 2X SDS loading buffer (4% SDS, 20% glycerol, 200 mM DTT, 0.2% bromophenol blue, and 100 mM Tris-HCl

pH 6.8) was added. Tubes were boiled at 100°C for 10 min and then centrifuged for 10 min at 4°C at 16,100*g*. The supernatants were then transferred to clean tubes.

The supernatants were separated using a 4 to 20% gradient SDS–polyacrylamide gel electrophoresis (PAGE) gel (Bio-Rad) then transferred to a polyvinylidene fluoride membrane. Membranes were blocked with 5% nonfat milk in TBS-T buffer (0.1% Tween-20 in tris-buffered saline) then sequentially incubated with primary antibodies and appropriate horseradish peroxidase (HRP)–conjugated secondary antibodies. Protein bands were detected with enhanced chemiluminescence (ECL) reagent (Bio-Rad, 1705060) using a Bio-Rad ChemiDoc system.

### *Recombinant SPRR2A protein expression and purification*

Human *SPRR2A* (NCBI accession: NM_005988.3) containing a C-terminal 6xHis tag was cloned into a pFastBac1 vector and heterologously expressed in Sf9 cells (Thermo Fisher). One liter of cells ($2.5 \times 10^6$ cells/ml) was infected with 10 ml of baculovirus at 28°C. Cells were cultured in suspension at 28°C and were harvested after 48 hours of infection. Harvested cells were resuspended in the buffer containing 25 mM Tris-HCl pH 8.0, 150 mM NaCl, and 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication. The mixture was pelleted by centrifugation at 10,000*g* for 30 min and the supernatant was loaded onto a $Ni^{2+}$ metal affinity column (Qiagen). Nonspecific contaminants were washed away with buffer containing 30 mM imidazole, and the protein was eluted in buffer containing 300 mM imidazole. The eluate was concentrated in a 3K cutoff Amicon Ultra centrifugal device (Millipore) and further purified by size exclusion chromatography on a Superdex 75 10/300 GL column (GE Healthcare Life Sciences) in 10 mM MES pH 5.5 and 25 mM NaCl.

SPRR2A protein was also expressed in *E. coli* BL21-CodonPlus (DE3)–RILP cells (Stratagene) to obtain protein that was free of disulfide bonds. The human *SPRR2A* gene was cloned into the pGEX-6P-1 expression vector (GE Healthcare Life Sciences). The amplicon was placed between the BamHI and XhoI restriction endonuclease sites with an N-terminal glutathione *S*-transferase (GST) tag followed by a PreScission protease cleavage site and a C-terminal stop codon. Expression of SPRR2A was induced with 0.4 mM isopropyl-β-D-galactoside (IPTG) for 12 hours at 18°C. Cells were harvested, resuspended in buffer containing 25 mM Tris-HCl pH 8.0 and 150 mM NaCl and lysed by sonication. The mixture was pelleted by centrifugation at 10,000*g* for 30 min, and the supernatant was loaded onto the Glutathione Sepharose 4B GST-tagged protein purification resin (GE Healthcare Life Sciences). After four washes with the resuspension buffer, PreScission protease (GE Healthcare Life Sciences) was added to the resin and incubated overnight at 4°C to remove the GST tag. The flow-through, which contained only the untagged SPRR2A protein, was collected and was further purified by size exclusion chromatography on a Superdex 75 10/300 GL column (GE Healthcare Life Sciences) in 10 mM MES pH 5.5 and 25 mM NaCl.

### *Mass spectrometry*

Proteins were analyzed by liquid chromatography–mass spectrometry, using a Sciex X500B Q-TOF mass spectrometer coupled to an Agilent 1290 Infinity II HPLC. The protocol was carried out as previously described (*43*), except for the controller settings: Ion source gas 1 30 psi, ion source gas 2 30 psi, curtain gas 35, CAD gas 7, temperature 300°C, spray voltage 5500 V, declustering potential 125 V, collision energy 10 V. Data were acquired from 400 to 2000 Da with a 0.5-s accumulation time.

### *Bacterial killing assays*

Bacterial killing assays were performed as previously described (*21*, *23*). Briefly, 10-ml bacterial cultures were grown to midlogarithmic phase and then pelleted and washed in 10 ml of standard assay buffer (10 mM MES pH 5.5 and 25 mM NaCl). Purified SPRR2A proteins were added in varying concentrations (0 to 10 μM) to ~$5 \times 10^6$ colony-forming units (CFU)/ml bacteria. Bacteria were incubated at 37°C for 2 hours, plated onto appropriate agar plates at various dilutions, and incubated overnight at 37°C. Surviving colonies were counted and calculated as a percentage of the remaining colonies in the buffer only sample.

### *Bacterial binding assays*

Bacterial cultures (10 ml) were grown to midlogarithmic phase. Cells were pelleted, washed, and resuspended in 1 ml of 90% ethanol. Bacteria were fixed for 20 min at room temperature, then pelleted and washed twice in PBS. The cells were resuspended in standard assay buffer (10 mM MES pH 5.5 and 25 mM NaCl), and 10 μM purified SPRR2A was added to ~$5 \times 10^8$ CFU/ml bacteria with a total volume of 200 μl. After a 10-min incubation at room temperature, 20 μl of each sample was collected as input (I). Samples were then centrifuged at 4°C for 10 min at 8000*g*. The supernatants (S) were collected to examine proteins that were not bound to bacteria. The pellets (P), which contained the bound protein, were washed twice with 200 μl of standard assay buffer by recentrifugation and then resuspended in 180 μl of buffer. The I, S, and P fractions were analyzed by SDS-PAGE and Coomassie blue staining.

### *Propidium iodide (PI) uptake assay*

Propidium iodide (PI) uptake assays were performed as previously described (*23*, *41*). Briefly, bacterial cultures were grown to midlogarithmic phase and then pelleted and washed in standard assay buffer (10 mM MES pH 5.5 and 25 mM NaCl). Bacteria were then diluted into standard assay buffer (~$5 \times 10^8$ cells/ml) containing 5.5 μg/ml of PI (Thermo Fisher, P3566). Bacterial samples (90 μl each) were added to black 96-well Costar plates (Fisher, 07-200-567) and placed into a Spectramax plate reader (Molecular Devices) that was preequilibrated to 37°C. After an initial reading, 10 μl of recombinant purified SPRR2A proteins at varying concentrations or BSA was added and fluorescence outputs (excitation, 535 nm; emission, 617 nm) were measured every 5 min for 1 hour. Bacterial permeabilization activity was measured against the maximum fluorescence output from the positive control (0.05% SDS).

### *Lipid strip assay*

Membrane lipid strips (Echelon, P-6002) were used following the manufacturer's protocol. Briefly, the lipid strips were blocked with blocking buffer (10 mM MES pH 5.5, 25 mM NaCl, 2% BSA, and 0.05% Tween-20) for 1 hour at room temperature. Recombinant purified SPRR2A proteins were diluted to 1 μg/ml in blocking buffer and incubated with the lipid strip overnight at 4°C. After washing three times with washing buffer (10 mM MES pH 5.5, 25 mM NaCl, and 0.05% Tween-20), the lipid strip was sequentially incubated with anti-SPRR2A antibody and HRP-conjugated secondary antibody. Protein dots were detected with ECL reagent (Bio-Rad, 1705060) using a Bio-Rad ChemiDoc system.

### *Liposome preparation*

Unilamellar liposomes were prepared using lipids from Avanti Polar Lipids: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (PC) (850457C), 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (PS) (840035C), 1,3-bis[1,2-dimyristoyl-sn-glycero-3-phospho]-glycerol, or cardiolipin (710332C). The lipids were prepared as previously described (*23*, *27*). Lipids dissolved in chloroform were mixed in defined molar ratios in glass tubes and then dried under a stream of $N_2$, followed by drying under vacuum overnight to ensure complete removal of organic solvents. Dried lipids were resuspended in standard assay buffer (10 mM MES pH 5.5, 25 mM NaCl), or standard assay buffer containing 5(6)-carboxyfluorescein (CF) (Sigma, 21877) or fluorescein-dextran 10 (FD10) (Thermo Fish, D1821). Lipids were transferred to a 2-ml freezing vial, subjected to five freeze-thaw cycles in liquid $N_2$, and stored at −80°C. Lipids were then thawed and passed through a 100-nm pore membrane using a mini-extruder kit (Avanti Polar Lipids, 610000). Dye-free or CF-loaded liposomes were purified on a PD-10 column (GE Healthcare Life Sciences), and liposomes loaded with FD10 were purified by size exclusion chromatography on a Superdex 200 Increase 10/300 GL column (GE Healthcare Life Sciences).

### Liposome binding assay

Recombinant purified SPRR2A (5 μM) was incubated with liposomes (500 μM lipids) at room temperature for 30 min in a total volume of 200 μl. Twenty-microliter samples were collected as input (I). Samples were then centrifuged in a Beckman Optima XE-90 ultracentrifuge at 4°C for 30 min at 100,000*g*. The supernatant (S) was collected to examine proteins not bound to the liposome. The pellets (P), which contained the bound protein, were washed twice with 200 μl of buffer by recentrifugation and then resuspended to 180 μl. The I, S, and P fractions were analyzed by SDS-PAGE and Coomassie blue staining.

### Liposome leakage assay

Prepared CF or FD10 loaded liposomes were diluted in standard assay buffer (10 mM MES pH 5.5 and 25 mM NaCl) to a final concentration of 1 mM. A QuantaMaster 300 fluorometer (Photon Technology International) was used to monitor fluorescence (excitation, 490 nm; emission, 516 nm). One hundred microliters of liposomes was used each time. After allowing establishment of a stable baseline for 200 s, 10 μl of recombinant purified SPRR2A proteins at varying concentrations was added, and the fluorescence was monitored for another 800 s. At the end time point, 10 μl of 10% *n*-octyl glucoside (OG) (Anatrace, O311) was added to completely disrupt the liposomes. Fluorescence was measured over time in seconds and is expressed as a percentage of total dye release by the detergent OG.

### Electron microscopy

For electron microscopy of bacteria, 10-ml bacterial cultures (*L. monocytogenes* or *E. faecalis*) were grown to midlogarithmic phase and then pelleted and washed in 10 ml of standard assay buffer (10 mM MES pH 5.5 and 25 mM NaCl). Bacteria were resuspended in 1 ml of standard assay buffer. Purified SPRR2A was added at a final concentration of 10 μM to 300 μl of resuspended bacteria and incubated at 37°C for 2 hours. Bacteria were then centrifuged for 10 min at 16,100*g*, resuspended in crosslinking reagent (4% paraformaldehyde and 5% glutaraldehyde in 0.1 M sodium phosphate buffer, pH 7.4), and incubated overnight at 4°C. After washing three times with the same buffer, bacterial pellets were embedded in 3% agarose, sliced into blocks (1 $mm^3$), and fixed with 1% osmium tetroxide and 0.8% potassium ferricyanide in 0.1 M sodium phosphate buffer for 1.5 hours at room temperature. The cells were then stained with 1% aqueous uranyl acetate for 1 hour. Cells were dehydrated stepwise through increasing ethanol concentrations and then transitioned into propylene oxide. The cells were permeated with Embed-812 resin and polymerized at 60°C overnight. Blocks were sectioned on a Leica Ultracut 7 ultramicrotome (Leica Microsystems) using a diamond knife (Diatome). The sections were then deposited onto copper grids and stained with 2% aqueous uranyl acetate and lead citrate. Images were captured on a Tecnai G2 spirit transmission electron microscope (Thermo Fisher) using a LaB6 source at 120 kV.

For electron microscopy of liposomes, recombinant purified SPRR2A (5 μM) was incubated with liposomes (500 μM lipids) at room temperature for 30 min. Aliquots of the mixture (5 μl) were transferred to carbon support films on electron microscopy grids (Electron Microscopy Sciences, FCF200-CU) and negatively stained with 2% uranyl acetate. Images were acquired on a Tecnai G2 Spirit transmission electron microscope (Thermo Fisher).

### DNA extraction for 16S rRNA analysis

For isolation of small intestinal luminal contents, a 2-cm section of ileum was cut, and luminal contents were flushed with 2 ml of ice-cold PBS into a preweighed 2-ml sterile freezing vial. The contents were pelleted at 16,100*g* for 10 min, the supernatants removed, and the pellets weighed before further processing. For isolation of luminal contents from the colon, a section of midcolon was cut open longitudinally. Whole fecal pellets were extracted and weighed. For analysis of tissue-associated bacteria, the same tissue samples that were used for analysis of lumenal contents were cut open longitudinally and washed in ice cold PBS until visibly clean. The whole tissue was then weighed before further processing. For DNA extraction from isolated luminal contents, DNA was extracted and purified using the FastDNA Spin Kit (MP Biomedicals 116560-200) following the manufacturer's protocol. For DNA extraction from tissue-associated bacteria, samples were added to Lysine Matrix E tubes (MP Biomedicals #116914050), and then 1 ml of lysis buffer (1X Tris-EDTA buffer, 0.5% SDS, and 200 μg/ml of proteinase K) was added. Tubes were incubated at 55°C for 1 hour, then added to a FastPrep-24 5G homogenizer. After homogenization, the lysates were centrifuged for 10 min at 4°C at 16,100*g*. The supernatants were then transferred to clean tubes, and DNA was extracted with 750 μl phenol:chloroform. The upper layer containing the DNA was precipitated with sodium acetate and ethanol and stored in TE buffer at −80°C.

### 16S qPCR analysis

16*S* qPCR analysis was performed as previously described (*22*, *23*). Briefly, to allow quantification of low copy number bacterial 16*S* DNA by qPCR, limited cycle number (LCN) PCRs were used to amplify the entire 16*S* gene. Twenty microliters of LCN PCRs was prepared using the HotStarTaq polymerase kit (Qiagen), 0.2 μM universal forward primer 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and reverse primer 1492R (5′-CGGTTACCTTGTTACGACTT-3′), and 500 ng of template DNA. After 16 thermocycles, the PCR products were diluted 1:10 into $H_2O$, and the diluted DNA samples were analyzed by qPCR using the SYBR Green kit (Thermo Fisher, 4309155) with taxon-specific primers (listed in table S3). PCR control reactions containing water were included to identify possible contamination. The 16*S* data were initially normalized to data from the water control and then normalized to the luminal content or tissue weight. The total number of 16*S* copies was determined using standard curves generated from quantified standard plasmids.

### 16S rRNA sequencing and data analysis

The hypervariable regions V3 and V4 of the bacterial 16*S* rRNA gene were sequenced and analyzed as previously described (*41*). 16*S* rRNA gene sequencing data are available from the Sequence Read Archive (SRA) under project ID PRJNA743545.

### Conventionalization of germ-free mice and LPS challenge

For conventionalization, ~50 mg of feces was collected from a conventional wild-type mouse and suspended in 1 ml of PBS. Fecal debris was pelleted, and 200 μl of the supernatant was used for oral gavage of each germ-free mouse. Mice were raised for 7 more days in contact with cage bedding collected from the SPF barrier facility at the UT Southwestern Medical Center. For LPS challenge, 8- to 10-week-old germ-free Swiss-Webster mice were given 500 μg of γ-irradiated LPS (Sigma-Aldrich, L4391) every 12 hours for 3 days through oral gavage. Mice were sacrificed 12 hours after the last treatment.

### Antibiotic treatment of mice

Conventionally raised mice were administered 200 μl of antibiotic cocktail water containing 2 mg/ml of neomycin, 2 mg/ml of gentamycin, 2 mg/ml of metronidazole, 2 mg/ml of streptomycin, and 1 mg/ml of vancomycin through oral gavage. Mice were raised for 7 more days using the drinking water containing 1 mg/ml of neomycin, 1 mg/ml of gentamycin, 1 mg/ml of metronidazole, 1 mg/ml of streptomycin, 0.5 mg/ml of vancomycin, and 10% sucrose. Microbiota depletion was verified by aerobic and anaerobic culture of fecal pellets.

### Intestinal permeability assay

Intestinal permeability assays were performed by treating mice with fluorescein isothiocyanate–dextran (FITC-dextran) (Sigma-Aldrich, FD4-1G) through oral gavage, which is normally too large to pass across the intestinal barrier. However, when there is intestinal damage, paracellular permeability is increased, and FITC-dextran penetrates into gut tissue and appears

in serum. Indomethacin (Sigma-Aldrich, I7378), a nonsteroidal anti-inflammatory drug (NSAID) that can induce intestinal damage in mice, was used as a positive control.

For the experimental group, wild-type and *Sprr2a*$^{-/-}$ mice were treated with 190 μl of 7% DMSO in PBS by oral gavage. For the positive control group, mice were treated with 190 μl of indomethacin (1.5 mg/ml in 7% DMSO in PBS) by oral gavage. After 1 hour, mice in both groups were treated with 190 μl of FITC-dextran (80 mg/ml in PBS) by oral gavage. Mice were sacrificed 4 hours later, and sera were collected. The sera were then centrifuged for 20 min at 4°C at 1740*g*, and the supernatants were collected. Serum FITC-dextran levels were measured by fluorescence microplate assay against a FITC-dextran standard curve using a Spectramax plate reader (Molecular Devices).

### L. monocytogenes *infection*

*L. monocytogenes* strain EGDe (erythromycin-resistant) was grown overnight in brain-heart infusion (BHI) broth containing 20 μg/ml erythromycin at 37°C. Wild-type and *Sprr2a*$^{-/-}$ mice were treated with erythromycin (0.4 mg per mice) by oral gavage the day before infection. Mice were inoculated with $1.0 \times 10^9$ *L. monocytogenes* for tissue dissemination experiments or $2.5 \times 10^9$ bacteria for survival experiments by oral gavage. For tissue dissemination experiments, mice were sacrificed 1 day after infection, and the mesenteric lymph nodes (MLN), liver, and spleen were collected and weighed. Tissues were homogenized in ice-cold PBS and the numbers of *L. monocytogenes* were counted by dilution plating on BHI-erythromycin agar plates. We calculated the limit of detection as [CFU (minimum)] × [dilution factor] / [tissue weight]. Data points below the limit of detection indicate that no *L. monocytogenes* colonies grew on any of the three replicate plates.

### *Recombinant IL-13 and IL-22 treatment*

Wild-type BALB/c mice were injected intraperitoneally with 1.5 μg of carrier-free recombinant mouse IL-13 (BioLegend, 575906) or IL-22 (BioLegend, 576206) every other day for a total of four treatments. Mice were sacrificed on the day after the last injection.

### Heligmosomoides polygyrus *infection*

*Heligmosomoides polygyrus* [*H. polygyrus*, recently renamed *H. bakeri* (*44*)] stocks were propagated as previously described (*45*, *46*), with a few exceptions. Briefly, B6.129S2(C)-*Stat6*$^{tm1Gru}$ (Jackson stock, 005977) (*38*) or wild-type C57BL/6 mice housed in a BSL2 animal facility were infected with 400 *H. polygyrus* L3 larvae by oral gavage. On day 13 of infection, grates were placed in the cages and feces were collected for the next 3 days. Feces were mashed and mixed with activated, washed charcoal (Sigma-Aldrich, C2764-500G), then plated on Whatman filter paper in petri dishes. PBS was added to the dishes until a pool formed around the filter paper. After 7 to 10 days of incubation in a humid box, the L3 larvae were collected with a modified Baermann apparatus filled with warm PBS. The tube of the Baermann apparatus was plugged with a stopper. A metal grate was placed at the top of the funnel, then one layer of muslin and one layer of Kimwipe (Fisher Scientific) were placed on top of that. The fecal mixture was spread on top of both layers, and the larvae were allowed to fall for 2 hours. The PBS from the apparatus was collected into a beaker and the larvae were concentrated by centrifugation at 300*g* for 10 min with no brake. After the larvae were concentrated, they were put through the Baermann apparatus and then concentrated by centrifugation a second time. The larvae were washed three times with PBS and stored at 4°C in PBS. A new stock was generated every 6 months.

For *H. polygyrus* infection experiments, wild-type, *Stat6*$^{-/-}$, and *Sprr2a*$^{-/-}$ mice were inoculated with 200 L3 *H. polygyrus* larvae through oral gavage in the BSL2 animal facility. After 2 weeks, mice were sacrificed, and the intestinal tissues were collected for further studies.

### H. polygyrus *worm and egg counting*

To quantify worm burden, small intestines and their luminal contents were collected and placed in 10 ml of PBS in Petri dishes at room temperature. The intestines were dissected longitudinally and worms associated with the intestinal tissue were collected by scraping with forceps. After 1 hour of incubation at 37°C, worms were counted manually using forceps.

For egg counting, 1 to 2 fecal pellets were collected and weighed. Three milliliters of water was added to the fecal pellets in 15-ml conical tubes and vortexed periodically until the pellets were completely dispersed. Just before each sample was added to the McMaster egg counter (Electron Microscopy Sciences, 63512-75), an equal volume of supersaturated NaCl solution was added to the tube. Eggs were counted under a microscope, and two independent counts were averaged. Egg numbers were normalized to the weight of fecal pellets.

### *Statistics*

We used two-tailed Student's *t* tests to determine the statistical significance of a difference between two treatments when a parametric test was appropriate. We used the Mann-Whitney *U* test for experiments requiring a nonparametric statistical test (e.g., Fig. 3E). For data having unequal variances among different experimental groups and following a lognormal distribution, a logarithmic transform was performed before the statistical analysis. The log-rank test was used in survival experiments where the null hypothesis was that there is no difference between experimental groups in the probability of lethal morbidity at any time point. Statistical details of experiments are provided in the figure legends, including how significance was defined and the statistical methods used. Data represent means ± standard errors of the mean (SEM). All statistical analyses were performed with GraphPad Prism software. For all tests, *P* values lower than 0.05 were considered statistically significant.

## REFERENCES AND NOTES


1. J. C. Clemente, L. K. Ursell, L. W. Parfrey, R. Knight, The impact of the gut microbiota on human health: An integrative view. *Cell* **148**, 1258–1270 (2012). doi: 10.1016/j.cell.2012.01.035; pmid: 22424233
2. A. Rapin, N. L. Harris, Helminth-bacterial interactions: Cause and consequence. *Trends Immunol.* **39**, 724–733 (2018). doi: 10.1016/j.it.2018.06.002; pmid: 29941203
3. G. Coakley, N. L. Harris, The intestinal epithelium at the forefront of host-helminth interactions. *Trends Parasitol.* **36**, 761–772 (2020). doi: 10.1016/j.pt.2020.07.002; pmid: 32713764
4. R. L. Gallo, L. V. Hooper, Epithelial antimicrobial defence of the skin and intestine. *Nat. Rev. Immunol.* **12**, 503–516 (2012). doi: 10.1038/nri3228; pmid: 22728527
5. S. Mukherjee, L. V. Hooper, Antimicrobial defense of the intestine. *Immunity* **42**, 28–39 (2015). doi: 10.1016/j.immuni.2014.12.028; pmid: 25607457
6. R. M. Maizels, M. Yazdanbakhsh, Immune regulation by helminth parasites: Cellular and molecular mechanisms. *Nat. Rev. Immunol.* **3**, 733–744 (2003). doi: 10.1038/nri1183; pmid: 12949497
7. J. E. Allen, R. M. Maizels, Diversity and dialogue in immunity to helminths. *Nat. Rev. Immunol.* **11**, 375–388 (2011). doi: 10.1038/nri2992; pmid: 21610741
8. W. C. Gause, T. A. Wynn, J. E. Allen, Type 2 immunity and wound healing: Evolutionary refinement of adaptive immunity by helminths. *Nat. Rev. Immunol.* **13**, 607–614 (2013). doi: 10.1038/nri3476; pmid: 23827958
9. D. Sorobetea, M. Svensson-Frej, R. Grencis, Immunity to gastrointestinal nematode infections. *Mucosal Immunol.* **11**, 304–315 (2018). doi: 10.1038/mi.2017.113; pmid: 29297502
10. L. V. Hooper *et al.*, Molecular analysis of commensal host-microbial relationships in the intestine. *Science* **291**, 881–884 (2001). doi: 10.1126/science.291.5505.881; pmid: 11157169
11. S. Gibbs *et al.*, Molecular characterization and evolution of the SPRR family of keratinocyte differentiation markers encoding small proline-rich proteins. *Genomics* **16**, 630–637 (1993). doi: 10.1006/geno.1993.1240; pmid: 8325635
12. A. Mueller *et al.*, Distinct gene expression profiles characterize the histopathological stages of disease in *Helicobacter*-induced mucosa-associated lymphoid tissue lymphoma. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1292–1297 (2003). doi: 10.1073/pnas.242741699; pmid: 12552104
13. F. J. Sun *et al.*, Decreased gastric bacterial killing and up-regulation of protective genes in small intestine in gastrin-deficient mouse. *Dig. Dis. Sci.* **48**, 976–985 (2003). doi: 10.1023/A:1023068116934; pmid: 12772799
14. J. B. Domachowske, C. A. Bonville, A. J. Easton, H. F. Rosenberg, Differential expression of proinflammatory cytokine genes in vivo in response to pathogenic and nonpathogenic pneumovirus infections. *J. Infect. Dis.* **186**, 8–14 (2002). doi: 10.1086/341082; pmid: 12089656
15. J. B. Vos *et al.*, A molecular signature of epithelial host defense: Comparative gene expression analysis of cultured bronchial epithelial cells and keratinocytes. *BMC Genomics* **7**, 9 (2006). doi: 10.1186/1471-2164-7-9; pmid: 16420688
16. S. Tamoutounour *et al.*, Keratinocyte-intrinsic MHCII expression controls microbiota-induced Th1 cell responses. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 23643–23652 (2019). doi: 10.1073/pnas.1912432116; pmid: 31672911
17. E. Candi, R. Schmidt, G. Melino, The cornified envelope: A model of cell death in the skin. *Nat. Rev. Mol. Cell Biol.* **6**, 328–340 (2005). doi: 10.1038/nrm1619; pmid: 15803139
18. A. L. Haber *et al.*, A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017). doi: 10.1038/nature24489; pmid: 29144463

19. M. Ghandi *et al*., Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019). doi: 10.1038/s41586-019-1186-3; pmid: 31068700
20. L. V. Hooper, T. S. Stappenbeck, C. V. Hong, J. I. Gordon, Angiogenins: A new class of microbicidal proteins involved in innate immunity. *Nat. Immunol.* **4**, 269–273 (2003). doi: 10.1038/ni888; pmid: 12548285
21. H. L. Cash, C. V. Whitham, C. L. Behrendt, L. V. Hooper, Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* **313**, 1126–1130 (2006). doi: 10.1126/science.1127119; pmid: 16931762
22. S. Vaishnava *et al*., The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255–258 (2011). doi: 10.1126/science.1209791; pmid: 21998396
23. D. C. Propheter, A. L. Chara, T. A. Harris, K. A. Ruhn, L. V. Hooper, Resistin-like molecule β is a bactericidal protein that promotes spatial segregation of the microbiota and the colonic epithelium. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11027–11033 (2017). doi: 10.1073/pnas.1711395114; pmid: 28973871
24. M. Zasloff, Antimicrobial peptides of multicellular organisms. *Nature* **415**, 389–395 (2002). doi: 10.1038/415389a; pmid: 11807545
25. E. Malik, S. R. Dennison, F. Harris, D. A. Phoenix, pH dependent antimicrobial peptides and proteins, their mechanisms of action and potential as therapeutic agents. *Pharmaceuticals* **9**, 67 (2016). doi: 10.3390/ph9040067; pmid: 27809281
26. T. Pelaseyed *et al*., The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. *Immunol. Rev.* **260**, 8–20 (2014). doi: 10.1111/imr.12182; pmid: 24942678
27. S. Mukherjee *et al*., Antibacterial membrane attack by a pore-forming intestinal C-type lectin. *Nature* **505**, 103–107 (2014). doi: 10.1038/nature12729; pmid: 24256734
28. S. C. Li, N. K. Goto, K. A. Williams, C. M. Deber, Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6676–6681 (1996). doi: 10.1073/pnas.93.13.6676; pmid: 8692877
29. M. E. Johansson *et al*., The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 15064–15069 (2008). zdoi: 10.1073/pnas.0803124105; pmid: 18806221
30. P. A. Knight *et al*., Expression profiling reveals novel innate and inflammatory responses in the jejunal epithelial compartment during infection with *Trichinella spiralis*. *Infect. Immun.* **72**, 6076–6086 (2004). doi: 10.1128/IAI.72.10.6076-6086.2004; pmid: 15385512
31. D. Voehringer *et al*., *Nippostrongylus brasiliensis*: Identification of intelectin-1 and -2 as Stat6-dependent genes expressed in lung and intestine during infection. *Exp. Parasitol.* **116**, 458–466 (2007). doi: 10.1016/j.exppara.2007.02.015; pmid: 17420014
32. L. J. Entwistle *et al*., Epithelial-cell-derived phospholipase A2 group 1B is an endogenous anthelmintic. *Cell Host Microbe* **22**, 484–493. e5 (2017). doi: 10.1016/j.chom.2017.09.006; pmid: 29024642
33. T. Sato *et al*., Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* **469**, 415–418 (2011). doi: 10.1038/nature09637; pmid: 21113151
34. M. Scocchi, A. Tossi, R. Gennaro, Proline-rich antimicrobial peptides: Converging to a non-lytic mechanism of action. *Cell. Mol. Life Sci.* **68**, 2317–2330 (2011). doi: 10.1007/s00018-011-0721-7; pmid: 21594684
35. D. Artis *et al*., RELMβ/FIZZ2 is a goblet cell-specific immune-effector molecule in the gastrointestinal tract. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13596–13600 (2004). doi: 10.1073/pnas.0404034101; pmid: 15340149
36. D. Ramanan *et al*., Helminth infection promotes colonization resistance via type 2 immunity. *Science* **352**, 608–612 (2016). doi: 10.1126/science.aaf3229; pmid: 27080105
37. O. Adachi *et al*., Targeted disruption of the *MyD88* gene results in loss of IL-1- and IL-18-mediated function. *Immunity* **9**, 143–150 (1998). doi: 10.1016/S1074-7613(00)80596-8; pmid: 9697844
38. M. H. Kaplan, U. Schindler, S. T. Smiley, M. J. Grusby, Stat6 is required for mediating responses to IL-4 and for development of Th2 cells. *Immunity* **4**, 313–319 (1996). doi: 10.1016/S1074-7613(00)80439-2; pmid: 8624821
39. S. A. Frese *et al*., The evolution of host specialization in the vertebrate gut symbiont *Lactobacillus reuteri*. *PLOS Genet.* **7**, e1001314 (2011). doi: 10.1371/journal.pgen.1001314; pmid: 21379339
40. T. S. Stappenbeck, L. V. Hooper, J. K. Manchester, M. H. Wong, J. I. Gordon, Laser capture microdissection of mouse intestine: Characterizing mRNA and protein expression, and profiling intermediary metabolism in specified cell populations. *Methods Enzymol.* **356**, 167–196 (2002). doi: 10.1016/S0076-6879(02)56932-9; pmid: 12418197
41. T. A. Harris *et al*., Resistin-like molecule α provides vitamin-A-dependent antimicrobial protection in the skin. *Cell Host Microbe* **25**, 777–788.e8 (2019). doi: 10.1016/j.chom.2019.04.004; pmid: 31101494
42. Y. Wang *et al*., The intestinal microbiota regulates body composition through NFIL3 and the circadian clock. *Science* **357**, 912–916 (2017). doi: 10.1126/science.aan0677; pmid: 28860383
43. M. H. Black *et al*., Bacterial pseudokinase catalyzes protein polyglutamylation to inhibit the SidE-family ubiquitin ligases. *Science* **364**, 787–792 (2019). doi: 10.1126/science.aaw7446; pmid: 31123136
44. J. Behnke, P. D. Harris, *Heligmosomoides bakeri*: A new name for an old worm? *Trends Parasitol.* **26**, 524–529 (2010). doi: 10.1016/j.pt.2010.07.001; pmid: 20729145
45. M. Camberis, G. Le Gros, J. Urban Jr., Animal model of *Nippostrongylus brasiliensis* and *Heligmosomoides polygyrus*. *Curr. Protoc. Immunol.* **Chapter 19**, 12 (2003). pmid: 18432905
46. C. J. Johnston *et al*., Cultivation of *Heligmosomoides polygyrus*: An immunomodulatory nematode parasite and its secreted products. *J. Vis. Exp.* **2015**, e52412 (2015). doi: 10.3791/52412; pmid: 25867600



## ACKNOWLEDGMENTS

We thank X. Zhan for advice on the generation of *Sprr2a*$^{-/-}$ mice; M. Spence and N. Salinas for assistance with mouse experiments; P. Doss, R. Roman, R. Jackson, and A. Darehshouri (UT Southwestern Electron Microscopy Core) for assistance with electron microscopy; A. Lemoff (UT Southwestern Proteomics Core) for assistance with mass spectrometry; B. Evers (UT Southwestern Histopathology Core) for assistance with mouse histology analysis; and E. Burstein for collaboration on patient sample collection. **Funding:** This work was supported by NIH grants R01 DK070855 (L.V.H.), R01 AI130020 (T.A.R.), and K99 DK120897 (Z.K.); Welch Foundation Grant I-1874 (L.V.H.); American Heart Association Grant 17SDG33670071 (T.A.R.); Cancer Prevention and Research Institute of Texas Grant RP200118 (T.A.R.); the Pew Charitable Trusts (T.A.R.); the Walter M. and Helen D. Bader Center for Research on Arthritis and Autoimmune Diseases (L.V.H.); and the Howard Hughes Medical Institute (L.V.H.). C.M.Z., Z.K., and M.P. were supported by NIH T32 AI005284. **Author contributions:** Z.H., C.Z., T.A.H.-T., T.A.R., and L.V.H. designed research. Z.H., C.Z., L.S.-D., C.M.Z., D.C.P., Z.K., Y.W., M.P., K.A.R., B.H., C.L.B., B.Z., and P.R. performed research. Z.H., C.Z., L.S.-D., and P.R. analyzed data. Z.H., C.Z., T.A.R., and L.V.H. wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** 16S rRNA gene sequencing data are available from the Sequence Read Archive (SRA) under project ID PRJNA743545. All other data are available in the main text or the supplementary materials.


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abe6723
Figs. S1 to S17
Tables S1 to S3
Reference (*47*)
MDAR Reproducibility Checklist

6 September 2020; resubmitted 2 April 2021
Accepted 9 September 2021
10.1126/science.abe6723

# Your Legacy to Science

AN ESTATE GIFT TO THE
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE





**Since 1848, our founding year,** the American Association for the Advancement of Science (AAAS) has been deeply committed to advancing science, engineering and innovation around the world for the benefit of all people.

By making AAAS a beneficiary of your will, trust, retirement plan or life insurance policy, you become a member of our 1848 Society, joining Thomas Edison, Alexander Graham Bell and the many distinguished individuals whose vision led to the creation of AAAS and our world-renowned journal, *Science*, so many years ago.

Unlike many of its peers, *Science* is not for-profit. Your estate gift would provide long-term financial stability and durable annual income that will support operations and competitive innovation for years to come. **This support is vital.**

***"As a teacher and instructor, I bear responsibility for the younger generations. If you have extra resources, concentrate them on organizations, like AAAS, that are doing work for all."***

**—Prof. Elisabeth Ervin-Blankenheim, 1848 Society member**

If you intend to include AAAS in your estate plans, provide this information to your lawyer or financial adviser:

**Legal Name:** American Association for the Advancement of Science
**Federal Tax ID Number:** 53-0196568
**Address:** 1200 New York Avenue, NW, Washington, DC 20005

If you would like more information on making an estate gift to AAAS, cut out and return the form below or send an email to philanthropy@aaas.org. Additional details are also available online at **www.aaas.org/1848Society**.

1848 SOCIETY AAAS

*cut here*

## Yes, I would like more information about joining the AAAS 1848 Society.

**PLEASE CONTACT ME AT:**

Name: _____

Address: _____

City: _____ State: _____ Zip code: _____ Country: _____

Email: _____ Phone: _____

**RETURN THIS FORM TO:**
AAAS Office of Philanthropy and Strategic Partnerships • 1200 New York Avenue, NW • Washington, DC 20005 USA

RESEARCH ARTICLES

MARTIAN GEOLOGY

# Perseverance rover reveals an ancient delta-lake system and flood deposits at Jezero crater, Mars

N. Mangold[1]*†, S. Gupta[2]†, O. Gasnault[3], G. Dromart[4], J. D. Tarnas[5], S. F. Sholes[5], B. Horgan[6], C. Quantin-Nataf[4], A. J. Brown[7], S. Le Mouélic[1], R. A. Yingst[8], J. F. Bell[9], O. Beyssac[10], T. Bosak[11], F. Calef III[5], B. L. Ehlmann[12], K. A. Farley[12], J. P. Grotzinger[12], K. Hickman-Lewis[13,14], S. Holm-Alwmark[15,16,17], L. C. Kah[18], J. Martinez-Frias[19], S. M. McLennan[20], S. Maurice[3], J. I. Nuñez[21], A. M. Ollila[22], P. Pilleri[3], J. W. Rice Jr.[9], M. Rice[23], J. I. Simon[24], D. L. Shuster[25], K. M. Stack[5], V. Z. Sun[5], A. H. Treiman[26], B. P. Weiss[5,11], R. C. Wiens[22], A. J. Williams[27], N. R. Williams[5], K. H. Williford[5,28]

Observations from orbital spacecraft have shown that Jezero crater on Mars contains a prominent fan-shaped body of sedimentary rock deposited at its western margin. The Perseverance rover landed in Jezero crater in February 2021. We analyze images taken by the rover in the 3 months after landing. The fan has outcrop faces, which were invisible from orbit, that record the hydrological evolution of Jezero crater. We interpret the presence of inclined strata in these outcrops as evidence of deltas that advanced into a lake. In contrast, the uppermost fan strata are composed of boulder conglomerates, which imply deposition by episodic high-energy floods. This sedimentary succession indicates a transition from sustained hydrologic activity in a persistent lake environment to highly energetic short-duration fluvial flows.

Mars is currently cold and hyperarid; liquid water is not stable at its surface. However, orbital and rover observations of features including valley networks, sedimentary fans, and ancient lake beds indicate that the planet once had a warmer, wetter climate (*1*–*3*). Uncertainties remain about the character, timing, and persistence of aqueous activity (and therefore potential habitability) on early Mars. The Mars 2020 mission, whose main component is the Perseverance rover, is the first step in a planned multimission campaign to return martian samples to Earth and examine them for potential biosignatures (*4*). The 45-km-diameter Jezero crater was selected as the landing site on the basis of orbital images, which showed geomorphic expressions of two sedimentary fan structures (western and northern) at the edges of the crater (*5*, *6*). These were inferred to be river delta deposits that formed in an ancient lake basin during the Late Noachian or Early Hesperian epochs on Mars [~3.6 to 3.8 billion years ago] (*5*–*9*) (Fig. 1 and fig. S1). Spectroscopic observations from orbit have detected phyllosilicates and carbonates, minerals indicative of past aqueous environments (*6*, *7*, *10*), in the crater. Rover investigations on the surface could provide insight into the evolution of Jezero's ancient lake system and the time scale of liquid water residence on the surface.

The Perseverance rover landed on the floor of Jezero crater on 18 February 2021. The landing site, informally named Octavia E. Butler, is ~2.2 km from the southeast-facing erosional scarp of the western fan deposits, a planned target for the mission (Fig. 1 and figs. S1 to S5). During the first 3 months of the mission, we obtained images of the western fan using the Mastcam-Z camera and the Remote Micro-Imager (RMI) of the SuperCam instrument (*11*–*14*) (Figs. 1 to 4; figs. S2 to S4, S6, and S7; tables S1 and S2). We use these long-distance images to investigate the stratigraphy and sedimentary characteristics of the fan deposits and interpret their implications for the ancient lake in Jezero crater.

## Kodiak butte

Images of a prominent butte (an isolated flat-topped hill) located ~1 km south of the main fan deposit (Fig. 1), which we informally named Kodiak, record ancient sedimentary processes at Jezero crater. Owing to the morphological similarity of Kodiak butte to the main fan exposures and the near-identical elevation of its top (*15*), we interpret Kodiak butte as an erosional remnant of an originally more extensive fan deposit. A mosaic of the east-southeast–facing wall of Kodiak (Figs. 1 and 2 and fig. S2) shows two main outcrop areas with three distinct sedimentary layer types: a series of inclined strata sandwiched between layers comprising horizontal strata, described in detail below. There is no evidence for later dislodgement or rotation of blocks, such as faults or slippage, and therefore we interpret the observed stratigraphy as reflecting the original depositional geometry.

Kodiak butte consists of two outcrop sections that expose five distinct stratigraphic bodies, which we designate k1 to k5 (Fig. 2). The unit k1 is 17 m thick vertically and extends horizontally at least 70 m to the northern butte margin visible from Perseverance (Fig. 2, A to C). The lowest visible part of k1 consists of plane-parallel horizontal to low-angle thinly bedded strata. These show recessive weathering, characteristic of readily eroded fine-grained lithologies (mudstones or sandstones). Overlying these is a ~10-m-thick series of strata composed of steeply inclined beds with apparently southward dips at angles up to 35°. Individual beds, defined by variations in erosion, have apparent thicknesses ranging from 10 to 50 cm. We infer their primary lithology to be finer-grained than a conglomerate, possibly sandstone, with scattered cobbles. A

[1]Laboratoire Planétologie et Géodynamique, Centre National de Recherches Scientifiques, Université Nantes, Université Angers, Unité Mixte de Recherche 6112, 44322 Nantes, France. [2]Department of Earth Science and Engineering, Imperial College London, London SW7 2AZ, UK. [3]Institut de Recherche en Astrophysique et Planétologie, Université de Toulouse, Université Paul Sabatier, Centre National de Recherches Scientifiques, Observatoire Midi-Pyrénées, 31400 Toulouse, France. [4]Laboratoire de Géologie de Lyon-Terre Planètes Environnement, Univ Lyon, Université Claude Bernard Lyon 1, Ecole Normale Supérieure Lyon, Centre National de Recherches Scientifiques, 69622 Villeurbanne, France. [5]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA. [6]Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN 47907, USA. [7]Plancius Research, Severna Park, MD 21146, USA. [8]Planetary Science Institute, Tucson, AZ 85719, USA. [9]School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA. [10]Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, Unité Mixte de Recherche 7590, Centre National de Recherches Scientifiques, Sorbonne Université, Museum National d'Histoires Naturelles, 75005 Paris, France. [11]Department of Earth, Atmospheric, and Planetary Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [12]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA. [13]Department of Earth Sciences, The Natural History Museum, South Kensington, London SW7 5BD, UK. [14]Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università di Bologna, I-40126 Bologna, Italy. [15]Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark. [16]Department of Geology, Lund University, 22362 Lund, Sweden. [17]Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark. [18]Department of Earth and Planetary Sciences, University of Tennessee, Knoxville, TN 37996, USA. [19]Instituto de Geociencias, Consejo Superior de Investigaciones Cientificas, Universidad Complutense Madrid, 28040 Madrid, Spain. [20]Department of Geosciences, Stony Brook University, Stony Brook, NY 11794, USA. [21]Johns Hopkins University Applied Physics Laboratory, Laurel, MD 20723, USA. [22]Space and Planetary Exploration Team, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. [23]Geology Department, College of Science and Engineering, Western Washington University, Bellingham, WA 98225, USA. [24]Center for Isotope Cosmochemistry and Geochronology, Astromaterials Research and Exploration Science, NASA Johnson Space Center, Houston, TX 77058, USA. [25]Department of Earth and Planetary Science, University of California, Berkeley, CA 94720, USA. [26]Lunar and Planetary Institute, Universities Space Research Association, Houston, TX 77058, USA. [27]Department of Geological Sciences, University of Florida, Gainesville, FL 32611, USA. [28]Blue Marble Space Institute of Science, Seattle, WA 98104, USA.
***Corresponding author. Email: nicolas.mangold@univ-nantes.fr**
†These authors contributed equally to this work.

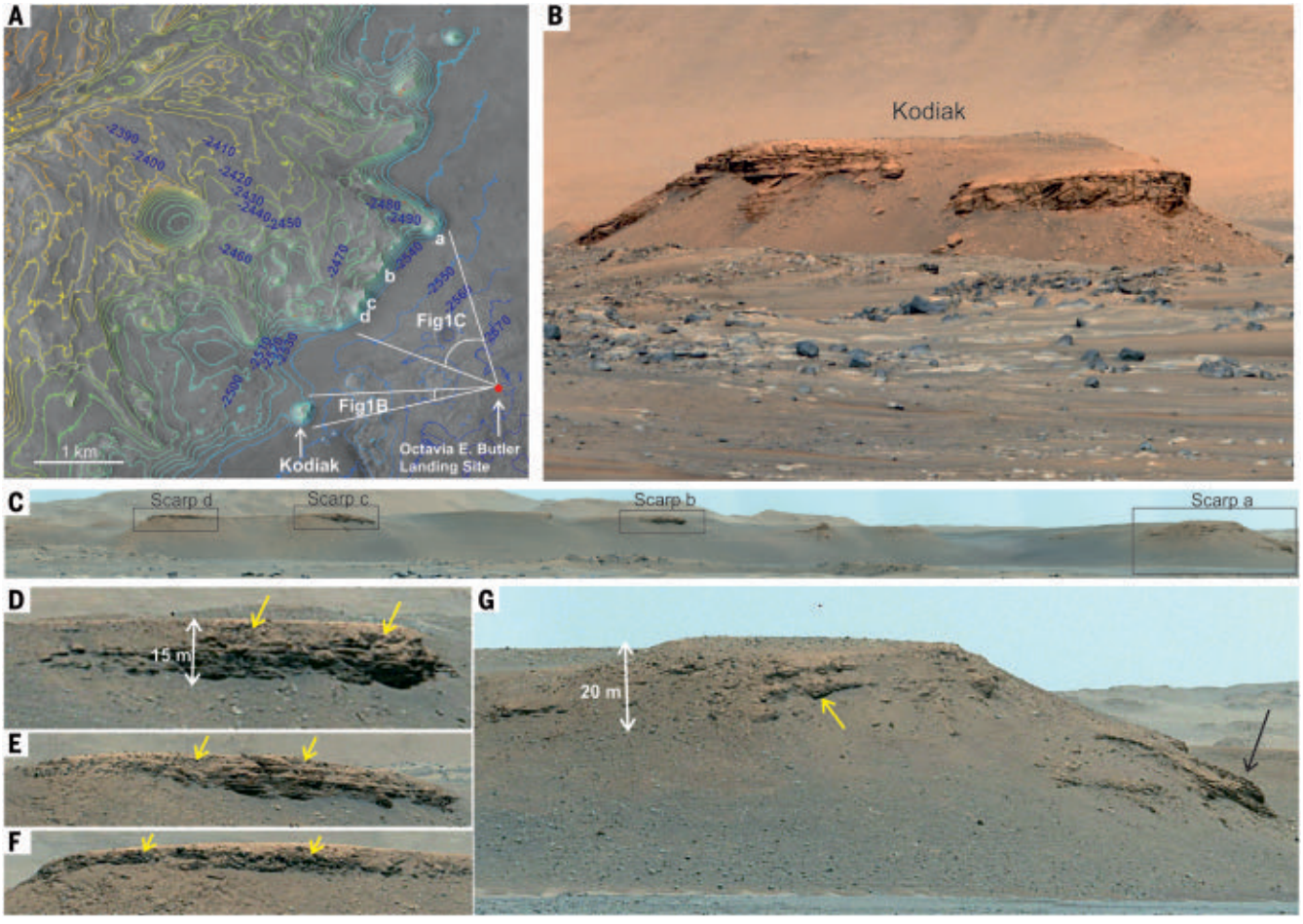



**Fig. 1. Orbital and rover context observations of the Jezero crater western fan.** (**A**) High Resolution Imaging Science Experiment (HiRISE) mosaic (*14*) with 10-m elevation contours from a digital elevation model (DEM) (*14*) showing the western fan inside Jezero crater and the landing site, informally named Octavia E. Butler (red dot). White arcs represent the fields of view of (B) and (C). (**B**) The butte informally named Kodiak, imaged from a distance of ~2.24 km by Mastcam-Z. (**C**) Mastcam-Z enhanced color mosaic of the delta front, taken from a ~2.20-km distance with black boxes indicating scarps of interest. (**D** to **G**) Each scarp viewed in the corresponding 110-mm focal length Mastcam-Z images. Yellow arrows indicate the location of boulder-rich material shown in Figs. 3 and 4. The black arrow in (G) indicates an exposure with dipping strata.

second unit of dipping strata (k2, 3 m thick) immediately overlies the uppermost strata of k1.

In the southern portion of Kodiak, sedimentary units (k3, 13 m thick) and (k4, 10 m thick) show similar geometries to those in k1 (Fig. 2, D to F). In its lower section, k3 consists of thinly bedded, gently dipping, and horizontal strata. These strata show recessive weathering, again indicating mudstones or sandstones, and pass upwards into a distinct 7-m-thick section of inclined beds that dip consistently to the south. Locally, these dipping beds contain isolated boulders and cobbles (up to 40 cm in diameter) (Fig. 2F). At their base, these beds show a downward asymptotic decrease in inclination and pass into the lowermost horizontal strata. Overlying the inclined beds across a sharp subhorizontal truncation surface, k4 shows low-angle to locally cross-stratified subhorizontal strata. The overlying unit k5 erosionally truncates k4 (Fig. 2E). Unit k5 consists of unsorted conglomerates, which contain boulders up to 1.5 m on the long axis, implying a marked change in depositional regime.

Inclined beds in k1 display a downward asymptotic decrease in apparent dip angle and pass gradually into underlying gently dipping and horizontal strata (Fig. 2). At the top, the transition from inclined beds to subhorizontal beds also shows a gradual change in dip (Fig. 2C). This geometric arrangement of strata shows that k1 consists of a single depositional unit with a tripartite architecture; we identify the lower gently dipping beds as bottomsets, the inclined beds as foresets, and the uppermost horizontal layers as topsets (see fig. S8D for a schematic diagram). We interpret the k3 inclined beds to be foresets that pass downward, similar to k1, with decreasing apparent dip angle into subhorizontal strata we interpret as bottomsets. The subhorizontal strata of k4 overlying k3 then represent topsets. The sharp discontinuity between k3 and k4 is distinct from that observed in k1, where the transition appears to be continuous.

We interpret this distinct tripartite bedding geometry (bottomsets, foresets, and topsets) of the units k1 to k4 as representing deposition in steeply fronted Gilbert-type deltas (see the supplementary text in the supplementary materials and fig. S8) (*16*, *17*). The thicknesses and lateral extents (>70 m) of the foreset units are too great to be explained by formation as dunes from underwater currents or as lateral accretion deposits in fluvial bars. The presence of cobbles and boulders in the foreset strata (Fig. 2F) is inconsistent with their formation as aeolian dunes. In a Gilbert delta, topset strata are fluvial deposits formed in delta top environments. The foreset strata represent deposits formed by gravity-driven flow processes on steeply dipping delta fronts. Bottomset strata represent finer-grained sediments deposited in areas immediately lakeward

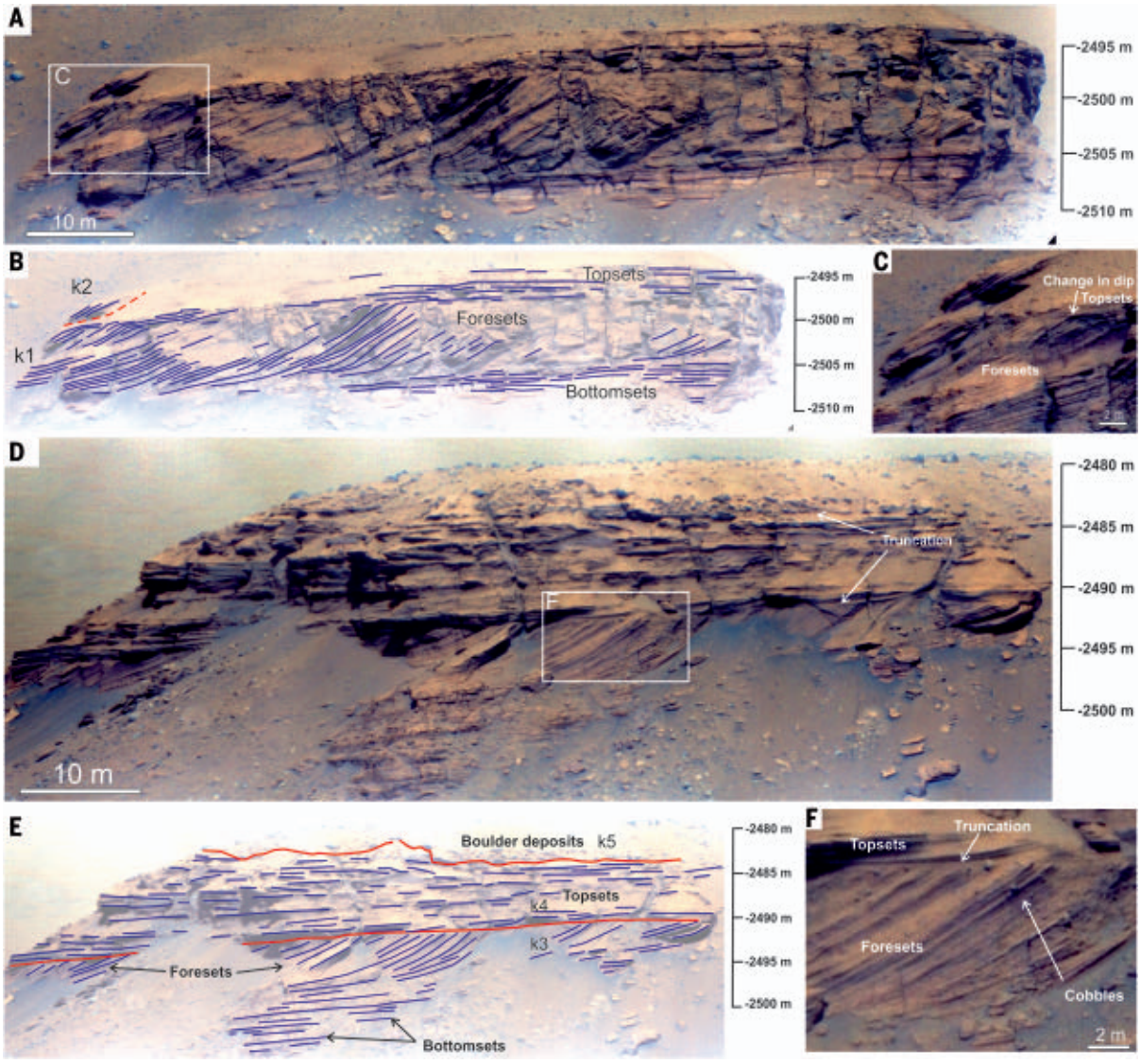



**Fig. 2. Stratigraphy of Kodiak butte.** (**A** and **D**) Zoomed images of the two scarps of Kodiak (see fig. S2 for wider context). Elevation scales were inferred from a HiRISE DEM (*14*) and have systematic uncertainties of ±2 m. White boxes indicate regions shown in more detail in other panels. (**B** and **E**) Interpreted line drawings of the main visible beds (blue lines for individual beds and red lines for discontinuities), overlain on the images from (A) and (D). Units k1 to k5 are labeled and discussed in the text. (**C**) Zoomed image of k1 showing the change in dip from subhorizontal beds (topsets) to inclined beds (foresets). (**F**) Zoomed image of the foresets in k3. This unit has a coarse texture with several cobble-size clasts (white arrow). The erosional truncation of k3 by k4 is labeled.

of the delta front. The transition from topset to foreset (the topset breakpoint) constrains the lake level at the time of deposition. The thickness of the foreset units provides a lower limit of 10 m water depth in this portion of the Jezero lake basin at that time. The bases of topset strata in units k1 and k4 are, respectively, at about −2500 and −2490 m elevation (below the reference equipotential), corresponding to past lake levels at the time of deposition (Fig. 5).

Elevation differences suggest that units k3 and k4 are stratigraphically higher, and hence younger, than k1 (Fig. 2). Examination of the exposures on both faces of the butte indicates a similar architecture with two differences: a discontinuity above k3 foresets (not present in k1), and the presence of the terminal boulder-rich unit k5 that truncates k4 topsets (absent above k1 or k2). These differences indicate that the k1 and k2 units on one side of the scarp and the k3 and k4 units on the other side are not the same stack of layers (fig. S2). This rules out the possibility of a fault being the explanation for the offset in elevation.

The orientation of foresets indicates an apparent southward progradation in this sector of the western fan (i.e., the delta advanced toward the south) during episodes of stationary or slowly decreasing lake level. The subhorizontal topset truncation of underlying foreset units between k3 and k4 may reflect a drop in lake level. In contrast, the stacking of delta units stratigraphically on top of one another indicates an overall lake level rise of ~10 m before the truncation by k4. Thus, the observed geometries in Kodiak indicate delta growth into a lake system with fluctuating lake levels.

Previous studies have proposed that Jezero crater hosted an open-lake system with the water level at an elevation of −2395 m (*5*); this inference is derived from the observations that the inlet valley (feeding the western fan) and the breaching valley (which dissects the eastern rim of the crater) have about the same elevation of −2395 m. However, our results indicate that the lake level during deposition of units k1 to k4 (about −2500 and −2490 m) was ~100 m below the inferred open-system lake level (Fig. 5). Thus, Jezero lake was closed (no outlet

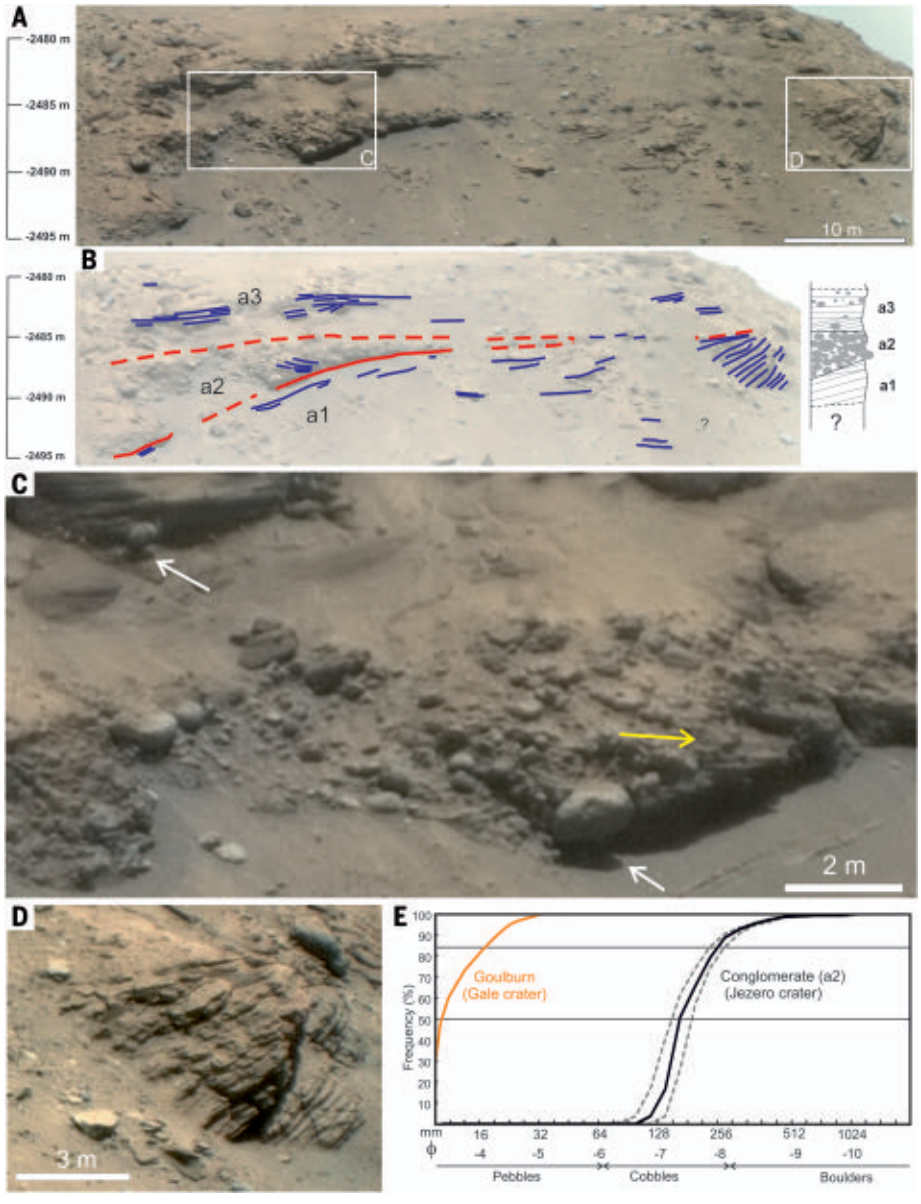



**Fig. 3. Stratigraphy of the western fan scarp a.** (**A**) RMI mosaic of the western fan scarp a (see Fig. 2C and fig. S3 for wider context). Elevation scale as in Fig. 2. White boxes indicate regions shown in more detail in other panels. (**B**) Interpreted line drawing of individual layers (blue lines) and main boundaries (red lines) between sedimentary bodies labeled a1 to a3. A simplified stratigraphic column of these three bodies is shown on the right. (**C**) Zoomed image of the boulder-bearing units a2 and a3. White arrows indicate the shadow cast beneath two boulders hanging from the bedrock. Right of the lowermost hanging boulder, an incipient oblique bedding is visible (yellow arrow). Unit a3 might be the result of an amalgamation of two or more depositional sequences. (**D**) Zoomed images of a1 showing dipping layers organized as cosets of dipping beds with an apparent dip of up to 30°. (**E**) Cumulative histogram, on a logarithmic scale (ϕ indicates a scale defined by $\log_2$ increments), of the measured sizes of 333 clasts (black) compared with the conglomerate Goulburn measured at Gale crater by the Curiosity rover (orange) (*35*). Dotted lines indicate the uncertainty around clast size measurements (*14*).

river) at the time of the delta progradation at Kodiak, which is a hydrological system conducive to short-term fluctuations in the lake level. Nevertheless, the overall stratigraphy indicates progradation of the western delta system and long-term lake level regression.

### The western fan

Images of the southeast-facing erosional front of the western fan expose sedimentary geometries within the uppermost fan deposits at several locations, at the top of ~60-m-tall scree-covered hillslopes (Fig. 1 and figs. S3 to S5). In a RMI mosaic (Fig. 3), the upper section of the northernmost hillslope exposes three sedimentary bodies (a1 to a3) that consist of conglomerates and finer-grained rocks (the grain size is not resolved). The lowermost unit, a1, has an apparent thickness of 7 m and is composed of 10- to 30-cm-thick tabular-bedded strata, which show an apparent dip to the southwest. At its northern margin, a1 exhibits steeply inclined beds (up to 30°) (Fig. 3D) that likely represent either lateral accretion sets formed in a large fluvial channel bar, or delta foresets.

A distinct coarse-grained lenticular unit, a2, overlies a1; it is ~30 m wide and asymmetric with a maximum thickness of 9 m at its southern edge, thinning to <1 m to the north. Unit a2 is dominated by unsorted, clast-supported conglomerates of cobbles and boulders (Fig. 3C). The deposit is structureless, locally displaying faint layering. Images do not show a preferred clast orientation or size segregation. The largest boulder, ~1.5 m on its long axis, casts a shadow below it, implying that it is embedded in the outcrop and therefore did not roll down from the upper slope. A shape assessment of 24 boulders shows that 13 are rounded and 11 are angular (*14*). Size measurements of 333 boulders and cobbles (figs. S6 and S7) indicate a distribution with a median size ($D_{50}$) of 16.4 ± 2.2 cm and a $D_{84}$ (84% of clasts are smaller) of 25.9 ± 2.2 cm (Fig. 3E) (*14*).

From unit a2's lack of sorting, large clast sizes, absence of well-developed stratification, and disorganized but clast-supported fabric, we infer that it was deposited from rapidly decelerating high velocity flood flows that can transport boulders. This interpretation is based on well-constrained observations of flood deposits on Earth (*18*, *19*). The rounding of some of the largest clasts indicates that they have undergone abrasion by collisional processes during fluvial transport. The lens-like shape of the conglomerate body a2 suggests that it is a channel fill. Assuming its dimensions represent the formative fluvial channel, the channel was 3 to 10 m deep. We estimate discharge rates using two methods: a Mars-modified version

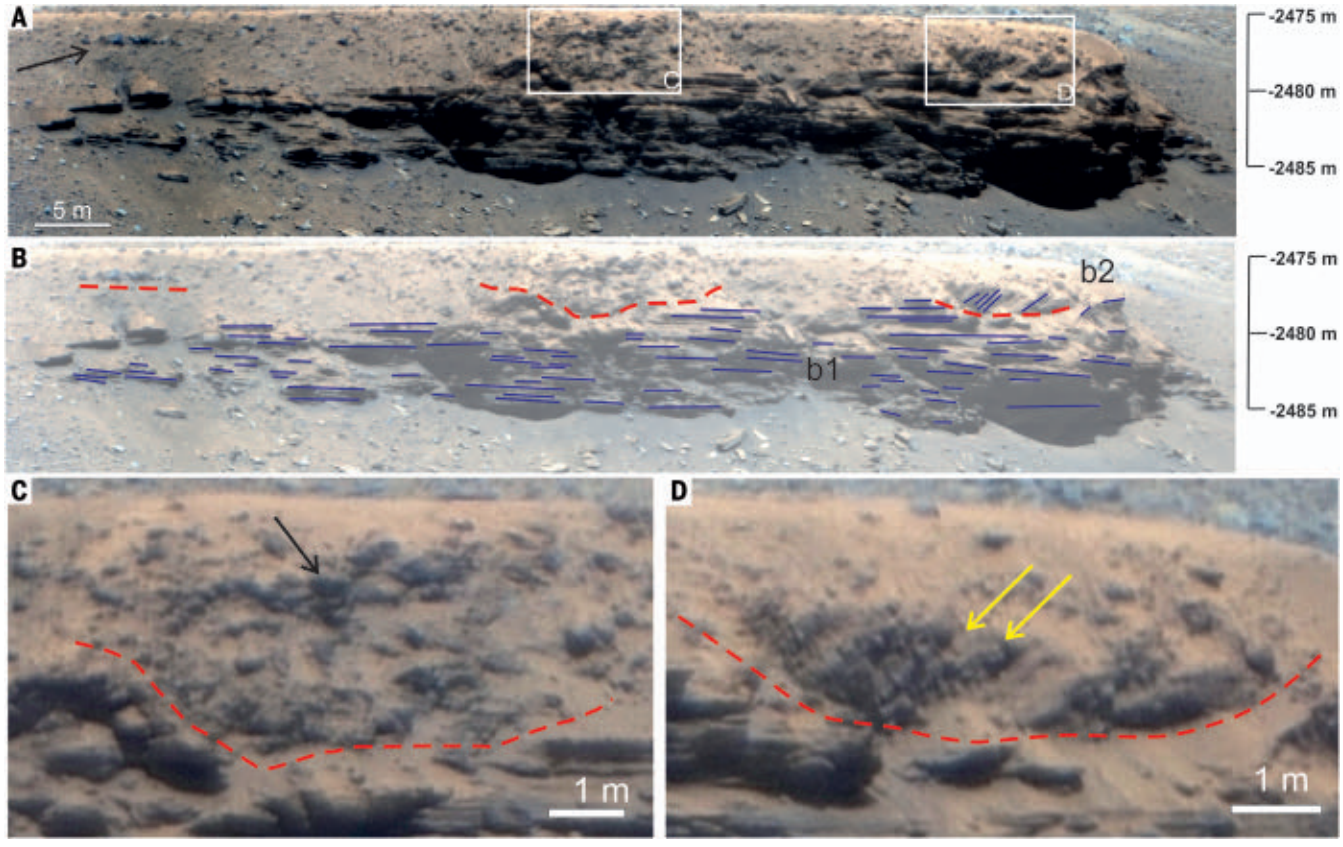



**Fig. 4. Stratigraphy of the western fan scarp b.** (**A**) Mosaic of five RMI frames of scarp b (localization in Fig. 1, mosaic in fig. S4). The black arrow at the top left indicates a thin bed with cobbles and boulders preserved in the scarp similar to the thinning of a2 in Fig. 3. Elevations scales are as in Fig. 2. White boxes indicate regions shown in more detail in other panels. (**B**) Line drawing interpretation of this scarp, showing individual beds (blue lines) and discontinuities (dotted red lines). Subhorizontal beds dominate the lower sedimentary unit b1. This unit displays a relatively fine-grained material, compared to overlying boulder-bearing conglomerates b2, which are present above a discontinuity (dashed red lines) interpreted as a truncation episode. (**C**) Zoomed image of a boulder conglomerate displaying a series of rounded boulders, piled up along a subhorizontal bed, 50 cm to 1 m in diameter, developed along a subhorizontal bed (black arrow). (**D**) Zoomed image of cobble-bearing conglomerates that were deposited as dipping beds (yellow arrows). The presence of bedding indicates the conglomerate exposure is not a residual lag lying on hillslopes.

of the Darcy-Weisbach equations for river flows (*20*, *21*) and the velocity threshold necessary to lift the largest clasts observed (*14*). Both methods give consistent results with velocities of 1.6 to 8.6 m $s^{-1}$ and discharge rates of 70 to 3000 $m^3$ $s^{-1}$ (table S3) (*14*).

Unit a3 overlies a2; a3 is generally finer-grained than a2, up to 10 m thick, and extends ~80 m laterally. Unit a3 shows horizontal to low-angle stratification, with some local cross-stratification. Unit a3 contains isolated cobbles and boulders; including a 50-cm-diameter boulder that is being eroded from the outcrop (Fig. 3C). On the basis of the presence of planar stratification and cross-stratification, we infer unit a3 to be a sandstone with outsized clasts. If the a2-a3 contact is gradational, then these units are part of the same depositional sequence, and a3 may record the waning stage of the fluvial flood flow. Alternatively, a3 could represent a second, lower-energy event in which the flux of boulders was reduced.

Stratigraphic relationships between a1 and underlying units are not well constrained because the exposure is debris covered (Fig. 1G). However, the lowest part of the hillslope does expose one set of beds that are inclined to the east (Fig. 1G). The inclined beds could be either delta foresets, as observed at Kodiak and hypothesized from orbital images (*22*, *23*), or they could represent a landslide block, as their dip is similar to the local slope.

Additional coarse-grained deposits are visible at scarps b to d (Fig. 1 and fig. S5). Scarp b is a 15-m-tall cliff that exposes two distinct sedimentary bodies (Figs. 1, C and D, and 4). The lower unit comprises 10- to 40-cm-thick parallel planar subhorizontal beds b1, which we infer to consist of sedimentary rocks that are finer-grained than conglomerate. Unit b1 is overlain by a unit b2, a conglomerate of unsorted cobbles and boulders (Fig. 4C) with varying degrees of rounding. Unit b2 shows faint inclined bedding (Fig. 4) and truncates unit b1 at a discontinuous, scoured contact at its base. The scarps at locations c and d (Fig. 1, E and F) expose units like those at b, with subhorizontal finer-grained strata overlain by coarser-grained, likely cobble-boulder conglomerates. The stratal geometry shown by these outcrops is similar to that observed in the uppermost section at Kodiak k5. Subhorizontal strata such as b1 could then represent delta topsets similar to k4.

The unconsolidated boulder-rich deposits observed at the scarp tops contain many rounded, scattered boulders (figs. S9 and S10 and supplementary text). We interpret these unconsolidated disorganized deposits as residual lags resulting from weathering of underlying boulder conglomerates and sandstones (Fig. 3). Comparison with geological maps constructed from orbital images (*14*) indicates that these unconsolidated deposits are part of the Delta Blocky unit, which includes much of the upper surface of the western fan and is defined by positive relief elongate ridges and the presence of numerous clasts. This unit has previously been interpreted as inverted fluvial channel-belt deposits (*8*, *15*, *22*, *23*) (fig. S1). On the basis of our rover images, we interpret the boulder-bearing units

a2, a3, b2, and k5 as fluvial deposits that represent locally preserved sections from these well-developed fluvial channel-belt deposits.

We use orbital and multispectral Mastcam-Z observations of the western fan exposures to investigate the mineralogy and provenance of the boulder conglomerates (figs. S11 and S12 and supplementary text). These data indicate that the boulder conglomerates and the blocky deposits are dominated by low calcium pyroxene (LCP), unlike other sections of the fan stratigraphy that are dominated by phyllosilicates and olivine (fig. S11). This interpretation is consistent with the source of the boulders and cobbles as being either the LCP-bearing crater rim of Jezero and/or the widespread exposures of LCP-rich crust >60 km upstream of Jezero crater (fig. S11) (*7*, *24*). An igneous rock source would be consistent with the boulders' massive shape and apparent lack of internal fabric. Substantial transport distances from distant sources are consistent with the presence of rounded boulders (*14*, *25*), whereas the source of angular boulders could be more proximal, such as the crater rim.

### Implications for hydrologic evolution and sample return

Our rover images constrain the hydrologic evolution of Jezero crater and potentially also the broader climate and habitability of early Mars. The delta architecture at Kodiak indicates deposition in a closed lake system, under fluctuating water levels and changing styles of flow during later stages. This indicates that the climate on Mars at that period (late Noachian or early Hesperian) was warm and humid enough to support a hydrologic cycle on the martian surface, at least episodically.

The presence of coarse-grained material (cobbles and boulders) in steep foresets is characteristic of Gilbert-type deltas prograding into deep lake systems (*16*, *26*, *27*) (fig. S6). The highest lake elevation recorded by the transition from topsets to foresets at Kodiak has an altitude of about −2490 m (Fig. 5), well below the previously proposed lake levels of −2395 and −2250 m based on the basin topography (*5*, *23*, *28*). The Kodiak delta deposits are located 5 km away from the outlet, and they correspond to a regression to lower lake levels, because they formed after a large part of the delta was already deposited. Our results do not exclude periods of higher standing lake levels in the crater but do imply that any such periods occurred before the one recorded at Kodiak. Our observations of Kodiak indicate that the delta front extended ~1 km further south than the main western fan scarp. Delta deposits could have originally extended further eastward as well.

The boulder conglomerates in units a2, b2, and k5 (Fig. 1) indicate repeated flood episodes of variable intensities. These deposits are distinct from the low- to moderate-energy fluvial deposits characteristic of river-dominated deltas (*19*). Their stratigraphic positions overlying delta deposits indicate that they are also unlikely to be sediment gravity flow deposits formed in a deep lacustrine setting. We cannot determine whether the boulder conglomerates were deposited when a lake still existed in Jezero crater. Their geometry is consistent with fluvial deposits on Earth that show downstream transition to gravel-to-sand Gilbert-type underwater foresets (*29*). The lowermost boulder conglomerates we observe are at an elevation of about −2490 m, similar to that of the lake level deduced from foresets at Kodiak. Therefore, these fluvial floods could have formed when the lake was around, or below, this level. Alternatively, the widespread boulder conglomerate deposits could represent a younger depositional system that overlies deltaic strata.

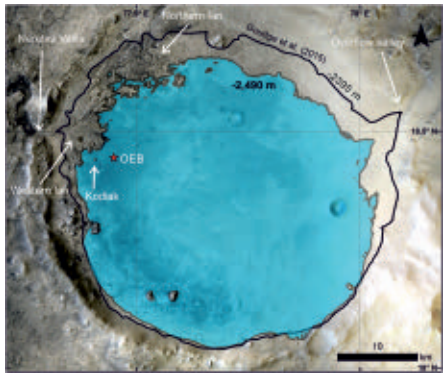



**Fig. 5. Inferred paleolake level inside Jezero crater at the time of Kodiak sediment deposition.** Blue shading indicates assumed lake level filled to the −2490 m gray contour following the uppermost elevation deduced from deltaic architecture at Kodiak (Fig. 2). The red star indicates the Octavia E. Butler (OEB) landing site of the Perseverance rover. The black outline of the implied earlier minimum water stand, corresponding to the overflow valley breach (*8*), is shown for comparison. Rocks present on the crater floor might not have been emplaced during the period of lake activity. Both western and northern fans are above the inferred lake surface, and the basin appears closed, 100 m below the breach to the east (labeled overflow valley). Background from the Context Camera mosaic (*14*).

Our results indicate a temporal transition in the energy regime of fluvial systems feeding the western fan, from sustained fluvial activity that built delta deposits prograding into the Jezero crater lake to episodes characterized by high discharge fluvial flows capable of mobilizing meter-scale boulders over transport distances of potentially tens of kilometers. Subhorizontal topset beds at Kodiak (and possibly b1) are relatively homogeneous deposits compared with the boulder conglomerates, and they are likely sandstones, consistent with deposition by sandy rivers. The presence of occasional boulders in the Kodiak foresets points to locally higher intensity flow conditions, but the boulder conglomerate units record much higher magnitude flood episodes. Local discharge rate estimates (70 to 3000 $m^3 \cdot s^{-1}$) for the floods are consistent with those previously estimated from braided fluvial channels observed upstream in Neretva Vallis (*9*). Nevertheless, these are late-stage deposits formed from more intermittent, energetic flows than the topsets they overlie, so our discharge rates cannot be used to estimate the formation time required for the entire delta fan.

The mechanism responsible for the flood events is unknown. The presence of rounded boulders demonstrates that substantial abrasion of clasts occurred during fluvial transport. This evidence, coupled with the presence of multiple flood episodes with similar boulder sizes as in unit a2, excludes megafloods such as those proposed for martian outflow channels (*30*). Flood episodes could have formed by a variety of processes (*18*, *31*), such as intense rainfall events, rapid snowmelt episodes [from either a climatic origin (*1*, *3*) or heating by volcanism or impact (*32*, *33*)], or through progressive building of glaciers and glacial lakes in the watershed creating episodic surges (*31*). Thus, the transition in flow intensity at Jezero crater may be related either to paleoclimatic shifts (global or regional) or to changes in watershed hydrology.

The Jezero crater deposits provide information which could be extrapolated to other paleolakes on Mars (*2*, *26*, *34*). Favorable climatic conditions for rivers and lakes are already known to have also been present at Gale crater (*2*, *35*). However, the conglomerates in Jezero crater require much higher energy environments than those in Gale crater, where the median clast size is <1 cm and the largest clasts are <10 cm (*35*) (Fig. 3E). A transition to drier conditions at Gale crater has been suggested to explain a change in mineralogy from clay- to sulfate-bearing minerals, and alternating eolian and fluvial deposits (*36*, *37*). However, in Gale crater, no fluvial flood deposits have been observed stratigraphically overlying the lacustrine deposits of the Murray formation (*2*), contrasting with the hydrologic evolution of Jezero crater.

Our results inform sampling strategies for Perseverance in Jezero crater (supplementary text). First, boulders >1 m in diameter provide an opportunity to analyze and collect samples from crustal rocks sourced from outside Jezero that must predate the rocks within the crater (*4*, *24*). These boulders likely contain records of the ancient martian interior. Second, the finer-grained bottomset strata, which are known from orbital data to contain Fe/Mg smectite

clays (*6*, *7*, *10*), have high potential to preserve organic matter or potential biosignatures (*38*–*40*).

## REFERENCES AND NOTES


1. R. A. Craddock, A. D. Howard, *J. Geophys. Res.* **107**, 5111 (2002).
2. J. P. Grotzinger *et al*., *Science* **350**, aac7575 (2015).
3. R. Wordsworth *et al*., *Nat. Geosci.* **14**, 127–132 (2021).
4. D. W. Beaty *et al*., *Meteorit. Planet. Sci.* **54**, S3–S152 (2019).
5. C. I. Fassett, J. W. Head III, *Geophys. Res. Lett.* **32**, L14201 (2005).
6. B. L. Ehlmann *et al*., *Nat. Geosci.* **1**, 355–358 (2008).
7. T. A. Goudge, J. F. Mustard, J. W. Head, C. I. Fassett, S. M. Wiseman, *J. Geophys. Res.* **120**, 775–808 (2015).
8. T. A. Goudge, R. A. Milliken, J. W. Head, J. F. Mustard, C. I. Fassett, *Earth Planet. Sci. Lett.* **458**, 357–365 (2017).
9. N. Mangold *et al*., *Astrobiology* **20**, 994–1013 (2020).
10. B. H. Horgan, R. B. Anderson, G. Dromart, E. S. Amador, M. S. Rice, *Icarus* **339**, 113526 (2020).
11. J. F. Bell 3rd *et al*., *Space Sci. Rev.* **217**, 24 (2021).
12. S. Maurice *et al*., *Space Sci. Rev.* **217**, 47 (2021).
13. R. C. Wiens *et al*., *Space Sci. Rev.* **217**, 4 (2021).
14. Materials and methods are available as supplementary materials.
15. K. M. Stack *et al*., *Space Sci. Rev.* **216**, 127 (2020).
16. G. K. Gilbert, *The Topographic Features of Lake Shores. 5th Annual Report* (United States Geological Survey, 1885), pp. 69–123.
17. H. Fayol, *Etudes sur le terrain houiller de Commentry. Bulletin de la Société de l'Industrie Minérale, XV* (Saint-Étienne, Imprimerie Théolier, 1886).
18. J. E. Costa, *Geol. Soc. Am. Bull.* **94**, 986–1004 (1983).
19. M. R. Leeder, *Sedimentology: Process and Product* (Chapman and Hamm, 1982).
20. E. S. Kleinhans, *J. Geophys. Res.* **110**, E12003 (2005).
21. L. Wilson, G. J. Ghatan, J. W. Head III, K. L. Mitchell, *J. Geophys. Res.* **109**, E09003 (2004).
22. T. A. Goudge, D. Mohrig, B. T. Cardenas, C. M. Hughes, C. I. Fassett, *Icarus* **301**, 58–75 (2018).
23. S. C. Schon, J. W. Head, C. I. Fassett, *Planet. Space Sci.* **67**, 28–45 (2012).
24. E. L. Scheller, B. L. Ehlmann, *J. Geophys. Res. Planets* **125**, e2019JE006190 (2020).
25. M. Attal, J. Lavé, *Geol. Soc. Am. Spe. Paper* **398**, 143–171 (2006).
26. G. G. Ori, L. Marinangeli, A. Baliva, *J. Geophys. Res.* **105**, 17629–17641 (2000).
27. S. Rohais, R. Eschard, F. Guillocheau, *Sediment. Geol.* **210**, 132–145 (2008).
28. F. Salese *et al*., *Astrobiology* **20**, 977–993 (2020).
29. I. G. Hwang, S. K. Chough, in *Coarse-Grained Deltas*, A. Colella, D. B. Prior, Eds., Special Publication No. 10 of the International Association of Sedimentologists (Wiley-Blackwell, 1990), pp. 235–254.
30. V. R. Baker, D. J. Milton, *Icarus* **23**, 27–41 (1974).
31. J. E. O'Connor, J. E. Costa, *The World's Largest Floods, Past and Present—Their Causes and Magnitudes: U.S. Geological Survey Circular 1254* (United States Geological Survey, 2004).
32. V. C. Gulick, V. R. Baker, *J. Geophys. Res.* **95**, 14325–14344 (1990).
33. N. Mangold *et al*., *Icarus* **220**, 530–551 (2012).
34. N. Cabrol, E. Grin, *Icarus* **142**, 160–172 (1999).
35. R. M. E. Williams *et al*., *Science* **340**, 1068–1072 (2013).
36. R. E. Milliken, J. P. Grotzinger, B. J. Thomson, *Geophys. Res. Lett.* **37**, L04201 (2010).
37. W. Rapin *et al*., *Geology* **49**, 842–846 (2021).
38. R. E. Summons *et al*., *Astrobiology* **11**, 157–181 (2011).
39. T. Bosak, K. R. Moore, J. Gong, J. P. Grotzinger, *Nat. Rev. Earth Environ.* **2**, 490–506 (2021).
40. J. L. Eigenbrode *et al*., *Science* **360**, 1096–1101 (2018).


## ACKNOWLEDGMENTS

We acknowledge the Mars 2020 project's management, engineering, and scientific teams for their diligent efforts in making this mission as effective as possible. We are grateful to Mars 2020 team members who participated in tactical and strategic science operations. We also thank the High Resolution Imaging Science Experiment (HiRISE) and CRISM instrument teams of the Mars Reconnaissance Orbiter (MRO) for the use of HiRISE images, and the Observatoire pour la Minéralogie, l'Eau, les Glaces et l'Activité (OMEGA) instrument team of the Mars Express mission for the use of OMEGA data. N.M., O.G., G.D., C.Q.-N., S.L.M., P.P., and S.M. acknowledge the Centre National de Recherches Scientifiques (CNRS) and the Centre National d'Etudes Spatiales (CNES) for the research infrastructures and collaborative networks enabling their participation to rover operations. The authors appreciated helpful suggestions from reviewers. **Funding:** Centre National d'Etudes Spatiales, France (N.M., O.G., G.D., C.Q.-N., S.L.M., P.P., S.M.); NASA Mars 2020 Project (J.D.T., S.F.S., B.H., J.F.B., K.A.F., K.H.W., K.M.S., R.C.W., B.L.E., S.M.M., R.A.Y., J.I.N.); NASA Planetary Science Division, Mars Program (J.I.S.); NASA M2020 Participating Scientist Program under Grant #80NSSC21K0332 (A.J.W.); NASA Mars 2020 Returned Sample Science Participating Scientist Program (RSSPS) award numbers 80NSSC20K0234 (T.B.) and 80NSSC20K0238 (B.P.W.); NASA Post-Doctoral program (JDT); UK Space Agency Aurora program (S.G.); UK Space Agency Aurora Research Fellowship (K.H.-L.); International Postdoc grant from the Swedish Research Council (grant no. 2017-06388) (S.H.-A.); Simons Foundation Collaboration on the Origins of Life, grant #327126 (T.B.). **Author contributions:** Conceptualization: N.M., S.G., and G.D. Methodology – data processing: O.G., P.P., S.L.M., J.F.B., J.I.N., M.R., A.M.O., B.H., C.Q.-N., J.D.T., R.A.Y., and L.C.K. Project administration: J.F.B., K.A.F., K.H.W., K.M.S., R.C.W., and S.M. Writing – original draft: N.M., S.G., G.D., A.J.B., B.H., B.W., J.F.B., O.G., and D.L.S. Writing – review & editing: N.M., S.G., O.G., G.D., J.D.T., S.F.S., B.H., R.A.Y., J.F.B., O.B., T.B., B.E., K.A.F., J.P.G., K.H.-L., S.H.-A., L.C.K., J.M.-F., S.M.L., J.I.N., J.W.R., M.R., J.I.S., D.L.S., K.M.S., V.Z.S., A.H.T., B.P.W., R.C.W., A.J.W., and K.H.W. Visualization: N.M., G.D., S.L.M., C.Q.-N., B.H., J.D.T., M.R., J.F.B., S.F.S., F.C., and N.R.W. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data used in this paper are available on the Planetary Data System (PDS). Tables S1 and S2 give links to PDS web pages for the Perseverance rover SuperCam and Mastcam-Z instruments and list the image numbers used in Figs. 1 to 4 and figs. S2 to S4, S6, S7, and S12. Data from the OMEGA instrument on Mars Express, used in fig. S11, are available at https://pds-geosciences.wustl.edu/mex/mex-m-omega-2-edr-flight-v1/mexomg-0001/data/ in the "gem04" and "gem22" directories. Data from the HiRISE instrument on MRO, used in Fig. 1 and figs. S1 and S9 to S11, are available at https://hirise-pds.lpl.arizona.edu/PDS/EDR/ESP/ORB_036600_036699/ESP_036618_1985/, https://hirise-pds.lpl.arizona.edu/PDS/EDR/ESP/ORB_037100_037199/ESP_037119_1985/, https://hirise-pds.lpl.arizona.edu/PDS/EDR/PSP/ORB_002300_002399/PSP_002387_1985/, and https://hirise-pds.lpl.arizona.edu/PDS/EDR/PSP/ORB_003700_003799/PSP_003798_1985/. The CRISM data used for fig. S11 are available at https://pds-geosciences.wustl.edu/mro/mro-m-crism-3-rdr-targeted-v1/mrocr_2101/trdr/2007/2007_029/hrl000040ff/hrl000040ff_07_if183l_trr3.img. The Context Camera image mosaic of Jezero used in Fig. 5 and fig. S1 is available at the United States Geological Survey https://astrogeology.usgs.gov/search/map/Mars/Mars2020/JEZ_ctx_B_soc_008_orthoMosaic_6m_Eqc_latTs0_lon0. The Entry, Descent, Landing (EDL) image used in fig. S5 is available at https://mars.nasa.gov/mars2020/multimedia/raw-images/. Our cobble size measurements, used to produce Fig. 3E and fig. S7, are provided in data S1.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abl4051
Materials and Methods
Supplementary Text
Figs. S1 to S12
Tables S1 to S3
References (*41*–*76*)
MDAR Reproducibility Checklist
Data S1

12 July 2021; accepted 21 September 2021
Published online 7 October 2021
10.1126/science.abl4051

DEVELOPMENTAL BIOLOGY

# Profiling cellular diversity in sponges informs animal cell type and nervous system evolution

**Jacob M. Musser[1]*, Klaske J. Schippers[1]†, Michael Nickel[1,2,3], Giulia Mizzon[4], Andrea B. Kohn[5], Constantin Pape[6], Paolo Ronchi[4], Nikolaos Papadopoulos[1], Alexander J. Tarashansky[7], Jörg U. Hammel[2,8], Florian Wolf[2], Cong Liang[9], Ana Hernández-Plaza[10], Carlos P. Cantalapiedra[10], Kaia Achim[1]‡, Nicole L. Schieber[6], Leslie Pan[1], Fabian Ruperti[1,11], Warren R. Francis[12], Sergio Vargas[12], Svenja Kling[1,13], Maike Renkert[1]§, Maxim Polikarpov[14,15], Gleb Bourenkov[14], Roberto Feuda[16], Imre Gaspar[1,17], Pawel Burkhardt[18], Bo Wang[7,19], Peer Bork[20], Martin Beck[20], Thomas R. Schneider[14], Anna Kreshuk[6], Gert Wörheide[3,12,21], Jaime Huerta-Cepas[10,20], Yannick Schwab[4,6], Leonid L. Moroz[5,22,23]*, Detlev Arendt[1,13]***

**The evolutionary origin of metazoan cell types such as neurons and muscles is not known. Using whole-body single-cell RNA sequencing in a sponge, an animal without nervous system and musculature, we identified 18 distinct cell types. These include nitric oxide–sensitive contractile pinacocytes, amoeboid phagocytes, and secretory neuroid cells that reside in close contact with digestive choanocytes that express scaffolding and receptor proteins. Visualizing neuroid cells by correlative x-ray and electron microscopy revealed secretory vesicles and cellular projections enwrapping choanocyte microvilli and cilia. Our data show a communication system that is organized around sponge digestive chambers, using conserved modules that became incorporated into the pre- and postsynapse in the nervous systems of other animals.**

Sponges represent a basal animal clade that lack neurons, muscles, and gut (Fig. 1A). They display canals for filter-feeding and waste removal (Fig. 1B and fig. S1) and are composed of three tissues: pinacocytes lining the canals, outer covering, and basal attachment layer (fig. S1, D to I); chambers of choanocytes with microvilli for food capture and motile cilia that drive water flow (fig. S1, J to O); and an inner mesohyl composed of stem, skeletogenic, and other mesenchymal cells. Despite their simple organization, sponges have genes that are usually expressed in neurons or muscles, including components of the pre- and postsynapse (*1*), and perform whole-body contractions that flush the canal system and expel debris (*2*). However, cells with integrative signaling functions are yet unknown.

### *Spongilla* is composed of 18 distinct cell types

Using whole-body single-cell RNA sequencing, we conducted a comprehensive survey

of genetically distinguishable cell types in 8-day-old freshwater demosponge *Spongilla lacustris*. Four capture experiments provided 10,106 *S. lacustris* cells expressing 26,157 genes (fig. S2, A to C).

Louvain graph clustering resolved 42 cell clusters with distinct expression signatures. Projection into two-dimensional expression space revealed a central cluster from which other clusters emanated (Fig. 1C and fig. S2D). This cluster expressed *Noggin*, *Musashi1*, and *Piwi-like*, demarcating stem cell–like archaeocytes (data S1) (*3*). We explored developmental relationships using partition-based graph abstraction (PAGA) (Fig. 1D), which assesses the degree of connectivity among clusters (*4*). This revealed connections along each arm of the t-distributed stochastic neighbor embedding (tSNE) plot and with many isolated clusters, suggesting developmental trajectories from archaeocytes to all other clusters. Supporting this, archaeocyte marker expression decreased with distance from the central cluster (Fig. 1E). Interconnecting clusters exhibited profiles intermediate between archaeocytes and peripheral clusters (Fig. 1E and figs. S4 to S10), with few specific markers (fig. S3A). By contrast, peripheral clusters showed distinctive expression profiles of transcription factors and effector genes (fig. S3 and data S1), which indicates differentiated cell types.

We next imaged cell clusters via single-molecule fluorescent in situ hybridization (smFISH) of selected marker genes. This resolved spatial relationships of stem cells, developmental progenitors, and differentiated cell types. Probes for the archaeocyte marker *Eef1a1* labeled large mesenchymal cells with a prominent nucleolus (fig. S11). Peripheral clusters 10 and 11 identified as choanocytes and apopylar cells by expression of actin-binding *Villin* and other choanocyte markers (*5*), whereas intermediate clusters 8 and 9 expressed proliferation markers, including *Pcna*, representing choanoblasts that often neighbored choanocyte chambers (fig. S11). Following similar strategies, we assigned 18 differentiated cell types (table S1) that represented a comprehensive molecular classification of *Spongilla* cell types consistent with previous morphological classification (*6*). For 13 previously described cell types, we use traditional nomenclature, or Greek translations (table S1). Five mesenchymal cell types were previously unrecognized, described here as myopeptidocytes, metabolocytes, and three different mesocytes.

### Interrelationships of cell types in sponges

Animal cell types are organized into families, which share expression of regulatory and effector genes and are identified by phylogenetic analysis (*7*). We first used weighted correlation network analysis to identify gene sets that covary across differentiated cell types (fig. S12A). This revealed 28 distinct gene sets, delineating individual cell types and cell type groups, which hinted at the presence of hierarchical organization. We validated this via a "treeness" score that was estimated for every combination of four cell types (tetrads), with close to 60% of tetrads exhibiting greater hierarchy than expected by chance, which indicated strong support for gene expression hierarchy that is independent of gene set or normalization method (fig. S12, B to D).

We then generated a cell type tree using neighbor-joining tree reconstruction (Fig. 2A), which revealed well-supported clades of differentiated cell types robust to bootstrapping and variable gene sets (fig. S12, E to G). Major cell type clades included extended pinacocyte and choanocyte families, as well as two families of mesenchymal cells, one of which contained amoebocytes and neuroid cells and another of which contained sclerocytes and mesocytes that are enriched for noncoding genes (fig. S3F). Notably, we observed mesenchymal cell types scattered across all families and internested with epithelial cell types. Evolutionary quantitative trait modeling revealed clade-specific genes and expression changes across the cell type tree (fig. S12H and data S1).

Assessing evolutionary conservation of cell types across sponges, we used self-assembling manifold mapping (SAMap) (*8*) to align *Spongilla* cells to a smaller single-cell dataset that was collected from adults of the demosponge *Amphimedon queenslandica* (*9*), which had a coarse identification of cell types based on expression signatures. We found consistent alignment of cell types within each family (Fig. 2B and fig. S13), albeit with higher resolution of annotated *Spongilla* pinacocyte- and choanocyte-related cell types. Newly identified *Spongilla* cell types aligned broadly with *Amphimedon* pinacocytes, as with metabolocytes, or were called an intermediate "choano-to-pinaco" cluster in *Amphimedon*, as with myopeptidocytes (Fig. 2B). For several *Spongilla* cell types, there were no *Amphimedon* counterparts, possibly because of incomplete cell type assignments in this dataset.

### Endymocytes are contractile and respond to nitric oxide signaling

We name each major cell type family to reflect its role within the organism. The family of endymocytes (ενδυμα: "lining, clothing") covers and shapes the sponge body (Fig. 3 and fig. S14). Incurrent pinacocytes constitute both layers of the tent, the vestibule lining, and the outer osculum layer. Apendopinacocytes line the excurrent canals and inner osculum layer. Basopinacocytes make up the basal epithelial layer attached to the substrate. Mesenchymal endymocytes were often found in proximity to pinacocytes and include collagen-secreting lophocytes, sclerophorocytes, and metabolocytes. Gene ontology (GO) analysis for endymocyte-specific markers found strong enrichment for Wnt and transforming growth factor–β (TGF-β) signaling and actomyosin-based contractility (figs. S14L and S15 and data S1), which is consistent with pinacocytes mediating whole-body contractions (*10*). Endymocytes also express contractile-cell master regulators *serum response factor* (*Srf*) and *Csrp1/2/3* (also known as *muscle LIM protein*) (*11*), which suggests a conserved regulatory module for actomyosin contractility (fig. S16).

In other demosponges, the signaling molecules γ-aminobutyric acid (GABA), glutamate, and nitric oxide (NO) alter whole-body contractions (*12*, *13*). We found paralogs of the metabotropic GABAB receptor used in distinct endymocytes (fig. S17). Pinacocytes also coexpressed *nitric oxide synthase*, which catalyzes

[1]Developmental Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany. [2]Friedrich-Schiller-Universität Jena, Institut für Zoologie und Evolutionsforschung mit Phyletischem Museum, Ernst-Haeckel-Haus und Biologiedidaktik, 07743 Jena, Germany. [3]GeoBio-Center, Ludwig-Maximilians-Universität München, 80333 München, Germany. [4]Electron Microscopy Core Facility, European Molecular Biology Laboratory, 69117 Heidelberg, Germany. [5]Whitney Laboratory for Marine Bioscience, University of Florida, St. Augustine, FL 32080, USA. [6]Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany. [7]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [8]Institute for Materials Physics, Helmholtz-Zentrum Hereon, 21502 Geesthacht, Germany. [9]Center for Applied Mathematics, Tianjin University, Tianjin 300072, China. [10]Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM) and Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), 28223 Madrid, Spain. [11]Collaboration for joint Ph.D. degree between EMBL and Heidelberg University, Faculty of Biosciences 69117 Heidelberg, Germany. [12]Department of Earth and Environmental Sciences, Paleontology & Geobiology, Ludwig-Maximilians-Universität München, 80333 München, Germany. [13]Centre for Organismal Studies (COS), University of Heidelberg, 69120 Heidelberg, Germany. [14]Hamburg Unit c/o DESY, European Molecular Biology Laboratory, Hamburg, 22607 Germany. [15]Department of Information Technology and Electrical Engineering, ETH Zurich, CH-8092 Zurich, Switzerland. [16]Department of Genetics and Genome Biology, University of Leicester, Leicester LE1 7RH, UK. [17]Department of Totipotency, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany. [18]Sars International Centre for Marine Molecular Biology, University of Bergen, 5008 Bergen, Norway. [19]Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305, USA. [20]Structural and Computational Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany. [21]Bayerische Staatssammlung für Paläontologie und Geologie (SNSB), 80333 München, Germany. [22]Department of Neuroscience and Brain Institute, University of Florida, Gainesville, FL 32610, USA. [23]McKnight Brain Institute, University of Florida, Gainesville, FL 32610, USA.
***Corresponding author. Email: jmmusser@gmail.com (J.M.M.); moroz@whitney.ufl.edu (L.L.M.); detlev.arendt@embl.de (D.A.)**
†Present address: Department of Tissue Dynamics and Regeneration, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany.
‡Present address: Faculty of Biological and Environmental Sciences, University of Helsinki, 00014 Helsinki, Finland.
§Present address: Centre for Biochemistry (BZH), University of Heidelberg, 69120 Heidelberg, Germany.

synthesis of NO, together with the NO receptor *Gucy1B2* (fig. S16). In the marine demosponge *Tethya wilhelma*, NO induces and modulates whole-body contractions (*14*) but only affected the osculum in the freshwater sponge *Ephydatia muelleri* (*13*). During the *S. lacustris* contraction cycle (fig. S18, A and B, and movie S1), the collapse of osculum and incurrent systems first causes excurrent canals to swell, which subsequently shrink to expel water through the osculum before returning to the resting state. We tested the response to NO by treating juvenile sponges with NOC-12, a donor of NO, and 1H-[1,2,4]Oxadiazolo[4,3-a]quinoxalin-1-one (ODQ), an inhibitor of the NO receptor (fig. S18, C to F, and movies S2 and S3). NO release immediately shrank the incurrent system, including incurrent canals and tent, and expanded the excurrent canals, as during initial stages of the contraction cycle (fig. S18C and movie S2). Subsequently, the incurrent system exhibited short, pulsed movements and returned to resting state only after NO removal (fig. S18C). The subtler effect of ODQ treatment involved expansion of incurrent, and possibly excurrent, canal systems (fig. S18D and movie S3). Serial combinations of NOC-12 and ODQ gave similar results (fig. S18, E and F), establishing a key role of NO in initiating and regulating the contraction cycle in *S. lacustris*, and possibly ancestrally in demosponges.

Beyond contraction, cell type–specific expression programs indicated endymocyte roles in photo- and mechanosensation, skeleton formation, glucose metabolism, and production of reactive oxygen for defense (fig. S15). Single orthologs of *Six* and *Pax* (*15*) are expressed in a complementary manner (fig. S16), with *Six*+ incurrent pinacocytes enriched for genes involved in circadian clock entrainment, and the blue-light sensor *cryptochrome*, which suggests light sensitivity, and *PaxB*+ apendopinacocytes expressing mechanosensory components and exhibiting short cilia (Fig. 3D and fig. S16). Next, we found basopinacocytes enwrapping spicules anchored at the base of the sponge (Fig. 3C and fig. S14, F and G) and sclerophorocytes forming clusters alongside mature spicules (Fig. 3, C and D, and fig. S14I), similar to *SoxB*+ spicule "transport cells" in *E. muelleri* (*16*). Both are enriched in genes that

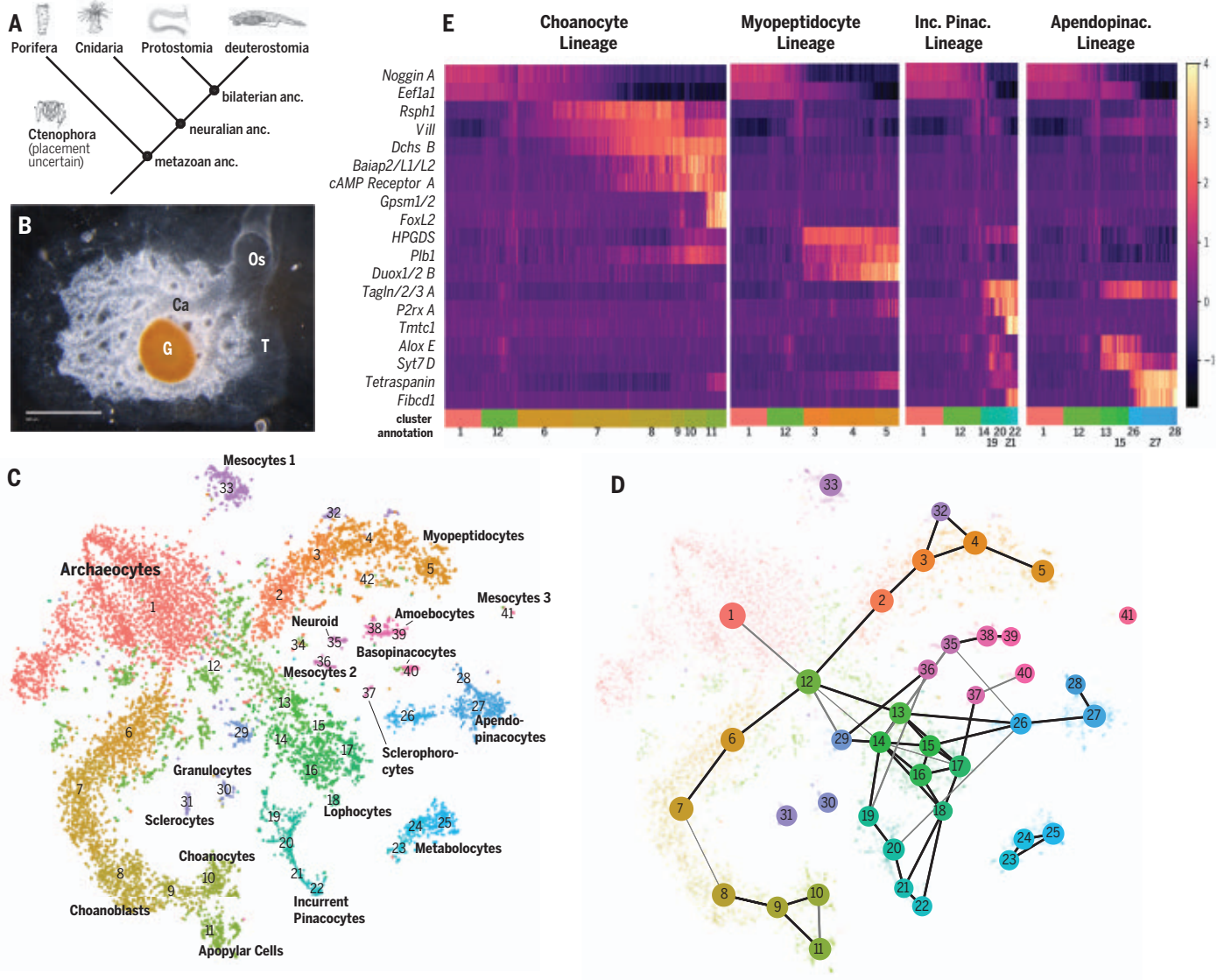



**Fig. 1. *S. lacustris* cell types from whole-body single-cell RNA sequencing.** (**A**) Animal phylogeny. anc., ancestor. (**B**) *S. lacustris* overhead view. Scale bar, 300 μm. Ca, excurrent canals; G, gemmule; Os, osculum; T, epithelial tent. (**C**) tSNE plot of 10,106 cells colored by 42 clusters. (**D**) PAGA connectivity plot showing connection strength (edge thickness) and log number of cells in cluster (node diameter). (**E**) Scaled expression values of differentially expressed markers for PAGA paths. Apendopinac., apendopinacocyte; Inc. Pinac., incurrent pinacocyte.

organize fibrillar collagen, and sclerophorocytes in secretory pathway genes (figs. S15 and S16). Multipolar metabolocytes exhibit wide extensions that may contact pinacocytes (Fig. 3D and fig. S14, J and K). They exhibit manifold metabolic activity (fig. S15), with specific upregulation of the complete set of enzymes that mediate glycolysis and the pentose phosphate pathway, and specific expression of *glycogen phosphorylase* and *pyruvate dehydrogenase* (fig. S16). Metabolocytes also express the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex known to produce reactive oxygen species for bacterial defense (figs. S15 and S16) (*17*).

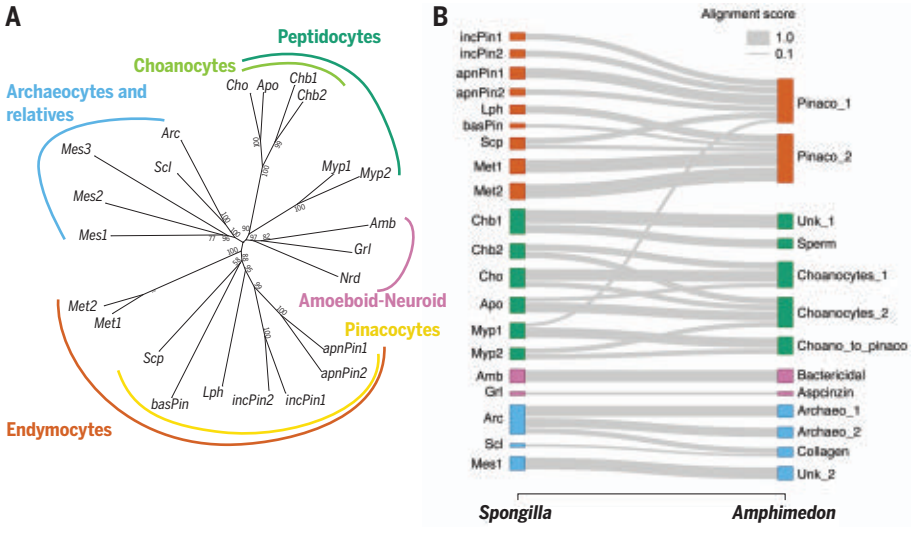



**Fig. 2. *Spongilla* cell type families and single-cell SAMap alignment to *Amphimedon*.** (**A**) *Spongilla* cell type tree with bootstrap support. (**B**) Sankey plot mapping *S. lacustris* to adult *A. queenslandica* cell types. Amb, amoebocytes; Apo, apopylar cells; Arc, archaeocytes; basPin, basopinacocytes; Cho, choanocytes; Grl, granulocytes; incPin, incurrent pinacocytes; Lph, lophocytes; Mes1, mesocytes 1; Mes2, mesocytes 2; Met, metabolocytes; Myp, myopeptidocytes; Nrd, neuroid cells; Scl, sclerocytes; Scp, sclerophorocytes.

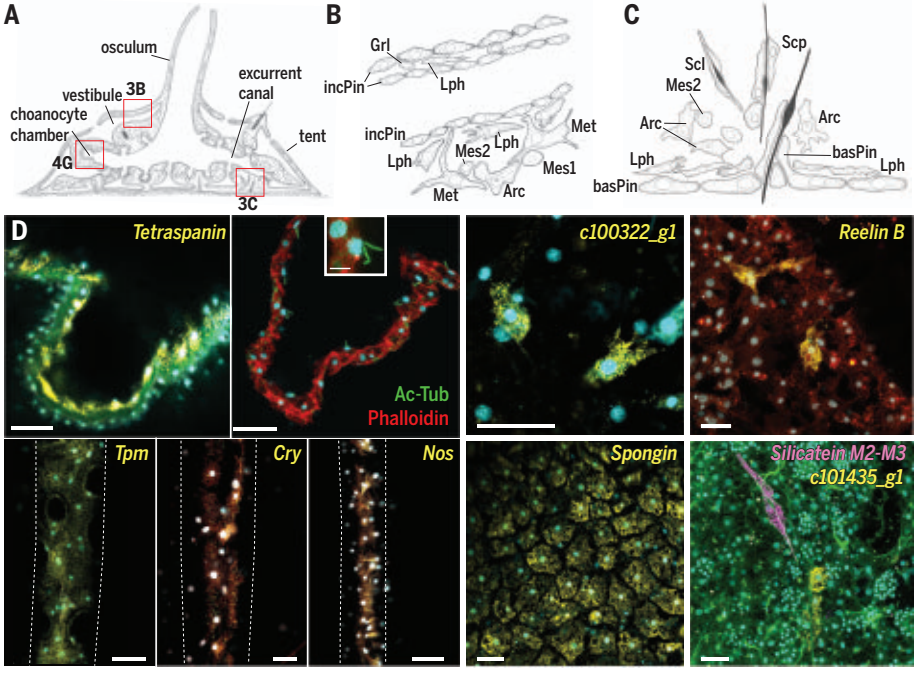



**Fig. 3. Endymocyte cell type family.** (**A**) Illustration of juvenile *S. lacustris*. Red boxes outline locations illustrated in other figure panels. (**B**) Drawing of incurrent pinacocytes (incPin) that make up the tent and vestibule, with adjacent mesenchymal cells: lophocytes (Lph), metabolocytes (Met), archaeocytes (Arc), mesocytes 1 (Mes1), and mesocytes 2 (Mes2). (**C**) Illustration of spicule production (sclerocytes, Scl), transport (sclerophorocytes, Scp), and anchoring (basopinacocytes, basPin). (**D**) smFISH of endymocyte markers. incPin, *Tpm*; apnPin, *Tetraspanin*; lph, *c100322_g1*; met, *Reelin B*; basPin, *Spongin*; Scp, *c101435_g1*. Membrane stains Fm-143Fx (red) and CellBrite Fix (green; also possibly stains collagen); nuclei 4′,6-diamidino-2-phenylindole (DAPI) stain (cyan). Scale bar, 30 μm.

### Digestive peptidocytes express "postsynaptic" scaffolding genes

The family of peptidocytes (πέπτειν, "to digest") share a role in digestion, which is reflected by phagocytic vesicle and lytic vacuole gene expression (figs. S19 and S20). They comprise choanocytes and apopylar cells, which form the excurrent pore of choanocyte chambers (Fig. 4, A, B, D, and G, and fig. S21). Myopeptidocytes are an abundant constituent of the mesohyl (Fig. 4C and fig. S21A) with long projections that contact each other and other cells, including choanocytes. They express actomyosin-based contractility modules (figs. S15 and S19) and three orthologs of *Mib2* (fig. S20) that regulates myofiber integrity (*18*), indicating a possible role for mesohyl in contractions.

In line with digestion functions (*19*), choanocytes express the transcription factors *Rbpj*, *Klf5*, and *NK homeobox 6* (fig. S20 and data S2 and S3) (*20*), and *sorting nexin* family members involved in macropinocytosis (figs. S19 and S20), an unregulated form of engulfment (*21*). The choanocyte subclade is enriched for genes of the receptive/postsynapse complex in animals with neurons (Fig. 5 and fig. S17), including "postsynaptic scaffold" (*homer*, *shank*, and *Baiap2*) and "*dystrophin-syntrophin* signaling complex" (*22*). Notably, scaffold genes such as *Baiap2* and *Eps8* are also active in organizing the actin skeleton of microvilli (*22*), and the single ortholog of *ezrin*, *radixin*, and *moesin*, known to stabilize the microvillar cytoskeleton, colocalized with F-actin in choanocyte microvilli (fig. S22). These data suggest a role of postsynapse-like scaffolding machinery in the choanocyte microvillar collar.

### The amoeboid-neuroid family: A role in cell clearance

The amoeboid-neuroid family comprises three small, tri- or multiangular cell types (Fig. 4, D to G, and fig. S21, B to G). The neuroid cells (*23*), also called "central cells" (*24*), associate with choanocyte chambers, whereas amoebocytes and granulocytes are dispersed in the mesohyl. We also observed granulocytes next to tent pinacocytes. Amoebocytes bear long extensions and were seen to engulf other mesohyl cells (Fig. 4F and fig. S21D), which is consistent with a macrophage-like role and classification as "bactericidal" cells in *A. queenslandica* (*9*). Family-specific transcription factors suggest a role in innate immunity. *Nuclear factor kappa B subunit* (*NFκB*), which controls cellular immune responses, and *high mobility group box 1/2/3* (*Hmgb1/2/3*)

(fig. S23) are universal sentinels of the nucleic acid–mediated innate immune response and activate key inflammatory pathways. Furthermore, amoebocytes specifically express the only *Nfat* ortholog, active in adaptive and innate immunity, and *Gata*, a conserved determinant of immunocytes (fig. S23) (*25*). We also noted family-specific expression of *engulfment and cell motility* (*Elmo*), which is essential for engulfing dying cells in worms (*26*), and the small guanosine triphosphatase *Rab21*, which is fundamental for phagosome formation (*27*).

### Communication between neuroid cells and choanocytes

Neuroid cells exhibited a secretory profile via up-regulation of the signal peptidase complex (fig. S19) and enrichment for "presynaptic" gene sets (also observed in endymocytes) (Fig. 5 and figs. S24 to S27) (*22*), suggesting communication via vesicle secretion. Notably, nearly all neuroid cells were associated with choanocyte chambers, with most chambers containing at least one neuroid cell in the middle of the chamber or in its lining (fig. S21, B and C). Corroborating historical descriptions (*23*, *24*), we observed multiple extensions from neuroid cells contacting individual choanocytes (fig. S21B).

To further characterize neuroid cell interactions, we devised a two-step imaging strategy, correlating confocal light microscopy and laboratory- or synchrotron-based x-ray imaging with focused ion beam scanning electron microscopy (FIB-SEM). These volumes are publicly accessible via the interactive FIJI plugin MoBIE (materials and methods). First, we acquired a lower-resolution FIB-SEM volume of an entire choanocyte chamber that contained two neuroid cells, identified by light microscopy. Using machine learning, we then segmented the entire volume (Fig. 6, A to C; fig. S28, A to C; and movies S4 and S5) and found one neuroid cell positioned centrally and another near the apopylar pore (Fig. 6, A and B), each with multiple protrusions directed toward different choanocyte collars and nearly always contacting and enwrapping one or more microvilli (Fig. 6, B and C, and fig. S28, A to C). Occasionally, extensions followed cilia that emerged straight and then bent, in contrast to normal undulatory cilia, which suggested that they may pause beating.

Second, we captured subcellular details of two additional neuroid cells via high-resolution FIB-SEM. Exploiting the deep penetration of x-rays (fig. S29), we combined the large field of view that was acquired via a laboratory-based x-ray microCT (microscopic x-ray computed tomography) with the smaller, higher-resolution field of view recordable on a synchrotron beamline. Registration of the synchrotron tomogram to the full sample block enabled precise trimming and acquisition of higher-resolution

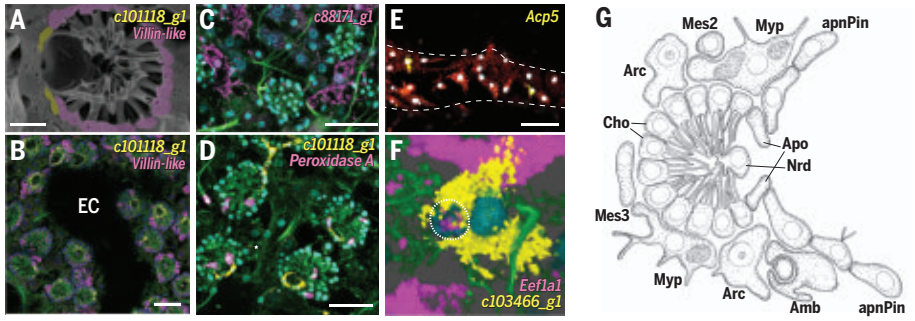



**Fig. 4. Peptidocyte and amoeboid-neuroid cell type families.** (**A**) SEM of choanocyte chamber. Scale bar, 10 μm. (**B** to **F**) smFISH of neuroid and amoeboid markers: choanocytes (Cho), *villin-like*; apopylar cells (Apo), *c101118_g1*; myopeptidocytes (Myp), *c88171_g1*; neuroid cells (Nrd), *Peroxidase A*; granulocytes (Grl), *Acp5*; and amoebocytes (Amb), *c103466_g1*. Dashed line indicates epithelial tent. Dotted line outlines archaeocyte being engulfed by amoebocyte. Membrane stains Fm-143Fx (red) and CellBrite Fix (green); nuclei DAPI stain (cyan). Scale bars, 30 μm. EC, excurrent canal. (**G**) Illustration of choanocyte chamber and neighboring mesenchymal cells.

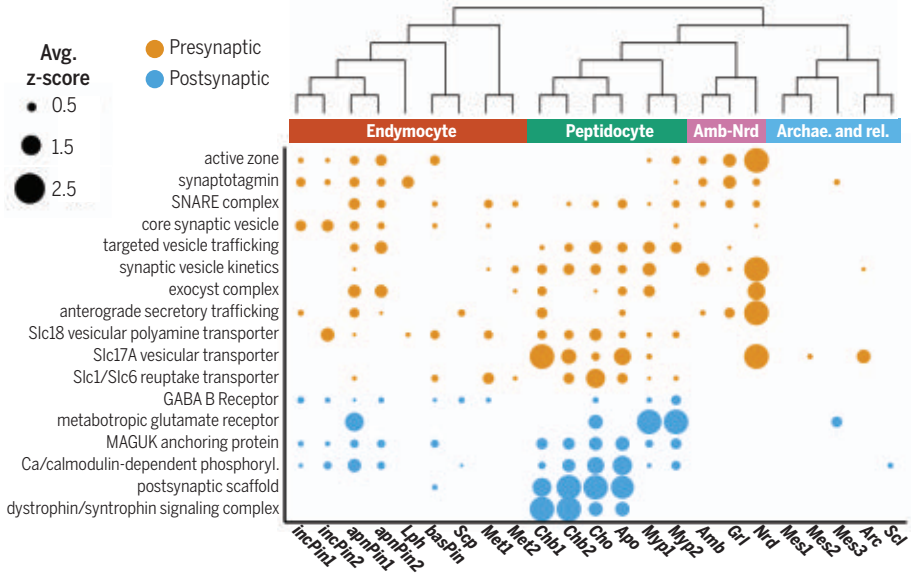



**Fig. 5. Expression of "pre-" and "postsynaptic" genes.** Dot plot showing expression of manually curated synaptic GO terms. Archae., Archaeocytes; rel., relatives.

FIB-SEM data. We term this method correlative x-ray electron microscopy (CXEM). In these volumes, we observed numerous secretory vesicles 50 to 500 nm in size throughout both neuroid cells, including in the cellular protrusions (Fig. 6D and fig. S28D), and two Golgi stacks per cell, each oriented toward the cellular protrusions (figs. S28, E and F). Larger vacuoles represented endolysosomes or phagosomes (fig. S28, G and H) and contained heterogeneous material, including a likely bacterial cell (fig. S28G, arrow), which is consistent with phagocytosis and engulfment gene expression. Lastly, both neuroid cells formed long invaginating pockets that encased cilia tips with shaft-like constrictions (Fig. 6E and fig. S28I). We never observed direct contact between neuroid cells and choanocyte collar or cilia that would have resembled synapses with targeted vesicle release. Together, these observations suggest a dual role for neuroid cells in intercellular communication and in the clearing of bacteria or cellular debris within choanocyte chambers.

### Discussion

A notable feature of gene and species evolution is a treelike history of common descent. We observed significant hierarchical organization among differentiated cell type expression programs in *Spongilla*, which was reflected by transcription factors and functional modules that are shared across subclades in the cell type tree. We propose that this reflects evolutionary relatedness rather than homoplasy (*28*) and that major sponge cell type families, such as

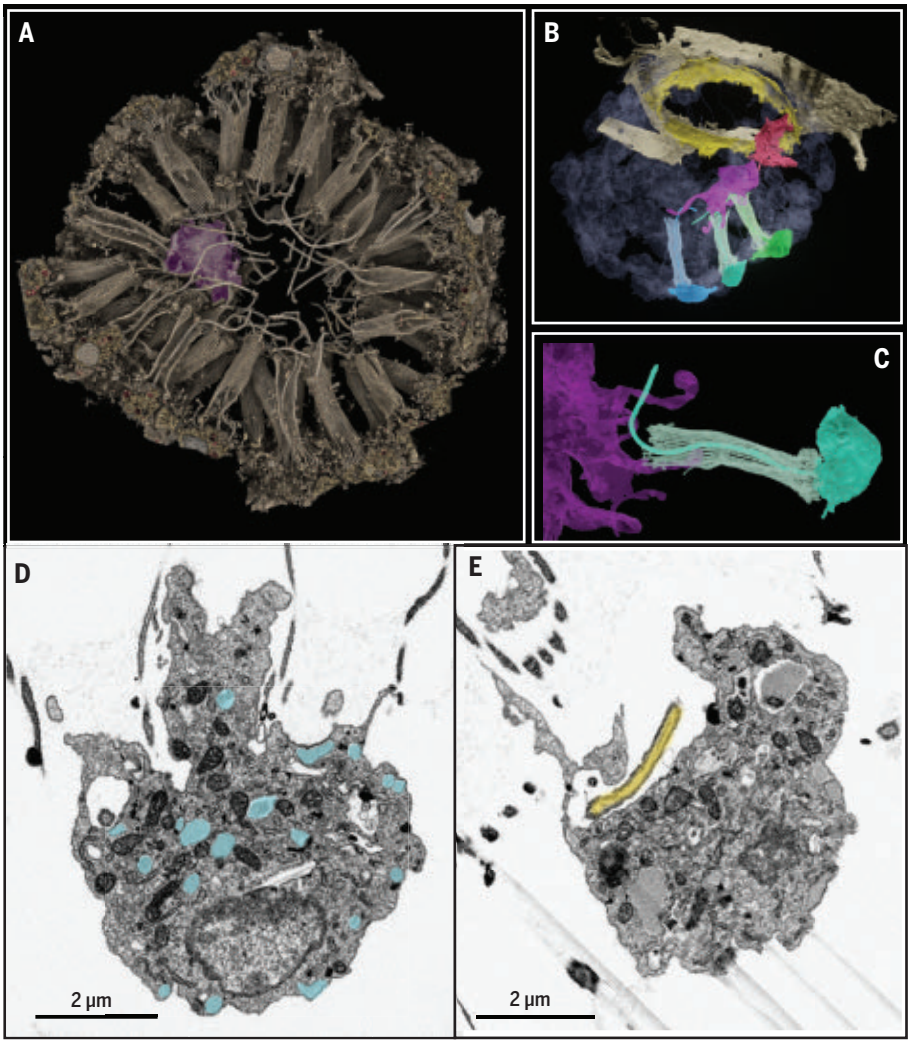



**Fig. 6. FIB-SEM of neuroid-choanocyte interactions.** (**A**) Rendered three-dimensional volume of choanocyte chamber with neuroid cell (violet). (**B**) Segmented volume showing two neuroid cells (violet and red) contacting cilia and microvillar collars of three choanocytes (blue, turquoise, and green) and apopylar cells (yellow). (**C**) Segmented neuroid cell (violet) with filopodia extending into the microvillar collar (turquoise). (**D**) High-resolution FIB-SEM section of neuroid cell 1 with secretory vesicles (cyan overlay). (**E**) High-resolution FIB-SEM section of neuroid cell 1 forming pocket around tip of choanocyte cilia (yellow overlay).

endymocytes and peptidocytes, originated via cell type diversification. These evolutionary lineages are distinct from developmental trajectories, although both histories can be congruent when closely related transcriptional programs arise from the same developmental precursors, as observed in this work.

For endymocytes, our experiments revealed coordination of the contraction cycle via NO signaling, which also triggers endothelial muscle relaxation in vertebrates (*29*). Endymocyte transcription factor signatures support an evolutionary link between sensory-contractile systems in sponges and other animals. In *Spongilla* and the calcisponge *Sycon ciliatum* (*30*), pinacocytes specifically express the sole sponge ortholog of *Msx*, which is known to specify sensory ectodermal regions and myocytes in other animals (*31*). *Msx* activates *atonal*, which plays a conserved role in mechanosensory neuron specification (*31*) and demarcates sponge pinacocytes together with its binding partner *Tcf4/E12/47*. Most notably, apendopinacocytes and incurrent pinacocytes express the single sponge orthologs of the *Pax* and *Six* homeodomain transcription factors, respectively, that also play conserved roles in sensory cell and myocyte specification (*15*). Our data are consistent with the hypothesis that sensory cells and myocytes arose by division of labor from shared evolutionary precursors that formed sensory-contractile, conducting epithelia in metazoan ancestors (*22*).

For peptidocytes, our analysis of cellular transcriptomes suggests digestive functions. Peptidocytes specifically express *Klf5*, which controls cytodifferentiation of intestinal epithelia (*32*), and choanocytes express *Nkx6*, *Rbpj*, and *Notch* that specify pancreatic cell types and absorptive enterocytes (*20*). *Nkx6* also demarcates pharyngeal and exocrine gland cells in sea anemone (*33*). Peptidocytes thus may have existed in metazoan ancestors, initially with intracellular digestion as observed in sponges and their unicellular relatives (*19*) and then acquiring external digestion, first as part of a digestive mucociliary sole and then incorporated into the gut (*34*).

For neuroid cells, our data suggest communicative functions within choanocyte chambers. Their close physical interaction with choanocyte microvillar collars and cilia, the expression of secretory genes, and the presence of secretory vesicles suggest a coordinative role in feeding, possibly by modulating ciliary arrest. *Spongilla* is a sponge with open architecture, in which incoming particles are continuously removed to prevent clogging. For this, the neuroid cells may stop water flow, engulf intruding cells and/or debris, and regulate microbial food uptake in response to quality and availability. Homology of neuroid cells to other animal cell types is unclear, as their transcription factor identity suggests no clear affinity other than limited resemblance to innate immune cells.

All three cell type families use conserved neural gene sets. This parallels the wider occurrence of neural modules across epithelial and fiber cells in the simple placozoans (*9*) and showcases possible parallel routes toward neuron evolution (*35*). In line with this, the contractile network and neurosecretory network hypotheses envisage the origin of neurons in distinct cellular contexts that resemble endymocytes and peptidocytes (*36*). Our work thus puts sponges center stage in elucidating nervous system evolution.

## REFERENCES AND NOTES

1. M. Srivastava *et al.*, *Nature* **466**, 720–726 (2010).
2. G. R. D. Elliott, S. P. Leys, *J. Exp. Biol.* **210**, 3736–3748 (2007).
3. A. Alié *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, E7093–E7100 (2015).
4. F. A. Wolf *et al.*, *Genome Biol.* **20**, 59–9 (2019).
5. J. F. Peña *et al.*, *Evodevo* **7**, 13 (2016).
6. N. Weissenfels, *Biologie und Mikroskopische Anatomie der Süßwasserschwämme (Spongillidae)* (Gustav Fischer, 1989).
7. C. Liang, FANTOM Consortium, A. R. Forrest, G. P. Wagner, *Nat. Commun.* **6**, 6066 (2015).
8. A. J. Tarashansky *et al.*, *eLife* **10**, e66747 (2021).
9. A. Sebé-Pedrós *et al.*, *Nat. Ecol. Evol.* **2**, 1176–1188 (2018).
10. M. Nickel, C. Scheer, J. U. Hammel, J. Herzen, F. Beckmann, *J. Exp. Biol.* **214**, 1692–1698 (2011).
11. J. M. Miano, X. Long, K. Fujiwara, *Am. J. Physiol. Cell Physiol.* **292**, C70–C81 (2007).
12. K. Ellwanger, A. Eich, M. Nickel, *J. Comp. Physiol.* **193**, 1–11 (2007).
13. G. R. D. Elliott, S. P. Leys, *J. Exp. Biol.* **213**, 2310–2321 (2010).
14. K. Ellwanger, M. Nickel, *Front. Zool.* **3**, 7 (2006).
15. A. Rivera *et al.*, *Evol. Dev.* **15**, 186–196 (2013).

16. S. Nakayama *et al*., *Curr. Biol.* **25**, 2549–2554 (2015).
17. T. Kawahara, M. T. Quinn, J. D. Lambeth, *BMC Evol. Biol.* **7**, 109–121 (2007).
18. M. Carrasco-Rando, M. Ruiz-Gómez, *Development* **135**, 849–857 (2008).
19. D. Laundon, B. T. Larson, K. McDonald, N. King, P. Burkhardt, *PLOS Biol.* **17**, e3000226–e22 (2019).
20. V. Gopalan *et al*., *Cancer Res.* **81**, 3958–3970 (2021).
21. J. T. H. Wang *et al*., *PLOS ONE* **5**, e13763–e13 (2010).
22. D. Arendt, The Evolutionary Assembly of Neuronal Machinery, *Curr. Biol.* **30**, R603–R616 (2020).
23. M. P. de Ceccatty, *Ann. Sci. Nat. Zool. Biol. Anim.* **17**, 203–288 (1955).
24. H. M. Reiswig, M. J. Brown, *Zoomorphologie* **88**, 81–94 (1977).
25. V. Hartenstein, L. Mandal, *BioEssays* **28**, 1203–1210 (2006).
26. T. L. Gumienny *et al*., *Cell* **107**, 27–41 (2001).
27. T. Khurana, J. A. Brzostowski, A. R. Kimmel, *EMBO J.* **24**, 2254–2264 (2005).
28. D. Arendt, P. Y. Bertucci, K. Achim, J. M. Musser, *Curr. Opin. Neurobiol.* **56**, 144–152 (2019).
29. R. M. J. Palmer, A. G. Ferrige, S. Moncada, *Nature* **327**, 524–526 (1987).
30. S. A. V. Fortunato *et al*., *Nature* **514**, 620–623 (2014).
31. Y. Li *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6352–E6360 (2017).
32. S. M. Bell *et al*., *Dev. Biol.* **375**, 128–139 (2013).
33. P. R. H. Steinmetz, A. Aman, J. E. M. Kraus, U. Technau, *Nat. Ecol. Evol.* **1**, 1535–1542 (2017).
34. D. Arendt, E. Benito-Gutierrez, T. Brunet, H. Marlow, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20150286–20150286 (2015).
35. L. L. Moroz, A. B. Kohn, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150041 (2016).
36. D. Arendt, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200347 (2021).
37. J. Musser, Gitlab repository zip for Profiling cellular diversity in sponges informs animal cell type and nervous system evolution, Zenodo (2021); https://doi.org/10.5281/zenodo.5094890

**ACKNOWLEDGMENTS**

We thank the Advanced Light Microscopy Facility, including S. Schnorrenberg, and the Electron Microscopy Core Facility at the EMBL for help with imaging and other support. We also thank G. P. Wagner and A. Dighe for valuable discussions about modeling gene expression on cell type trees, which served as the impetus for designing our quantitative trait modeling method to identify cell-type clade genes. **Funding:** This work was supported by grants from the European Research Council (BrainEvoDevo 294810 and NeuralCellTypeEvo 788921 to D.A.); by the Marie Curie COFUND program from the European Commission (to J.M.M.); by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 764840 IGNITE (to F.R.); by National Programme for Fostering Excellence in Scientific and Technical Research (grant PGC2018-098073-A-I00 MCIU/AEI/FEDER, UE; to J.H.-C.); by Severo Ochoa Centres of Excellence Programme from the State Research Agency (AEI) of Spain [grant SEV-2016-0672 (2017–2021) to C.P.C.); by Research Technical Support Staff Aid (grant PTA2019-017593-I / AEI / 10.13039/501100011033 to A.H.-P.); by LMU Munich's Institutional Strategy LMUexcellent within the framework of the German Excellence Initiative (to G.W.); by the Baden-Wuerttemberg Stiftung (to C.P.); and by the NIH and NSF (grants R01NS114491, 1146575, 1557923, 1548121 and 1645219) and Human Frontiers Science Program (to L.L.M.). **Author contributions:** Conceptualization: J.M.M., M.N., L.L.M., and D.A.; Methodology: J.M.M., M.N., A.B.K., Y.S., L.L.M., D.A., M.P., T.R.S., and G.B.; Software: J.M.M., M.N., C.P., N.P., A.J.T., C.L., A.H.-P., W.R.F., and J.H.-C.; Validation: J.M.M., K.J.S., S.K., and M.R.; Formal analysis: J.M.M., M.N., C.P., N.P., A.J.T., C.L., A.H.-P., W.R.F., and J.H.-C.; Investigation: J.M.M., K.J.S., M.N., G.M., A.B.K., P.R., J.U.H., F.W., L.P., F.R., K.A., N.L.S., S.K., M.R., I.G., and R.F.; Writing—Original draft: J.M.M. and D.A.; Writing—Review and editing, J.M.M., K.J.S., M.N., L.L.M., and D.A.; Visualization: J.M.M., K.J.S., M.N., G.M., C.P.C., P.R., N.P., A.J.T., F.W., A.H.-P., and D.A.; Supervision: J.M.M., M.N., P.Bu., B.W., P.Bo., M.B., A.K., T.R.S., G.W., J.H.-C., Y.S., L.L.M., and D.A.; Funding acquisition, L.L.M. and D.A. **Competing interests:** The authors declare no competing interests. **Data and materials availability**: Raw and processed RNAseq datasets generated for this study are available from NCBI GEO (accession number GSE134912). Custom analysis scripts, de novo transcriptome, proteome, and phylome are available on GitLab (https://git.embl.de/musser/profiling-cellular-diversity-in-sponges-informs-animal-cell-type-and-nervous-system-evolution) and archived at Zenodo (*37*).

**SUPPLEMENTARY MATERIALS**

science.org/doi/10.1126/science.abj2949
Materials and Methods
Figs. S1 to S29
Table S1
References (*38*–*94*)
Movies S1 to S5
Data S1 to S3

4 May 2021; accepted 23 September 2021
10.1126/science.abj2949

BIOSYNTHESIS

# Modular polyketide synthase contains two reaction chambers that operate asynchronously

**Saket R. Bagde[1,2], Irimpan I. Mathews[3], J. Christopher Fromme[2]*, Chu-Young Kim[1,4]***

**Type I modular polyketide synthases are homodimeric multidomain assembly line enzymes that synthesize a variety of polyketide natural products by performing polyketide chain extension and β-keto group modification reactions. We determined the 2.4-angstrom-resolution x-ray crystal structure and the 3.1-angstrom-resolution cryo–electron microscopy structure of the Lsd14 polyketide synthase, stalled at the transacylation and condensation steps, respectively. These structures revealed how the constituent domains are positioned relative to each other, how they rearrange depending on the step in the reaction cycle, and the specific interactions formed between the domains. Like the evolutionarily related mammalian fatty acid synthase, Lsd14 contains two reaction chambers, but only one chamber in Lsd14 has the full complement of catalytic domains, indicating that only one chamber produces the polyketide product at any given time.**

Polyketide natural products are bioactive molecules that are widely used and effective in medicine. This class of molecules includes erythromycin (an antibiotic), rapamycin (an immunosuppressant), and epothilone (a chemotherapeutic). Polyethers, a subgroup of polyketide natural products, are characterized by the presence of multiple cyclic ether groups in their structure. Natural polyethers vary in the number, size, and arrangement of the cyclic ether groups they contain. However, all polyethers are thought to be generated in nature through a common three-stage biosynthetic scheme (*1*): Stage 1 is the construction of the polyketide backbone by modular polyketide synthases (PKSs), stage 2 is the stereoselective epoxidation of the polyene intermediate by a monooxygenase, and stage 3 is the formation of the hallmark cyclic ether groups by one or more epoxide hydrolases.

Modular PKSs synthesize polyketides by performing successive Claisen-like condensation reactions (*2*, *3*). They are responsible for the biosynthesis of polyether polyketide, nonpolyether polyketide, and polyketide-nonribosomal peptide hybrid natural products. PKS modules minimally contain three functional domains, ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP), which together perform polyketide chain extension. One or more noncondensing domains, namely ketoreductase (KR), dehydratase (DH), and enoylreductase (ER), can transform the β-keto group formed during the condensation step to a hydroxyl, alkene, and methylene, respectively. Multiple modules act successively in an assembly-line-like fashion in which each module performs a single round of chain extension, followed by β-keto group modification reaction, and then transfers the growing polyketide chain to the next module. The final PKS module in the biosynthesis pathway typically contains a thioesterase (TE) domain that catalyzes release of the fully extended polyketide product.

Modular PKSs are structurally and functionally homologous to the mammalian and metazoan fatty acid synthase (*4*–*7*). However, fatty acid synthase performs iterative rounds of chain extension and produces a fully reduced alkyl chain, whereas each module of modular PKS performs a single-chain extension cycle and generates a reduced product. Additionally, only modular PKSs contain an N-terminal docking domain (DD) that facilitates inter-PKS communication. In both modular PKSs and mammalian fatty acid synthases, the growing alkyl chain remains attached to the enzyme until it is fully extended and processed. The growing chain is covalently attached to the phosphopantetheine (P-pant) group of the ACP domain. During the reaction cycle, ACP constantly changes its position, which enables all catalytic domains present in these mega

[1]Department of Chemistry and Biochemistry, The University of Texas at El Paso, El Paso, TX 79968, USA. [2]Department of Molecular Biology and Genetics/Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY 14853, USA. [3]Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Stanford University, Menlo Park, CA 94025, USA. [4]Border Biomedical Research Center, The University of Texas at El Paso, El Paso, TX 79968, USA.
***Corresponding author. Email: ckim7@utep.edu (C.-Y.K.); jcf14@cornell.edu (J.C.F.)**

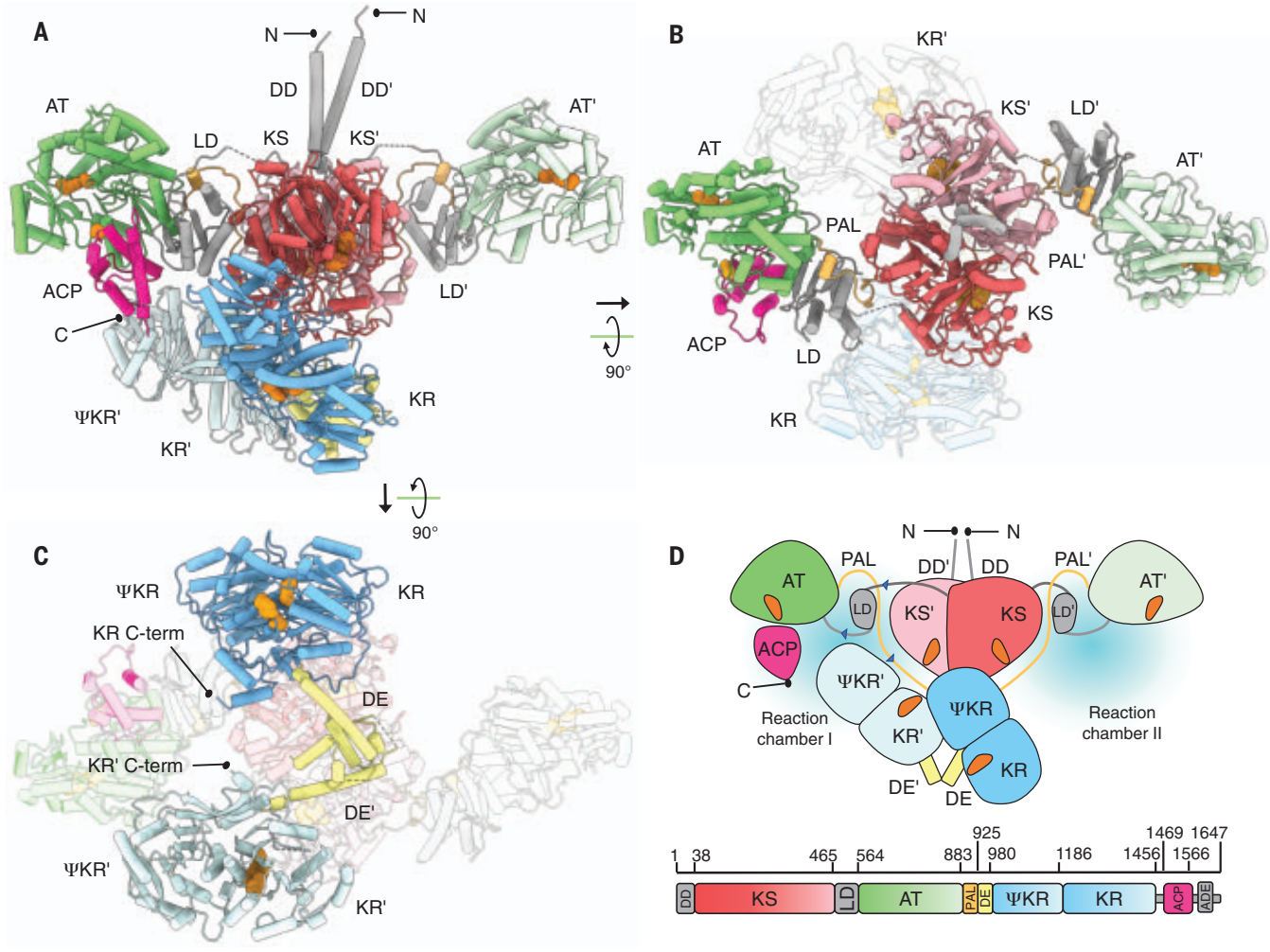



**Fig. 1. Crystal structure of Lsd14.** (**A** to **C**) Side view (A), top view (B), and bottom view (C) of Lsd14. Active site residues of each domain are shown in orange. (**D**) Cartoon representation and linear organization of Lsd14. ΨKR, KR structural subdomain.

enzymes to act on the growing chain. Therefore, the movement of ACP and the specific interactions that ACP forms with other domains is critical for polyketide and fatty acid production. Previous crystallographic studies of modular PKSs were conducted on proteolytic fragments and recombinantly expressed individual domains. Although these investigations have yielded important information on each domain, the three-dimensional organization of multiple domains and the nature of domain-domain interactions in an intact modular PKS are poorly defined. To address this knowledge gap, we selected the lasalocid A antibiotic biosynthesis pathway from *Streptomyces lasalocidi* as a model system for studying polyether biogenesis (fig. S1A). This pathway consists of seven modular PKSs (Lsd11 to Lsd17) that act sequentially to construct the dodecaketide backbone of lasalocid A from five malonyl-CoA, four methylmalonyl-CoA, and three ethylmalonyl-CoA units (*8*). Additionally, a flavin-dependent monooxygenase, Lsd18, converts the two *E*-olefins in the polyketide backbone into epoxides (*9*), and an epoxide hydrolase, Lsd19, transforms the epoxides into cyclic ether groups (*10*). We have previously reported the characterization of Lsd19 (*11*, *12*), and here we present the structural study of the Lsd14 modular PKS.

We first determined the crystal structure of *apo*-Lsd14 at 2.4-Å resolution, which revealed that Lsd14's eight catalytic domains form two unequal reaction chambers. This was an unexpected finding because all previous modular PKS models depict a symmetric architecture. In our crystal structure, ACP is docked to the AT, which is expected to occur during the transacylation step (fig. S1B). Because we could only crystallize Lsd14 in this conformation, we turned to cryo–electron microscopy (cryo-EM) to study alternative conformations of Lsd14. This effort yielded the 3.1-Å-resolution cryo-EM structure of *holo*-Lsd14 in which ACP is docked to the KS, which is expected to occur during the condensation step (fig. S1B). The Lsd14 cryo-EM structure also has an asymmetric architecture, further validating that the two reaction chambers adopt different conformations. Our work highlights the complementary nature of protein x-ray crystallography and cryo-EM techniques.

### X-ray crystal structure of apo-Lsd14 stalled at the transacylation step

We solved the x-ray crystal structure of *apo*-Lsd14 at 2.4-Å resolution (Fig. 1, fig. S2, and table S1, PDB 7S6B). Lsd14 is a single-module PKS composed of a pair of KS, AT, KR, and ACP domains. Seven of the eight functional domains that constitute the Lsd14 homodimer are clearly visible in the electron density map and could be modeled. One ACP domain was not detected, presumably because it is not locked into a single position. Additionally, parts of the AT-to-KR and KR-to-ACP linker were not visible in the electron density map (table S2), so the polypeptide chain identity of KR and ACP remained unassigned. We labeled

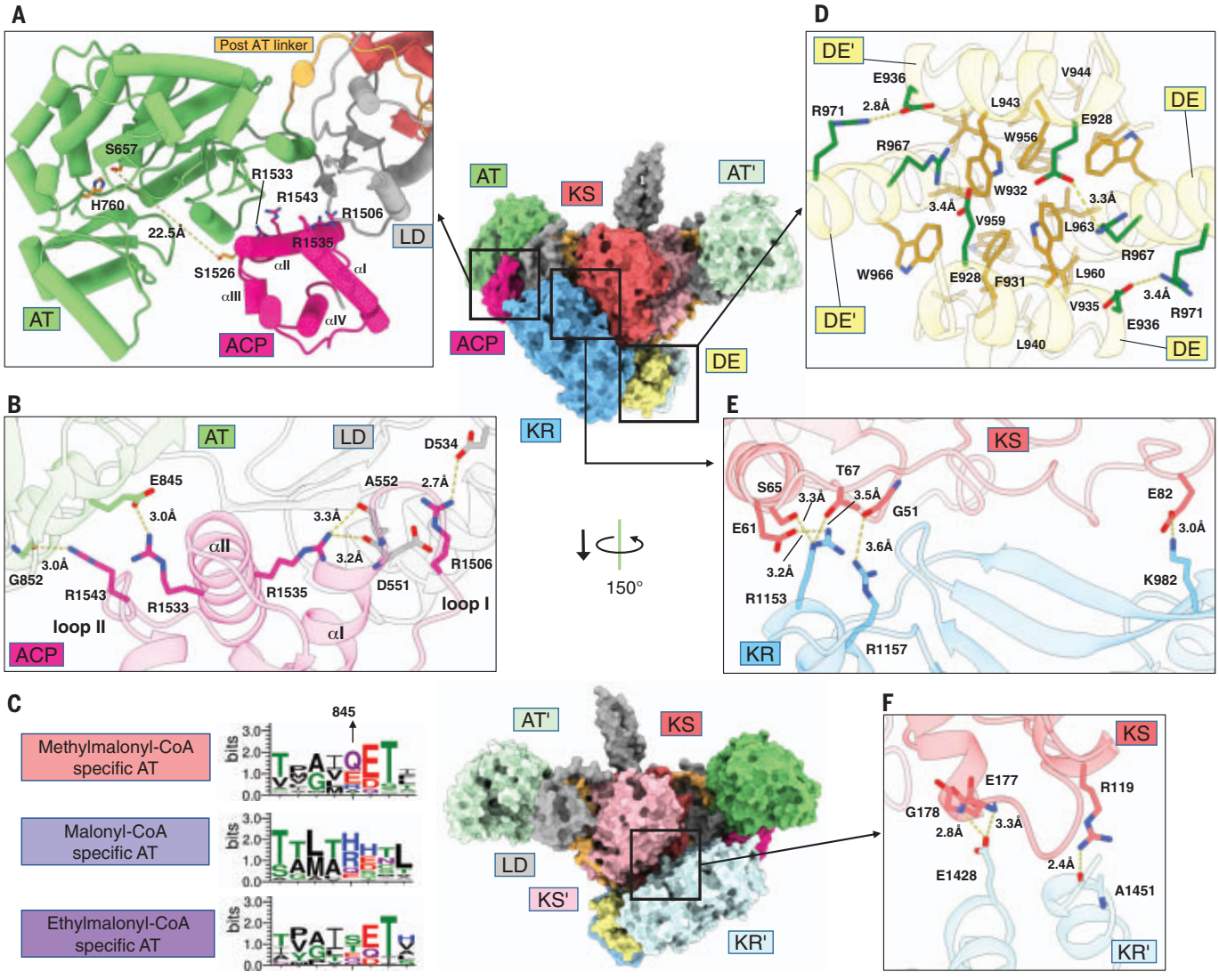



**Fig. 2. Interdomain interactions of *apo*-Lsd14 trapped in the transacylation step.** (**A**) ACP docking site. (**B**) Interactions at the AT-ACP interface. (**C**) WebLogo image showing relative frequency of amino acids at position 845 in AT sequences. (**D**) Interactions at the DE. (**E** and **F**) Interactions at the KR-KS and KR′-KS interfaces.

the equivalent domains in Lsd14 with and without the prime symbol (′) for clarity. However, only DD-KS-LD-AT-PAL (chain A) and DD′-KS′-LD′-AT′-PAL′ (chain B) have been crystallographically confirmed to reside on the same polypeptide chain.

Lsd14 has a compact homodimeric structure, with the dimer interface formed by DD/DD′ (~524 Å$^2$), KS/KS′ (~1952 Å$^2$), and pre-KR dimerization element (DE), DE/DE′ (~826 Å$^2$) (Fig. 1). The post-ACP dimerization element (ADE), which is not visible in the current structure, may provide additional contact surface area. The two AT domains lie on opposite ends of the KS dimer and are connected to the KS by a highly ordered linker domain (LD). The post-AT linker (PAL) loops back to the KS before connecting with the KR (fig. S3, A to E). The KS-LD-AT/KS′-LD′-AT′ dimer, hereby designated as (KS-LD-AT)$_2$, forms an extended structure with a twofold symmetry that mirrors the previously reported crystal structures of the (KS-LD-AT)$_2$ fragments from 6-deoxyerythronolide B synthase (DEBS) module 3 (PDB 2QO3), DEBS module 5 (PDB 2HG4), and curacin synthase module CurL (PDB 4MZ0) (fig. S3F) (*13*–*15*). The two KR domains are located below the (KS-LD-AT)$_2$ platform, thus creating two separate reaction chambers, reminiscent of the porcine fatty acid synthase architecture (*4*, *5*). However, Lsd14's two reaction chambers are substantially different from each other because of how the KR domains are positioned. In chamber I, the KS′, AT, and KR′ active site entrances all face the center of the reaction chamber, where ACP is located. However, in chamber II, only the KS and AT′ active site entrances face the center of the reaction chamber. Furthermore, ACP is docked to the AT chamber I, whereas no ACP is present in chamber II. In this configuration, transacylation, condensation, and β-keto group reduction can only take place in chamber I. Three structural features support this hypothesis. First, the Lsd14 sequence contains an ADE that places the second ACP also in chamber I, close to the observed ACP. Second, the C termini of both KR domains are located in chamber I, and because KR and ACP are tethered by a 12-residue linker, both ACPs likely reside in chamber I. Third, entrance to the KR active site is pointed away from chamber II, and therefore reduction cannot take place in this reaction chamber. Although chain elongation and β-keto group modification are expected to take place only in chamber I, chamber II can attain the same domain configuration as chamber I through an ~260° rotation of

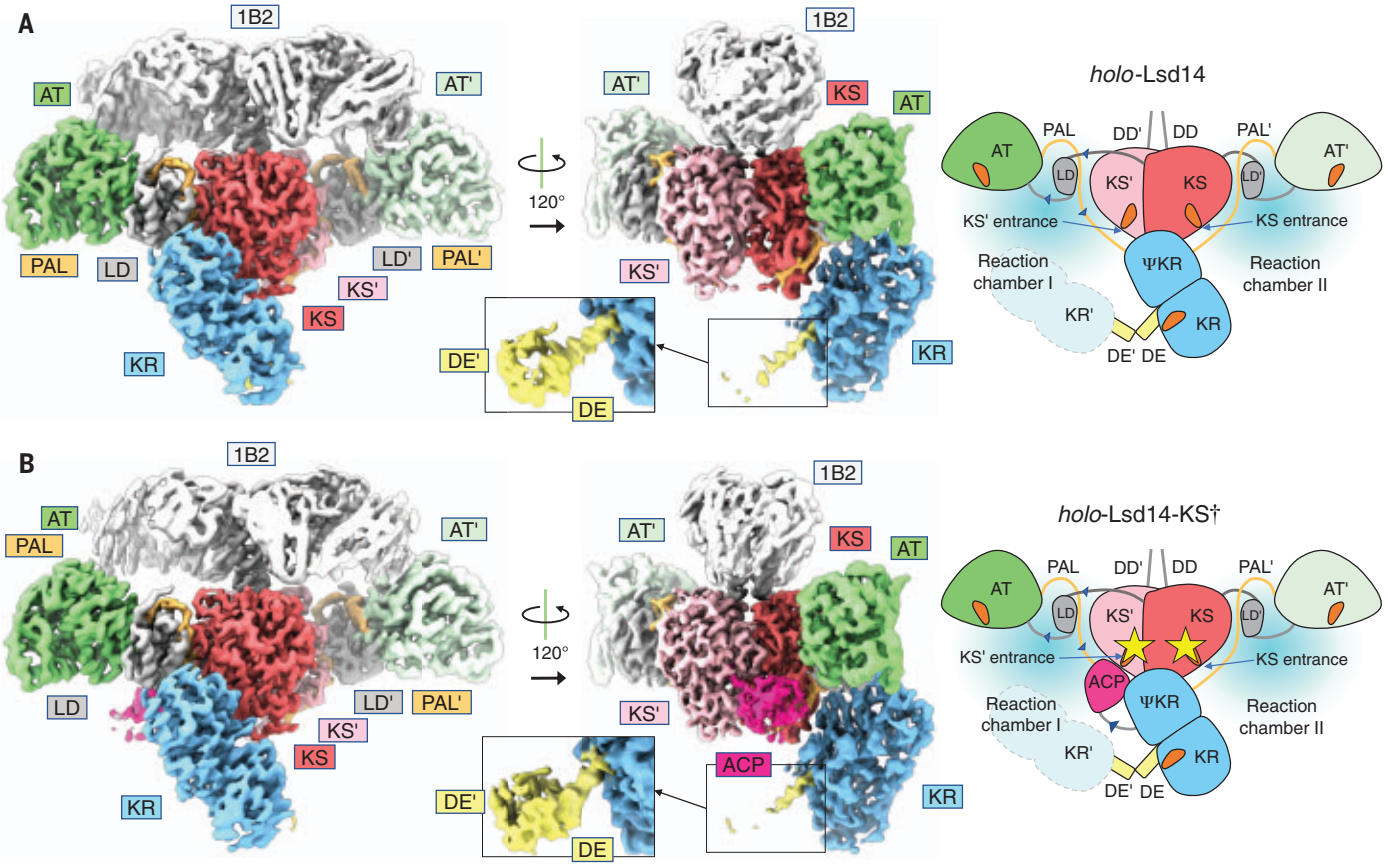



**Fig. 3. Cryo-EM maps of *holo*-Lsd14 with the DEBS docking domain bound to the 1B2 Fab fragment with and without treatment with KS substrate analog.** Composite maps made by combining unsharpened focused maps are shown. (**A** and **B**) *holo*-Lsd14-DD*+1B2 (A) and *holo*-Lsd14-DD*-KS†+1B2 [*holo*-Lsd14-DD*+1B2 treated with KS substrate analog; probable acylation is denoted by a yellow star on the KS active sites (B)]. The map threshold value is 14 and 12 for (A) and (B), respectively. Magnified version shows map for DE/DE′ interface at lower map threshold values of 7.7 and 5.5 for (A) and (B), respectively. A cartoon representation depicting the Lsd14 domain organization and location of active sites (orange) is shown on the right. Fab has been omitted from the cartoon for clarity.

$(KR)_2$. Overall, the asymmetric domain organization in Lsd14 indicates that its two reaction chambers operate asynchronously, and only one chamber performs polyketide chain synthesis at any given time.

### Interactions at the AT-ACP interface

In reaction chamber I, ACP is docked to the AT. Because *apo*-Lsd14 crystals were used for this study, ACP lacks the P-pant group modification. ACP residue S1526, which would bear the ~20-Å-long P-pant group in *holo*-Lsd14, is pointed at the catalytic residue S657 of AT (S1526 to S657 distance = 22.5 Å) (Fig. 2A). Multiple hydrogen bond and salt bridge interactions are present at the ACP docking site (Fig. 2B). We tested the requirement of these interactions by performing a multiple turnover assay using recombinantly expressed $(DD\text{-}KS\text{-}LD\text{-}AT)_2$ and ACP fragments of Lsd14 (fig. S4, A and B) (*16*). Among residues interacting at the AT-ACP interface, ACP residues are highly conserved across modular PKSs, whereas AT residue E845 varies according to the identity of the partner acyl-CoA (Fig. 2C and fig. S5). Typically, ATs specific for methylmalonyl-CoA and ethylmalonyl-CoA have an acidic or a polar residue at position 845, whereas ATs specific for malonyl-CoA have a basic residue. This structural variation may allow the ACP to dock to the AT in a slightly different orientation depending on the identity of the extender unit found on the AT. Our Lsd14 crystal structure contains a second AT-ACP interface, but this interaction appears to involve the AT′ from a crystallographic symmetry mate and seems to be artificial (fig. S4B and fig. S6, A and B).

The tertiary structure of AT and AT′ in Lsd14 are nearly identical (RMSD = 0.3 Å). Furthermore, the structure of ACP in the Lsd14 crystal structure is highly similar to the solution nuclear magnetic resonance structure of standalone ACP2 from DEBS (PDB 2JU2, RMSD = 1.5 Å) (fig. S6C) (*17*). In addition to the *cis*-AT-ACP complex in the current work, crystal structures of two *trans*-AT-ACP complexes, DSZS AT-ACP1 (PDB 5ZK4) (*18*) and VinK-VinL (PDB 5CZD) (*19*), have been reported. Unexpectedly, the AT-ACP interaction is not conserved in these three structures (fig. S7), suggesting that ACP has latitude in how it docks to the AT.

### Architecture and specific interactions of the DE-KR domains

We confirmed the KR activity of Lsd14 using a colorimetric KR activity assay (fig. S8). The KR domains dimerize through the DE, forming a single $(DE\text{-}KR)_2$ unit in which KR and KR′ lie parallel to each other in a head-to-head manner, but their active site entrances face opposite directions (Fig. 1C). The two DE domains form a six-helix bundle that is held together by multiple hydrophobic and hydrogen bond interactions (Fig. 2D and figs. S9A and S10). DE/DE′ helices I and III make specific interactions with KR/KR' in the crystal structure (fig. S9, B and C, and fig. S10). $(DE\text{-}KR)_2$ is docked to the KS/KS′ dimer in such a way that it occupies chamber I more than chamber II (Fig. 1). Furthermore, the active site entrance of KR′ points

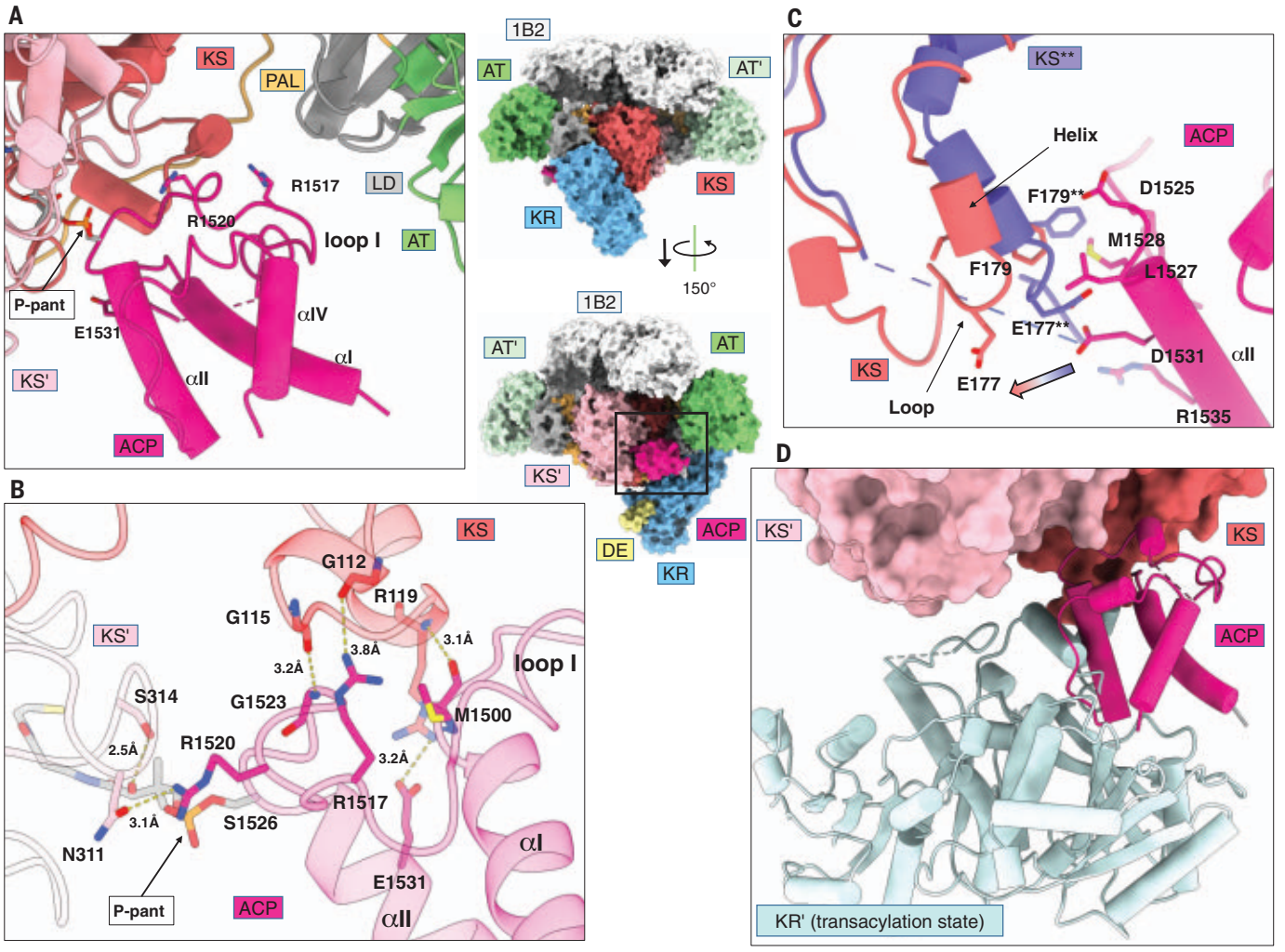



**Fig. 4. Interdomain interactions of *holo*-Lsd14 trapped in the condensation step.** (**A**) ACP docks in a cleft formed between the KS dimer and the LD-AT and the P-pant group stretches into the KS′ active site. (**B**) Interactions at the ACP-KS-KS′ interface. (**C**) Rearrangement of the KS helix + loop upon ACP binding. KS domain from the *holo*-Lsd14-DD*+1B2 (purple) superimposed on the KS domain from the *holo*-Lsd14-DD*-KS†+1B2 structure (salmon). Direction of the movement of KS helix + loop motif upon ACP binding is indicated by the multicolored arrow. (**D**) Region where ACP binds to KS during the condensation step partially overlaps with the region where KR′ binds to KS during the transacylation step.

toward KS′ in reaction chamber I, whereas the active site entrance of KR faces away from either reaction chamber (fig. S11). $(DE\text{-}KR)_2$ is a prominent, symmetry-breaking element in the Lsd14 architecture and has important functional implications that are described further below. $(DE\text{-}KR)_2$ appears to be mobile, as indicated by the relatively high B-factors calculated for KR/KR′ and DE/DE′ residues compared with the rest of the protein (fig. S12). The overall organization of $(DE\text{-}KR)_2$ in Lsd14 is similar to the crystal structure of the isolated $(DE\text{-}KR)_2$ from the amphotericin PKS (PDB 4L4X) and the spinosyn PKS (PDB 4IMP) (*20*, *21*). However, the relative positioning of the KR domains in spinosyn PKS differs from the other two structures because of flexing of helix III of the DE (fig. S9, D and E).

In reaction chamber I, both KR and KR′ contact the KS and establish multiple hydrogen bond and salt bridge interactions (Fig. 2, E and F). KS residues E61/S65/T67, and E82 form two regions with negative electrostatic surface potential that interact with positively charged surfaces formed by KR residues R1157/R1153 and K982, respectively (fig. S13).

### Cryo-EM structure of holo-Lsd14-Fab

To investigate Lsd14's domain organization in other states of the PKS reaction cycle, we conducted cryo-EM analysis of *holo*-Lsd14. Our initial attempts resulted in extensive dissociation of the Lsd14 homodimer into monomeric species on cryo-EM grids (fig. S14A). Our efforts to preserve the Lsd14 dimer integrity through buffer exchange, ligand addition, and chemical cross-linking failed to produce a stable Lsd14 homodimer suitable for cryo-EM study. We predicted that Fab 1B2, which was previously used by the Khosla group for structural study of the DEBS PKS (*22*), and also in the accompanying article (*23*), would bind and stabilize the Lsd14 homodimer because the residues involved in Fab binding are conserved in Lsd14 and DEBS. However, Fab 1B2 failed to bind to Lsd14, which prompted us to prepare an Lsd14 variant, Lsd14-DD*, which contains the DD of DEBS module 3. 1B2 readily bound to Lsd14-DD*, and the resulting complex, Lsd14-DD*+1B2, was suitable for cryo-EM analysis (fig. S14B). We also found that addition of methylmalonyl-CoA, the native AT substrate, was essential for visualizing an ordered KR domain. These strategies produced a 3.1-Å-resolution cryo-EM map of *holo*-Lsd14-DD*+1B2 (Fig. 3A). The final structure contains two 1B2 molecules,

with each Fab making contact with DD, KS, and AT, similar to the crystal structure of DEBS module 3 (DD-KS-LD-AT)$_2$ bound to 1B2 Fab (*22*) (fig. S15). The Lsd14 (KS-LD-AT)$_2$ in the cryo-EM structure has the same overall extended conformation as that in the Lsd14 crystal structure. Our in vitro enzyme assay showed that 1B2 did not affect the rate of transacylation (fig. S4D). Unlike in the Lsd14 crystal structure, only one KR is observed, and no ACP is visible in the cryo-EM structure. The local resolution of the DE/DE′ interface in the cryo-EM maps is low, likely because of the flexibility of the DE. However, visualization of the cryo-EM maps at a lower threshold reveals that the DE dimerizes in the cryo-EM maps, indicating that the (DE-KR)$_2$ is intact in the cryo-EM sample, although KR′ domain is unresolved (fig. S16). The interface between KR and KS is identical in the Lsd14 crystal and cryo-EM structures (fig. S17). We also imaged the *apo* form of Lsd14-DD*+1B2 under the same condition and found that the maps were essentially indistinguishable (fig. S18).

### Cryo-EM structure of holo-Lsd14-Fab stalled at the condensation step

Next, we treated *holo*-Lsd14-DD*+1B2 with the KS substrate analog 2-acetaminoethyl-thio-3-oxobutanoate and subjected this sample to cryo-EM analysis (fig. S14D). We obtained a cryo-EM map for this sample, referred to as *holo*-Lsd14-DD*-KS†+1B2, at an overall resolution of 3.1 Å (Fig. 3B). In this second cryo-EM structure, ACP is docked to KS in reaction chamber I. This interaction is required for the condensation step of the PKS cycle. The organization of KS/KS′, AT/AT′, and KR in *holo*-Lsd14-DD*-KS†+1B2 is the same as that of *holo*-Lsd14-DD*+1B2. We determined that the ACP in this structure is connected to KR, not KR′, based on the partially visible map density for the KR-ACP linker and linker length constraints (fig. S19).

The active site of KS′ in reaction chamber I and that of KS in chamber II can be accessed through their dedicated side entrances. In the *holo*-Lsd14-DD*-KS†+1B2 structure, ACP is docked at the KS′ side entrance in a cleft between the KS-KS′ dimer and the LD-AT domain, with its P-pant group stretched toward the KS′ active site (Fig. 4A). Residues in loop I and helix II of the ACP make specific interactions with a loop region in KS (G112-R119) and KS′ (N311) (Fig. 4B). The ACP and KS residues involved in these interactions are highly conserved in modular PKSs (fig. S20). The Lsd14 KS-ACP interface is consistent with previously reported biochemical results for DEBS (*24*), confirming the importance of ACP loop I and helix II for the ACP-KS interaction.

There is clear map density for the P-pant group attached to S1526 of ACP (fig. S21A). KS residue S314 forms a hydrogen bond with the carbonyl oxygen atom of the P-pant group (Fig. 4B). This serine was previously predicted to stabilize the binding of methylmalonyl-ACP to KS in DEBS (*25*). We also observed extra map density extending from the terminal sulfur atom of both the P-pant group and the catalytic cysteine (C210) of the KS domain (fig. S21A). The extra map density on the P-pant arm is likely the methylmalonyl moiety, the product of AT-catalyzed transacylation reaction. The extra map density at C210 may be caused by acylation because it was not observed in the *holo*-Lsd14-DD*+1B2 map, although the map density in this region is insufficient to accurately model this feature (fig. S21B).

Comparison of the region of KS that interacts with the ACP in *holo*-Lsd14-DD*-KS†+1B2 with the equivalent region in *holo*-Lsd14-DD*+1B2 revealed that the helix + loop motif within the KS (175-184) undergoes reorganization upon binding to ACP (Fig. 4C). The ACP-KS docking mode in the *holo*-Lsd14-DD*-KS†+1B2 structure is similar to that observed in the 7.1-Å-resolution cryo-EM structure of the CTB1-iterative PKS fragment (*26*) (fig. S22). R119 of KS, which forms a salt bridge with E1531 of ACP at the Lsd14 ACP-KS interface, was shown to be important for nor-toralactone production by CTB1 PKS (R461 in CTB1) (*26*). The *cis*-ACP-KS interfaces of Lsd14 and CTB1 are distinct from the previously observed *trans*-ACP-KS interfaces: *cis*-ACPs dock to the KS dimer by orienting loop I toward the KS, whereas *trans*-ACPs orient loop I away from the KS dimer (fig. S22).

In the Lsd14 crystal structure, ACP is docked to the AT (Fig. 2A), whereas KR′ is docked to the KS (Fig. 2F). In the holo-Lsd14-DD*-KS†+1B2 cryo-EM structure, ACP is docked to the KS/KS′ dimer at the KS′ active site entrance. Overlaying these two structures shows that the location of ACP in the cryo-EM structure overlaps with location of KR′ in the crystal structure (Fig. 4D). This indicates that the ACP cannot directly translocate from AT to KS′ upon completion of the transacylation reaction because the ACP docking site on KS′ is sterically blocked by KR′. Therefore, KR′ must first undock from the KS, which would enable the ACP to dock to the KS′ and allow condensation reaction to ensue. This process would ensure that ACP gains access to KS′ only after transacylation has taken place.

### Discussion

Both the crystal and cryo-EM structures of Lsd14 show an intriguing asymmetrical architecture that contrasts with previous PKS models derived from chemical cross-linking, domain complementation, cryo-EM, and small-angle x-ray scattering experiments, which depicted a perfect, or near-perfect, twofold symmetry (*22*, *27*–*29*). In our Lsd14 structures, ACP is found only in one of the two reaction chambers docked either to the AT or to the KS. The second ACP is not visible in the experimental maps, which suggests that it is not docked to any domain. Additionally, the presence of the ADE in the Lsd14 sequence suggests that the two ACPs remain close together. On the basis of these observations, we propose a pendulum clock model for Lsd14. In this model, (DD-KS-LD-AT)$_2$ is static while (DE-KR-ACP)$_2$ traverses between the two reaction chambers through a swinging motion with the DE as the fulcrum such that the two reaction chambers are used nonconcurrently (fig. S23). We hypothesize that chamber I performs extender unit transacylation, polyketide chain elongation, and β-keto group reduction, whereas chamber II becomes transacylated with the growing polyketide chain from the upstream PKS. Next, the (DE-KR-ACP)$_2$ pendulum swings to the opposite side and performs the same set of reactions that was just completed in chamber I. Alternatively, the two reaction chambers may be used stochastically, but only one polyketide product would be formed at a time.

PKS engineering has the potential to generate countless new polyketides for drug discovery. It has already been demonstrated that domain substitution, insertion, and deletion in modular PKSs leads to the production of new natural product analogs (*30*). However, the overall yield is drastically decreased, and mixed products are often formed. These deficiencies are most likely caused by protein misfolding and suboptimal domain-domain interactions resulting from introduction of non-native domains. Our Lsd14 structures show that the constituent domains in a modular PKS are elaborately organized and interact with one another through numerous specific interactions. To obtain the best result, domain manipulations in PKSs must be executed in a manner that preserves the native protein fold and interaction surfaces. Our work provides a high-resolution blueprint for designing minimally invasive domain manipulations. However, the current work only provides insight into the structure and function of a single-module PKS consisting of KS, AT, KR, and ACP domains. If we are to realize routine rational PKS engineering, future work is needed to reveal the workings of PKSs that also contain DH, ER, and TE domains and those with multiple modules on the same polypeptide chain. Furthermore, we must elucidate the structural basis of intermodule and inter-PKS communication.

Our study illustrates how x-ray crystallography and cryo-EM may be used concurrently to extract the maximum amount of protein structural information. However, we could not use native Lsd14 for cryo-EM analysis because the PKS dimers disintegrated into monomers when placed on the cryo-EM grid. We solved this problem by substituting the N-terminal docking domain in Lsd14 with the counterpart

from a DEBS PKS and then applying the Fab fragment that binds to the DEBS docking domain. By comparison, our crystal structure is that of the native full-length Lsd14 and thus accurately reflects the PKS as it is found in nature. Comparison of the crystal and cryo-EM Lsd14 structures shows that Fab binding only causes a slight movement of the docking domain, indicating that the cryo-EM Lsd14 structure is biologically relevant. The application of Fab to stabilizing the PKS homodimer may facilitate future structural studies of other modular PKSs using cryo-EM.

## REFERENCES AND NOTES


1. D. E. Cane, W. D. Celmer, J. W. Westley, *J. Am. Chem. Soc.* **105**, 3594–3600 (1983).
2. J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature* **348**, 176–178 (1990).
3. S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, L. Katz, *Science* **252**, 675–679 (1991).
4. T. Maier, S. Jenni, N. Ban, *Science* **311**, 1258–1262 (2006).
5. T. Maier, M. Leibundgut, N. Ban, *Science* **321**, 1315–1322 (2008).
6. E. J. Brignole, S. Smith, F. J. Asturias, *Nat. Struct. Mol. Biol.* **16**, 190–197 (2009).
7. F. M. C. Benning *et al.*, *ACS Nano* **11**, 10852–10859 (2017).
8. A. Migita *et al.*, *Biosci. Biotechnol. Biochem.* **73**, 169–176 (2009).
9. A. Minami *et al.*, *J. Am. Chem. Soc.* **134**, 7246–7249 (2012).
10. Y. Shichijo *et al.*, *J. Am. Chem. Soc.* **130**, 12230–12231 (2008).
11. K. Hotta *et al.*, *Nature* **483**, 355–358 (2012).
12. F. T. Wong *et al.*, *J. Am. Chem. Soc.* **137**, 86–89 (2015).
13. Y. Tang, A. Y. Chen, C.-Y. Kim, D. E. Cane, C. Khosla, *Chem. Biol.* **14**, 931–943 (2007).
14. Y. Tang, C.-Y. Kim, I. I. Mathews, D. E. Cane, C. Khosla, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 11124–11129 (2006).
15. J. R. Whicher *et al.*, *Chem. Biol.* **20**, 1340–1351 (2013).
16. B. J. Dunn, D. E. Cane, C. Khosla, *Biochemistry* **52**, 1839–1841 (2013).
17. V. Y. Alekseyev, C. W. Liu, D. E. Cane, J. D. Puglisi, C. Khosla, *Protein Sci.* **16**, 2093–2107 (2007).
18. A. Miyanaga *et al.*, *J. Am. Chem. Soc.* **140**, 7970–7978 (2018).
19. A. Miyanaga, S. Iwasawa, Y. Shinohara, F. Kudo, T. Eguchi, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1802–1807 (2016).
20. J. Zheng, S. K. Piasecki, A. T. Keatinge-Clay, *ACS Chem. Biol.* **8**, 1964–1971 (2013).
21. J. Zheng, C. D. Fage, B. Demeler, D. W. Hoffman, A. T. Keatinge-Clay, *ACS Chem. Biol.* **8**, 1263–1270 (2013).
22. X. Li *et al.*, *J. Am. Chem. Soc.* **140**, 6518–6521 (2018).
23. D. P. Cogan *et al.*, *Science* **374**, 729–734 (2021).
24. S. Kapur, A. Y. Chen, D. E. Cane, C. Khosla, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22066–22071 (2010).
25. T. Robbins, J. Kapilivsky, D. E. Cane, C. Khosla, *Biochemistry* **55**, 4476–4484 (2016).
26. D. A. Herbst *et al.*, *Nat. Chem. Biol.* **14**, 474–479 (2018).
27. J. Staunton *et al.*, *Nat. Struct. Biol.* **3**, 188–192 (1996).
28. C. M. Kao, R. Pieper, D. E. Cane, C. Khosla, *Biochemistry* **35**, 12363–12368 (1996).
29. S. Dutta *et al.*, *Nature* **510**, 512–517 (2014).
30. M. Klaus, M. Grininger, *Nat. Prod. Rep.* **35**, 1070–1081 (2018).


## ACKNOWLEDGMENTS


We thank C. Khosla for providing the 1B2 Fab expression plasmid. x-ray diffraction data were collected at the Stanford Synchrotron Radiation Lightsource (SSRL), Advanced Light Source (ALS), and Advanced Photon Source (APS). Use of the SSRL, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences under contract no. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). ALS is a U.S. DOE Office of Science User Facility under contract no. DE-AC02-05CH11231. APS is a DOE Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. We acknowledge the Cornell Center for Materials Research (CCMR), notably K. Spoth and M. Silvestry-Ramos, for access and support of electron microscopy sample preparation and data collection (NSF MRSEC program, DMR-1719875). We also thank the staff of the Proteomics and Metabolomics Facility at Cornell Institute of Biotechnology for conducting HPLC-MS. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS, NIH, or DOE. **Funding:** This work was supported by National Institutes of Health grants R01GM138990 to C.-Y.K., R35GM136258 to J.C.F., and 5U54MD007592. **Author contributions:** C.-Y.K. conceived and supervised the project. S.R.B. expressed, purified, and crystallized Lsd14. S.R.B. and I.I.M. conducted x-ray diffraction experiments. S.R.B., I.I.M., and C.-Y.K. solved the Lsd14 crystal structure. S.R.B. performed cryo-EM data acquisition and analysis. J.C.F. supervised the cryo-EM experiments. S.R.B. and C.-Y.K. wrote the manuscript with input from all authors. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All atomic coordinates and structure factors have been deposited in the Protein Data Bank (accession codes: 7S6B, 7S6C, and 7S6D). All cryo-EM maps have been deposited in the Electron Microscopy Data Bank (accession codes: EMD-24862, EMD-24863, EMD-24864, EMD-24865, EMD-24866, EMD-24867, EMD-24868, EMD-24869, EMD-24870, EMD-24871, EMD-24872, EMD-24873, EMD-24874, EMD-24875, and EMD-24880). All other data are presented in the main text or the supplementary materials.


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abi8532
Materials and Methods
Supplementary Text
Figs. S1 to S28
Tables S1 to S5
References (*31*–*56*)
MDAR Reproducibility Checklist


5 April 2021; accepted 21 September 2021
10.1126/science.abi8532


# REPORTS

BIOSYNTHESIS

# Mapping the catalytic conformations of an assembly-line polyketide synthase module


**Dillon P. Cogan[1]†, Kaiming Zhang[2,3]†, Xiuyuan Li[1], Shanshan Li[2,3], Grigore D. Pintilie[2], Soung-Hun Roh[2,4], Charles S. Craik[5], Wah Chiu[2,6]*, Chaitan Khosla[1,7,8]***



**Assembly-line polyketide synthases, such as the 6-deoxyerythronolide B synthase (DEBS), are large enzyme factories prized for their ability to produce specific and complex polyketide products. By channeling protein-tethered substrates across multiple active sites in a defined linear sequence, these enzymes facilitate programmed small-molecule syntheses that could theoretically be harnessed to access countless polyketide product structures. Using cryogenic electron microscopy to study DEBS module 1, we present a structural model describing this substrate-channeling phenomenon. Our 3.2- to 4.3-angstrom-resolution structures of the intact module reveal key domain-domain interfaces and highlight an unexpected module asymmetry. We also present the structure of a product-bound module that shines light on a recently described "turnstile" mechanism for transient gating of active sites along the assembly line.**


Polyketide synthases (PKSs) are multifunctional enzymes that synthesize numerous complex polyketides in nature. Mining of microbial and metazoan genomes continues to add to the surplus of untapped and potentially bioactive polyketide products and hybrids thereof. Among the small fraction (<5%) of characterized polyketides and their derivatives, some important medicines (*1*) and agrochemicals (*2*) have emerged. Many PKSs are composed of multiple homodimeric modules that individually catalyze C–C bond formation and optional modification of elongated intermediates. Such multimodular PKS arrays operate as enzymatic assembly lines by orchestrating the transfer of intermediates across modules in an ordered linear sequence. An example of a prototypical assembly-line PKS is the 6-deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythraea* (Fig. 1) (*3*, *4*). A representative catalytic cycle of the first module of this assembly line is described in fig. S1.

Given the homodimeric nature of assembly-line PKSs, we had previously sought to trace, through functional analysis of hybrid DEBS


[1]Department of Chemistry, Stanford University, Stanford, CA 94305, USA. [2]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [3]MOE Key Laboratory for Cellular Dynamics, Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, China. [4]Department of Biological Sciences, Institute of Molecular Biology & Genetics, Seoul National University, Seoul 151-742, Korea. [5]Department of Pharmaceutical Chemistry, University of California–San Francisco, San Francisco, CA 94158, USA. [6]Division of CryoEM and Bioimaging, Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Stanford University, Menlo Park, CA 94025, USA. [7]Department of Chemical Engineering, Stanford University, Stanford, CA 94305, USA. [8]Stanford ChEM-H, Stanford, CA 94305, USA.
***Corresponding author. Email: khosla@stanford.edu (C.K.); wahc@stanford.edu (W.C.)**
†These authors contributed equally to this work.

proteins, the trajectory of a growing polyketide chain as it is channeled from one covalently bound state to the next (*5*). Our data led to a model in which polyketide translocation occurred through acyl carrier protein (ACP) → ketosynthase (KS) substrate channeling within the same subunit (intramolecularly), and chain elongation occurred through KS→ACP substrate channeling across subunits (intermolecularly). In the present work, we used single-particle cryogenic electron microscopy (cryo-EM) to structurally validate this long-standing model for substrate channeling.

We present near-atomic-resolution structures of DEBS module 1 (M1, 3.2 to 4.3 Å) with backbone and side-chain resolvability in each of the relevant catalytic domains, including the ACP. In addition to validating the substrate-channeling model (Fig. 1), these structures shine light on hitherto

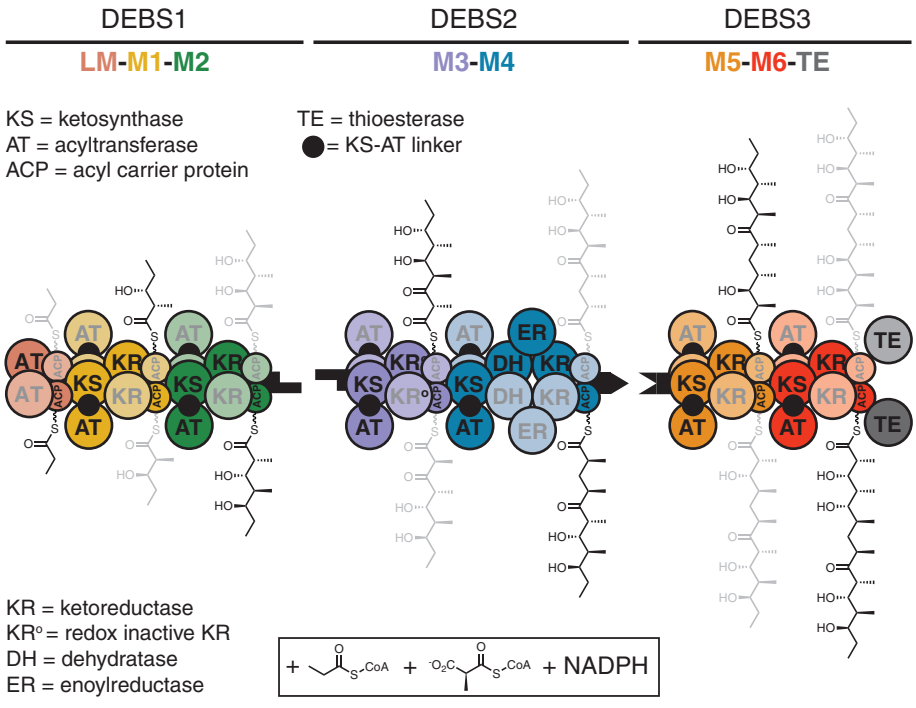



**Fig. 1. Model for channeling of biosynthetic intermediates by the DEBS assembly line.** DEBS includes three homodimeric polypeptides (DEBS 1 to 3) with a collective molecular mass of 2.1 MDa and harbors six elongation modules (M1 to M6) flanked by a loading module (LM) and a terminal TE domain. When provided propionyl-CoA (starter unit), (2*S*)-methylmalonyl-CoA (extender units), and NADPH, the assembly line synthesizes 6-deoxyerythronolide B, the macrocyclic aglycone of the antibiotic erythromycin. Two growing polyketide chains are synthesized in parallel by two sets of active sites of the homodimeric DEBS; individual monomeric copies of each module are distinguished by heavy and light shading. Earlier experiments established that translocation of the polyketide intermediate from one module to the next involves intrapolypeptide ACP→KS channeling, whereas polyketide chain elongation involves interpolypeptide KS→ACP channeling across the dimer interface (*5*). The cryo-EM structures solved in this work provide a clear mechanistic rationale for this observation. Shape-complementary black tabs at the end of each protein represent docking domains that facilitate intermodular polyketide translocation (*40*).

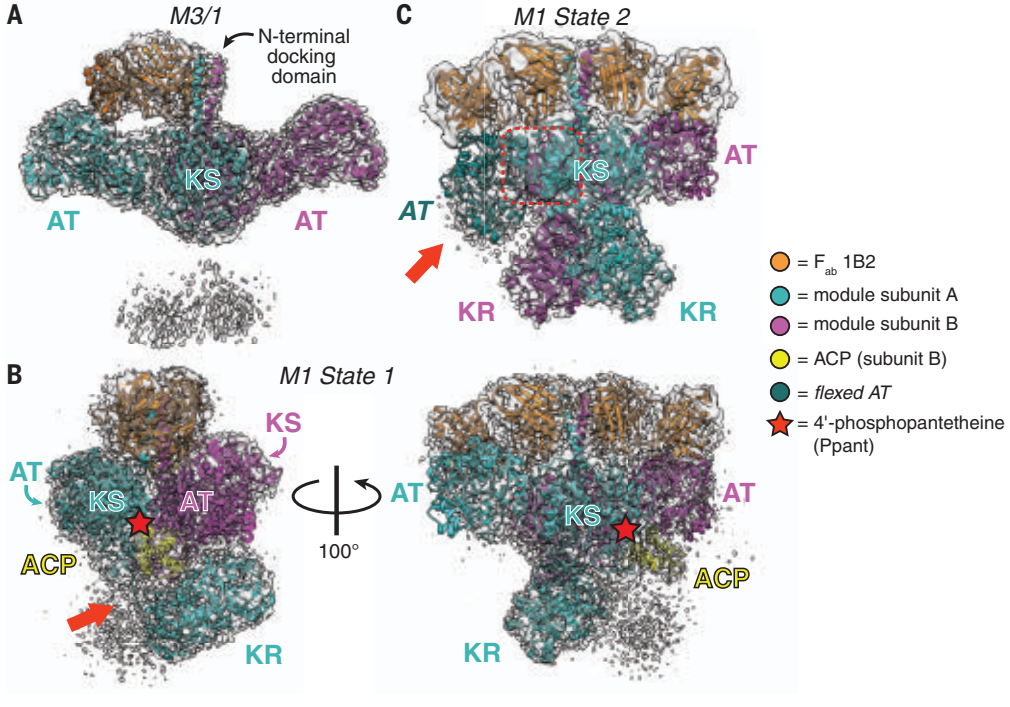



**Fig. 2. Snapshots of three asymmetric cryo-EM structures.** (**A**) The 3.2-Å-resolution structure of a hybrid module composed of domains from DEBS M3/1 (map threshold = 0.28). (**B** and **C**) Two distinct states of DEBS M1. (B) The 3.2-Å-resolution structure of *State 1* in which the P-pant cofactor (red star) of one ACP domain is bound in the KS active site of the other monomer (map threshold = 0.53). The red arrow highlights the resolved AT-KR linker (figs. S12, S13, and S14A and supplementary text). (C) The 4.1-Å-resolution structure of *State 2* featuring two symmetric KR domains in addition to an atypically flexed AT domain (red arrow). Dashed red box is the KS-AT linker that is not conformationally altered by AT flexing (fig. S17 and movie S1; map threshold = 0.35). Domain acronyms are defined in Fig. 1. Note that the color scheme (bottom) adopted for this figure is different from that of Fig. 1 because a different color palette was needed to clearly distinguish the module subunits, Fabs, and ACP.

uncharacterized domain-domain interfaces. Perhaps most unexpectedly, they also reveal module asymmetry not observed in previous structures of intact PKS modules (*6*–*9*). To further characterize a recently discovered "turnstile" mechanism (*10*), we also report the cryo-EM structure of DEBS M1 in its putatively *turnstile-closed* state, corresponding to the product-bound module. The resulting structure offers a plausible basis for KS active site shielding and vectorial biosynthesis by PKS assembly lines.

In preparation for structural analysis, we adapted a previous protocol for buffer optimization by using protein melt-curve analysis paired with catalytic activity measurements (described in the materials and methods) (*11*). DEBS M1 and M3 fused to the C-terminal thioesterase (TE) domain were selected as model specimens, given the extensive catalytic (*12*–*14*) and structural (*15*–*18*) characterization of M3 and the ease of access to the substrates of M1. The addition of citrate to the buffer elevated thermostability, catalytic activity, and oligomeric uniformity (figs. S2 to S6), so we used it as an additive for single-particle cryo-EM analysis. In addition to small-molecule buffer additives, antigen-binding fragments (Fabs) of antibodies can be used as tags (*19*, *20*) to augment the sample rigidity for high-resolution structure determination. Our initial target module was a recently characterized hybrid DEBS module, M3/1, comprising fragments from M3 and M1 fused to the TE domain (i.e., KS3-AT3-KR1-ACP1-TE) (*21*), because all of its catalytic domains have been previously crystallographically characterized at atomic resolution (*15*–*17*, *22*). M3/1 was complexed with a noninhibitory Fab (1B2) that is specific for the N-terminal docking domain (*15*) and purified in a buffer containing 100 mM citrate. The Fab-free module was also prepared under otherwise identical conditions (fig. S7). Combined with grid-freezing parameter optimization, diagnostic cryo-EM micrographs of the vitrified samples revealed Fab-bound modules to have markedly improved particle quality compared with Fab-free modules. This effect is likely exerted by the conformational selectivity of Fab 1B2 for modules in the dimeric state, because we previously measured numerous contacts between individual Fabs and each of the modular subunits at the dimeric interface (*15*). Substantiating the utility of this Fab, a companion study also used it to successfully study a PKS module from the lasalocid assembly line by single-particle cryo-EM (*23*). Collection of additional cryo-EM micrographs of the M3/1–Fab complex followed by particle picking and two-dimensional class averaging revealed particles with features resembling an extended KS–acyltransferase (AT) didomain (fig. S8). Further reconstruction without imposed symmetry resulted in a 3.2-Å-resolution map that could be fitted with two M3 KS-AT didomains bound to a single Fab molecule (Fig. 2A, fig. S8, and table S1), but neither the ketoreductase (KR) and ACP domains of M1 nor the C-terminal TE domains could be resolved. The single Fab occupancy contrasts with the symmetric binding of two Fab heterodimers observed crystallographically (*15*), likely because of a bent conformation in the coiled-coil docking domain (detailed in fig. S9). A local resolution range of ~3.2 to 5 Å between the N-terminal docking domain and the AT-KR linker of M3/1 allowed for unambiguous assignment of the extended conformation to its KS-AT didomain (fig. S8). In regions beyond the AT-KR linker, only two low-resolution lobes of density could be observed, which are assumed to be the flexible KR domains because of their domain connectivity (Fig. 2A). We speculated that the invisible KR-ACP-TE segment of this hybrid module might be disordered because of the lack of complementary interactions at its KS:KR interface.

On the basis of the above observations, we next purified DEBS M1 complexed to the same Fab. This module harbors a fully native KS-AT-KR-ACP segment fused to the N-terminal docking domain of M3 and the C-terminal TE domain (for improved expression and solubility). Single-particle cryo-EM analysis of the vitrified M1-Fab complex provided three structurally distinct and asymmetric maps with resolutions ranging from 3.2 to 4.1 Å, each with the expected Q-scores to reflect residue resolvability (figs. S10 to S12) (*24*). Each map reveals the KS-AT core bound to two Fab copies (fig. S10), as well as additional density for what appears to be one or both KRs. One map includes a clearly resolved ACP, although no features suggest a discernible TE domain.

The best-resolved map of the M1-Fab complex, obtained at 3.2-Å resolution and corresponding to one of the three classes, featured a single KR and a single ACP; this is hereafter referred to as the asymmetric *State 1*. Further model fitting and real-space refinement revealed an extended, homodimeric KS-AT core, albeit with reduced local resolution at the flanks of both AT domains, which may be altered relative to canonical KS-AT structures defined through x-ray crystallography (Fig. 2B, fig. S11, and table S1). The AT-KR linker could be fully traced to accurately assign the subunit to which the

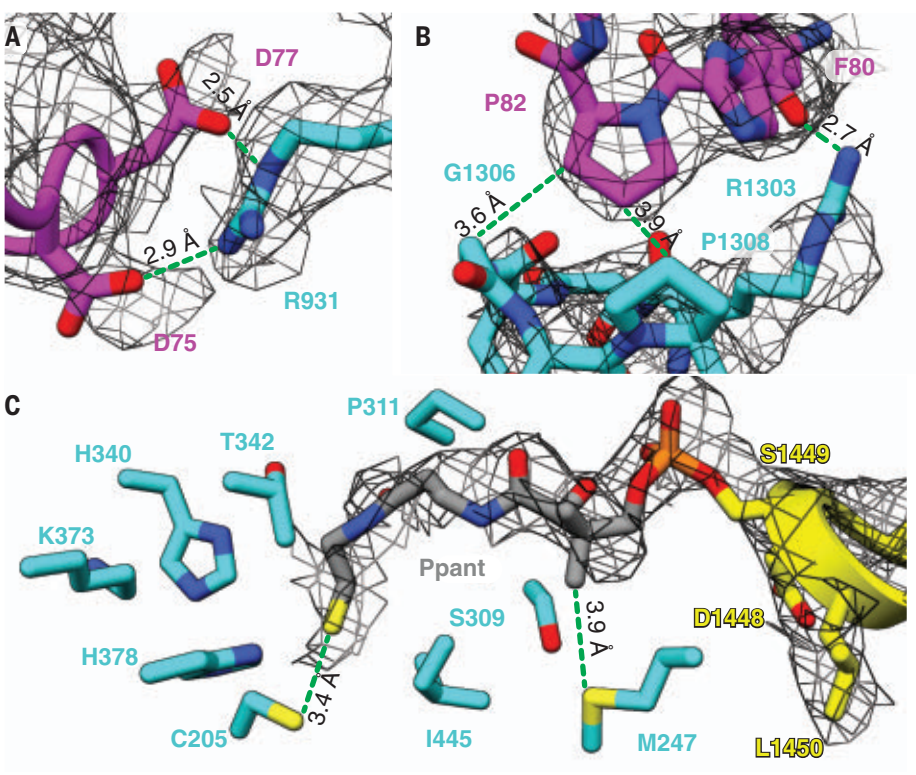



**Fig. 3. The KS:KR interface and the P-pant arm of the ACP domain in the cryo-EM density map of *State 1* of DEBS M1.** (**A** and **B**) Selected residues at the KS:KR interface. These residues are conserved in homologous PKS modules (fig. S15). (**C**) The P-pant arm is attached to Ser1449 of the ACP domain and extends to interact with Cys205 in the KS active site.

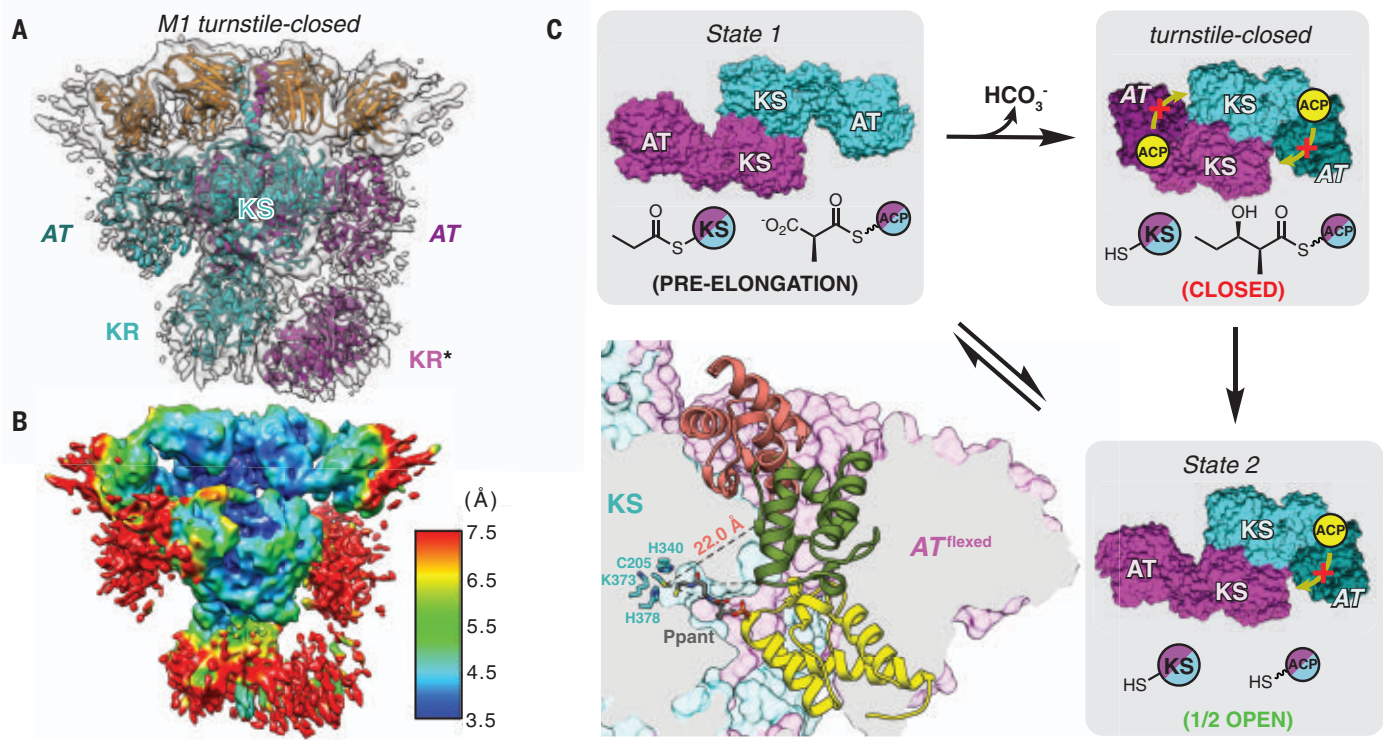



**Fig. 4. Cryo-EM analysis of DEBS M1 in its *turnstile-closed* state.** (**A**) The 4.3-Å-resolution cryo-EM structure of M1 in its post-elongation (*turnstile-closed*) state was solved, as described in the text (map threshold = 0.26). *The KR domain of subunit B (magenta) has been modeled here, although it is absent from the deposited coordinates because of orientation ambiguity and reduced local resolution [~7.5 Å in (B); PDB 7M7J]. (**B**) Resolution map for the structure in (A) (map threshold = 0.26). (**C**) Bottom left panel, ACP inaccessibility of the KS active site in this state. Accessibility of the cleft resulting from AT flexing was evaluated using the KS-ACP interaction pose observed in *State 1* (yellow ACP) as well as two prior models for KS-ACP interaction during intermodular polyketide translocation (*6*, *7*), (*29*, *30*) (salmon and green ACPs, respectively). Although the yellow and green poses are disfavored for steric reasons, the P-pant arm of the salmon pose is insufficiently long to access the KS active site (see fig. S23 for details). Remaining panels show a model for asymmetric turnstile closing of M1. Upon chain elongation, the pre-elongation state (*State 1*) undergoes a conformational change into its *turnstile-closed* state (A) to occlude the KS active site from an upstream or intramodular ACP, shown in yellow (movie S3). The duration of the *turnstile-closed* state is long enough to allow the ACP-bound diketide to undergo KR-catalyzed reduction and translocation to the downstream module, after which one catalytic subunit of the module returns to its translocation-competent state (*State 2*). Note that only the KS and AT catalytic domains are shown for clarity, and the KS and ACP acylation states are depicted for both modular subunits behaving equivalently but asynchronously.

visible KR is attached (Fig. 2B; figs. S12, S13, and S14A; and supplementary text). This KR domain is buttressed by a patch of non-canonically structured KS residues (Asp75 to Arg86; Fig. 2B and fig. S14, A and B). Its orientation is rotated by ~90° compared with our previous low-resolution model of DEBS M3 (*17*). Interacting residue pairs of the KS:KR intermolecular interface could be accurately visualized (local resolution ~3.2 to 5 Å; Fig. 3, A and B, and figs. S11 and S14B), providing an unprecedented glimpse into quaternary structural interactions between the KS-AT catalytic core and the reductive segment of an assembly-line PKS module. Mapping of the sequences of 300 homologous PKS modules onto the structure of DEBS M1 indicated considerable diversity at this domain-domain interface (fig. S14C) (*25*). More detailed comparative analysis between the six elongation modules of DEBS and 12 additional modules drawn from the closest homologs of DEBS M1 revealed that the KS:KR interface is more conserved across orthologs of DEBS M1 than other DEBS modules (fig. S15), perhaps reflecting a fitness advantage to the conservation of this interface in modules with shared biosynthetic chemistry. This protein interface also revealed a pronounced electrostatic complementarity (fig. S15E), thereby offering a foothold into structure-based engineering of heterologous KS:KR interfaces. Although previous domain-swapping efforts have exploited variable interdomain linker regions to define heterologous junctions (*26*–*28*), no studies to our knowledge have attempted to modify domain-internal residues that comprise the KS:KR interface observed in our *State 1* structure.

Perhaps the most noteworthy feature of *State 1* is that one of the ACP domains is observed in the KS-AT cleft in a manner consistent with our previous model for polyketide elongation (*29*, *30*). Earlier mutational analysis identified loop 1 residues of the ACP as being the most important for elongation; the same residues comprise the chief ACP contacts in *State 1*. Interdomain interactions are observed between the ACP and both KS domains of the homodimer as well as the KS-AT linker (fig. S16). Closer inspection of the KS:ACP interface revealed a tunnel of continuous density between the modified Ser (S1449) of the ACP and KS catalytic Cys (C205) separated by ~16 Å, consistent with the 4′-phosphopantetheine (P-pant) cofactor of the ACP embedding into the KS active site (Fig. 3C and fig. S16). Modeling of the P-pant arm and its congruence with the density is supported by Q-score analysis (fig. S12A). In addition to C205, the P-pant thiol is in close proximity to other KS active site residues (i.e., H340, K373, and H378),

consistent with placing a 2(*S*)-methylmalonyl extender unit favorably for catalysis. We therefore propose that *State 1* is descriptive of an asynchronously operating elongation mode of a homodimeric PKS module.

Another map, designated *State 2*, from the same M1-Fab particle dataset was resolved at 4.1-Å resolution and featured both KRs oriented symmetrically about the $C_2$ axis of the homodimeric module. This class was initially fitted from relevant portions of the *State 1* structure in addition to a second KR copy (Fig. 2C, figs. S10 and S11, and table S1). Subunit tracing in the *State 2* map favors a model in which the KRs interact with the KS of their own subunit, in contrast to the intermolecular KS:KR pairing observed in *State 1*. Thus, the reductive segment of an assembly-line PKS module must be capable of significant (~180°) $C_2$ rotational motions. Real-space refinement of the *State 2* model revealed another notable feature: One of the AT domains is flexed downward relative to the canonical crystallographic structure of the KS-AT core (Fig. 2C and fig. S17A). Unrestrained fitting suggests that the observed flexing of the AT domain hinges on two unstructured linkers, Thr551–Gln555 and Arg860–Ser864 (fig. S17A and movie S1) (*31*, *32*), although local resolution at the AT was reduced (fig. S11). This flexing occurs within the boundaries of the AT domain, leaving the KS-AT linker unaltered, a key distinction from the previously reported "arched" conformation of pikromycin synthase module 5 (fig. S17, A and B) (*6*, *7*). Furthermore, this flexing constricts the KS-AT cleft, thereby precluding the ability of the ACP domain to gain access to the KS active site as defined by *State 1* (fig. S18) and offering a clue into the structural basis for the experimentally observed gating of the KS (i.e., the turnstile mechanism) (*10*). Taking a reductionist, two-state-model approach, we therefore postulate that *State 2* is descriptive of the translocation and/or transacylation mode of a PKS module (fig. S1).

Although the associated map of the *State 1* structure does not resolve the KR-ACP linker with sufficient clarity to identify the subunit to which the structurally defined ACP is attached, we predicted it would be attached to subunit B (colored magenta in Fig. 2B) in accordance with earlier evidence that polyketide chain elongation occurs across subunits (*5*). To resolve this ambiguity, we further classified the M1 dataset using heterogeneous refinement to identify candidate classes with better structural resolution near the N terminus of the ACP domain. One class average meeting this criterion was selected and further processed to obtain a 4.1-Å-resolution map of M1. This model, hereby designated as *State 1′*, is reminiscent of *State 1* but reveals additional density corresponding to the second KR (figs. S10 to S12 and S19 and table S1). Although the KR and ACP domains are not as well resolved, a continuous region of density can be traced between these two domains, allowing inference of the subunit connectivity of the single ACP domain observed in the higher-resolution *State 1* structure (Fig. 2B and fig. S19); in turn, this structural model verifies that polyketide elongation occurs intermolecularly (*5*). As a corollary, the KR connected to the ACP must undergo significant rotational and translational motion relative to its position in *State 2* to position the ACP into its KS-AT cleft for elongation (fig. S19). These findings reinforce the notion that the lower torso region of the module (i.e., C-terminal to the AT) exhibits a high degree of conformational flexibility.

Integrating the above structural data, a model for the catalytic cycle of DEBS M1 emerges with four key features. First, the transition from its translocation and/or transacylation state (*State 2*) to its elongation state (*State 1*) requires the flexible KR domain of the homodimeric module to undergo rotational and translational motion. Second, chain elongation requires intermolecular docking of the ACP domain from one subunit onto the KS domain from the other subunit. Third, chain elongation is asynchronously catalyzed by the two KS-ACP pairs of the homodimeric module. Finally, post-elongation "turnstile" closing of M1 involves a conformational change in the KS-AT core that transiently occludes the KS active site from the ACP domain of the upstream loading module to allow the newly formed ACP-bound diketide to undergo β-ketoreduction and be translocated to M2.

To test this model, we prepared DEBS M1 in its diketide-bound form (fig. S1). To prevent chain release, the construct lacked a C-terminal TE domain (see the supplementary materials for details). The resulting protein was complexed with Fab 1B2, incubated with its native substrates for a sufficient duration to complete one catalytic cycle, promptly vitrified, and subjected to cryo-EM analysis. The resulting 4.3-Å-resolution cryo-EM map of the diketide-linked module (Fig. 4A, figs. S20 and S21, and table S1) revealed a distinct structural change in which both AT domains resided in a flexed orientation analogous to the flexed AT observed in *State 2* of M1 (Fig. 2C).

As before, the conformation of the flexed ATs in the product-bound state is accompanied by an ~8000-$Å^3$ reduction in the KS-AT cleft volumes, thereby blocking accessibility to acyl-ACP substrates for repeated elongation (fig. S18). Thus, the product-bound state appears to provide a structural rationale for turnstile closing, albeit with modest support from the poorly resolved AT domains (~7.5 Å, Fig. 4B; see fig. S22 and movie S2 for visualization of the map at various thresholds and movie S3 for visualization of turnstile closing). On the basis of this observation, we have therefore assigned the product-bound structure of M1 as the *turnstile-closed* state. To address whether the *turnstile-closed* state can support intermodular polyketide translocation, we evaluated two previous models for ACP binding during translocation (Fig. 4C and described in fig. S23). The results from model superpositions are consistent with a KS-AT cleft that is too constricted for ACP entry during translocation, providing an additional layer of support for the structural basis of turnstile closing. The catalytic cycle is summarized in Fig. 4C, and its conformational dynamic aspects are detailed in fig. S24.

At present, hundreds of assembly-line PKSs have been functionally annotated (*33*), and in silico prediction tools (*34*–*36*) are gaining widespread acceptance as more systems are sequenced and characterized (*37*). Although the elementary reactions of substrate loading and elongation are conserved across each assembly line, frequent reports that defy the DEBS paradigm continue to expand our sense of the enzymology that is possible both on (*38*) and off (*39*) the assembly line. Despite these advances, the fundamental question of how assembly-line PKSs coordinate their vectorial biosynthetic chemistry across multiple modules in sequence remains unanswered. A satisfactory answer will require high-resolution spatiotemporal information on PKS modular dynamics.

## REFERENCES AND NOTES


1. D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **83**, 770–803 (2020).
2. H. A. Kirst, *J. Antibiot. (Tokyo)* **63**, 101–111 (2010).
3. T. Robbins, Y.-C. Liu, D. E. Cane, C. Khosla, *Curr. Opin. Struct. Biol.* **41**, 10–18 (2016).
4. C. Khosla, Y. Tang, A. Y. Chen, N. A. Schnarr, D. E. Cane, *Annu. Rev. Biochem.* **76**, 195–221 (2007).
5. C. M. Kao, R. Pieper, D. E. Cane, C. Khosla, *Biochemistry* **35**, 12363–12368 (1996).
6. S. Dutta *et al.*, *Nature* **510**, 512–517 (2014).
7. J. R. Whicher *et al.*, *Nature* **510**, 560–564 (2014).
8. M. A. Skiba *et al.*, *ACS Chem. Biol.* **13**, 1640–1650 (2018).
9. J. Wang *et al.*, *Nat. Commun.* **12**, 867 (2021).
10. B. Lowry, X. Li, T. Robbins, D. E. Cane, C. Khosla, *ACS Cent. Sci.* **2**, 14–20 (2016).
11. A. Chari *et al.*, *Nat. Methods* **12**, 859–865 (2015).
12. R. S. Gokhale, S. Y. Tsuji, D. E. Cane, C. Khosla, *Science* **284**, 482–485 (1999).
13. A. Y. Chen, N. A. Schnarr, C.-Y. Kim, D. E. Cane, C. Khosla, *J. Am. Chem. Soc.* **128**, 3067–3074 (2006).
14. B. Lowry *et al.*, *J. Am. Chem. Soc.* **135**, 16809–16812 (2013).
15. X. Li *et al.*, *J. Am. Chem. Soc.* **140**, 6518–6521 (2018).
16. Y. Tang, A. Y. Chen, C.-Y. Kim, D. E. Cane, C. Khosla, *Chem. Biol.* **14**, 931–943 (2007).
17. A. L. Edwards, T. Matsui, T. M. Weiss, C. Khosla, *J. Mol. Biol.* **426**, 2229–2245 (2014).
18. S. C. Tsai *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 14808–14813 (2001).
19. L. Griffin, A. Lawson, *Clin. Exp. Immunol.* **165**, 285–291 (2011).

20. S. Wu *et al.*, *Structure* **20**, 582–592 (2012).
21. D. P. Cogan *et al.*, *J. Am. Chem. Soc.* **142**, 14933–14939 (2020).
22. A. T. Keatinge-Clay, R. M. Stroud, *Structure* **14**, 737–748 (2006).
23. S. R. Bagde, I. I. Mathews, J. C. Fromme, C.-Y. Kim, *Science* **374**, 723–729 (2021).
24. G. Pintilie *et al.*, *Nat. Methods* **17**, 328–334 (2020).
25. H. Ashkenazy *et al.*, *Nucleic Acids Res.* **44**, W344–W350 (2016).
26. M. Klaus, M. Grininger, *Nat. Prod. Rep.* **35**, 1070–1081 (2018).
27. S. Yuzawa *et al.*, *ACS Synth. Biol.* **6**, 139–147 (2017).
28. A. Hagen *et al.*, *ACS Synth. Biol.* **5**, 21–27 (2016).
29. S. Kapur, A. Y. Chen, D. E. Cane, C. Khosla, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22066–22071 (2010).
30. S. Kapur *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4110–4115 (2012).
31. L. G. Trabuco, E. Villa, K. Mitra, J. Frank, K. Schulten, *Structure* **16**, 673–683 (2008).
32. L. G. Trabuco, E. Villa, E. Schreiner, C. B. Harrison, K. Schulten, *Methods* **49**, 174–180 (2009).
33. S. A. Kautsar *et al.*, *Nucleic Acids Res.* **48** (D1), D454–D458 (2020).
34. K. Blin *et al.*, *Nucleic Acids Res.* **47** (W1), W81–W87 (2019).
35. P. Cimermancic *et al.*, *Cell* **158**, 412–421 (2014).
36. M. A. Skinnider, N. J. Merwin, C. W. Johnston, N. A. Magarvey, *Nucleic Acids Res.* **45** (W1), W49–W54 (2017).
37. A. Nivina, K. P. Yuet, J. Hsu, C. Khosla, *Chem. Rev.* **119**, 12524–12547 (2019).
38. A. T. Keatinge-Clay, *Chem. Rev.* **117**, 5334–5366 (2017).
39. E. J. N. Helfrich, J. Piel, *Nat. Prod. Rep.* **33**, 231–316 (2016).
40. R. W. Broadhurst, D. Nietlispach, M. P. Wheatcroft, P. F. Leadlay, K. J. Weissman, *Chem. Biol.* **10**, 723–731 (2003).

### ACKNOWLEDGMENTS

We thank J. Perrino and D.-H. Chen (Stanford University) for assistance with initial negative-stain and cryo-EM condition screening. **Funding:** This work was supported by the National Institutes of Health (grants R01GM087934 and R35GM141799 to C.K.; grant F32GM136039 to D.P.C.; and grants R01GM079429, P41GM103832, S10OD021600, and U24GM129541 to W.C.). **Author contributions:** C.K., W.C., D.P.C., and K.Z. conceived of the experimental design and project aims. C.S.C. provided key reagents and advice on best practices for their utilization. D.P.C. and K.Z. collected all of the non–cryo-EM experimental data, which were analyzed by all authors. K.Z., X.L., and S.-H.R. conducted preliminary experiments that informed the experiments described in this manuscript. K.Z. and S.L. performed the cryo-EM data collection and processing steps that led to the final deposited cryo-EM maps. K.Z. and D.P.C. refined the models. G.D.P. conducted the Q-score analysis and generated movies S2 and S3. C.K., W.C., and C.S.C. supervised all experiments. D.P.C. and C.K. wrote the initial manuscript, which was then revised and edited by all authors. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** All cryo-EM maps and atomic coordinates have been deposited in the Protein Data Bank under accession codes 7M7E, 7M7F, 7M7G, 7M7H, 7M7I, and 7M7J and in the Electron Microscopy Data Bank under accession codes EMD-23710, EMD-23711, EMD-23712, EMD-23713, EMD-23714, and EMD-23715. All other data are present in the main text or the supporting materials.

### SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abi8358
Materials and Methods
Supplementary Text
Figs. S1 to S24
Tables S1 to S3
References (*41*–*69*)
Movies S1 to S3
MDAR Reproducibility Checklist

2 April 2021; accepted 9 September 2021
10.1126/science.abi8358

ORGANIC CHEMISTRY

# Carbonyl 1,2-transposition through triflate-mediated α-amination

**Zhao Wu, Xiaolong Xu, Jianchun Wang, Guangbin Dong***

**To date, it remains challenging to selectively migrate a carbonyl oxygen within a given molecular scaffold, especially to an adjacent carbon. In this work, we describe a simple one- or two-pot protocol that transposes a ketone to the vicinal carbon. This approach first converts the ketone to the corresponding alkenyl triflate, which can then undergo the palladium- and norbornene-catalyzed regioselective α-amination and ipso-hydrogenation enabled by a bifunctional hydrogen and nitrogen donor. The resulting "transposed enamine" intermediate can subsequently be hydrolyzed to produce the 1,2-carbonyl–migrated product. This method allows rapid access to unusual bioactive analogs through late-stage functionalization.**

The specific position of a carbonyl group in a compound can strongly influence that compound's biological and physical properties, as well as its strategic use as a synthetic intermediate. For example, transposition of the C3-OH to the adjacent C2 position in ursolic acid (Fig. 1A) resulted in a 13-fold potency boost for inhibiting glycogen phosphorylase (*1*). Likewise, the compound derived from the C2-carbonyl isomer of nortropinone showed a ninefold increase in activity against aminoglycoside-induced hearing loss compared with the corresponding C3 analog (*2*). From a synthesis-planning viewpoint, efficient carbonyl 1,2-transposition methods could also simplify the construction of complex target molecules by allowing strategic bond disconnections with more accessible substrates; for example, the total syntheses of cascarillone and lycoraminone would be streamlined with a more efficient carbonyl 1,2-migration strategy (Fig. 1B) (*3*, *4*).

Carbonyl 1,2-transposition has not been a trivial process (Fig. 1C) (*5*–*7*). Currently, the most general strategy uses α-functionalizations of ketones to introduce a carbonyl surrogate, followed by a series of downstream transformations (*8*–*10*). Alternative tactics involve forming a three-membered ring or a 1,2-dione intermediate, which can be desymmetrized to afford the formal carbonyl-migrated products. Besides the need for a long synthetic sequence, substrate specificity and regioselectivity are additional concerns associated with the desymmetrization approaches (*11*–*14*). Hence, a general, regioselective, and straightforward carbonyl 1,2-transposition method with broad functional group (FG) tolerance remains an important goal. Given the ease of forming alkenyl sulfonates (e.g., triflates) from carbonyl compounds, we conceived the idea of developing a triflate-mediated α-amination process (*15*) to access a transposed enamine intermediate, which could undergo in situ hydrolysis to generate the carbonyl 1,2-migrated product (Fig. 1D). The key enamine-forming step involves simultaneous addition of a hydride to the ipso position (where the triflate was located) and a nitrogen substituent to the vicinal alkenyl carbon (the α position). We aimed to achieve α-amination and ipso-hydrogenation of alkenyl triflates by means of cooperative catalysis with palladium and norbornene (Pd/NBE) (*16*–*19*).

Pd/NBE cooperative catalysis, which was originally discovered by Catellani (*20*), has been extensively developed for arene functionalization since 1997 (*16*–*19*). By contrast, its use in functionalizing nonaromatic substrates has been very rare (*21*, *22*); in particular, it has not been used to introduce heteroatom substituents to alkenyl substrates. To realize the proposed α-amination and ipso-hydrogenation of alkenyl triflates through Pd/NBE catalysis, two major challenges must be addressed (Fig. 2). First, after the oxidative addition with Pd(0) and migratory insertion of the resulting alkenyl-Pd(II) species into NBE to avoid the undesired 3-*exo*-trig cyclization that produces side product **B** (a known major competing pathway) (*23*), the NBE cocatalyst must have a rigid substituent such as an amide moiety at the C2 position (the carbon bonded to Pd) to strongly favor formation of the key alkenylnorbornyl palladacycle (ANP) (*21*). Although the added steric hindrance on NBE may not hamper the reaction with less bulky electrophiles (e.g., primary alkyl halides), it could considerably retard the reaction when using more sterically demanding amine electrophiles. A sluggish amination would cause the premature hydride terminations—which form side products **A**

Department of Chemistry, University of Chicago, Chicago, IL 60637, USA.
***Corresponding author. Email: gbdong@uchicago.edu**

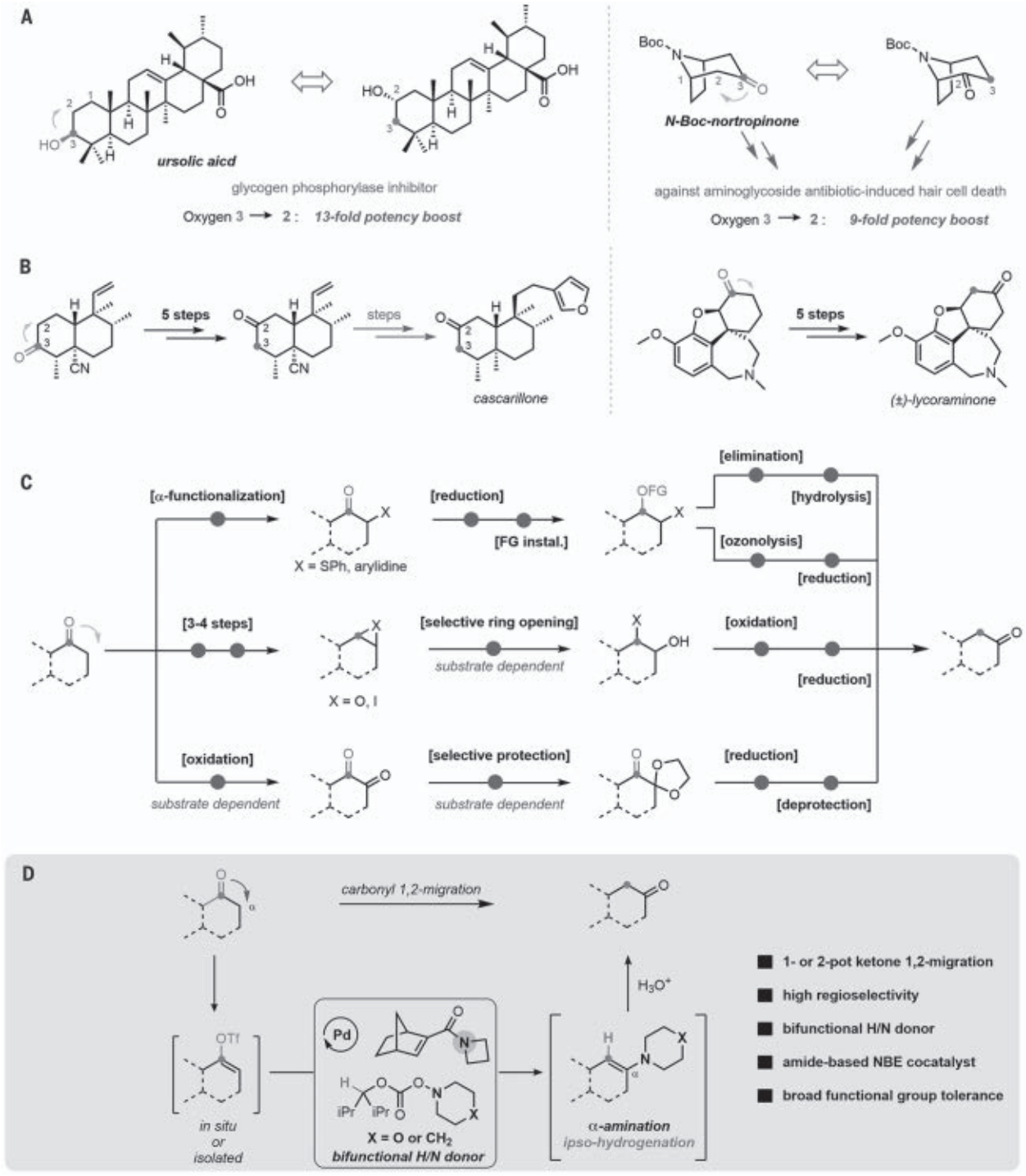



**Fig. 1. Carbonyl 1,2-transposition.** (**A**) Known examples of carbonyl 1,2-migration in medicinal chemistry to produce bioactive analogs with enhanced potency. Boc, *tert*-butoxycarbonyl. (**B**) Carbonyl 1,2-transposition as the key strategy in total syntheses of natural products cascarillone and lycoraminone. (**C**) Multistep FG manipulation is often required to relocate ketone carbonyl to the adjacent carbon. (**D**) One- or two-pot ketone 1,2-migration developed here: α-amination and ipso-reduction of alkenyl triflates through the Pd/NBE catalysis. iPr, isopropyl; OTf, triflate.

and **C**—to outcompete the desired pathway. Second, even if the C–H amination does take place (*24*), the resulting enamine moiety would coordinate strongly to the Pd(II) species through either the nitrogen lone-pair electrons or the electron-rich π bond, in contrast to the analogous reaction with carbon-based electrophiles. As a result, the extrusion of NBE (the β-carbon elimination step) would be more difficult, which could lead to side product **D**. Therefore, judicious choice of the NBE cocatalyst and careful design of the amine reagent would likely be the key to balance the requirements for each elementary step and address the aforementioned challenges.

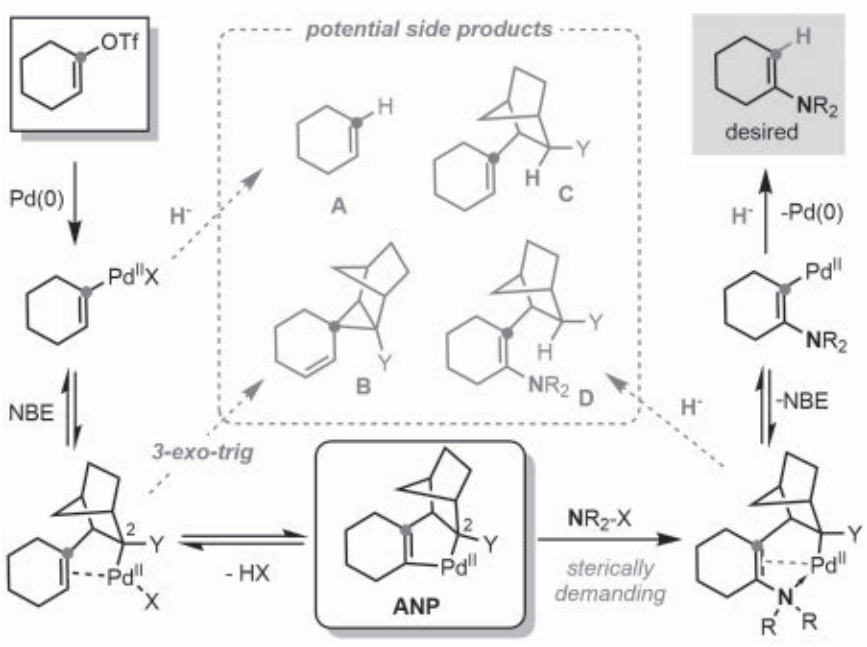



**Fig. 2. Challenges and mechanistic considerations of the Pd/NBE-catalyzed α-amination of alkenyl triflates.** Cyclopropanation and premature reduction of the NBE-containing intermediates are major side reactions. ANP, alkenyl-norbornyl palladacycle; R, alkyl group; X, X-type ligands; Y, a substituent on the NBE; 3-exo-trig, three-membered ring formation via exocyclic ring closure at $sp^2$ atoms.

We hypothesized that it would be beneficial to have moderate steric hindrance on the NBE—sufficient to inhibit the cyclopropanation pathway and promote the β-carbon elimination but not to inhibit the reaction between the alkenyl-norbornyl palladacycle and the electrophile (avoiding side products **A** and **B**). Also, a bulky hydride source, particularly one that could slowly release hydride, would be desirable to minimize early hydride terminations (avoiding side products **C** and **D**). To test the hypothesis, a range of NBE cocatalysts were surveyed, and bifunctional carbonate-based reagents that contain an electrophilic amine moiety (*25*, *26*) and a masked hydride (*27*) were designed. When α-tetralone (**1**) was used as the model substrate, the desired β-tetralone product (**3**) was ultimately obtained in 88% yield in one pot after careful evaluation of various reaction parameters (tables S1 to S3). Through sequential addition of the reactants, the alkenyl triflate intermediate (**2**) was formed in situ and underwent the subsequent transposed amination and hydrolysis without the need for isolation or change of the reaction vessel. The key was to identify triflation conditions that are efficient and compatible with the Pd/NBE catalysis. With triflic anhydride as the triflation reagent, the use of $Cs_2CO_3$ as the base in a mixed solvent of toluene and 1,4-dioxane proved to be optimal, affording alkenyl triflate **2** in near-quantitative yield (table S1, entry 1). Toluene alone afforded a lower conversion (table S1, entry 2). Although the bulky pyridine base was efficient for the triflate formation, it was detrimental for the subsequent amination step, likely as a result of compatibility issues (table S1, entries 3 and 4). Clean reactions and good overall efficiency were also observed when using lithium tetramethylpiperidide (LiTMP) as the base to form the nonaflate intermediate that reacted smoothly in the Pd/NBE catalysis step (table S1, entry 5). This offers an alternative one-pot condition (see below).

Under the optimal triflation conditions, various electrophilic amine sources were investigated for the α-amination and ipso-hydrogenation step. The common electrophilic morpholino benzoate (**R1**), together with an exogenous hydride source, produced **3** in only 24% yield while forming a substantial amount of the direct ipso-reduction (**A**-type) side product (table S1, entry 6). Through tethering with a bulky alcohol fragment, the readily accessible, morpholine-derived carbonates **R2** to **R6** produced much-improved yields and suppressed the side reactions, with the reagent derived from 2,4-dimethyl-3-pentanol (**R4**) showing superior results. Carbonate reagents derived from other amines (**R7** to **R9**) were also tested; the piperidine-based one (**R8**) was found to be optimal in this case, as it resulted in full conversion and negligible side products compared with the others (table S2). Control experiments indicated that both the palladium and the NBE are essential to this reaction, and diminished yield was observed in the absence of the pyridone additive (table S1, entries 7 to 9) (*28*). Among all of the structurally modified NBEs (*29*) that were investigated (table S1, entry 10), the azetidine-amide NBE **N10** proved to be the most effective. A comparable yield was observed when lowering the **N10** loading (table S1, entry 11), showing some catalytic activity. Simple NBE **N1** and C1-substituted **N4** generated substantial cyclopropane **B** and enamine **D**-types of side products, consistent with our previous findings (*21*). The NBEs with C5 and C6 disubstitution (**N2** and **N3**) and C2-ester groups (**N5**) proved to be inferior for this transformation. The *N*-methyl (*N*-Me) amide NBE **N7** only afforded moderate yield while forming NBE-incorporated side products, and the bulkier dimethyl and diethyl-amide NBEs led to substantially decreased reactivity with more side product **A** formation and homocoupling. The high reactivity and selectivity of **N10** is more evident in the more challenging substrates, such as nonconjugated alkenyl triflate **18a** (table S4). In addition, the Buchwald Ph-DavePhos ligand (Ph, phenyl) was most efficient (*21*). Though the slightly less electron-rich Ph-JohnPhos (**L2**) produced somewhat lower yield, regular DavePhos (**L3**) and $PPh_3$ (**L4**) showed much-diminished reactivity (table S1, entry 12) (*30*).

With the optimized conditions in hand, we further examined the scope of this carbonyl 1,2-transposition reaction (Fig. 3). A range of α-tetralone derivatives bearing substituents on both aliphatic (**5** and **6**) and aromatic (**8** to **12** and **14**) rings afforded the desired β-tetralones in good yields via the one-pot protocol. Chromanone (**6**) also reacted, albeit in somewhat lower yields, probably owing to the coordinative capability of the heteroatoms. Besides α-tetralones, substrates derived from five-membered 1-indanone (**15**), seven-membered 1-benzosuberone (**16**), and a substituted benzocycloheptenone (**17**) also delivered the desired deconjugated products. This one-pot protocol was particularly beneficial when the corresponding alkenyl triflate intermediates were sensitive to air, such as the precursor to **8**. For thiochromanone (**7**) and the substrate that contained a highly electron-deficient nitroarene (**13**), much higher overall yields were obtained when the alkenyl triflate intermediates were purified. In addition, nonconjugated ketones also proved to be suitable substrates. Although the one-pot procedure (using LiTMP as the base) provided the desired carbonyl 1,2-shifted product (e.g., **18**) in a useful yield, the two-step procedure generally resulted in higher overall efficiency. Cyclohexanones bearing substituents at the γ position were converted to the corresponding β-substituted

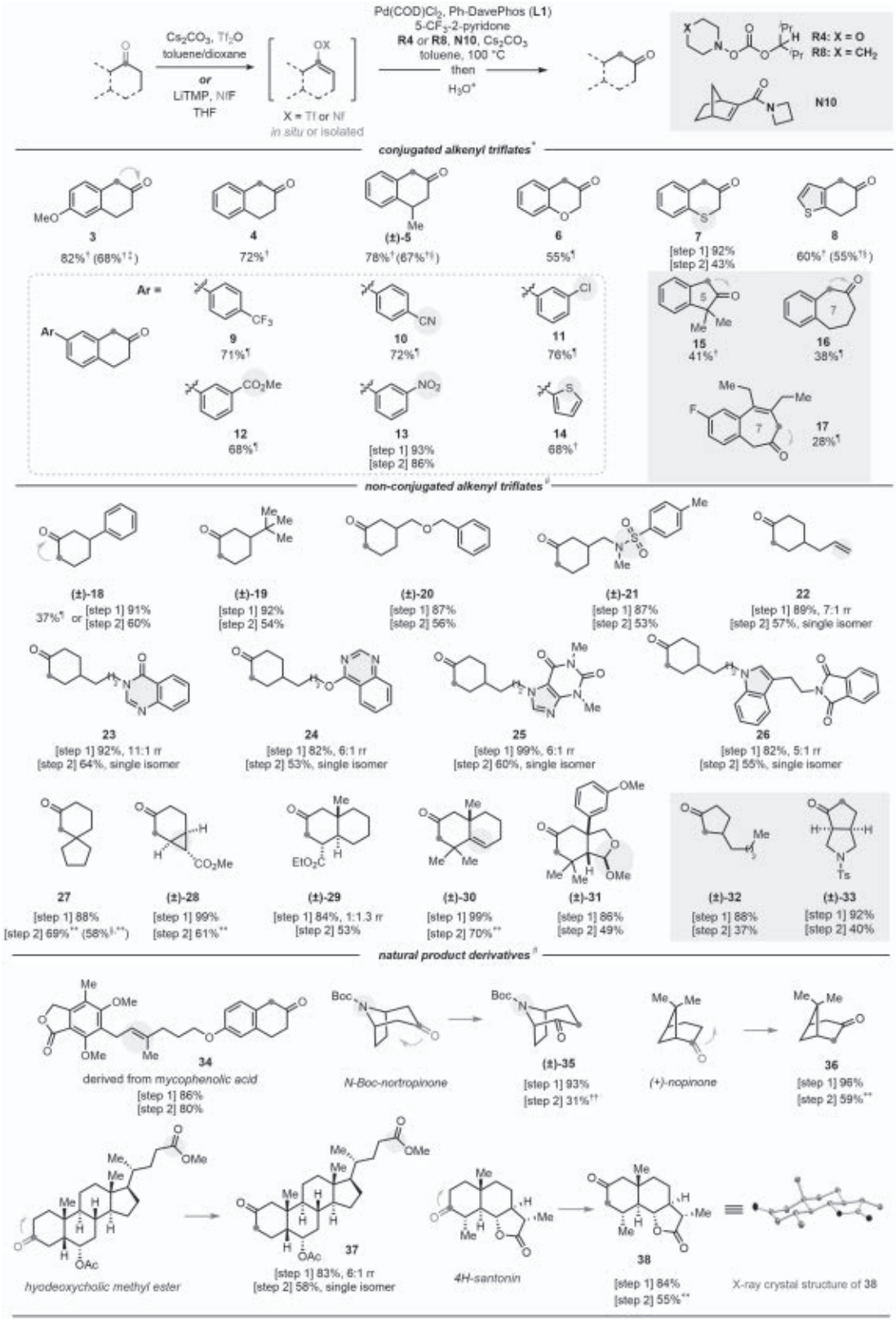



**Fig. 3. Substrate scope.** Reaction conditions for alkenyl triflate: alkenyl triflate (0.10 mmol), **R4** or **R8** (0.22 mmol), $Pd(COD)Cl_2$ (0.01 mmol), **L1** (0.012 mmol), **N10** (0.10 mmol), 5-$CF_3$-2-pyridoine (0.02 mmol), $Cs_2CO_3$ (0.20 mmol), toluene (1.5 ml), 1,4-dioxane (0.5 ml), and 100°C for 24 hours. Isolated average yields of two runs are reported. *R8 was used. †$Cs_2CO_3$ and $Tf_2O$ were used. ‡2.0 mmol scale with 5 mol % $Pd(COD)Cl_2$ and 50 mol % **N10**. §1.0 mmol scale. ¶LiTMP and NfF were used. #**R4** was used. **3,5-di-$CF_3$-2-pyridone was used. ††15 mol % $Pd(COD)Cl_2$ was used. See supplementary materials for details. Et, ethyl; Nf, nonaflyl; OAc, acetate; rr, regioisomeric ratio; THF, tetrahydrofuran; Ts, tosyl.

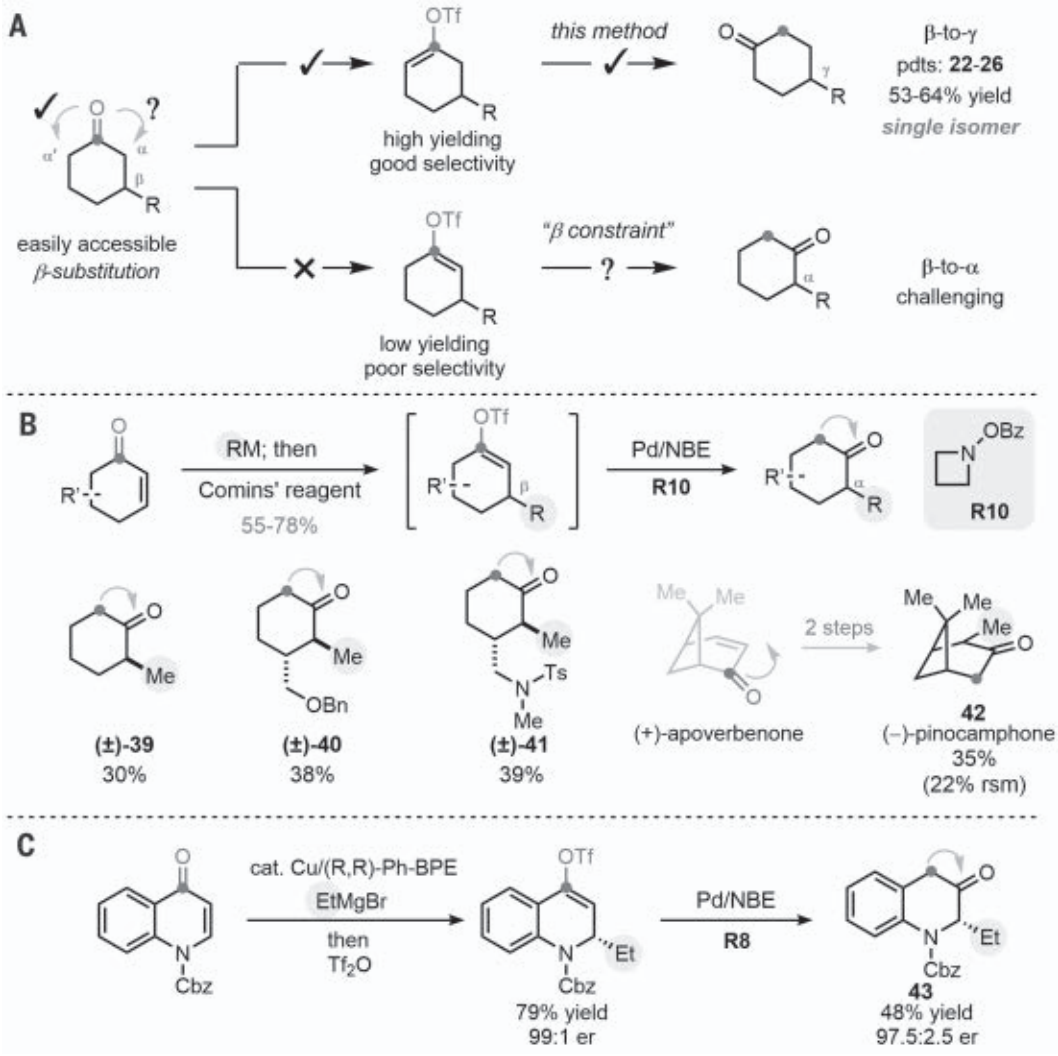



**Fig. 4. Surmounting the β-constraint via conjugate addition.** (**A**) β-substituted cyclohexanones can be transformed to γ-substituted analogs by using this method; however, migration to the more sterically congested α-positions is more challenging. (**B**) Two-step "β-to-α" migration protocol is achieved through conjugate addition or alkenyl triflate formation, followed by the Pd/NBE catalysis. (**C**) Asymmetric conjugate addition or ketone 1,2-migration cascade provides the access to α-stereocenter. BPE, bis(phospholano) ethane; cat., catalyst; Cbz, benzyloxycarbonyl; er, enantiomeric ratio; Bn, benzyl; OBz, benzoate; pdts, products; RM, organometallic reagents; rsm, recovered starting material.

cyclohexanones (**18** to **21**) in good yields, complementary to enone conjugate addition chemistry. Conversely, β-substituted cyclohexanones could be transformed to the γ-substituted cyclohexanones in good yields (**22** to **26**), which are nontrivial to prepare through conventional methods. Whereas the regioselectivity of the enolization step varied from 4.7:1 to 11:1 (favoring deprotonation at the less sterically hindered side), the amination occurred almost exclusively to yield the major (less bulky) isomers in all cases (**22** to **26** and **37**). Notably, a terminal alkene (**22**) remained intact, and no isomerization was observed under these conditions. Cyclohexanones with α-quaternary centers (**27**, **30**, and **31**), α,β-cyclopropane (**28**), and α-ester (**29**) substituents also reacted to generate the desired products that are difficult to access otherwise. Moreover, α-substituted cyclopentanone **32** and the *5-5* bicyclic ketone (**33**) derived from the Pauson-Khand reaction proved competent substrates. A variety of FGs, including aryl chloride (**11**), nitrile (**10**), nitro (**13**), benzyl ether (**20**), sulfonamide (**21**), olefin (**22**, **30**, and **34**), acetal (**31**), ester (**12** and **28**), lactone (**34** and **38**), thiophene (**8** and **14**), quinazoline (**23** and **24**), theophylline (**25**), and indole (**26**), were all compatible with these conditions. Moreover, alkenyl triflates derived from natural products, such as tropinone (**35**), nopinone (**36**), hyodeoxycholic acid (**37**), and α-santonin (**38**) underwent smooth reactions to afford the corresponding carbonyl-migrated analogs.

As mentioned, converting β-substituted cyclohexanones to the γ-substituted analogs is straightforward because of the favored enolization process; however, migrating the carbonyl in the other direction and forming the α-substituted products would be more difficult because the corresponding kinetically disfavored enolates are generally formed in low yields and with poor selectivity (Fig. 4A). In addition, the presence of the β-substituent in the desired alkenyl triflate can hinder the C–H palladation step, similar to the "meta constraint" observed for aromatic substrates (*31*). Therefore, realizing such a "β-to-α" transformation is difficult.

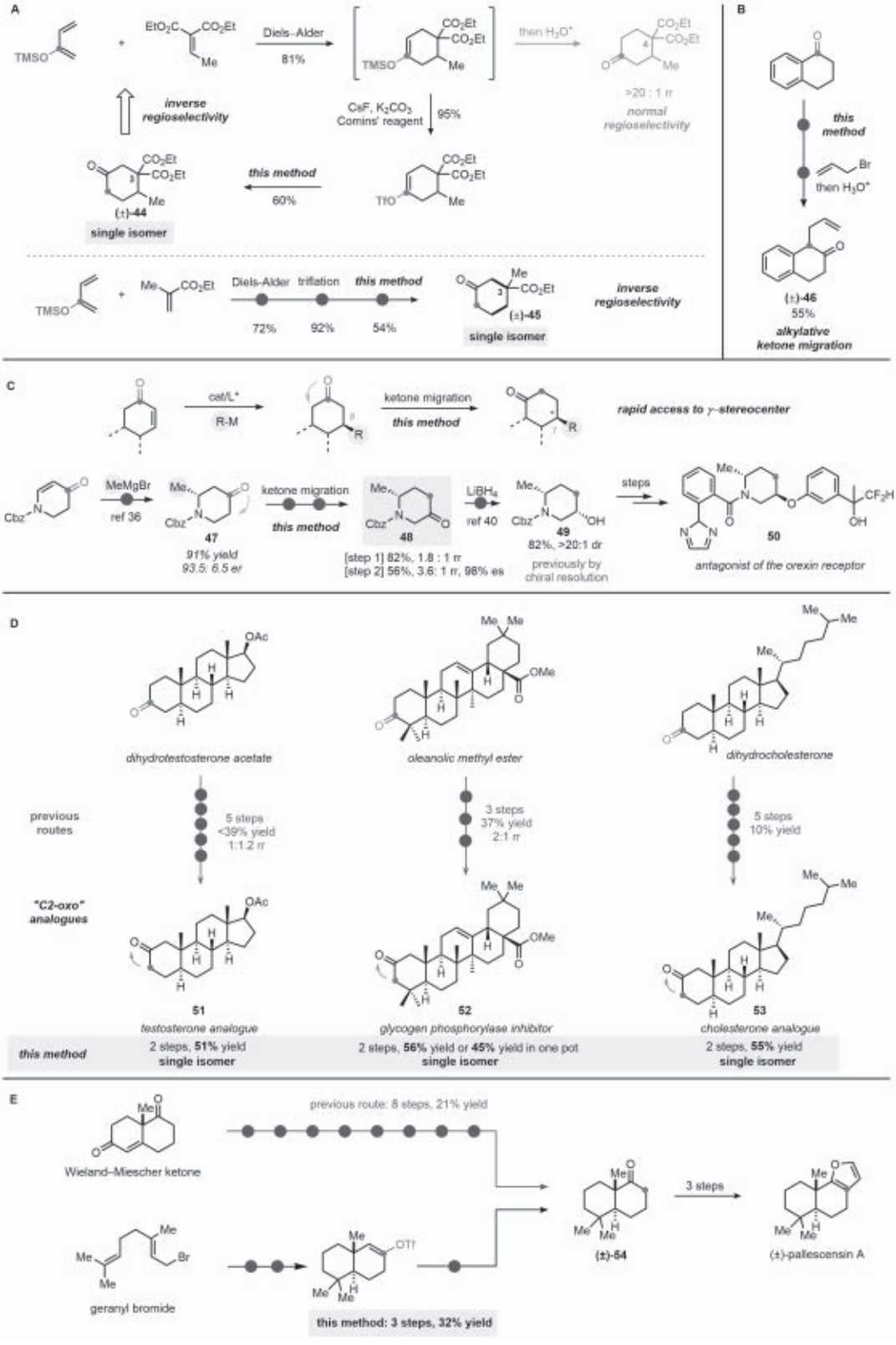



**Fig. 5. Synthetic applications.** (**A**) Synthesis of Diels-Alder products with inverse regioselectivity was achieved by the regioselective ketone 1,2-migration. TMS, trimethylsilyl. (**B**) Alkylative 1,2-carbonyl migration was accomplished by trapping the enamine intermediate with allyl bromide. (**C**) Rapid access to γ-stereocenters was realized through asymmetric conjugate addition that was followed by the carbonyl 1,2-transposition, which led to a formal enantioselective synthesis of an antagonist of the orexin receptor **50**. dr, diastereomeric ratio; es, enantiomeric specificity; L, ligand; M, metal. (**D**) "C2-oxo" analogs of three steroids were synthesized with this method, which shows substantially improved efficiency and selectivity. (**E**) Synthesis of *trans*-decalone **54** was achieved in a much-shortened sequence, representing a concise formal synthesis of pallescensin A.

Stimulated by this challenge, we conceived an alternative enone-based strategy (Fig. 4B). Conjugate addition of alkyl nucleophiles to cyclohexenone substrates, followed by quenching with Comins' reagent, generated the desired alkenyl triflates in moderate to good yields (*32*). Next, considering the β-steric hindrance, we postulated that a less bulky amine electrophile should be beneficial for the C–H amination step. The use of an azetidine-derived electrophile (**R10**), which is expected to be even smaller than dimethylamine,

was found to be optimal (for detailed optimization, see table S5). The β-to-α migrated products (**40** and **41**) were formed with complete retention of the relative stereochemistry. These α,β-dialkylated products such as **40** and **41** have not been reported previously and could serve as useful synthetic intermediates. Moreover, a two-step protocol converted (+)-apoverbenone to (−)-pinocamphone (**42**). Furthermore, this strategy could be particularly useful for preparing enantioenriched α-alkylated ketones, as direct asymmetric nonallylic α-alkylation of ketones remains challenging (*33*–*35*). For example, by taking advantage of the well-established enantioselective conjugate alkyl addition (*36*), α-ethylated ketone **43** was obtained in excellent enantioselectivity with the ketone 1,2-migration strategy (Fig. 4C).

Finally, the general synthetic utility of this method was explored (Fig. 5). Diels-Alder reactions with electron-rich dienes represent one of the most widely used methods to construct cyclohexanones and have been applied in numerous total syntheses (*37*). For instance, when Danishefsky's diene reacts with Michael acceptors, cyclohexanones with the electron-withdrawing group at the C4 position are obtained (*38*). In this work, through trapping the silyl-enol-ether intermediate in situ as the corresponding alkenyl triflate, the Diels-Alder products with inverse regioselectivity (an electron-withdrawing group placed at the C3 position) were obtained in good efficiency (Fig. 5A). This sequence offers a distinct and modular approach to access multisubstituted cyclohexanones that would otherwise be inaccessible by the Diels-Alder strategy. Besides protons, other electrophiles can also be used to trap the enamine intermediate. For example, an alkylative ketone 1,2-migration was realized by using allyl bromide as the electrophile to generate the *ipso* C–C bond–forming product (**46**) (Fig. 5B). In addition, asymmetric conjugate addition followed by this carbonyl 1,2-migration can introduce a γ-stereocenter in an enantioselective manner. This approach is exemplified in the asymmetric synthesis of an antagonist of the orexin receptor (**50**), which was previously prepared through a chiral resolution approach (Fig. 5C) (*39*, *40*). In this approach, the alcohol intermediate (**49**) can be accessed through a known diastereoselective reduction of enantioenriched ketone **48** (*40*) that was synthesized via the sequence of asymmetric conjugate addition (*36*) and ketone 1,2-migration with good overall yield.

The steroids dihydrotestosterone acetate and dihydrocholesterone were previously transformed to their bioactive "C2-oxo" analogs in five steps with <39% (*41*) and 10% (*12*) overall yields, respectively (Fig. 5D). The former synthesis also suffered from poor regioselectivity (1:1.2), which complicated the product purification. With the same starting materials, this carbonyl 1,2-transposition method allows a two-step synthesis of these products as single regioisomers in good yields (**51** and **53**). Similarly, the "C2-oxo" oleanolic acid derivative, which is a glycogen phosphorylase inhibitor, was previously prepared through a three-step sequence with a 37% overall yield yet low selectivity (*42*). By comparison, our method can deliver the same product with fewer steps, higher overall yield, and complete regioselectivity (**52**). Finally, *trans*-decalone **54**, a key intermediate in the total synthesis of pallescensin A (*43*), was previously constructed from Wieland-Miescher ketone in eight steps with 21% total yield (Fig. 5E) (*44*). By comparison, starting from the readily available geranyl bromide, we synthesized the same ketone intermediate **54** in only three steps with 32% overall yield. Given the convenience and broad FG tolerance of this ketone transposition method, we anticipate its widespread application in complex molecule synthesis and the efficient preparation of FG-transposed analogs, which should benefit medicinal chemistry research.

## REFERENCES AND NOTES


1. P. Zhang *et al*., *J. Nat. Prod.* **72**, 1414–1418 (2009).
2. S. Chowdhury *et al*., *J. Med. Chem.* **61**, 84–97 (2018).
3. H. Iio, Y. Matsumoto, K. Shimokata, K. Shibata, T. Tokoroyama, *J. Chem. Soc. Perkin Trans. 1* **1989**, 1360–1361 (1989).
4. A. G. Schultz, Y. K. Yee, M. H. Berger, *J. Am. Chem. Soc.* **99**, 8065–8067 (1977).
5. V. V. Kane, V. Singh, A. Martin, D. L. Doyle, *Tetrahedron* **39**, 345–394 (1983).
6. D. G. Morris, *Chem. Soc. Rev.* **11**, 397–434 (1982).
7. T. Nakai, T. Mimura, *J. Synth. Org. Chem. Jpn.* **35**, 964–978 (1977).
8. B. M. Trost, K. Hiroi, S. Kurozumi, *J. Am. Chem. Soc.* **97**, 438–440 (1975).
9. J. E. Bridgeman *et al*., *J. Chem. Soc. C* **1970**, 244–250 (1970).
10. J. E. Bridgeman, E. R. H. Jones, G. D. Meakins, J. Wicha, *Chem. Commun.* **1967**, 898 (1967).
11. J. W. Cornforth, J. M. Osbond, G. H. Phillipps, *J. Chem. Soc.* **1954**, 907–912 (1954).
12. M. Hasan *et al*., *Liebigs Ann.* **1995**, 889–896 (1995).
13. S. H. Pines, *Org. Process Res. Dev.* **8**, 708–724 (2004).
14. S. Thorén, *Acta Chem. Scand.* **24**, 93–98 (1970).
15. J. R. Dehli, J. Legros, C. Bolm, *Chem. Commun.* **2005**, 973–986 (2005).
16. J. Ye, M. Lautens, *Nat. Chem.* **7**, 863–870 (2015).
17. N. Della Ca', M. Fontana, E. Motti, M. Catellani, *Acc. Chem. Res.* **49**, 1389–1400 (2016).
18. J. Wang, G. Dong, *Chem. Rev.* **119**, 7478–7528 (2019).
19. H. G. Cheng, S. Chen, R. Chen, Q. Zhou, *Angew. Chem. Int. Ed.* **58**, 5832–5844 (2019).
20. M. Catellani, F. Frignani, A. Rangoni, *Angew. Chem. Int. Ed.* **36**, 119–122 (1997).
21. J. Wang, Z. Dong, C. Yang, G. Dong, *Nat. Chem.* **11**, 1106–1112 (2019).
22. Z. Wu, N. Fatuzzo, G. Dong, *J. Am. Chem. Soc.* **142**, 2715–2720 (2020).
23. M. Catellani, G. P. Chiusoli, *J. Organomet. Chem.* **233**, C21–C24 (1982).
24. Y. Park, Y. Kim, S. Chang, *Chem. Rev.* **117**, 9247–9301 (2017).
25. S. Zhu, N. Niljianskul, S. L. Buchwald, *J. Am. Chem. Soc.* **135**, 15746–15749 (2013).
26. Y. Miki, K. Hirano, T. Satoh, M. Miura, *Angew. Chem. Int. Ed.* **52**, 10830–10834 (2013).
27. Z. Dong, J. Wang, Z. Ren, G. Dong, *Angew. Chem. Int. Ed.* **54**, 12664–12668 (2015).
28. P. Wang *et al*., *J. Am. Chem. Soc.* **138**, 9269–9276 (2016).
29. R. Li, G. Dong, *J. Am. Chem. Soc.* **142**, 17859–17875 (2020).
30. M. Lautens, S. Piguel, *Angew. Chem. Int. Ed.* **39**, 1045–1046 (2000).
31. J. Wang, Y. Zhou, X. Xu, P. Liu, G. Dong, *J. Am. Chem. Soc.* **142**, 3050–3059 (2020).
32. E. M. Ferreira, B. M. Stoltz, *J. Am. Chem. Soc.* **125**, 9578–9579 (2003).
33. R. Cano, A. Zakarian, G. P. McGlacken, *Angew. Chem. Int. Ed.* **56**, 9278–9290 (2017).
34. A. G. Doyle, E. N. Jacobsen, *J. Am. Chem. Soc.* **127**, 62–63 (2005).
35. J. Streuff, D. E. White, S. C. Virgil, B. M. Stoltz, *Nat. Chem.* **2**, 192–196 (2010).
36. Y. Guo, S. R. Harutyunyan, *Angew. Chem. Int. Ed.* **58**, 12950–12954 (2019).
37. K. C. Nicolaou, S. A. Snyder, T. Montagnon, G. Vassilikogiannakis, *Angew. Chem. Int. Ed.* **41**, 1668–1698 (2002).
38. S. Danishefsky, *Acc. Chem. Res.* **14**, 400–406 (1981).
39. M. P. Christopher *et al*., "Substituted Imidazopyridines as HDM2 Inhibitors." US Patent 20140179680 A1 (2014).
40. D. Chen *et al*., *Angew. Chem. Int. Ed.* **55**, 12632–12636 (2016).
41. R. L. Clarke, *J. Org. Chem.* **28**, 2626–2633 (1963).
42. J. Hao, P. Zhang, X. Wen, H. Sun, *J. Org. Chem.* **73**, 7405–7408 (2008).
43. J. S. Foot, A. T. Phillis, P. P. Sharp, A. C. Willis, M. G. Banwell, *Tetrahedron Lett.* **47**, 6817–6820 (2006).
44. K. Hatzellis, G. Pagona, A. Spyros, C. Demetzos, H. E. Katerinopoulos, *J. Nat. Prod.* **67**, 1996–2001 (2004).


## ACKNOWLEDGMENTS


S. Ochi and X. Liu are acknowledged for x-ray crystallography. **Funding:** Financial support from the University of Chicago and NIGMS (1R01GM124414-01A1) is acknowledged. X.X. thanks ShanghaiTech University for a fellowship. **Author contributions:** Z.W., X.X., and J.W. discovered the reaction. Z.W., J.W., and G.D. conceived the idea. Z.W., X.X., and J.W. conducted the experimental investigation. Z.W. and G.D. wrote the manuscript. G.D. directed the research. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Metrical parameters for the structure of **38** are available free of charge from the Cambridge Crystallographic Data Centre (www.ccdc.cam.ac.uk/) under reference number CCDC 2062517. All other data are available in the main text or the supplementary materials.


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abl7854
Materials and Methods
Supplementary Text
Tables S1 to S5
Spectral Data
X-ray Data
References (*45*–*76*)

4 August 2021; accepted 29 September 2021
10.1126/science.abl7854

ORGANIC CHEMISTRY

# Scalable Birch reduction with lithium and ethylenediamine in tetrahydrofuran

James Burrows, Shogo Kamo, Kazunori Koide*

The Birch reduction dearomatizes arenes into 1,4-cyclohexadienes. Despite substantial efforts devoted to avoiding ammonia and cryogenic conditions, the traditional, cumbersome, and dangerous procedure remains the standard. The Benkeser reduction with lithium in ethylenediamine converts arenes to a mixture of cyclohexenes and cyclohexanes; this is operationally easier than the Birch reduction but does not afford 1,4-cyclohexadienes. Here, we report a Birch reduction promoted by lithium and ethylenediamine (or analogs) in tetrahydrofuran at ambient temperature. Our method is easy to set up, inexpensive, scalable, rapid, accessible to any chemical laboratory, and capable of reducing both electron-rich and electron-deficient substrates. Our protocol is also compatible with organocuprate chemistry for further functionalization.

Dearomatization is widely used in chemical synthesis (*1*). The Birch reduction dearomatizes arenes into 1,4-cyclohexadienes with lithium, sodium, or potassium in liquid ammonia at ≤−33°C (Fig. 1A) (*2*, *3*) and has been employed throughout the pharmaceutical industry (*4*, *5*), perfumery industry (*6*, *7*), and academia (*8*–*11*).

Liquid ammonia must be prepared with specialized equipment and carefully dissipated after the reaction is complete. Both steps are time consuming; for example, removal of 1 L of liquid ammonia (850 L as gas) can take up to 12 hours (*12*), and as much as 7.5 L of liquid ammonia per mole of substrate may be needed (*5*, *13*). Even on a 3.5-mmol scale, the Birch process requires 7 hours from setting up equipment to the completion of biphasic extraction (*14*). These logistical challenges make it difficult to perform multiple Birch reductions in parallel. Also, the liquid ammonia solvent has long been deemed necessary to solubilize alkali metals to form the solvated electron.

To overcome these challenges, researchers have developed ammonia-free conditions. For example, the Benkeser group used lithium and neat ethylamine, ethylenediamine, or a mixture of primary and secondary amines, providing a mixture of over-reduced products, and did not use any other solvents (Fig. 1B) (*15*–*17*). Arenes could be reduced to the Birch-type products with lithium in a mixture of methylamine and isopropanol, but overreduction appeared inevitable (*18*). Benzoic acid was reduced to benzaldehyde in 25% yield in the presence of lithium, methylamine, and ammonium nitrate (*19*). The benefit of ethylenediamine as a solvent for dissolving metal reductions was also demonstrated by others (*20*). The Dolby group reduced three substates to the corresponding Birch-type products in 45% to quantitative yield using lithium, ethylenediamine, *n*-propylamine, and *t*-butanol (*4*). This method was moderately successful in one instance (*21*) and was not effective in the *N*-detosylation of a challenging substrate (*22*). Donohoe and House reported the reduction of electron-deficient arenes and heterocycles using di-*tert*-butylbiphenyl ($1000/mol; Sigma-Aldrich) and lithium at −78°C (Fig. 1C) (*23*). Their method was highly oxygen sensitive and as lengthy as the standard Birch procedure (*14*). An's method (Fig. 1D) requires sodium and 3 to 9 equivalents of 15-crown-5 ($1579/mol; Sigma-Aldrich) and is limited to electron-rich or neutral substrates (*24*). The Baran group described an electrochemical reduction of electron-rich arenes (Fig. 1E) with 3.5 to 10 equivalents of tri(pyrrolidin-1-yl)phosphine oxide ($5040/mol; Sigma-Aldrich) and 3 equivalents of 1,3-dimethylurea ($5/mol; Sigma-Aldrich), both of which must be removed from the product by column chromatography (*13*). Their 0.45-mol scale reaction took 3 days in a flow reactor without tri(pyrrolidin-1-yl)phosphine oxide (*13*). The Sugai group treated arenes with lithium and ethylenediamine in tetrahydrofuran (THF) or $Et_2O$ but did not isolate 1,4-cyclohexadiene products (*25*, *26*) and indicated that THF might be a ligand for a lithium ion (*25*).

Despite these efforts, the original, cumbersome, and dangerous Birch protocol remains the current standard (*14*, *27*). Because of the inconvenient procedure, Birch reductions are often avoided in favor of safer and often

Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, USA.
*Corresponding author. Email: koide@pitt.edu

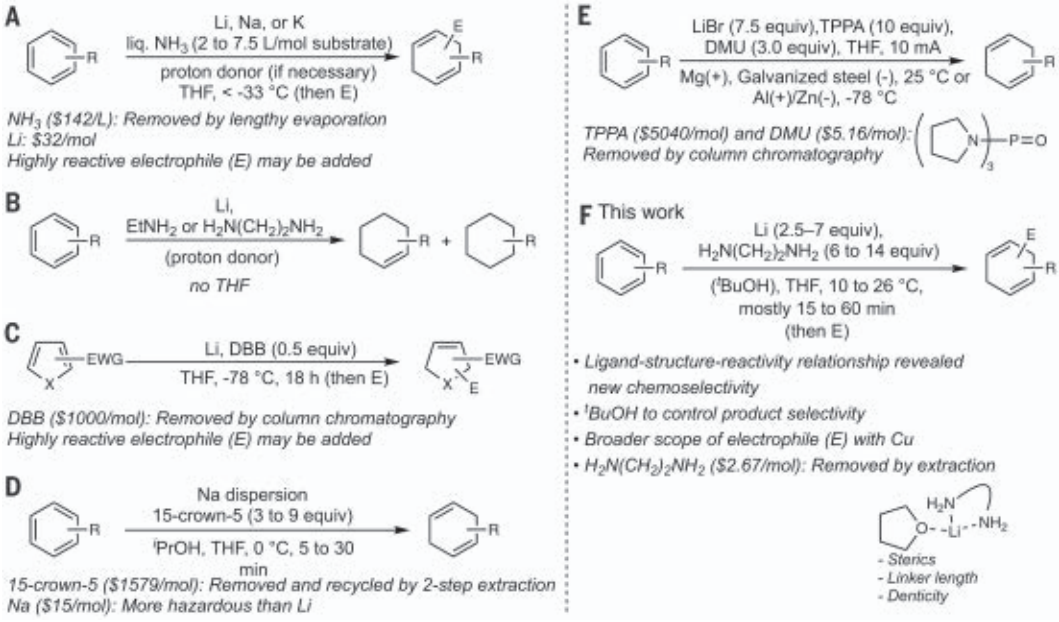



**Fig. 1. Previous Birch reductions and this work.** (**A**) General Birch reduction. (**B**) Benkeser's ammonia-free reduction. (**C**) Donohoe's ammonia-free Birch reduction. (**D**) An's ammonia-free Birch reduction. (**E**) Baran's electrochemical reduction. (**F**) This work. liq., liquid; EWG, electron-withdrawing group; DBB, 4,4′-di-*tert*-butylbiphenyl; TPPA, tri(pyrrolidin-1-yl)phosphine oxide; DMU, 1,3-dimethylurea.

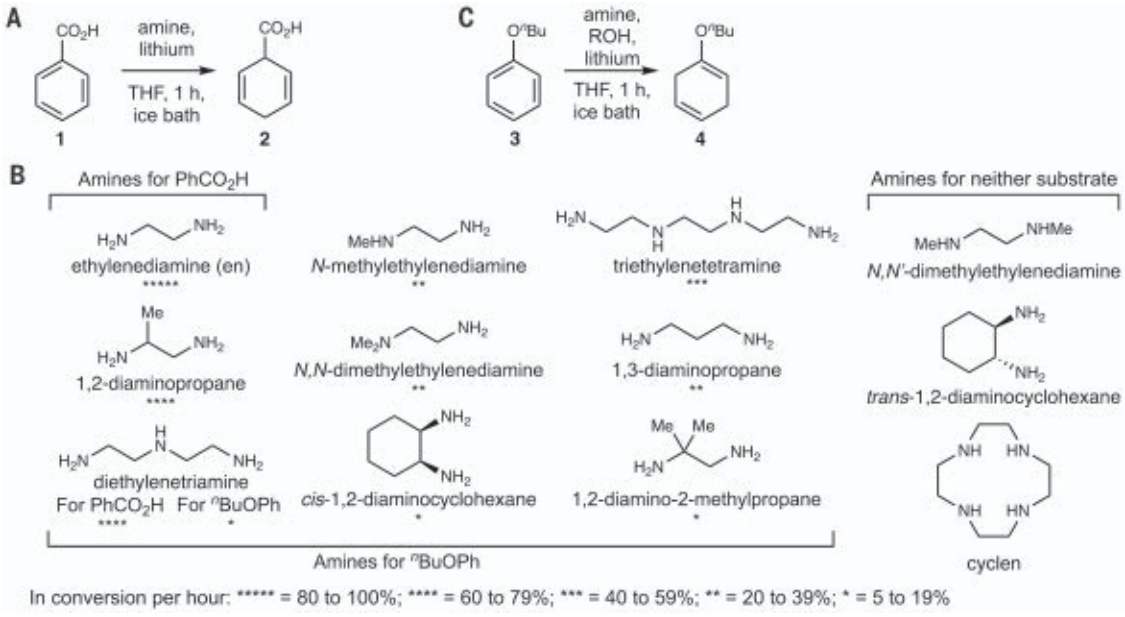



**Fig. 2. Summary for the optimization study.** (**A**) Reduction of $PhCO_2H$ (**1**). (**B**) Amine structures and efficiency. (**C**) Reduction of *n*-butoxybenzene (**3**).

lengthier synthetic schemes (*28*, *29*). Therefore, there remains a need for a Birch reduction protocol that is fast and effective for both electron-rich and -deficient arenes without ammonia, specialized equipment, or expensive additives. Here, we report such a Birch reduction with ethylenediamine as a ligand ($2.67/mol; Sigma-Aldrich) and lithium in THF (Fig. 1F). We also report influence of the amine structure on the selectivity, including inverse electron-demand chemoselectivity. Finally, we propose more active roles for amines, alcohols, and solvents than previously considered, providing a platform for controlling the chemoselectivity.

Benzoic acid ($PhCO_2H$, **1**; Fig. 2A) was chosen as our starting model substrate because of the deficiency of currently reported conditions for the reduction of electron-deficient arenes. First, we evaluated a protocol in which a balloon was filled with ammonia gas and attached to a flask containing lithium and the substrate in THF (*30*) to find that diene **2** was obtained in 83% yield (table S1, entry 1). However, this method was not effective for electron-rich substrates, typically resulting in incomplete reactions. Consequently, we began to investigate alternative amine-based ligands (Fig. 2B) that could be broadly applicable, inexpensive, and easy to handle while also affording the desired Birch reduction products. With 5.0 equivalents of lithium and 1.0 or 2.5 equivalent(s) of ethylenediamine, diene **2** was produced in 4 or 83% yield, respectively, after 6 hours (entries 2 and 3). Using 5.0 equivalents of ethylenediamine could lower the necessary amount of lithium to 2.5 equivalents and the time to 1 hour (90% yield; entry 4).

The reaction did not proceed without ethylenediamine (entry 5). Also, the combination of ethylenediamine and lithium was essential as there was no reduction when sodium metal was employed (entry 6). We then began to investigate whether the reaction could be improved further by fine-tuning the linker length and denticity of the ligand. 1,3-Diaminopropane gave no product (entry 7). Diethylenetriamine was as effective as ethylenediamine, providing **2** in 86% yield (entry 8), but triethylenetetramine was ineffective (entry 9). Other 1,2-diamines (*N*-methylethylenediamine, *N*,*N*′-dimethylethylenediamine, *trans*-1,2-diaminocyclohexane, and *cis*-1,2-diaminocyclohexane) failed to promote the reduction (entries 10 to 13). Although 1,2-diaminopropane had similar reactivity as ethylenediamine, affording diene **2** in 85% yield (entry 14), the reaction did not progress with 1,2-diamino-2-methylpropane (entry 15). Also, no reduction occurred with cyclen (entry 16). Although diethylenetriamine was as effective as ethylenediamine, we continued to use ethylenediamine (ethylenediamine, $2.67/mol, versus diethylenetriamine, $7.59/mol). The reaction could be scaled up to 10 g (82 mmol), resulting in 95% isolated yield (entry 17).

We proceeded to optimize the reaction conditions for an electron-rich system using *n*-butoxybenzene ($^n$BuOPh, **3**; Fig. 2C) as a model substrate, lithium (2.5 equivalents), and amine (5 equivalents) in THF on ice. This substrate was not reduced without alcohol present (table S2, entry 1). This is consistent with the known mechanism in which electron-rich arenes cannot accept the second electron unless the radical anion intermediate is protonated to form the corresponding radical species (*31*). Also, we did not observe any of the dealkylated phenol by-product that was reported in the method developed by Sugai (*25*). With methanol, ethanol, isopropanol, *t*-butanol, (*2*, *32*) and *t*-amyl alcohol, diene **4** was produced in 33, 58, 62, 75, and 68% yields, respectively, with the over-reduced product **5** in 4 to 11% yields (entries 2 to 6). To study the importance of the acidity of the alcohol, we tested 2,2,2-trifluoroethanol and 1,1,1,3,3,3-hexafluoroisopropanol, which afforded yields of 52 and 26%, respectively (entries 7 and 8). Use of 1,3-diaminopropane and diethylenetriamine diminished yields to 33 and 9%, respectively (entries 9 and 10). The yield was increased to 51% with triethylenetetramine (entry 11; as compared with entry 9 in table S1). The effects of the *t*-butanol/ethylenediamine ratio are described in fig. S1. The reaction did not proceed when sodium was used in lieu of lithium (entry 12). With *N*-methylethylenediamine, *N*,*N*′-dimethylethylenediamine, and *N*,*N*-dimethylethylenediamine, diene **4** was produced in 33, 3, and 27% yields, respectively (entries 13 to 15). When employing cyclen, no reduction occurred (entry 16). *Trans*-1,2-diaminocyclohexane was ineffective even after 3 hours (entry 17), and *cis*-1,2-diaminocyclohexane promoted the reduction, albeit more slowly than ethylenediamine (entries 18 and 19). Use of 1,2-diaminopropane and 1,2-diamino-2-methylpropane produced diene **4** in 65 and 11% yields (entries 20 and 21). To fully consume the starting material, the equivalents of lithium and ethylenediamine were increased to 3 and 6, respectively, to form diene **4** in 85% yield (entry 22).

After obtaining the optimal reaction conditions for $PhCO_2H$ and $^n$BuOPh, we investigated the substrate scope. All of the experiments were performed on 2.5- to 10-mmol scales and took only 0.25 to 3 hours, including preparation and workup. $PhCO_2H$ and its analogs were reduced to the corresponding products **2**, **6**, **7**, **8**, **9**, and **10** in mostly high yields (Fig. 3A). We also

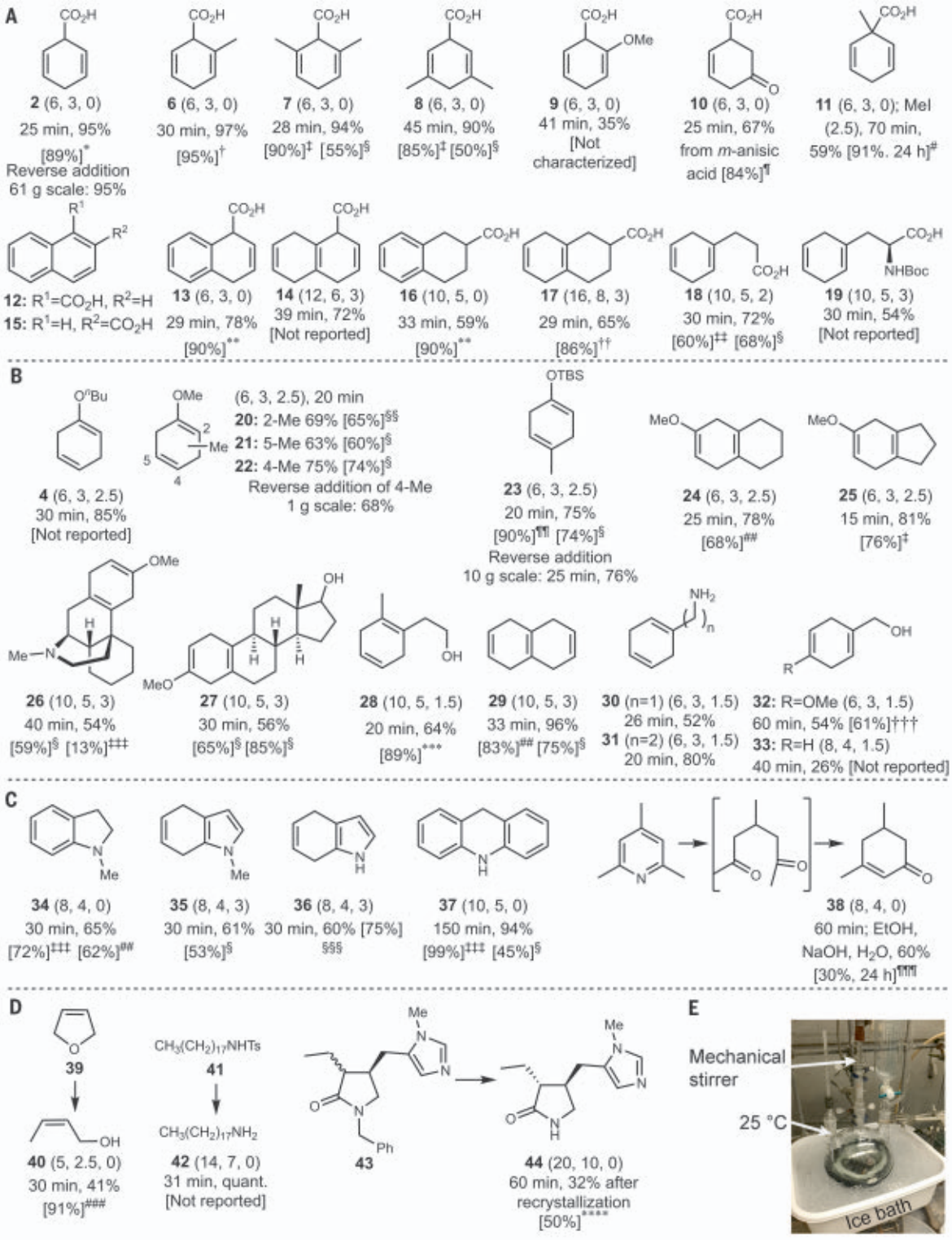



**Fig. 3. Scope of the current method.** The numbers in the parentheses are the equivalents of ethylenediamine, lithium, and t-butanol, in this order. All reactions were performed in THF on ice bath. Previous yields with other methods are shown in brackets. (**A**) Products of the Birch reduction of carboxylic acids. (**B**) Products of the Birch reduction of aryl ethers and phenyl derivatives. (**C**) Products of the Birch reduction of assorted N-heterocycles. (**D**) Assorted dissolving-metal reduction reaction products. (**E**) Photograph of setup for 0.5-mol reduction of $PhCO_2H$. Symbols indicate the following references: *, (*58*); †, (*59*); ‡, (*60*); §, (*13*); ¶, (*61*); #, (*62*); **, (*33*); ††, (*34*); ‡‡, (*63*); §§, (*64*); ¶¶, (*65*); ##, (*66*); ***, (*35*); †††, (*37*); ‡‡‡, (*24*); §§§, (*67*); ¶¶¶, (*40*); ###, (*6*); ****, (*4*).

found that quenching the reduction of $PhCO_2H$ with methyl iodide gave the corresponding alkylated product **11** in 59% yield. The streamlined process allowed us to expeditiously modify stoichiometries and probe scarcely investigated reactivities; for example, carboxylic acid **12** was reduced to the typical Birch reduction product **13** in 78% yield without *t*-butanol. With *t*-butanol, triene **14** was produced in 72% yield. Carboxylic acid **15** was reduced to acid **16** or diene **17** without or with *t*-butanol in 59 or 65% yield, respectively (*33*, *34*). Hydrocinnamic acid was converted to the corresponding diene **18** in 72% yield. *N*-Boc-L-phenylalanine was reduced to diene **19** in 54% yield.

Next, we subjected $^{n}$BuOPh and 2-, 3-, and 4-methylanisole to the reaction conditions to form 1-alkoxy-1,4-cyclohexadienes **4**, **20**, **21**,

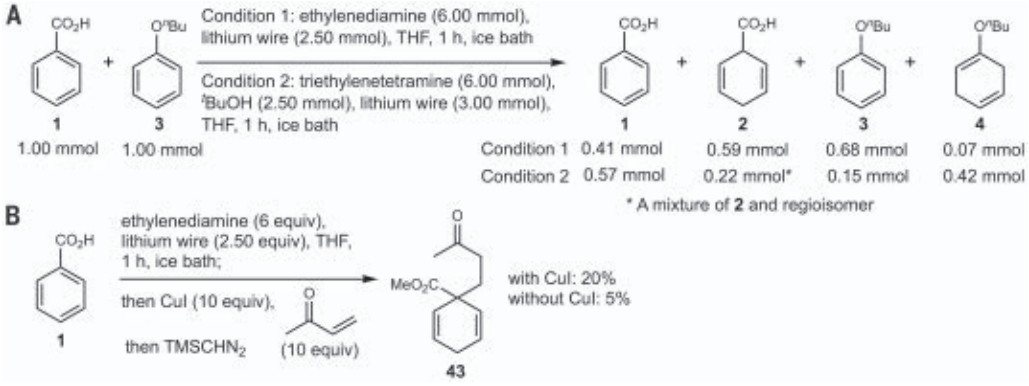



**Fig. 4. Tuning selectivity and downstream reactivity.** (**A**) Normal chemoselectivity with ethylenediamine and reversed chemoselectivity with triethylenetetramine. (**B**) The Birch reduction–cuprate addition reaction.

and **22** in 85, 69, 63, and 75% yields, respectively (Fig. 3B). The Baran group demonstrated the industrial application of the reduction of 4-TBSO-toluene to form the silyl enol ether **23** in 74% yield after 16 hours in batch (*13*). The current method produced the same product on a similar scale in 75% yield after 20 min. 6-Methoxy-1,2,3,4-tetrahydronaphthalene was reduced to methyl ether **24** in 78% yield. A similar transformation was equally efficient (81% yield) to afford methyl ether **25**, which was previously used in synthetic studies for (–)-daphlongamine H (*11*). Dextromethorphan (cough suppressant) and estrone-3-methyl ether were reduced to **26** and **27** in 54 and 56% yields, respectively. The demethylated phenol products were also not observed with dienes **20** to **22** and **24** to **27**. The Birch reduction of 2-(*o*-tolyl)ethanol to form **28** was the first step in the total synthesis of atractyligenin (*35*). Here, our method produced **28** in 64% yield. Naphthalene was reduced to triene **29** in 94% yield. Our method reduced benzylamine and phenethylamine to dienes **30** and **31** in 52 and 80% yields, respectively. The benzylic hydroxy group is generally lost in the Birch reduction (*36*), but *p*-methoxybenzylic alcohols may be converted to the corresponding reduced alcohols (*37*, *38*). Our method reduced *p*-methoxybenzyl alcohol to diene **32** in 54% yield. Benzyl alcohol was reduced to diene **33** in 26% yield under our conditions without losing the benzylic hydroxy group. This diene was previously prepared in two steps (*39*).

*N*-Methylindole was converted to **34** or **35** without or with *t*-butanol in 65 or 56% yield (Fig. 3C), whereas other methods afforded only one of the two products (*13*, *24*). Indole and acridine were reduced to pyrrole **36** and dihydroacridine **37** in 60 and 94% yields, respectively. The transformation of pyridines to cyclohexenones is useful but underutilized (*40*, *41*). Although our standard reaction conditions with 2,4,6-collidine produced cyclohexenone **38** in 27% yield, the simple procedure enabled rapid screenings to discover that *t*-butanol was unnecessary, improving the yield to 60%.

Previously, reductive ring openings of cyclic allylic ethers were performed at –78°C (*6*) or at ambient temperature for 48 hours (*7*). Here, our method reductively opened 2,5-dihydrofuran (**39**; Fig. 3D) to produce Z-allylic alcohol **40** in 41% yield after 30 min on ice (lower yield is because of the volatility of the alcohol) (*6*). Tosyl amide **41** was deprotected under our reaction conditions to form amine **42** in quantitative yield. The *N*-debenzylation of **43** (a mixture of diastereomers) to form amide **44** was of industrial interest (*4*) and was accomplished in 32% yield after recrystallization. Failed substrates are shown in fig. S2.

We found that the reaction proceeded faster with a greater stir rate (fig. S3A). Also, the current method is compatible with trace water and air, as the use of distilled and degassed THF only mildly improved the yield (78 versus 75%) (fig. S3A). For the development of scalable procedures, it was desirable to perform the reaction at higher concentrations. We found (fig. S3, B and C) that the $PhCO_2H$ and $^n$BuOPh concentrations could be increased up to 0.8 M, which also accelerated the reduction. Although the yield of diene **2** was unaffected, the increased reaction rate led to monoolefin **5** and the 1,3-cyclohexadiene with $^n$BuOPh (fig. S3B). To further improve the scalability, we suspended lithium in THF then cooled the flask on ice, after which a solution of ethylenediamine and $PhCO_2H$ in THF was added. This procedure was equally effective when we scaled up this reaction to 61 g (0.50 mol) (Fig. 3E). Monitoring the internal temperature revealed that the reaction proceeded at ~10°C; therefore, we decided to keep the internal temperature in the 10° to 26°C range. The 0.50-mol scale reaction took 1 hour including preparation and workup to obtain diene **2** in 95% yield. A similar reverse-addition protocol was applied to 4-methylanisole and 4-OTBS toluene. It was necessary to add *t*-butanol to the suspension last to suppress both the overreduction and isomerization of the desired product to the 1,3-cyclohexadiene. 4-Methylanisole was reduced to diene **22** in 68% yield, and 4-OTBS toluene (10-g scale) was reduced to diene **23** in 76% yield. Both of these experiments (setup plus reaction plus isolation of the products) also took 1 hour, which is substantially shorter than literature precedent (1 to 2 days) (*13*).

Generally, electron-withdrawing groups increase reduction rates, whereas electron-donating groups decrease them (*42*–*44*). Specifically, the Birch reduction of benzoate is more than 61 times as fast as that of anisole under traditional conditions (*45*). In Fig. 2A, we summarize the relative reactivity with asterisks, which suggested that it might be possible to reduce an electron-rich arene in preference to an electron-deficient arene. We performed a reduction with an equimolar mixture of $PhCO_2H$ and $^n$BuOPh, ethylenediamine, and lithium without *t*-butanol to obtain acid **2** and ether **4** in 59 and 7% yields, respectively (Fig. 4A), which is consistent with the literature (*45*). The same conditions with *t*-butanol gave a mixture of acid **2** along with other intractable products. We then exploited our earlier results by replacing ethylenediamine with triethylenetetramine to discover that the electron-rich arene $^n$BuOPh was more reactive than the electron-deficient arene $PhCO_2H$ (43% consumption of $PhCO_2H$ versus 85% consumption of $^n$BuOPh), affording **2** and **4** in a 1:2 ratio.

The liquid ammonia solvent in the Birch reduction has hampered productive interception of the carbanion intermediate. For example, a Birch alkylation with methyl vinyl ketone failed because ammonia caused the polymerization of the ketone (*46*). The large excess of amine solvent can also interfere with added metals for cross-coupling reactions. This would not be circumvented under Dolby's conditions with ethylenediamine and *n*-propylamine (*4*). Given the use of only 6 equivalents of ethylenediamine under our conditions, we hypothesized that the Birch reduction could be coupled with cuprate chemistry. After the reduction of $PhCO_2H$ under our reaction conditions, CuI and methyl vinyl ketone were added (Fig. 4B). Under these unoptimized conditions, ketone **45** was generated in 20% yield while forming a quaternary carbon.

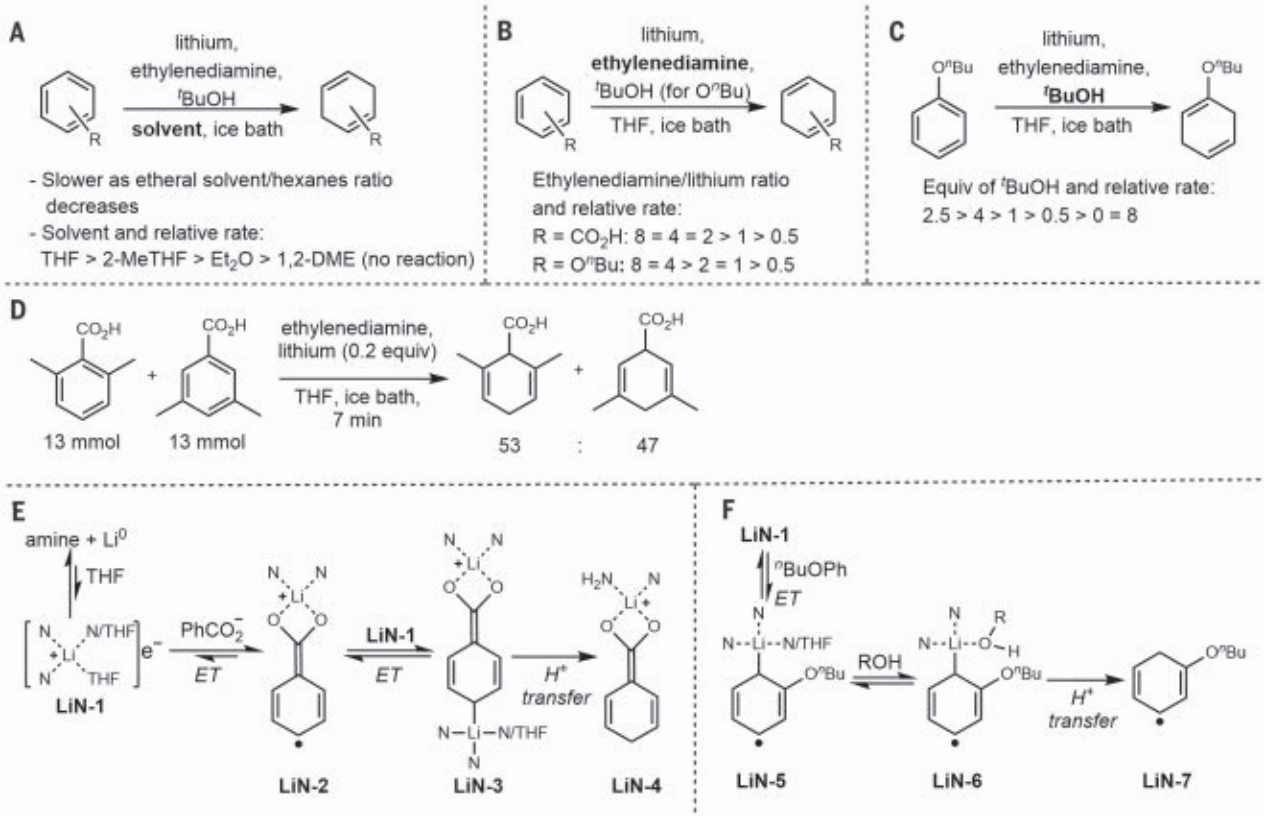



**Fig. 5. Summary of kinetic studies and proposed mechanisms.** (**A**) Rate dependence on solvent for the reduction of $PhCO_2H$ or $^{n}$BuOPh. (**B**) Dependence on ethylenediamine/lithium ratio for the reduction of $PhCO_2H$ or $^{n}$BuOPh. (**C**) Dependence on *t*-butanol for the reduction of $^{n}$BuOPh. For all data, the yields were determined by proton nuclear magnetic resonance ($^{1}$H NMR) spectroscopic analysis using 1-methoxyadamantane as an internal standard. (**D**) Competition experiment. (**E**) Proposed mechanism for $PhCO_2H$. (**F**) Proposed mechanism for $^{n}$BuOPh. Hydrogen atoms, linkers, and methyl groups on ligands are omitted for clarity. ET, electron transfer.

We then investigated the kinetics for both $^{n}$BuOPh and $PhCO_2H$ reduction under various conditions (fig. S3, D to K). First, we found that the reduction rates depended upon THF concentrations as well as the choice of ethereal solvent (Fig. 5A and fig. S3, D to G). The decreasing polarity of the reaction medium most likely affects the stability and solubility of the radical anion that is usually stabilized by ammonia (*47*). Next, as the ethylenediamine/lithium ratio increased (Fig. 5B and fig. S3, H and I), so did the reaction rate (fig. S3, J and K). However, the reduction of $^{n}$BuOPh produced increasing amounts of 1-butoxy-1,3-cyclohexadiene as well as monoolefin **5**. This outcome is similar to the results observed when we only increased the initial concentration of the reaction mixture. Finally, the reduction rate steadily increased when changing from 0 equivalents to 2.5 equivalents of *t*-butanol (Fig. 5C). However, when increasing the equivalents of *t*-butanol past 2.5, the reduction rate decreased (fig. S4A).

From the data presented here, we propose that the reduction of $PhCO_2H$ proceeds according to Fig. 5E; first, lithium(0) is dissolved through the coordination of the amine ligand and THF to create LiN-1. Second, an electron transfer occurs to give radical anion LiN-2. Subsequently, another electron transfer occurs to afford trianion LiN-3, which may be in equilibrium with higher-order aggregates (*48*). Finally, this species is protonated to form LiN-4. Figure 5F shows our hypothesized mechanism for the reduction of $^{n}$BuOPh. An electron is transferred from LiN-1 to the substrate to form radical anion LiN-5. Next, *t*-butanol binds the lithium to give LiN-6, which triggers the rate-determining intramolecular protonation to form the radical species LiN-7. In Fig. 5, E and F, the lithium dissolution and electron transfer are in equilibrium (*45*).

To understand the ligand's effect on reactivity, we first considered the dissolution of lithium(0). If the dissolution step accounts for the structure-reactivity relationship, the effective amines should dissolve lithium faster than ineffective amines (Fig. 2A). Our qualitative experiments with lithium and ethylenediamine, *cis*-, or *trans*-1,2-diaminocyclohexane in THF without arenes showed that although ethylenediamine partially dissolved lithium, the other two amines did not. This is distinct from the fast dissolution of lithium in the presence of arene substrates. Therefore, dissolution alone cannot account for the structure-reactivity relationship.

Second, how do the ligand structures affect the electron transfer processes? We reason that as the denticity of the ligand increases from ethylenediamine to diethylenetriamine then to triethylenetetramine, the amino groups displace the benzoate of LiN-2 with nitrogens, disrupting the electron transfer step, particularly if this is an inner-sphere electron transfer. Currently, it is unclear how many nitrogen atoms are bound to lithium in each intermediate, but the failure with cyclen suggests that when four amino groups are bound, such a complex appears unreactive. Steric effects of amines warrant further studies.

Third, how do the amines influence the rate-determining protonation step (*49*, *50*) for the reduction of $^{n}$BuOPh? Organolithum's carbon is protonated faster with 1,2-diamines than with 1,3-diamines (*48*). Therefore, we suggest that the protonations of LiN-3 and LiN-6 are faster with ethylenediamine than with 1,3-diaminopropane.

Fourth, we considered how the alcohol affects the protonation and product distribution in our reduction. Figure S4 indicates that the alcohol may play a more substantial role than only a proton donor. For example, if *t*-butanol intermolecularly protonates radical anion LiN-5, the rate should be linearly proportional to the alcohol concentration. Instead, we observed a bell-shaped trend (fig. S4B), which indicates

that protonation may occur intramolecularly through LiN-6. The slight preference between related substrates with different steric environments (Fig. 4D) bodes well with this hypothesis. Notably, the reaction mixture containing $^{n}$BuOPh turned light blue with 8 equivalents of *t*-butanol, although the desired reduction did not occur. This suggests that excess alcohol may outcompete amino groups on the lithium at an earlier stage of the reaction, forming less-reductive solvated electrons, similar to work with $SmI_2$ (*51*). A mass effect may have obscured the additional role of *t*-butanol in the past; traditionally, the amine has been used in greater excess than the alcohol, outcompeting the alcohol for coordination to the lithium.

When <1 equivalents of *t*-butanol were present in the reduction of $^{n}$BuOPh, the monoolefin was formed in ~20% yield. This is similar to the Benkeser reduction without alcohol (Fig. 1B) (*15*–*17*, *52*–*54*). Although the addition of an alcohol under the Benkeser-type conditions gave Birch-type products (*4*, *18*, *55*), these findings have not garnered widespread use. The alcohol is necessary to synthesize Birch products by protonating both the organolithiated species (LiN-5 or LiN-6) and the lithium amide in the reaction mixture (*18*). The protonation of the lithium amide then hinders the isomerization of the 1,4-diene to the 1,3-diene, which slows the formation of the monoolefin. Potential effects of *t*-butoxide would warrant further investigation.

Literature has shown that more acidic alcohols (e.g., methanol and ethanol) give faster reductions but lower yields than bulkier alcohols (e.g., isopropanol and *t*-butanol) because of an off-reaction with lithium to create $H_2$ (*45*, *50*). Although our data mostly support such a notion, we wish to consider other factors based on the data with trifluoroethanol (52%), methanol (33%), and ethanol (58%) (table S2) combined with the structural requirements of the amine (Fig. 2A), including optimal bite angle (*56*) (ethylenediamine versus 1,2-diamino-2-methypropane). For example, fig. S5 describes how the equilibrium between a monomer and higher-order aggregates of various ligated lithium intermediates can be affected by the amine ligand among other factors.

The switch of the solvent from an amine to an ethereal solvent (THF) was essential for this work. Altundas's conditions (ammonia gas in a balloon, lithium, and THF) (*30*) suggested that the amine might not be needed as a solvent. 1,2-Dimethoxyethane was ineffective as the solvent, which indicates that only one molecule of THF binds to a lithium ion to form reactive species. The role of THF as a ligand for the alkali metal ion most likely had not been considered before because the ethereal solvent was previously used in smaller amounts than the amine solvent.

The method discussed in this paper could reverse the chemoselectivity for the reduction of $PhCO_2H$ and $^{n}$BuOPh by two orders of magnitude with triethylenetetramine (61-fold difference under the standard Birch reduction conditions in favor of $PhCO_2H$ and twofold difference under our conditions in favor of $^{n}$BuOPh). More broadly, the structure-reactivity relationship indicates the potential for (reverse) chemoselective reduction in synthesis. To control the selectivity, inner- and outer-sphere electron transfer processes may be considered (*22*, *24*). Our work also suggests a broader role for the alcohol than previously considered, including the product selectivity with naphthalene and indole systems. Also, this study gives a platform to investigate solvated electrons at room temperature.

In addition to the theoretical advancements, the practicality of the technology should render the lithium-mediated reduction and deprotection more accessible to a broader scientific community and more amenable to the time-economic synthesis of complex molecules (*57*). Finally, the scope of the Birch reduction may be expanded by combining the chemistry of organolithium with other organometallic chemistry.

## REFERENCES AND NOTES


1. C. J. Huck, D. Sarlah, *Chem* **6**, 1589–1603 (2020).
2. A. J. Birch, *J. Chem. Soc.* 430–436 (1944).
3. C. B. Wooster, K. L. Godfrey, *J. Am. Chem. Soc.* **59**, 596–597 (1937).
4. M. E. Garst *et al.*, *J. Org. Chem.* **65**, 7098–7104 (2000).
5. D. K. Joshi, J. W. Sutton, S. Carver, J. P. Blanchard, *Org. Process Res. Dev.* **9**, 997–1002 (2005).
6. T. Kobayashi, H. Tsuruta, *Synthesis* **1980**, 492–493 (1980).
7. C. Chapuis, D. Skuy, C.-A. Richard, *Helv. Chim. Acta* **102**, e1900097 (2019).
8. E. J. Corey, A. G. Myers, *J. Am. Chem. Soc.* **107**, 5574–5576 (1985).
9. H.-J. Zhang *et al.*, *Angew. Chem. Int. Ed.* **55**, 11638–11641 (2016).
10. X. Zhu, C. C. McAtee, C. S. Schindler, *J. Am. Chem. Soc.* **141**, 3409–3413 (2019).
11. C. L. Hugelshofer, V. Palani, R. Sarpong, *J. Org. Chem.* **84**, 14069–14091 (2019).
12. L.-F. Tietze, T. Eicher, *Reactions and Syntheses in the Organic Chemistry Laboratory* (University Science Books, 1989).
13. B. K. Peters *et al.*, *Science* **363**, 838–845 (2019).
14. T. J. Donohoe, R. E. Thomas, *Nat. Protoc.* **2**, 1888–1895 (2007).
15. R. A. Benkeser, C. Arnold Jr., R. F. Lambert, O. H. Thomas, *J. Am. Chem. Soc.* **77**, 6042–6045 (1955).
16. R. A. Benkeser, R. E. Robinson, D. M. Sauve, O. H. Thomas, *J. Am. Chem. Soc.* **77**, 3230–3233 (1955).
17. R. A. Benkeser *et al.*, *J. Org. Chem.* **29**, 1313–1316 (1964).
18. R. A. Benkeser, M. L. Burrous, J. J. Hazdra, E. M. Kaiser, *J. Org. Chem.* **28**, 1094–1097 (1963).
19. A. O. Bedenbaugh, J. H. Bedenbaugh, W. A. Bergin, J. D. Adkins, *J. Am. Chem. Soc.* **92**, 5774–5775 (1970).
20. L. Reggel, R. A. Friedel, I. Wender, *J. Org. Chem.* **22**, 891–894 (1957).
21. F. Saito, J. Becker, P. R. Schreiner, *J. Org. Chem.* **85**, 4441–4447 (2020).
22. J. J. Gaston *et al.*, *J. Org. Chem.* **86**, 9163–9180 (2021).
23. T. J. Donohoe, D. House, *J. Org. Chem.* **67**, 5015–5018 (2002).
24. P. Lei *et al.*, *Org. Lett.* **20**, 3439–3442 (2018).
25. T. Shindo, Y. Fukuyama, T. Sugai, *Synthesis* **2004**, 692–700 (2004).
26. C. Hiraoka *et al.*, *Tetrahedron Asymmetry* **17**, 3358–3367 (2006).
27. V. K. Tiwari, D. R. Powell, S. Broussy, D. B. Berkowitz, *J. Org. Chem.* **86**, 6494–6503 (2021).
28. D. Huang, A. W. Schuppe, M. Z. Liang, T. R. Newhouse, *Org. Biomol. Chem.* **14**, 6197–6200 (2016).
29. R. G. Harvey, *Synthesis* **1970**, 161–172 (1970).
30. A. Altundas, A. Menzek, D. D. Gültekin, M. Karakaya, *Turk. J. Chem.* **29**, 513–518 (2005).
31. P. W. Rabideau, *Tetrahedron* **45**, 1579–1603 (1989).
32. A. J. Birch, *J. Chem. Soc.* **1946**, 593 (1946).
33. A. R. Murthy, N. S. Sundar, G. S. R. S. Rao, *Tetrahedron* **38**, 2831–2836 (1982).
34. L. N. Mander, R. H. Prager, J. V. Turner, *Aust. J. Chem.* **27**, 2645–2656 (1974).
35. A. K. Singh, R. K. Bakshi, E. J. Corey, *J. Am. Chem. Soc.* **109**, 6187–6189 (1987).
36. A. J. Birch, *J. Chem. Soc.* **1945**, 809–813 (1945).
37. H. Iio, M. Isobe, T. Kawai, T. Goto, *Tetrahedron* **35**, 941–948 (1979).
38. R. A. Archer *et al.*, *J. Org. Chem.* **42**, 2277–2284 (1977).
39. Y. Zong *et al.*, *Angew. Chem. Int. Ed.* **60**, 15286–15290 (2021).
40. A. J. Birch, *J. Chem. Soc.* **1947**, 1270 (1947).
41. S. Danishefsky, P. Cain, *J. Org. Chem.* **40**, 3606–3608 (1975).
42. P. W. Rabideau, Z. Marcinow, in *Organic Reactions* (Wiley, 2004), pp. 1–334.
43. A. J. Birch, A. L. Hinde, L. Radom, *J. Am. Chem. Soc.* **102**, 4074–4080 (1980).
44. A. J. Birch, A. L. Hinde, L. Radom, *J. Am. Chem. Soc.* **102**, 3370–3376 (1980).
45. A. P. Krapcho, A. A. Bothner, *J. Am. Chem. Soc.* **81**, 3658–3666 (1959).
46. G. S. R. S. Rao, H. Ramanathan, K. Raj, *J. Chem. Soc. Chem. Commun.* **1980**, 315–316 (1980).
47. K. Brezina, P. Jungwirth, O. Marsalek, *J. Phys. Chem. Lett.* **11**, 6032–6038 (2020).
48. J. L. Rutherford, D. Hoffmann, D. B. Collum, *J. Am. Chem. Soc.* **124**, 264–271 (2002).
49. H. E. Zimmerman, *Acc. Chem. Res.* **45**, 164–170 (2012).
50. A. Greenfield, U. Schindewolf, *Ber. Bunsenges. Phys. Chem.* **102**, 1808–1814 (1998).
51. M. Shabangi, R. A. Flowers II, *Tetrahedron Lett.* **38**, 1137–1140 (1997).
52. R. A. Benkeser, R. E. Robinson, D. M. Sauve, O. H. Thomas, *J. Am. Chem. Soc.* **76**, 631–632 (1954).
53. R. A. Benkeser, R. F. Lambert, P. W. Ryan, D. G. Stoffey, *J. Am. Chem. Soc.* **80**, 6573–6577 (1958).
54. R. A. Benkeser, R. K. Agnihotri, M. L. Burrous, *Tetrahedron Lett.* **1**, 1–3 (1960).
55. R. A. Benkeser, J. A. Laugal, A. Rappa, *Tetrahedron Lett.* **25**, 2089–2092 (1984).
56. R. M. Beesley, C. K. Ingold, J. F. Thorpe, *J. Chem. Soc.* **107**, 1080–1106 (1915).
57. Y. Hayashi, *J. Org. Chem.* **86**, 1–23 (2021).
58. K. D. Ashtekar, M. Vetticatt, R. Yousefi, J. E. Jackson, B. Borhan, *J. Am. Chem. Soc.* **138**, 8114–8119 (2016).
59. G. S. R. S. Rao, K. V. Bhaskar, *J. Chem. Soc., Perkin Trans. 1* **1993**, 2333–2337 (1993).
60. B. K. Peters *et al.*, *J. Am. Chem. Soc.* **138**, 11930–11935 (2016).
61. M. Biffin, A. Moritz, D. Paul, *Aust. J. Chem.* **25**, 1329–1334 (1972).
62. T. Bykova, N. Al-Maharik, A. M. Z. Slawin, D. O'Hagan, *Org. Biomol. Chem.* **14**, 1117–1123 (2016).
63. J. P. Cole *et al.*, *J. Am. Chem. Soc.* **142**, 13573–13581 (2020).
64. P. F. Schuda, S. J. Potlock, H. Ziffer, *Tetrahedron* **43**, 463–468 (1987).
65. J. Liu *et al.*, *J. Am. Chem. Soc.* **139**, 14470–14475 (2017).
66. M. J. Costanzo, M. N. Patel, K. A. Petersen, P. F. Vogt, *Tetrahedron Lett.* **50**, 5463–5466 (2009).
67. S. Bayindir, N. Saracoglu, *RSC Advances* **6**, 72959–72967 (2016).


## ACKNOWLEDGMENTS

We thank the Koide group members and L. Burrows (National Energy and Technology Laboratory) for their critical comments science.org/doi/10.1126/science.abk3099 on the manuscript. **Funding:** This study was supported by US National Science Foundation CHE-1955758 (to K.K.) and a Uehara Memorial Foundation postdoctoral fellowship (to S.K.). **Author contributions:** Conceptualization: J.B. and K.K. Investigation: J.B., S.K., and K.K. Funding acquisition: K.K. Supervision: K.K. Writing – original draft: J.B. Writing – review and editing: J.B., S.K., and K.K. **Competing interests:** J.B. and K.K. are inventors on US nonprovisional patent application 63/080.205, submitted by the University of Pittsburgh, which covers the use of lithium and the amines shown in this manuscript in ethereal solvents. **Data and materials availability:** All data are available in the main text or the supplementary materials.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abk3099
Materials and Methods
Figs. S1 to S6
Tables S1 and S2
NMR Spectra
References (*68*–*77*)

5 July 2021; accepted 20 September 2021
10.1126/science.abk3099

AEROSOL CHEMISTRY

# A surface-promoted redox reaction occurs spontaneously on solvating inorganic aerosol surfaces

Xiangrui Kong[1]*, Dimitri Castarède[1], Erik S. Thomson[1], Anthony Boucly[2], Luca Artiglia[2], Markus Ammann[2], Ivan Gladich[3]*, Jan B. C. Pettersson[1]*

**A surface-promoted sulfate-reducing ammonium oxidation reaction was discovered to spontaneously take place on common inorganic aerosol surfaces undergoing solvation. Several key intermediate species—including elemental sulfur ($S^0$), bisulfide ($HS^-$), nitrous acid (HONO), and aqueous ammonia [$NH_{3(aq)}$]—were identified as reaction components associated with the solvation process. Depth profiles of relative species abundance showed the surface propensity of key species. The species assignments and depth profile features were supported by classical and first-principles molecular dynamics calculations, and a detailed mechanism was proposed to describe the processes that led to unexpected products during salt solvation. This discovery revealed chemistry that is distinctly linked to a solvating surface and has great potential to illuminate current puzzles within heterogeneous chemistry.**

Gas–particle interfaces play essential roles in the atmosphere and directly influence many atmospheric processes (*1*), including gas uptake (*2*), halogen chemistry (*3*), ozone depletion (*4*) and heterogeneous ice nucleation (*5*). Because interfacial processes take place on molecular scales, classical bulk thermodynamic theories are often insufficient to describe interfaces. For example, hygroscopic substrates are understood to solvate at a deliquescence relative humidity (DRH), but modern observational techniques have revealed that reversible water adsorption and ion solvation already take place at a relative humidity (RH) lower than the DRH (*6*, *7*). Explanations for such surface-specific behavior can be rooted in strictly accounting for the distinctive physicochemical characteristics of surfaces (*8*–*10*).

Redox reactions are essential for many key processes in the atmosphere and regulate the formation of gas molecules and aerosol particles. However, the current redox chemistry framework cannot explain the chemical cycles of some key atmospheric components, such as nitrous acid (HONO) (*11*) and sulfate (*12*). Unexpected redox reactions may take place on surfaces and interfaces, but interfacial processes are challenging to characterize and have often been overlooked in atmospheric chemistry. In this study, a surface-promoted redox reaction during salt solvation was discovered with the use of ambient-pressure x-ray photoelectron spectroscopy. Classical and first-principles molecular dynamics (FPMD) simulations were performed to understand and characterize the solvation process. Ammonium sulfate is a common atmospheric compound and was used as the salt in this study, as a proxy for atmospheric inorganic aerosol (*13*).

Figure 1 shows the binding energies measured using x-ray photoelectron spectroscopy (XPS) for sulfur 2p and nitrogen 1s electrons at dry (RH = 3%), pre-deliquescence (RH = 48%), and full deliquescence (RH = 78%, including transient and steady state) conditions. At the lowest RH, the only features observed were the expected sulfate ($SO_4^{2-}$) doublet with the sulfur $2p_{3/2}$ peak around 167.5 eV (Fig. 1C, spectrum α) and the ammonium ($NH_4^+$) peak at ~401.0 eV (*14*). As the RH was increased to 48%, additional doublets of reduced sulfur species ($S^r$), including elemental sulfur ($S^0$ with $2p_{3/2}$ around 163.2 eV) (*14*) and bisulfide ($HS^-$ with $2p_{3/2}$ around 162.5 eV), became visible (Fig. 1D, spectrum β). Two additional nitrogen-containing components—nitrous acid (HONO) at ~404 eV and dissolved ammonia ($NH_3$) (*15*) at ~399 eV (*14*)—also appeared with increasing RH (Fig. 1F, spectrum δ). At RH = 78%, where full deliquescence occurred (*16*), a full conversion from $SO_4^{2-}$ to $S^0$ was observed (Fig. 1E, spectrum γ) during the deliquescing transient period (fig. S1). As the deliquescence process proceeded and a new steady state was reached, the $SO_4^{2-}$ reappeared and only a small amount of $HS^-$ remained on the surface. Identification of the spectroscopic signatures of different sulfur and nitrogen species has been supported and/or guided by both molecular dynamics (MD) and extensive density of states (DOS) calculations: MD trajectories were consistent with the formation and stability of the identified experimental species, and their calculated spectroscopic signatures showed excellent agreement with the experimental peak assignments [see the supplementary materials (SM) for a detailed discussion, as well as tables S2 and S3]. Notably, a series of measurements was performed to evaluate the influence of x-ray beam damage, and the results confirmed that beam-induced artifacts and adventitious carbon had not contributed to and/or influenced the observations [see SM for details (supplementary text section "Exclusion of Contributions from Radiation Damage and Adventitious Carbon"), as well as figs. S2 to S5). In addition, oxygen gas was dosed to check its influence on the redox reactions, but no effects from oxygen were observed (fig. S6).

During the initial stage of deliquescence at RH = 78%, all nitrogen species completely vanished from the spectrum (Fig. 1B). Photoemission wide-range spectra measured during deliquescence (fig. S7) showed the absence of nitrogen species and a simultaneous increase of $S^r$ during the salt solvation. Notably, this phenomenon, during which the deliquesced surface was pure water mixed with $S^r$, was transient, and the $SO_4^{2-}$, $NH_4^+$, and $NH_{3(aq)}$ species reappeared in 1 hour, once the system was stabilized (fig. S1). Our MD simulations (Fig. 2B and fig. S9) showed that a water submonolayer on top of the ammonium sulfate crystal was sufficient to solvate $SO_4^{2-}$ and $NH_4^+$ ions from the crystal into the adsorbed liquid-like layer. Thus, the temporary complete depletion observed by means of XPS during the transient period indicated that additional chemical mechanisms were at play.

A sulfate-reducing ammonium oxidation (SRAO) reaction scheme, involving $NH_4^+$ and $SO_4^{2-}$ in aqueous systems, has been hypothesized previously to explain nitrogen depletion and sulfate reduction in anaerobic environments (*17*) and has been widely applied in water treatment contexts (*18*). At present, the reaction scheme with Gibbs free energy ($\Delta G^0$) is described as (*17*)

$$3SO_4^{2-} + 6NH_4^+ \rightarrow 3S^0 + 3N_2 + 12H_2O \ (\Delta G^0 = -135.9\ \mathrm{kJ/mol}) \quad (1)$$

which consists of three subreaction steps (*19*)

$$3SO_4^{2-} + 4NH_4^+ \rightarrow 3S^{2-} + 4NO_2^- + 4H_2O + 8H^+ \ (\Delta G^0 = +1714.1\ \mathrm{kJ/mol}) \quad (2)$$

$$2NO_2^- + 2NH_4^+ \rightarrow 2N_2 + 4H_2O \ (\Delta G^0 = -720.2\ \mathrm{kJ/mol}) \quad (3)$$

$$3S^{2-} + 2NO_2^- + 8H^+ \rightarrow 3S^0 + N_2 + 4H_2O \ (\Delta G^0 = -1129.8\ \mathrm{kJ/mol}) \quad (4)$$

Although the overall reaction (1) is thermodynamically favorable, it is not spontaneous because not all of the elementary steps are

[1]Atmospheric Science Research Division, Department of Chemistry and Molecular Biology, University of Gothenburg, SE-41296 Gothenburg, Sweden. [2]Laboratory of Environmental Chemistry, Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland. [3]Qatar Environment and Energy Research Institute, Hamad Bin Khalifa University, P.O. Box 31110, Doha, Qatar.
*Corresponding author. Email: kongx@chem.gu.se (X.K.); igladich@hbku.edu.qa (I.G.); janp@chem.gu.se (J.B.C.P.)

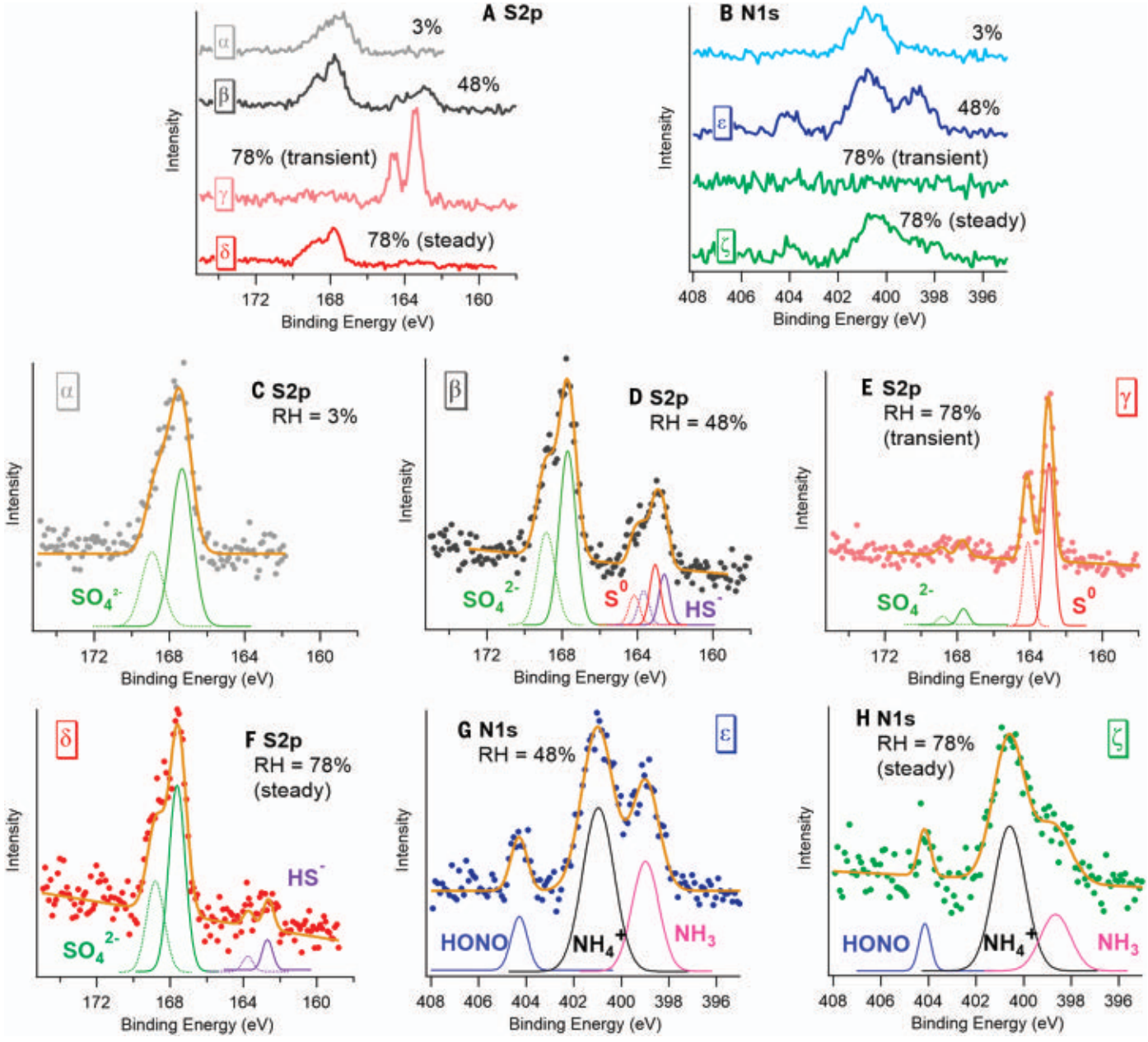



**Fig. 1. Measured and fitted XPS spectra at RHs of 3, 48, and 78%.** The photon energy was set to 435 eV for sulfur XPS and 665 eV for nitrogen XPS. The probing depth was ~0.8 nm (*30*). The binding energies were aligned with aliphatic C 1s at 284.8 eV. Greek symbols denote the fitted spectra shown in (C) to (H). (**A**) Measured S 2p and (**B**) N 1s XPS peaks. At RH = 3%, a $NH_4^+$ peak (B) and a $SO_4^{2-}$ doublet (**C**) were observed. At RH = 48%, $S^0$ and $HS^-$ doublets (**D**) and HONO and $NH_3$ peaks (**G**) appeared. At RH = 78%, where full deliquescence occurred, most sulfate transferred to $S^0$ (**E**), and no nitrogen species could be found (B). When deliquescence was complete, the sulfur region was again dominated by sulfate (**F**), and nitrogen species reappeared (**H**).

favored (*17*). Thus, biological catalysts (e.g., bacteria) are often used to facilitate SRAO reactions in industrial applications (*18*). However, owing to a lack of direct experimental evidence, a detailed mechanism has not been fully elucidated. Taken together, results from our XPS, FPMD, and DOS analyses indicated that the two intermediate species ($S^{2-}$ and $NO_2^-$) appeared to be protonated—i.e., $HS^-$ and HONO. Thus, the relevant reactions observed become

$$3SO_4^{2-} + 4NH_4^+ \rightarrow 3HS^- + 4HONO + 4H_2O + H^+ \quad (5)$$

$$2HONO + 2NH_4^+ \rightarrow 2N_2 + 4H_2O + 2H^+ \quad (6)$$

$$3HS^- + 2HONO + 3H^+ \rightarrow 3S^0 + N_2 + 4H_2O \quad (7)$$

All relevant surface species ($HS^-$, $S^0$, and HONO) were detected during salt solvation. Although $N_2$ was not found, owing to its low solubility, the observed nitrogen depletion during the transient period indirectly indicated that $N_2$ was likely formed and quickly escaped from the surface. In the atmosphere, the escape of $N_2$ to the gas phase is also expected to occur because of the low solubility of $N_2$ and the spontaneity of reactions (4) and (7). The reactions [(5) to (7)] stopped when the new steady state (full solvation) was reached, and only in the transient period (during deliquescence) did it appear that the SRAO-type reactions were boosted at the

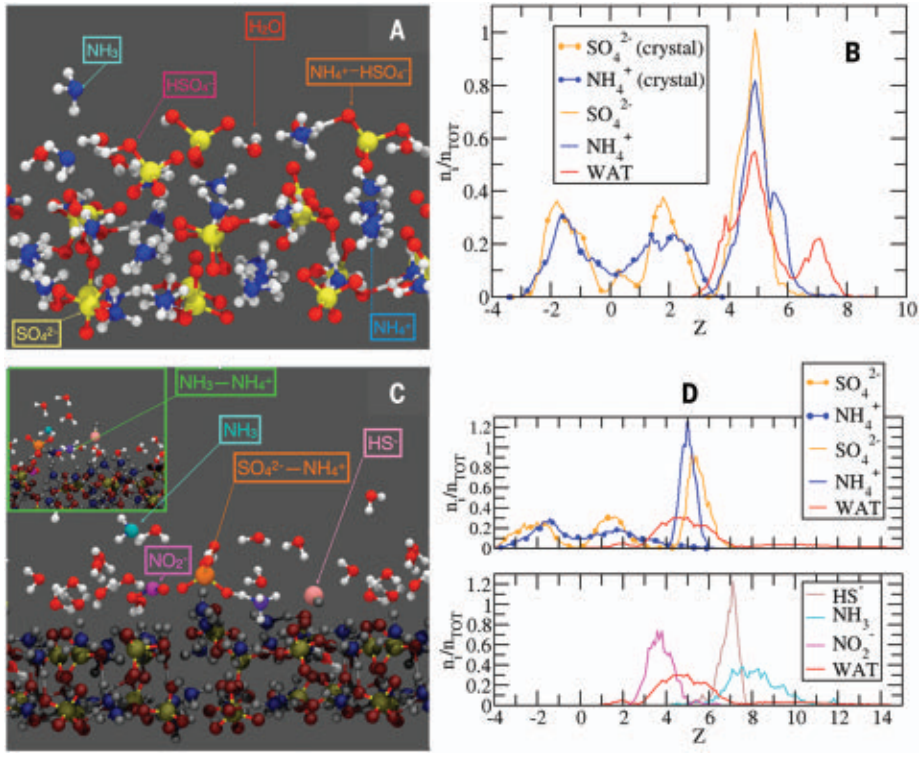



**Fig. 2. FPMD results.** (**A**) Snapshot from a 40-ps FPMD of ammonium and sulfate ions solvated in an adsorbed eight–water molecule (WAT) sublayer on top of the ammonium sulfate crystal [configuration a (conf-a) in the SM]. Atom color code: N, blue; S, yellow; O, red; and H, white. (**B**) Probability distribution profile, $n_i/n_{TOT}$, normalized to unity as a function of the *z*-coordinate perpendicular to the liquid–crystal interface, obtained by collecting the center-of-mass position of each species over the MD trajectory. Dotted blue and orange lines denote sulfate and ammonium ions that initially belong to the crystal substrate. (**C**) Snapshot from a 40-ps FPMD [configuration c (conf-c) FPMD in the SM] of ammonia and ammonium, sulfate, sulfide, and nitrite ions in a 16–water molecule layer adsorbed on top of the crystal. Orange, ammonium sulfate ion pairing; pink, hydrogen sulfide formation; magenta, nitrite; and cyan, ammonia. (Inset) A different time frame from the same trajectory, which shows proton shuffling between $NH_4^+$ and $NH_3$ (green) in the adsorbed water layer. Shadowed atoms are those that belonged to the initial ammonium sulfate crystal structure. (**D**) Same as (B) but collected over the 40-ps conf-c FPMD trajectory (see SM). For visual clarity, the density profile is split into two panels.

salt–vapor interface. During efflorescence, no $S^r$ formation was observed with decreasing RH (fig. S5; from RH = 78 to 8%), which indicates that this mechanism is associated with surface solvation but not crystallization.

When RH was raised from 3 to 48% (Fig. 1, D and G), all SRAO-relevant products ($S^0$, $HS^-$, HONO) were detected, but during transient conditions at DRH (RH = 78%; Fig. 1, B and E) only the final products ($S^0$ and water) were detected on the surface. This difference was likely because at high RH there was abundant molecular water and sufficient interactions between intermediate species to complete the reaction. Furthermore, a layer of pure water on a $(NH_4)_2SO_4$ surface was not expected to be stable, and salt surface dissolution would proceed instead. Indeed, the gradual salinization of the water layer was recorded during a 1-hour period (fig. S1). The slow salinization was most likely a result of the competition between the solvation and the SRAO-like reaction, and when the system reached steady state the SRAO-like reaction stopped and the aqueous phase was stabilized as a saturated solution. The MD simulation trajectories (Fig. 2, C and D, and figs. S9 and S11) showed that ammonium and sulfate ions from the crystal structure started to be solvated in the adsorbed water quite rapidly (within 40 ps). The density profiles for ammonium (blue dotted line), sulfate ions (orange dotted line), and water (red line) overlap in Fig. 2, B and D, supporting the concept that a first adsorbed water layer was sufficient to initiate the dissolution of $(NH_4)_2SO_4$ from the crystal into the liquid layer. Trajectory movie S3 shows that ammonium and sulfate ions initially located in the crystal structure interacted with the adsorbed species and began to be solvated in the water layer. Finally, Fig. 2, C and D, indicates the propensity of nitrite ions to reside in the inner part of the adsorbed water layer, close to the crystal structure interface: The FPMD showed that the hectic proton dynamic in this environment facilitated HONO formation by protonation of nitrite ions. When HONO was formed, it remained stable over all of the MD trajectory lengths (figs. S11 and S14D).

Dissolved $NH_3$ was also detected and was formed from $NH_4^+$ ions after deprotonation. The $NH_{3(aq)}$ has a higher surface propensity than $NH_4^+$ ions (*15*), as confirmed in this study by depth profiling and MD simulation results (Fig. 2D and SM). Notably, the MD results suggest that the ammonia surface propensity was not limited to the fully deliquesced steady state but also occurred in the water-restricted environment of approximately one water layer adsorbed on the ammonium sulfate surface (Fig. 2D and SM). Nevertheless, $NH_{3(aq)}$ could be stabilized in the condensed phase by forming clusters with $NH_4^+$ (*20*). This was also observed in the MD trajectory (Fig. 2C, inset) and may be relevant in this case because $NH_4^+$ was highly concentrated at the interface in the pre-deliquescence stages. Ammonia and ammonium ion water-mediated clustering was also observed in FPMD trajectories (movie S1C). The proposed pathway and reaction mechanisms are illustrated in Fig. 3.

Depth profiles of the ratios between elements and species (i.e., as a function of electron kinetic energy) are shown in Fig. 4. The idealized elemental ratios for pure ammonium sulfate and ammonium bisulfate are marked by dashed and dotted lines, respectively.

1) $Sr:SO_4^{2-}$ ratio. First, the ratios of $S^r$ and $NH_3$ with respect to their primary ions ($SO_4^{2+}$ and $NH_4^+$) are shown in Fig. 4A ($S^r:SO_4^{2+}$) and Fig. 4B ($NH_3:NH_4^+$). Because $S^r$ was formed during salt solvation, its abundance depended on the deliquescence process. Thus, $S^r$ was abundant at RH = 48%, when water likely partially covered the surface and the SRAO-like reactions for $S^r$ formation were favored. As the system gradually approached steady state (as shown by the time-lapse XPS survey in fig. S1), the aqueous solution at RH = 78% showed much less $S^r$, because the formed $S^r$ was diluted by the condensing water and moved away.

2) $NH_{3(aq)}:NH_4^+$ ratio. $NH_{3(aq)}$ was relatively surface-enriched relative to $NH_4^+$, because $NH_{3(aq)}$ has a higher surface propensity than $NH_4^+$ ions (*15*) in aqueous solutions. At RH = 48%, the $NH_{3(aq)}$ might be stabilized in the condensed phase by $NH_4^+$ ions (*20*). MD results supported the $NH_3$ enrichment at the vapor/adsorbed liquid layer (even in the pre-deliquesced state at a relatively low RH; Fig. 2D) and its stabilization in the condensed phase by clustering with ammonium ions (Fig. 2C, inset).

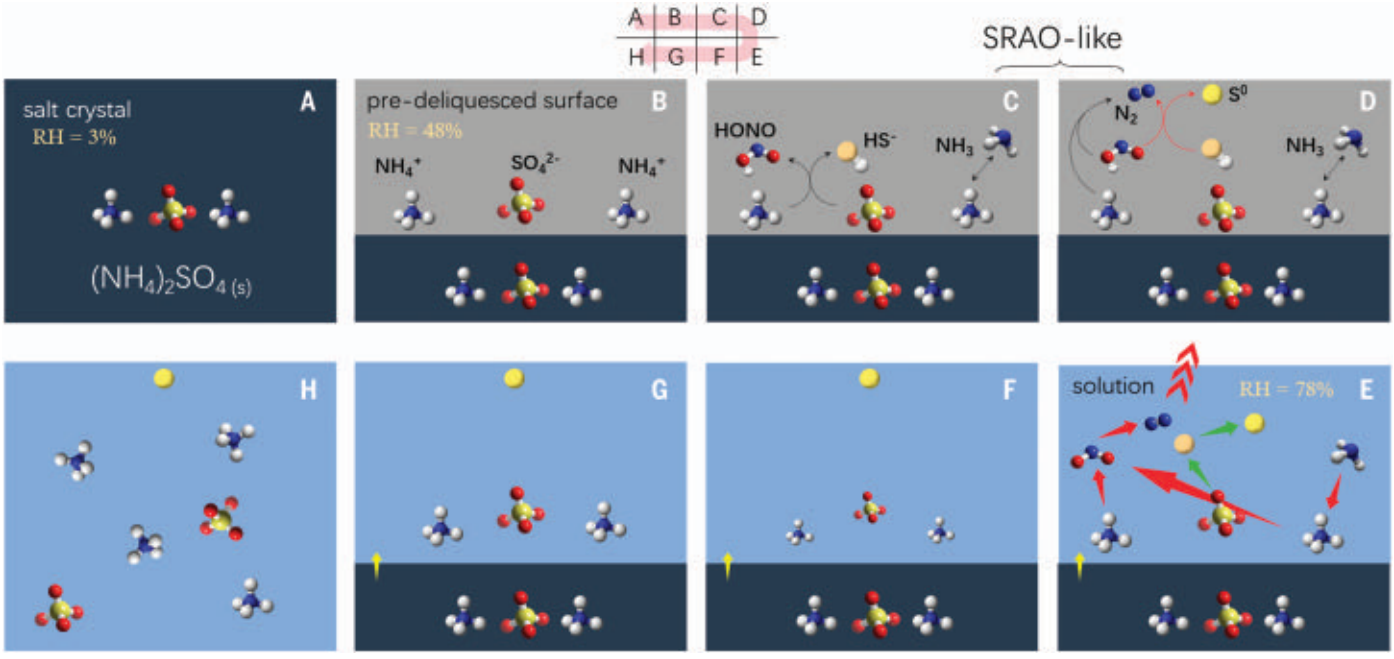



**Fig. 3. Schematic view of the outlined mechanism, with $H_2O$ molecules and protons omitted.** (**A**) Dry salt crystal. (**B**) Pre-deliquescence, when ions are released from salt crystals. (**C**) SRAO-like process; reaction (3). (**D**) SRAO-like process; reaction (4). (**E**) During full deliquescence, the escape of $N_2$ drives SRAO reactions. (**F**) SRAO end products—i.e., $H_2O$, $S^0$, and $N_2$ (escaped)—dominate. (**G**) Gradually salinized solution. (**H**) Saturated solution after full salt dissolution.

3) S:N ratios. The $S_{tot}$:$N_{tot}$ ratio in Fig. 4C lies between the reference dashed and dotted lines for $(NH_4)_2SO_4$ and $NH_4HSO_4$, indicating that the sample was stoichiometrically equivalent to a mixture of the two salts. The mixture of $(NH_4)_2SO_4$ and $NH_4HSO_4$ salts in the pre-deliquesced state was also supported by the MD trajectories, which showed vivid proton dynamics on the adsorbed water layer, with ion pairing and proton transfer from $NH_4^+$ to $SO_4^{2-}$, forming $HSO_4^-$, and $NH_4HSO_4$ and $(NH_4)_2SO_4$ clustering (Fig. 2, A and C; fig. S14B; and movie S2). A closer look at the $SO_4^{2-}$:$NH_4^+$ ratio in Fig. 4D (i.e., the primary ions) showed similar trends to those in Fig. 4C, which suggests that the primary ions dominated the S and N species. In Fig. 4E ($S^r$:$NH_4^+$), the surface enrichment of $S^r$ was the main reason for the relatively high ratio at the first point at RH = 48%.

4) S:O ratios. In the comparison of sulfur to oxygen (Fig. 4F), the dry sample showed the expected S:O ratio because $SO_4^{2-}$ was the only source of sulfur. At RH = 48%, even though the $SO_4^{2-}$:O ratios decreased (Fig. 4G) because of water dilution and SRAO-like reactions, the overall $S_{tot}$:$O_{tot}$ ratios increased because of the existence of $S^r$ (Fig. 4H). In the aqueous system at RH = 78%, the S:O ratio showed the lowest values because of the contributions from condensed water.

5) N:O ratios. The N:O ratio in Fig. 4I was generally stable over the entire probed depth except for the uppermost point of the aqueous solution, which was caused by slight surface enrichment of both $NH_4^+$ (Fig. 4J) and $NH_{3(aq)}$ (Fig. 4K) with respect to oxygen. The surface-enriched $NH_3$ and $NH_4^+$ indicated that the nitrogen depletion during deliquescence totally recovered in the fully solvated state. At RH = 48%, the $NH_4^+$:O ratio (Fig. 4J) showed a gentle surface nitrogen-depletion trend, which might be due to the slow but ongoing SRAO-like reactions that generated $N_2$ gas. Otherwise, the $N_{tot}$:$O_{tot}$ ratios located between the two reference lines were consistent with the $S_{tot}$:$N_{tot}$ ratios in Fig. 4C.

Redox reactions are essential in the atmosphere because they are responsible for many processes that govern the formation of gas molecules and aerosol particles. In this study, a redox reaction that spontaneously occurs on typical inorganic aerosol surfaces during water solvation was discovered and characterized, from both experimental and theoretical perspectives. The limited hydration at the initial solvation stage of the ammonium sulfate crystal promoted the SRAO-like reaction, favoring transfer and accommodation of protons from ammonium ions and only partially solvating sulfate ions (see SM for discussion of FPMD results). Because this surface chemistry mechanism had not been identified previously, new implications and applications are likely and will benefit from this more complete picture of microscopic surface processes.

In the context of atmospheric science, sulfate and ammonium are the primary components of secondary inorganic aerosols, and the thermodynamic properties of the relevant aerosol particles have been well characterized (*21*). It has previously been indicated that a spontaneous oxidation pathway from $SO_2$ to sulfate may exist on salt surfaces (*22*), which implies that the identified surface mechanism may play a role in atmospheric heterogeneous chemistry. The sulfur reduction is an essential component of the full sulfur chemistry cycle. Also, the revealed mechanism may complement the autocatalyzed sulfur chemistry pathway of elemental sulfur aerosol formation in planetary atmospheres (*23*). Furthermore, this mechanism may contribute to the formation of HONO, which is a key tropospheric intermediate. Recently, it is has been reported that the traditionally considered HONO formation pathway—the heterogeneous reaction of $NO_2$ on water surfaces—is not sufficient to explain the observed high HONO concentrations in the troposphere (*11*), and an additional surface catalysis mechanism may have a critical role (*24*). In this study, HONO was formed as an intermediate product on the ammonium sulfate surface, indicating an unconventional pathway of HONO formation directly from inorganic aerosol particles during surface solvation. Ammonium sulfate particles have also been recognized as active ice nuclei for cirrus cloud formation (*25*). As with deliquescence, deposition ice nucleation is initiated by surface water adsorption, which means that it too may be influenced by the observed processes and products, such as HONO, $HS^-$, and

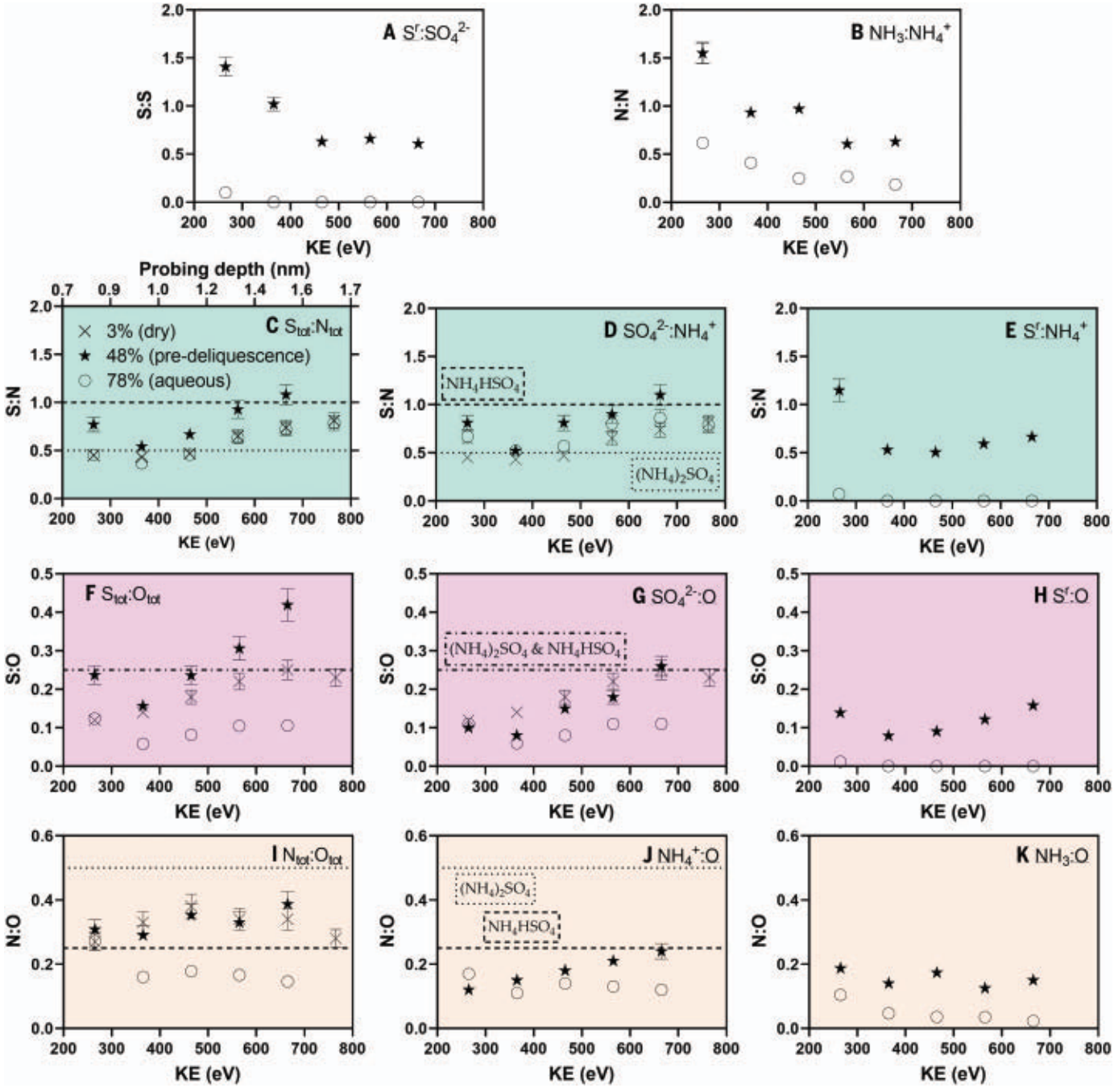



**Fig. 4. Depth profiles of the species ratios.** (**A** to **K**) The specific species are indicated in each panel. Background color indicates the elemental ratios (S:N in green, S:O in purple, N:O in beige). Three RH conditions are included: RH = 3% for dry conditions (crosses), 48% for pre-deliquesced conditions (stars), and 78% for aqueous solution in steady state (circles). The electron inelastic mean free path (IMFP) in a sulfate salt ($Na_2SO_4$) is calculated by QUASES-IMFP-TPP2M version 3.0 (*30*), and the probed depth (attenuation length) shown on the upper *x* axis in (C) is calculated as IMFP · cos(30°), owing to the geometry of the setup. Note the uncertainty created by the electron attenuation length (*31*). Dotted lines indicate the elemental ratio in pure $(NH_4)_2SO_4$, dashed lines show the ratio of pure $NH_4HSO_4$, and dashed-dotted lines show the identical ratio for both $(NH_4)_2SO_4$ and $NH_4HSO_4$. Error bars indicate SDs estimated according to the variations of individual XPS spectra. Depth profiles from the MD simulations can be found in Fig. 2 and fig. S9.

$S^0$. For example, such a complicated surface environment may explain why ammonium sulfate acts as an efficient ice-nucleating particulate (*5*, *25*–*28*) when compared with other hygroscopic salts such as NaCl (*29*).

Furthermore, because the reactions are spontaneous and therefore cost-efficient, potential industrial applications may be developed on the basis of this concept. For example, SRAO has been widely used as an efficient approach for wastewater treatment, because it simultaneously converts ammonium and sulfate into $N_2$ gas and elemental sulfur without secondary pollution (*18*). A major difference is that the currently used SRAO reaction

requires biocatalysts to overcome the energy barrier in the aqueous environment, whereas the mechanism reported here was spontaneous during surface solvation. This surprising finding adds to the current understanding of heterogeneous surface chemistry and creates new challenges for further investigations.

## REFERENCES AND NOTES


1. C. George, M. Ammann, B. D'Anna, D. J. Donaldson, S. A. Nizkorodov, *Chem. Rev.* **115**, 4218–4258 (2015).
2. T. Huthwelker, M. Ammann, T. Peter, *Chem. Rev.* **106**, 1375–1444 (2006).
3. W. R. Simpson, S. S. Brown, A. Saiz-Lopez, J. A. Thornton, R. von Glasow, *Chem. Rev.* **115**, 4035–4062 (2015).
4. J. B. Burkholder, R. A. Cox, A. R. Ravishankara, *Chem. Rev.* **115**, 3704–3759 (2015).
5. C. Hoose, O. Möhler, *Atmos. Chem. Phys.* **12**, 9817–9854 (2012).
6. M. E. Wise, S. T. Martin, L. M. Russell, P. R. Buseck, *Aerosol Sci. Technol.* **42**, 281–294 (2008).
7. X. Kong *et al*., *J. Phys. Chem. C* **124**, 5263–5269 (2020).
8. D. A. Bruzewicz *et al*., *J. Chem. Phys.* **134**, 044702 (2011).
9. H. Hansen-Goos, E. S. Thomson, J. S. Wettlaufer, *Planet. Space Sci.* **98**, 169–181 (2014).
10. D. Castarède, E. S. Thomson, *Atmos. Chem. Phys.* **18**, 14939–14948 (2018).
11. M. T. C. Martins-Costa, J. M. Anglada, J. S. Francisco, M. F. Ruiz-López, *J. Am. Chem. Soc.* **142**, 20937–20941 (2020).
12. G. Wang *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13630–13635 (2016).
13. R. J. Huang *et al*., *Nature* **514**, 218–222 (2014).
14. A. V. Naumkin, A. Kraut-Vass, C. J. Powell, NIST Standard Reference Database 20, Version 3.4, NIST X-ray Photoelectron Spectroscopy Database (National Institute of Standards and Technology, 2003); https://srdata.nist.gov/xps/.
15. C. F. Fu, S. X. Tian, *J. Phys. Chem. C* **117**, 13011–13020 (2013).
16. Y. M. Shul'ga *et al*., *Synth. Met.* **70**, 1381–1382 (1995).
17. F. Fdz-Polanco *et al*., *Water Sci. Technol.* **44**, 77–84 (2001).
18. E. Rikmann *et al*., *Biodegradation* **23**, 509–524 (2012).
19. F. Fdz-Polanco *et al*., *Water Res.* **35**, 1111–1114 (2001).
20. P. Hvelplund *et al*., *J. Phys. Chem. A* **114**, 7301–7310 (2010).
21. C. A. Colberg, B. P. Luo, H. Wernli, T. Koop, T. Peter, *Atmos. Chem. Phys.* **3**, 909–924 (2003).
22. J. Turšič, A. Berner, B. Podkrajšek, I. Grgić, *Atmos. Environ.* **38**, 2789–2795 (2004).
23. M. Kumar, J. S. Francisco, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 864–869 (2017).
24. Y. Suleymanov, *Science* **371**, 139 (2021).
25. J. P. D. Abbatt *et al*., *Science* **313**, 1770–1773 (2006).
26. J. E. Shilling, T. J. Fortin, M. A. Tolbert, *J. Geophys. Res.* **111**, D12204 (2006).
27. D. J. Cziczo, J. P. D. Abbatt, *J. Geophys. Res.* **104**, 13781–13790 (1999).
28. K. J. Baustian, M. E. Wise, M. A. Tolbert, *Atmos. Chem. Phys.* **10**, 2307–2317 (2010).
29. X. R. Kong *et al*., *Tellus B* **70**, 1463806 (2018).
30. S. Tanuma, C. J. Powell, D. R. Penn, *Surf. Interface Anal.* **21**, 165–176 (1994).
31. N. Ottosson, M. Faubel, S. E. Bradforth, P. Jungwirth, B. Winter, *J. Electron Spectrosc. Relat. Phenom.* **177**, 60–70 (2010).
32. X. Kong *et al*., Data for "A surface-promoted redox reaction occurs spontaneously on solvating inorganic aerosol surfaces," Version 2, Zenodo (2021); https://doi.org/10.5281/zenodo.5222334.


## ACKNOWLEDGMENTS

X.K. thanks P. Corral Arroyo for inspiring discussions during the beamtime at the SLS. I.G. thanks F. Pietrucci, F. Jensen, and M. Carignano for fruitful discussions. **Funding:** This work is supported by the Swedish Research Council VR (2015-04212). X.K. acknowledges support from the National Natural Science Foundation of China (41975160) and the Swedish Foundation for International Cooperation in Research and Higher Education (CH2019-8361). E.S.T. is supported by the Swedish Research Councils Formas (2017-00564) and VR (2020-03497) and by the Swedish Strategic Research Area MERGE. For high-performance computing (HPC) resources and services, we acknowledge the Research Computing group at Texas A&M University at Qatar, founded by the Qatar Foundation for Education, Science and Community Development, and use of Qatar Environment and Energy Research Institute (QEERI) HPC under Project ID HPC-P20003. **Author contributions:** X.K., E.S.T., and J.B.C.P. initialized the concept. X.K. analyzed the data and wrote the experimental part of the paper. I.G. performed the calculations and wrote the theoretical part of the paper. X.K., L.A., and E.S.T. coordinated the beamtime. X.K., D.C., A.B., and E.S.T. carried out the beamtime experiments. L.A. and M.A. supported the beamtime and data analysis. All authors discussed the results and contributed to data interpretation and paper revisions. **Competing interests:** None declared. **Data and materials availability:** All data needed to evaluate the conclusions in this paper are present in the paper or in the supplementary materials. All data presented in this paper are deposited at Zenodo (*32*).

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abc5311
Materials and Methods
Supplementary Text
Figs. S1 to S15
Tables S1 to S3
References (*33*–*73*)
Movies S1 to S5

7 December 2020; resubmitted 14 July 2021
Accepted 19 August 2021
10.1126/science.abc5311

ORGANIC CHEMISTRY

# Enantioselective catalytic 1,2-boronate rearrangements

**Hayden A. Sharma†, Jake Z. Essman†, Eric N. Jacobsen***

**A strategy that facilitates the construction of a wide variety of trisubstituted stereocenters through a catalytically accessed common chiral intermediate could prove highly enabling for the field of synthetic chemistry. We report the discovery of enantioselective, catalytic 1,2-boronate rearrangements for the synthesis of α-chloro pinacol boronic esters from readily available boronic esters and dichloromethane. The chiral building blocks produced in these reactions can undergo two sequential stereospecific elaborations to generate a wide assortment of trisubstituted stereocenters. The enantioselective reaction is catalyzed by a lithium-isothiourea-boronate complex, which is proposed to promote rearrangement through a dual–lithium-induced chloride abstraction orchestrated by Lewis basic functionality on the catalyst scaffold.**

Organoboron compounds are valuable reagents for the synthesis of a wide variety of molecules with useful biological or material properties (*1*–*4*). In particular, the stereospecific elaboration of enantioenriched alkylboron reagents is a powerful strategy for the construction of molecules bearing carbon- and heteroatom-substituted stereocenters (*5*–*7*). Many of these elaboration reactions leverage the characteristic reactivity of tetracoordinate boronate derivatives bearing α-leaving groups, which are prone to 1,2-rearrangement pathways that are stereoinvertive at the carbon bearing the leaving group and stereoretentive at the migrating carbon (Fig. 1A). These rearrangements find substantial utility in stereoselective synthesis, as exemplified by the venerable Matteson homologation reaction in which boronic esters undergo a one-carbon chain extension via the formation and rearrangement of dichloromethyl-substituted boronates (Fig. 1B) (*8*–*10*). This process can be rendered stereoselective through the use of chiral diol auxiliaries on boron. The α-chloro boronic ester products of these diastereoselective rearrangement reactions are configurationally stable and can undergo a second nucleophilic addition and rearrangement with high stereospecificity to afford a variety of chiral secondary boronic esters bearing chiral diol substitution on boron.

We envisioned that the development of a catalytic, enantioselective rearrangement of pinacol-substituted dichloromethyl boronates could form the basis for a modular and highly general strategy for the synthesis of trisubstituted stereocenters. Such a reaction would afford chiral building blocks bearing chlorine and pinacol boronic ester substituents on a defined stereogenic carbon center (*11*) that could be elaborated enantiospecifically through a second rearrangement sequence and a carbon-boron bond transformation (Fig. 1C). In addition to requiring only catalytic levels of a chiral source to achieve absolute stereocontrol, this strategy would leverage recent, powerful methods for the elaboration of carbon-boron bonds that primarily make use of pinacol boronic esters (*5*). The catalytic, enantiodetermining event would involve commitment to only one of the three ultimate substituents, thereby enabling immense versatility and potential generality in the enantioselective construction of trisubstituted carbon stereocenters from dichloromethane as the common starting material. Here, we report the discovery of a catalytic method for controlling

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA.
***Corresponding author. Email: jacobsen@chemistry.harvard.edu**
†These authors contributed equally to this work.

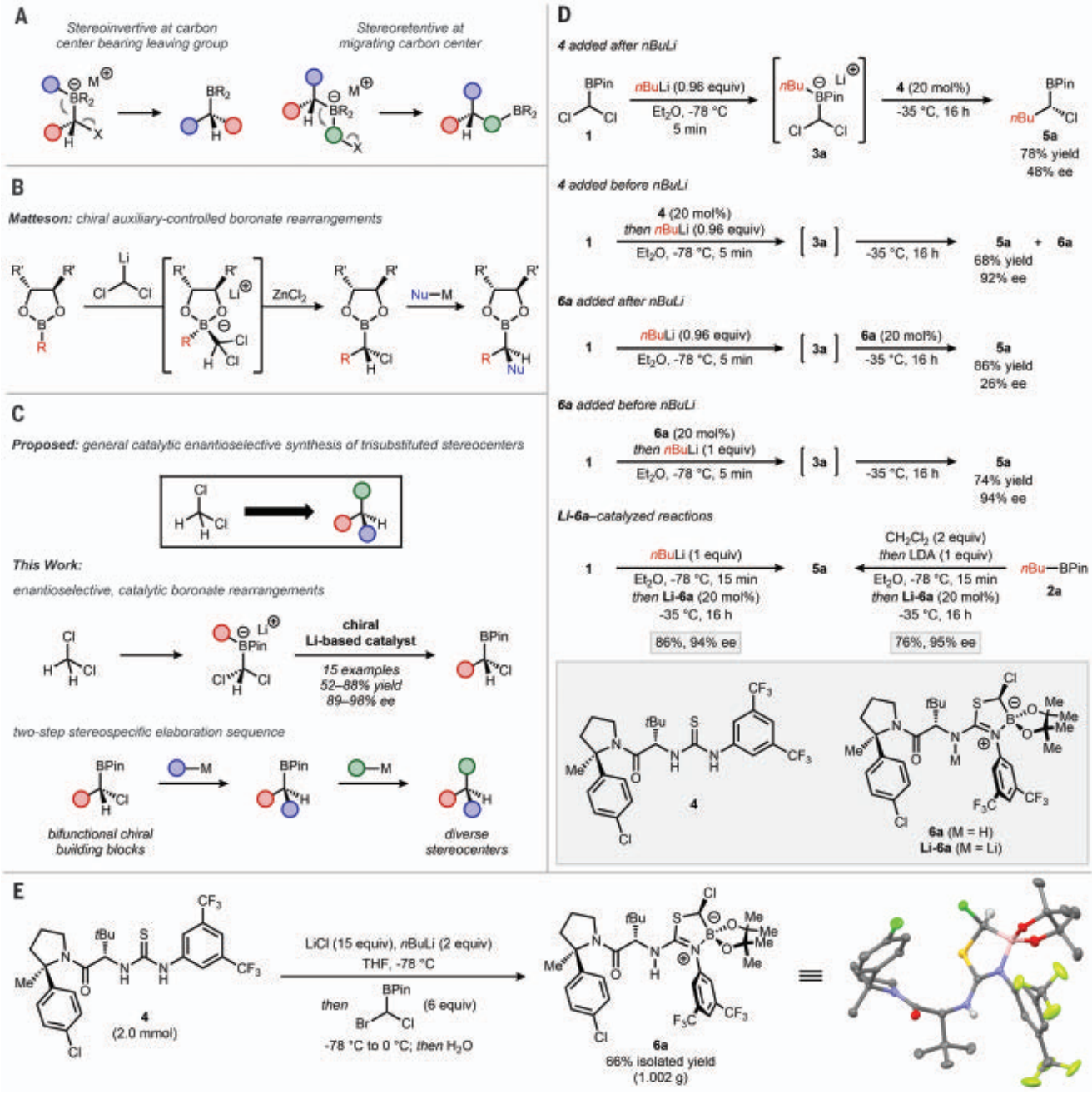



**Fig. 1. Concepts for controlling 1,2-boronate rearrangements and reaction optimization.** (**A**) Stereochemical outcomes of 1,2-boronate rearrangements. M, metal; X, leaving group. (**B**) Diastereoselective Matteson homologation using a chiral diol auxiliary on boron. Nu, nucleophile. (**C**) Proposed general three-step strategy for the construction of tertiary stereocenters from boronic esters, dichloromethane, and organometallic nucleophiles. Pin, pinacolato. (**D**) Reaction optimization and discovery of lithium-isothiourea-boronate catalyst **Li-6a**. Reactions were run at 0.125-mmol scale. Yield values reflect product quantification by gas chromatography analysis relative to an internal standard. *n*Bu and *t*Bu, normal and *tert*-butyl; Et, ethyl; Me, methyl. (**E**) Gram-scale synthesis and x-ray crystal structure of isothiourea-boronate precatalyst **6a**. Hydrogen atoms have been removed for presentation.

enantioselective 1,2-boronate rearrangements using an unexpected catalyst, and its application to the synthesis of molecules featuring a diverse range of trisubstituted stereocenters (Fig. 1C).

Numerous approaches toward the synthesis of stereodefined alkylboronic esters have leveraged rearrangement strategies, drawing from the broad generality of 1,2-boronate rearrangements toward various classes of migrating groups, leaving groups, and boron ligands. In addition to Matteson's seminal work using chiral diol auxiliaries, Aggarwal and co-workers have developed stoichiometric chiral lithium carbenoid reagents that enable the enantioselective synthesis of pinacol boronic ester products via stereospecific formation and rearrangement of chiral boronate derivatives (*12*–*14*). Morken and others have reported robust, highly enantioselective catalytic rearrangements of prochiral alkenylboronates

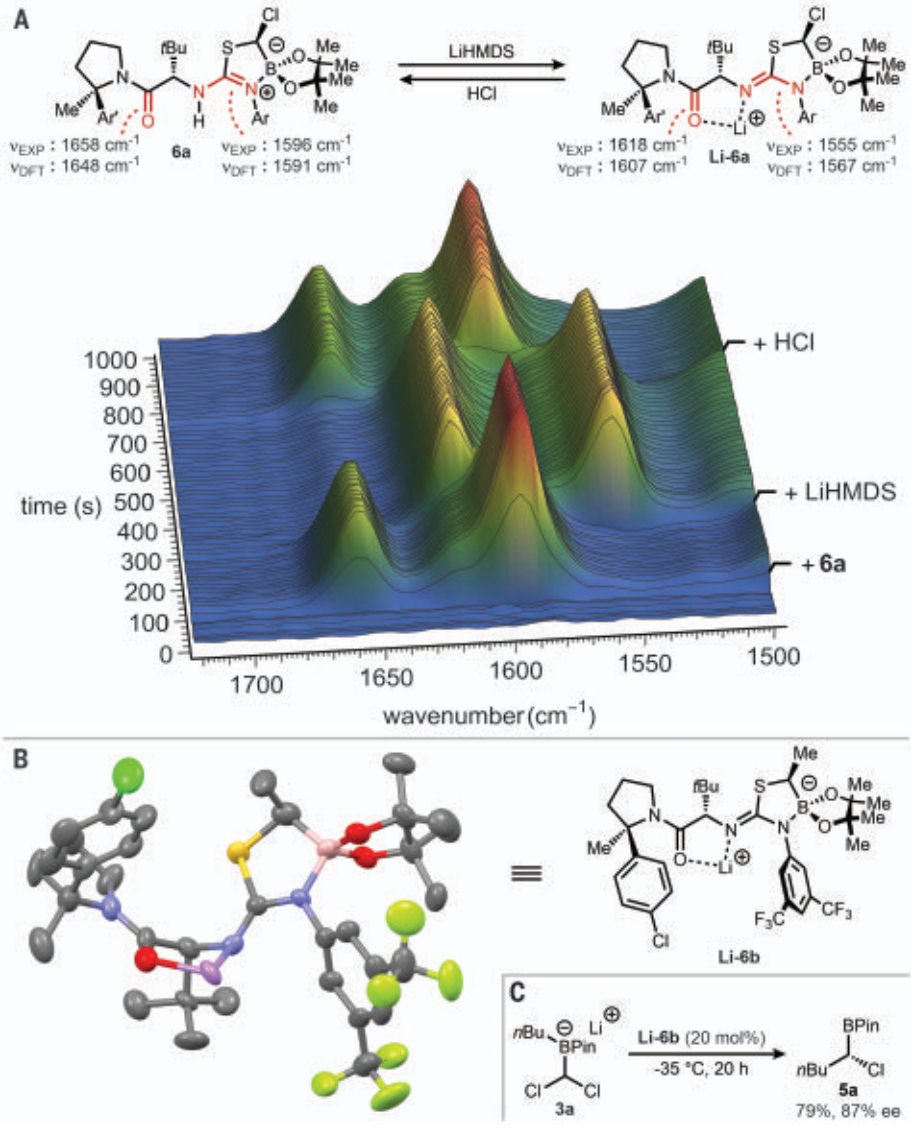



**Fig. 2. Structural characterization of lithium-isothiourea-boronate catalysts.** (**A**) ReactIR study of the deprotonation and reprotonation of **6a**, consistent with the formation of an amide-lithium-isothiourea chelate. Predicted IR stretching frequencies were calculated from structures optimized at the PCM($Et_2O$)-M06-2X/6-31+G(d,p) level of theory, with explicit solvation of the lithium cation in **Li-6a** by dimethyl ether. See supplementary materials for computational details. (**B**) X-ray crystal structure of **Li-6b** confirming the presence of the proposed amide-lithium-isothiourea chelate in the solid state. Hydrogen atoms and a molecule of pinacol coordinated to Li have been removed for presentation. (**C**) Assessment of the performance of catalyst analog **Li-6b** in the rearrangement of lithium boronate **3a**. Yield value refers to gas chromatography yield relative to internal standard.

that provide access to a variety of boronic ester products (*15*–*19*). A distinct approach that does not proceed via a boronate rearrangement yet leverages a similar retrosynthetic disconnection to the sequence proposed here has been developed by Fu and co-workers, involving the enantioconvergent nickel-catalyzed cross-coupling of racemic α-halo boronic ester derivatives with alkylzinc nucleophiles (*20*, *21*). In Fu's method, both substituents geminal to the boronic ester in the product are defined in the enantiodetermining reaction; as a result, the products accessible by this strategy are dictated by the intersecting scope of both organozinc reagents and α-halo boronic esters that can be effectively engaged by the catalyst. In contrast, the sequence proposed here would install only a single substituent in the enantioselective catalytic reaction, followed by the enantiospecific introduction of the second substituent with a scope not dictated by the chiral catalyst. The closest efforts toward an enantioselective, catalytic variant of the Matteson homologation specifically are those reported by Jadhav and Man in the chiral Lewis acid–mediated rearrangement of a dichloromethyl-substituted *n*-butyl boronate, wherein up to 88% enantiomeric excess (ee) was obtained using superstoichiometric quantities of a chiral bis(oxazoline) ligand (*22*, *23*).

We hypothesized initially that the rearrangement of prochiral dichloromethyl boronates might be enabled by anion-abstraction catalysis. In particular, weakly Brønsted acidic dual–hydrogen bond donor catalysts have been applied broadly and successfully to promote stereoselective nucleophilic substitution reactions via leaving group–abstraction pathways (*24*–*26*). The rearrangement of lithium boronate substrate **3a**, prepared by addition of *n*-butyllithium to dichloromethyl boronic acid pinacol ester (**1**), was selected as a model reaction (Fig. 1D). The arylpyrrolidine-*tert*-leucine–derived thiourea **4** was observed to promote the formation of the α-chloro boronic ester product **5a** with modest levels of enantioselectivity (48% ee) when added to a preformed solution of **3a**. However, the same reaction proceeded with up to 92% ee when thiourea **4** was present in the reaction mixture during the generation of **3a** from **1** and *n*-butyllithium. Under these strongly basic conditions, thiourea **4** was found to undergo deprotonation followed by *S*-alkylation by displacement of a chloride on the starting material **1** or product **5a** to generate a mixture of isolable isothiourea-boronate adducts (see supplementary materials for additional details) (*27*, *28*).

The major adduct derived from **4** isolated from the highly enantioselective "**4** added before *n*BuLi" reaction conditions was isothiourea-boronate derivative **6a**. This species promoted the rearrangement reaction in only 26% ee when added to a preformed solution of boronate **3a**, but in 94% ee when introduced to the reaction mixture before *n*-butyllithium. We hypothesized that the active, highly enantioselective catalyst might be **Li-6a**, generated cleanly under the latter conditions by deprotonation of **6a** by *n*-butyllithium but only in small quantities in the presence of preformed boronate **3a** alone. Indeed, the lithiated isothiourea-boronate species **Li-6a** generated quantitatively by mixing **6a** and lithium hexamethyldisilazide (LiHMDS) in a 1:1 ratio proved to be a highly effective catalyst for the model reaction. Thus, rearrangement of **3a** prepared either by addition of *n*-butyllithium to **1** or by addition of dichloromethyllithium [generated in situ via deprotonation of dichloromethane by lithium diisopropylamide (LDA)] to *n*-butyl boronic ester **2a** was promoted by **Li-6a** in 94 to 95% ee.

Bench-stable isothiourea-boronate precatalyst **6a** was synthesized independently on

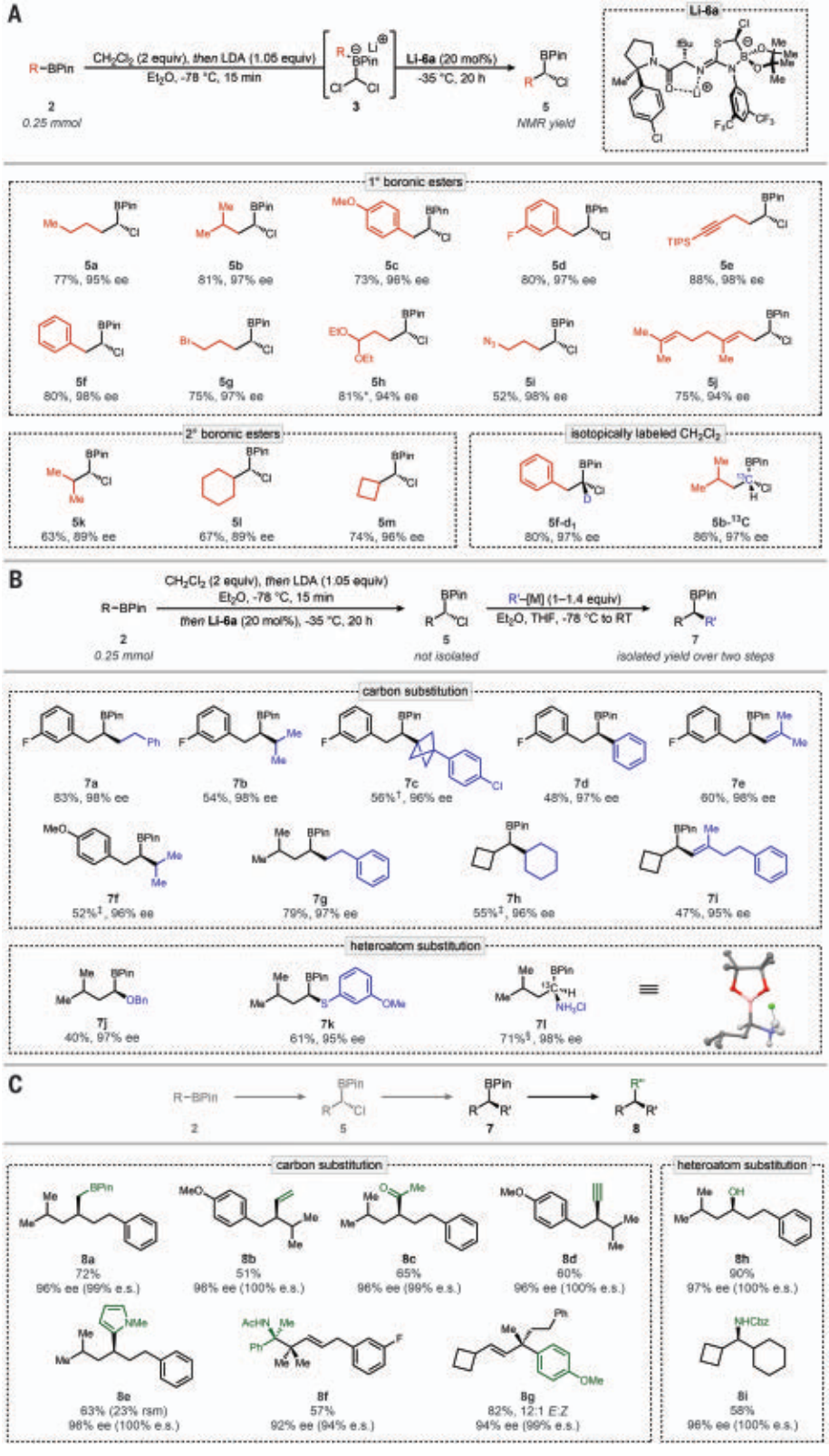



**Fig. 3. Synthetic application.** (**A**) Substrate scope of rearrangement catalyzed by **Li-6a**. Yield values refer to NMR yields relative to internal standard. The absolute configuration of rearrangement products **5** was determined by x-ray crystallographic analysis of **7l** and of the secondary boronic ester generated by addition of phenethylmagnesium bromide to **5l** (see supplementary materials). (**B**) Telescoped carbon-chlorine bond elaborations of rearrangement products with organometallic nucleophiles. Yield values refer to isolated yields over two steps. (**C**) Carbon-boron bond elaborations of secondary boronic ester products. Yield and ee values refer to isolated yields over one step and enantiomeric excess of products **8**, respectively; e.s. is the reaction enantiospecificity based on the ee observed for the products **7** from the catalytic reaction; rsm is recovered starting material. See supplementary materials for reaction conditions. Ph, phenyl; Cbz, benzyloxycarbonyl. *Lithium boronate substrate prepared with additional equivalents of $CH_2Cl_2$. †Yield refers to NMR yield. ‡Reactions carried out at 0.5-mmol scale. §Yield refers to isolated yield over three steps.

gram scale via kinetic resolution of excess racemic bromochloromethyl boronic acid pinacol ester by deprotonated thiourea **4** to enable further synthetic and mechanistic studies (Fig. 1E). Upon lithiation of precatalyst **6a** with LiHMDS, shifts were observed in the infrared (IR) frequencies of both the isothiourea N–C–N stretch (1596 $cm^{-1}$ to 1555 $cm^{-1}$, $\delta = -41\ cm^{-1}$) and the amide C–O stretch (1658 $cm^{-1}$ to 1618 $cm^{-1}$, $\delta = -40\ cm^{-1}$), indicative of perturbations to both functional groups (Fig. 2A). These shifts were fully reversible by addition of one equivalent of HCl. The observed IR spectral changes were interpreted as being due to deprotonation of

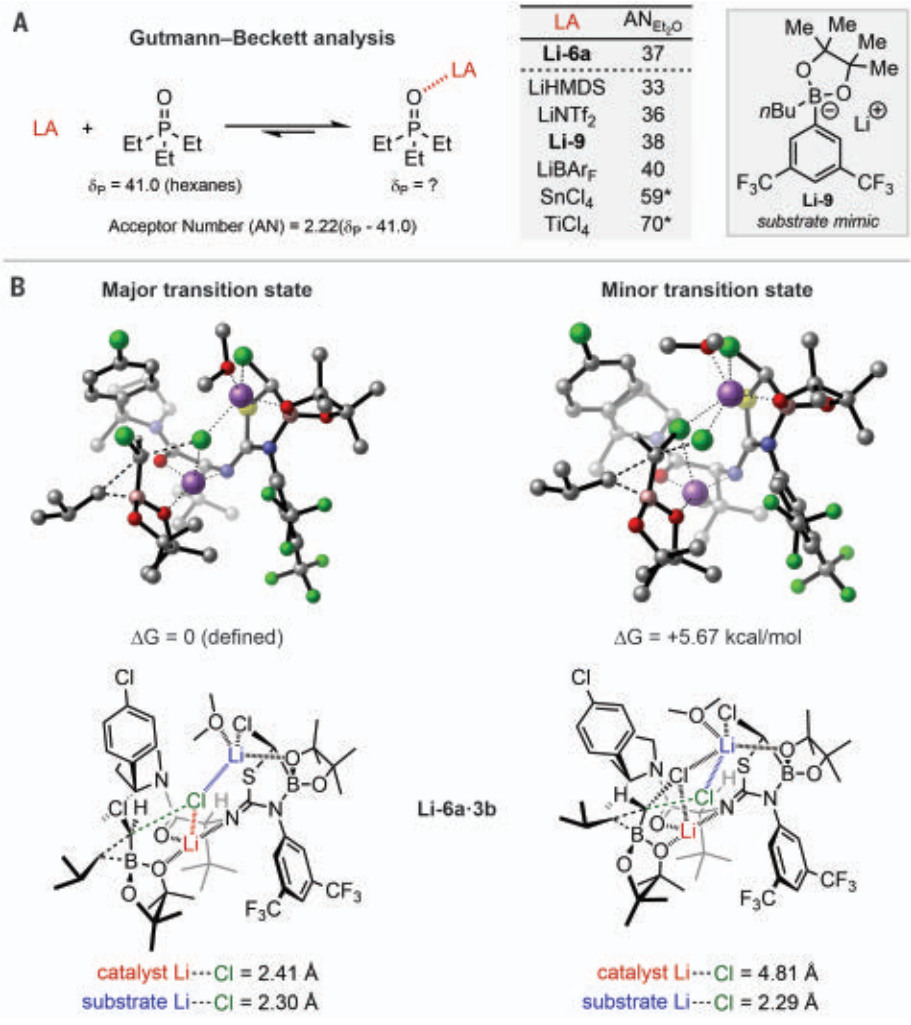



**Fig. 4. Mechanistic analysis and computational study.** (**A**) Gutmann-Beckett analysis of **Li-6a** in comparison to other Lewis acids. The analysis was performed by addition of a Lewis acid to a solution of triethylphosphine oxide in diethyl ether using a solution of triphenylphosphine in $C_6D_6$ in a sealed capillary tube as an internal standard, with the acceptor number (AN) determined from the deshielding effect of the additive on the $^{31}P$ NMR resonance. (**B**) DFT computational study of the putative enantiodetermining major and minor transition states for the developed rearrangement reaction proceeding through a chloride-abstraction mechanism. Calculations were carried out at the SMD($Et_2O$)-DLPNO-CCSD(T)/ma-def2-TZVP//M06-2X/6-31G(d) level of theory, with explicit solvation of the lithium cations by dimethyl ether where appropriate. See supplementary materials for computational details. *Literature values (*40*).

the N-H of **6a** by LiHMDS with chelation of the lithium counterion by the exocyclic isothiourea nitrogen and the amide oxygen. In line with experimental observations, density functional theory (DFT) calculations on **6a** and the proposed chelate structure of **Li-6a** predict a similar shift for both the isothiourea N–C–N stretch (1591 $cm^{-1}$ to 1567 $cm^{-1}$, $\delta = -24\ cm^{-1}$) and the amide C–O stretch (1648 $cm^{-1}$ to 1607 $cm^{-1}$, $\delta = -41\ cm^{-1}$). An x-ray crystal structure analysis of methyl analog **Li-6b** provided additional experimental support for the chelate model (Fig. 2B). Complex **Li-6b** displayed spectral characteristics similar to those of **Li-6a** but proved more amenable to crystallization; it also afforded similar but slightly diminished enantioinduction in the rearrangement of the model boronate **3a** (Fig. 2C). Taken together, these observations point to a catalyst structure featuring a five-membered lithium chelate with the cation located in a highly dissymmetric pocket defined by the 4-chlorophenyl group, the *tert*-butyl group, the isothiourea-boronate heterocycle, and the 3,5-bis(trifluoromethyl)phenyl group, the last of which is rotated out of the plane of the heterocyclic core.

With the identification and structural characterization of **Li-6a** as a highly enantioselective catalyst, the scope of migrating groups for the catalytic rearrangement was explored (Fig. 3A). Excellent levels of enantioselectivity were achieved in rearrangements of boronates accessed by addition of dichloromethyllithium to readily available primary alkyl (**5a–j**) and unhindered secondary alkyl (**5k–m**) pinacol boronic ester derivatives bearing a variety of substitution and branching patterns. Compounds bearing mildly Lewis basic (**5c**,**h**,**i**) or base-sensitive (**5e**,**g**) functionality underwent the rearrangement effectively. Aryl or alkenyl pinacol boronic esters, substrates bearing more strongly Lewis basic substitution, and tertiary alkylboronates either failed to undergo rearrangement or did so with variable levels of enantioselectivity (see supplementary materials for additional information and discussion regarding the substrate scope). Compounds bearing isotopically labeled stereocenters could be readily generated by using commercially available $CD_2Cl_2$ (**5f-d$_1$**) and $^{13}CH_2Cl_2$ (**5b-$^{13}$C**).

As noted above, the products of rearrangement of dichloromethyl boronates are bifunctional building blocks that can be elaborated sequentially and stereospecifically to a broad array of stereodefined structural motifs (Fig. 3B). To illustrate the synthetic utility of the enantioselective catalytic rearrangement reaction in the preparation of such compounds, we developed a two-step, telescoped procedure entailing the **Li-6a**–catalyzed rearrangement followed by a second, enantiospecific rearrangement to afford enantioenriched secondary boronic ester products directly. Grignard reagents bearing various substitution patterns were engaged in the second step to generate boronic ester products bearing primary alkyl (**7a**,**g**), secondary alkyl (**7b**,**f**,**h**), cyclic tertiary alkyl (**7c**), aryl (**7d**), and alkenyl (**7e**,**i**) substitution in good yields over two steps and with complete enantiospecificity. Addition of heteroatomic nucleophiles to enantioenriched chloride building block **5b** afforded the corresponding ether (**7j**) and thioether (**7k**) products with high levels of enantiospecificity. Similarly, **5b-$^{13}$C** was elaborated to the α-boryl amine (**7l**) as the basis for an efficient route to an isotopically labeled analog of the proteasome inhibitor drug bortezomib (*29*).

The purified secondary boronic ester products obtained as outlined above were further elaborated to compounds bearing a diverse array of trisubstituted stereocenters using precedented carbon-boron bond transformations (Fig. 3C) (*30*–*37*). Boronic ester intermediates **7f–h** were thus transformed directly into products bearing a broad range of C–C, C–N, and C–O bonds in high yields and enantiospecificities (**8a–e**,**h**,**i**). Additionally, allylboronic ester intermediates **7e** and **7i** were elaborated selectively at the γ-position to afford fully substituted products bearing an α-tertiary amine (**8f**) and a quaternary stereocenter (**8g**),

respectively. In this manner, readily available aliphatic boronic esters could be converted through a three-step sequence into a variety of highly enantioenriched products bearing stereocenters with diverse substitution patterns.

The effectiveness of lithium-isothiourea-boronate **Li-6a** in enantioselective 1,2-boronate rearrangements raises intriguing questions about the mode of catalysis and origin of enantioinduction. Given the absence of any potential H-bond donor groups on the catalyst, the Lewis acidic lithium center represents the logical agent for chloride abstraction. Quantification by Gutmann-Beckett analysis (*38*–*40*) revealed that **Li-6a** possesses only moderate Lewis acidity, similar to that of other lithium salts and substantially lower than that of traditional Lewis acid catalysts such as $TiCl_4$ (Fig. 4A). Notably, lithium boronate **Li-9**, which was designed to serve as a substrate mimic that cannot undergo rearrangement, displayed a slightly higher acceptor number than that obtained with **Li-6a**. Given the observed levels of enantioselectivity under catalytic conditions, **Li-6a** must be more reactive than the lithium boronate substrate ***3*** itself in promoting the rearrangement. Although these observations do not rule out **Li-6a** acting as a chiral Lewis acid, the similar degree of Lewis acidity exhibited by lithium boronate substrate mimic **Li-9** precludes Lewis acidity alone as the only factor responsible for rate acceleration in the catalyst-mediated rearrangement.

Valuable insight into the basis for the high reactivity of **Li-6a** in mediating the boronate rearrangement reaction was provided by DFT computational studies. In the initial investigations, a substrate:catalyst stoichiometry of 1:1 in the enantiodetermining transition states was assumed. The lowest-energy transition state structures reveal networks of attractive interactions involving the lithium ions and the boronate moieties of both the substrate and the catalyst. In the putative enantiodetermining transition state leading to the major enantiomer in the **Li-6a**–catalyzed rearrangement of isobutyl-substituted lithium boronate ***3b*** (Fig. 4B), the anterior boronate oxygen and α-boryl chloride of **Li-6a** chelate the lithium ion of ***3b*** (depicted in blue in the graphic) while the posterior boronate oxygen of substrate ***3b*** is anchored by the lithium ion of the catalyst (depicted in red). In this geometry, both lithium ions are poised to promote abstraction of the departing chloride (shown in green) by a cooperative mechanism, with the Li–Cl bond distances for the catalyst and substrate lithium ions being 2.41 Å and 2.30 Å, respectively. In contrast, in the lowest-energy transition state structure leading to the minor product enantiomer, the departing chloride resides proximal to the substrate lithium (2.29 Å) but distant from the catalyst lithium (4.81 Å); additionally, the substrate and catalyst lithium ions are both bound to the spectator chloride of the substrate. The modestly lower level of enantioselectivity observed in the rearrangement mediated by catalyst analog **Li-6b** (Fig. 2C) is consistent with the proposed attractive interaction between the α-boryl chloride of **Li-6a** and the substrate lithium ion being beneficial but not obligatory in the enantiodetermining event. Although a deeper analysis must await fuller characterization of the aggregation state and relative stoichiometry of the product-determining transition state, these results suggest a role for both the catalyst and substrate lithium ions in promoting chloride departure. Taken together, the nuclear magnetic resonance (NMR) and DFT studies are consistent with lithium-isothiourea-boronate species **Li-6a** promoting the rearrangement reaction through a dual–lithium-induced chloride abstraction, aided by Lewis basic functionality on the rigid catalyst scaffold that orients both the catalyst and substrate lithium ions in a precise geometry. This analysis raises the intriguing possibility that this class of structurally well-defined lithium-based catalysts may mediate enantioselective reactions of different classes of lithium boronates and, more broadly, other alkali metal salts.

## REFERENCES AND NOTES


1. L. Xu *et al*., *Advances in Organoboron Chemistry Towards Organic Synthesis* (Thieme, 2020).
2. E. Fernández, A. Whiting, Eds., *Synthesis and Application of Organoboron Compounds* (Springer, 2015).
3. N. Miyaura, Y. Yamamoto, in *Comprehensive Organometallic Chemistry III*, D. M. P. Mingos, R. H. Crabtree, Eds. (Elsevier, 2007), pp. 145–244.
4. D. G. Hall, *Boronic Acids: Preparation and Applications in Organic Synthesis, Medicine and Materials* (Wiley, 2012).
5. C. Sandford, V. K. Aggarwal, *Chem. Commun.* **53**, 5481–5494 (2017).
6. F. A. Carey, R. J. Sundberg, in *Advanced Organic Chemistry: Part B: Reactions and Synthesis*, F. A. Carey, R. J. Sundberg, Eds. (Springer, 2007), pp. 783–860.
7. T. Chinnusamy, K. Feeney, C. G. Watson, D. Leonori, V. K. Aggarwal, in *Comprehensive Organic Synthesis II*, P. Knochel, Ed. (Elsevier, ed. 2, 2014), pp. 692–718.
8. D. S. Matteson, *J. Org. Chem.* **78**, 10009–10023 (2013).
9. D. S. Matteson, *Tetrahedron* **54**, 10555–10607 (1998).
10. D. S. Matteson, R. Ray, *J. Am. Chem. Soc.* **102**, 7590–7591 (1980).
11. I. Gazić Smilović *et al*., *Angew. Chem. Int. Ed.* **51**, 1014–1018 (2012).
12. J. L. Stymiest, G. Dutheuil, A. Mahmood, V. K. Aggarwal, *Angew. Chem. Int. Ed.* **46**, 7491–7494 (2007).
13. J. L. Stymiest, V. Bagutski, R. M. French, V. K. Aggarwal, *Nature* **456**, 778–782 (2008).
14. E. Beckmann, V. Desai, D. Hoppe, *Synlett* **2004**, 2275–2280 (2004).
15. L. Zhang *et al*., *Science* **351**, 70–74 (2016).
16. M. Chierchia, C. Law, J. P. Morken, *Angew. Chem. Int. Ed.* **56**, 11870–11874 (2017).
17. S. Namirembe, J. P. Morken, *Chem. Soc. Rev.* **48**, 3464–3474 (2019).
18. Z. Tao, K. A. Robb, J. L. Panger, S. E. Denmark, *J. Am. Chem. Soc.* **140**, 15621–15625 (2018).
19. S. Panda, J. M. Ready, *J. Am. Chem. Soc.* **139**, 6038–6041 (2017).
20. J. Schmidt, J. Choi, A. T. Liu, M. Slusarczyk, G. C. Fu, *Science* **354**, 1265–1269 (2016).
21. S.-Z. Sun, R. Martin, *Angew. Chem. Int. Ed.* **57**, 3622–3625 (2018).
22. P. K. Jadhav, H.-W. Man, *J. Am. Chem. Soc.* **119**, 846–847 (1997).
23. K. Smith, B. A. Saleh, M. B. Alshammari, G. A. El-Hiti, M. C. Elliott, *Org. Biomol. Chem.* **19**, 4279–4284 (2021).
24. A. J. Bendelsmith, S. C. Kim, M. Wasa, S. P. Roche, E. N. Jacobsen, *J. Am. Chem. Soc.* **141**, 11414–11419 (2019).
25. D. A. Kutateladze, D. A. Strassfeld, E. N. Jacobsen, *J. Am. Chem. Soc.* **142**, 6951–6956 (2020).
26. Y. Park *et al*., *Science* **355**, 162–166 (2017).
27. M. Biedrzycki, W. H. Scouten, Z. Biedrzycka, *J. Organomet. Chem.* **431**, 255–270 (1992).
28. D. S. Matteson, G. D. Schaumberg, *J. Org. Chem.* **31**, 726–731 (1966).
29. S. P. A. Hinkes, C. D. P. Klein, *Org. Lett.* **21**, 3048–3052 (2019).
30. R. P. Sonawane *et al*., *Angew. Chem. Int. Ed.* **50**, 3760–3763 (2011).
31. A. P. Pulis, D. J. Blair, E. Torres, V. K. Aggarwal, *J. Am. Chem. Soc.* **135**, 16054–16057 (2013).
32. Y. Wang, A. Noble, E. L. Myers, V. K. Aggarwal, *Angew. Chem. Int. Ed.* **55**, 4270–4274 (2016).
33. M. Odachowski *et al*., *J. Am. Chem. Soc.* **138**, 9521–9532 (2016).
34. J. L.-Y. Chen, V. K. Aggarwal, *Angew. Chem. Int. Ed.* **53**, 10992–10996 (2014).
35. B. Potter, E. K. Edelstein, J. P. Morken, *Org. Lett.* **18**, 3286–3289 (2016).
36. E. K. Edelstein, A. C. Grote, M. D. Palkowitz, J. P. Morken, *Synlett* **29**, 1749–1752 (2018).
37. S. Yu, C. Jing, A. Noble, V. K. Aggarwal, *Angew. Chem. Int. Ed.* **59**, 3917–3921 (2020).
38. U. Mayer, V. Gutmann, W. Gerger, *Monatsh. Chem.* **106**, 1235–1257 (1975).
39. M. A. Beckett, G. C. Strickland, J. R. Holland, K. Sukumar Varma, *Polymer* **37**, 4629–4631 (1996).
40. M. A. Beckett, D. S. Brassington, S. J. Coles, M. B. Hursthouse, *Inorg. Chem. Commun.* **3**, 530–533 (2000).


## ACKNOWLEDGMENTS

We thank S.-L. Zheng (Harvard University) for determination of x-ray crystal structures, and M. Levin and P. Vojáčková for helpful discussions. **Funding:** NIH grant GM043214 (E.N.J.). **Author contributions:** H.A.S. and E.N.J. conceived the work, H.A.S. and J.Z.E. designed and conducted the experiments, E.N.J. directed the research, and all authors wrote the manuscript. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** Crystallographic data for compounds **4**, **6a**, **6b**, **Li-6b**, **6f**, **7l**, and **7v** are available free of charge from the Cambridge Crystallographic Data Centre under references CCDC 2111740, CCDC 2111743, CCDC 2111741, CCDC 2111739, CCDC 2111744, CCDC 2111742, and CCDC 2111745, respectively. All other data are available in the main text or the supplementary materials.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm0386
Materials and Methods
Supplementary Text
Figs. S1 to S20
Tables S1 to S25
References (*41*–*69*)

21 August 2020; accepted 29 September 2021
10.1126/science.abm0386

AIR POLLUTION

# Abating ammonia is more cost-effective than nitrogen oxides for mitigating $PM_{2.5}$ air pollution

Baojing Gu[1,2]*, Lin Zhang[3]*, Rita Van Dingenen[4], Massimo Vieno[5], Hans JM Van Grinsven[6], Xiuming Zhang[7], Shaohui Zhang[8,9], Youfan Chen[3], Sitong Wang[1], Chenchen Ren[10], Shilpa Rao[11], Mike Holland[12], Wilfried Winiwarter[9,13], Deli Chen[7], Jianming Xu[1,2], Mark A. Sutton[5]*

Fine particulate matter ($PM_{2.5}$, particles with a mass median aerodynamic diameter of less than 2.5 micrometers) in the atmosphere is associated with severe negative impacts on human health, and the gases sulfur dioxide, nitrogen oxides, and ammonia are the main $PM_{2.5}$ precursors. However, their contribution to global health impacts has not yet been analyzed. Here, we show that nitrogen accounted for 39% of global $PM_{2.5}$ exposure in 2013, increasing from 30% in 1990 with rising reactive nitrogen emissions and successful controls on sulfur dioxide. Nitrogen emissions to air caused an estimated 23.3 million years of life lost in 2013, corresponding to an annual welfare loss of 420 billion United States dollars for premature death. The marginal abatement cost of ammonia emission is only 10% that of nitrogen oxides emission globally, highlighting the priority for ammonia reduction.

Air pollution from $PM_{2.5}$ (fine particulate matter with a mass median aerodynamic diameter <2.5 μm) has been estimated to cause millions of premature deaths annually in recent years (*1*, *2*). Therefore, mitigating $PM_{2.5}$ pollution is a high priority for environmental protection in many countries such as China (*3*), India (*4*), the United States (*5*), and the member states of the European Union (EU) (*6*). The most cost-effective abatement measures need to be identified to balance environmental protection and economic development. This is particularly relevant for atmospheric emissions of reactive nitrogen ($N_r$), which are generally driven by fossil fuel combustion in power plants and transport and by the production of food and energy (*7*–*9*). Previous quantifications of the health impacts of nitrogen oxides ($NO_x$) and ammonia ($NH_3$) emissions from $PM_{2.5}$ pollution have not been conducted on a global scale because of differences in atmospheric chemistry, population density, and exposure characteristics relative to the pollution consequences of $N_r$ emissions in different countries (*7*, *9*–*11*). In addition, differences in gross domestic product (GDP) and social preferences affect economic values of mortality (*1*). Economic growth and higher living standards increase healthy life expectancy and also increase the willingness to pay for actions that ultimately reduce the health risks of air pollution (*1*). Thus, a generic method is needed to ensure that the costs of mortality from $N_r$ emissions are comparable across countries and regions and reflective of local conditions.

Here, we developed and applied a metric that we call the "N-share" of $PM_{2.5}$ pollution, which is the contribution of $N_r$ compounds to total $PM_{2.5}$ concentration determined by modeling with and without $N_r$ emission. We applied three atmospheric chemistry transport models, EMEP-WRF (*12*), TM5-FASST (*13*), and GEOS-Chem (*14*), which include emissions of sulfur dioxide ($SO_2$), $NO_x$, $NH_3$, volatile organic compounds, and primary $PM_{2.5}$ (see the materials and methods). The N-share is different from the mass fraction of $N_r$ within $PM_{2.5}$ because $N_r$ emissions are not only the precursors of $PM_{2.5}$ but also affect the chemical reactions that lead to $PM_{2.5}$ formation (*15*). By combining the calculated values of N-share with the estimated global burden of disease derived from $PM_{2.5}$ pollution (*1*), we were able to estimate welfare loss associated with $N_r$ emissions. Finally, we applied the GAINS model (*16*) to estimate the implementation costs of $N_r$ air pollution abatement (see the materials and methods), which has the advantage of being able to quantify the abatement cost of each measure in different countries, allowing us to compare the relative welfare gains of reducing different forms of $N_r$ emission. $NO_x$ is an important precursor to ozone that also has a negative impact on human health (*9*), so we also considered ozone formation in the analysis. By contrast, we have not separately analyzed the direct health impacts of $NO_2$.

Although the overall N-share of $PM_{2.5}$ pollution generally increased from 1990 to 2013, we note substantial regional variation, with increases in Asia, South America, and South Africa and a decrease in Europe (Fig. 1 and Table 1). We found that $NH_3$ emission made a larger contribution to $PM_{2.5}$ than $NO_x$ emission globally and in most countries, indicating that $PM_{2.5}$ is more strongly $NH_3$ limited than $NO_x$ limited (*17*, *18*). The N-share caused by $NH_3$ emissions contributed an estimated 25% (range, 20 to 31%) to $PM_{2.5}$ pollution in 1990, increasing to 32% (25 to 39%) in 2013, whereas $NO_x$ emission contributed 17% (14 to 20%) in 1990, increasing to 28% (23 to 33%) in 2013. These changes agree well with the widespread increase of $N_r$ emissions and the decrease of $SO_2$ emission in many areas of the world from 1990 to 2013 (fig. S5) (*19*).

The N-shares of total $N_r$ emissions are much smaller than the sum of N-shares from $NH_3$ and $NO_x$ separately (Fig. 1) because of the interactions between $NH_3$ and $NO_x$ during secondary $PM_{2.5}$ formation. Alkaline $NH_3$ can form aerosols with the acidic products of $NO_x$ and $SO_2$ emission, so that reduction of $NH_3$ emission also tends to reduce the contribution of $NO_x$ and $SO_2$ to $PM_{2.5}$ formation (*15*, *18*, *20*), which also explains the larger N-share of $PM_{2.5}$ pollution arising from $NH_3$ compared with $NO_x$. In this study, we treated the health effects of $NH_3$-derived $PM_{2.5}$ the same as we did those of $NO_x$-derived $PM_{2.5}$ because evidence of differential harm between chemical species is lacking (see the supplementary materials) (*21*).

For the period 1990–2013, we estimate that total years of life lost (YLL) caused by $PM_{2.5}$ pollution derived from $N_r$ emissions increased from 19.5 to 23.3 million globally, with the $NH_3$ and $NO_x$ contributions increasing from 16.3 to 19.3 and 11.4 to 16.2 million, respectively (table S1 and fig. S6). We used YLL to quantify premature mortality because this metric is considered robust and versatile in capturing health impacts (*22*). Every premature death represents an individual story, with variation between age at death and expected lifespan, which can be well captured by the YLL (*23*). When expressed as YLL per gigogram of N (where 1 Gg = $10^9$ g) emission (YLL/N), higher values were found in developing or transition economies such as Asia, East Europe, and some African countries, which have high $PM_{2.5}$ pollution and lower GDP per capita (Fig. 2 and fig. S6). Higher $PM_{2.5}$ pollution increases the YLL, whereas lower GDP per capita normally indicates lower access to and quality of medical care (table S2) (*24*), as well as greater insecurity in access to food and water that can also increase the YLL/N. By contrast, lower YLL/N is found in

[1]College of Environmental and Resource Sciences, Zhejiang University, Hangzhou 310058, China. [2]Zhejiang Provincial Key Laboratory of Agricultural Resources and Environment, Zhejiang University, Hangzhou 310058, China. [3]Laboratory for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China. [4]European Commission, Joint Research Centre, 21027 Ispra (VA), Italy. [5]UK Centre for Ecology & Hydrology, Edinburgh Research Station, Bush Estate, Penicuik, Midlothian EH26 0QB, UK. [6]PBL Netherlands Environmental Assessment Agency, 2500 GH The Hague, Netherlands. [7]School of Agriculture and Food, The University of Melbourne, Melbourne, Victoria 3010, Australia. [8]School of Economics and Management, Beihang University, 100091 Beijing, China. [9]International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria. [10]Department of Land Management, Zhejiang University, Hangzhou 310058, China. [11]Norwegian Institute of Public Health, N-0213 Oslo, Norway. [12]Ecometrics Research and Consulting, Reading RG8 7PW, UK. [13]Institute of Environmental Engineering, University of Zielona Góra, PL 65-417 Zielona Góra, Poland.
*Corresponding author. Email: bjgu@zju.edu.cn (B.G.); zhanglg@pku.edu.cn (L.Z.); ms@ceh.ac.uk (M.A.S.)

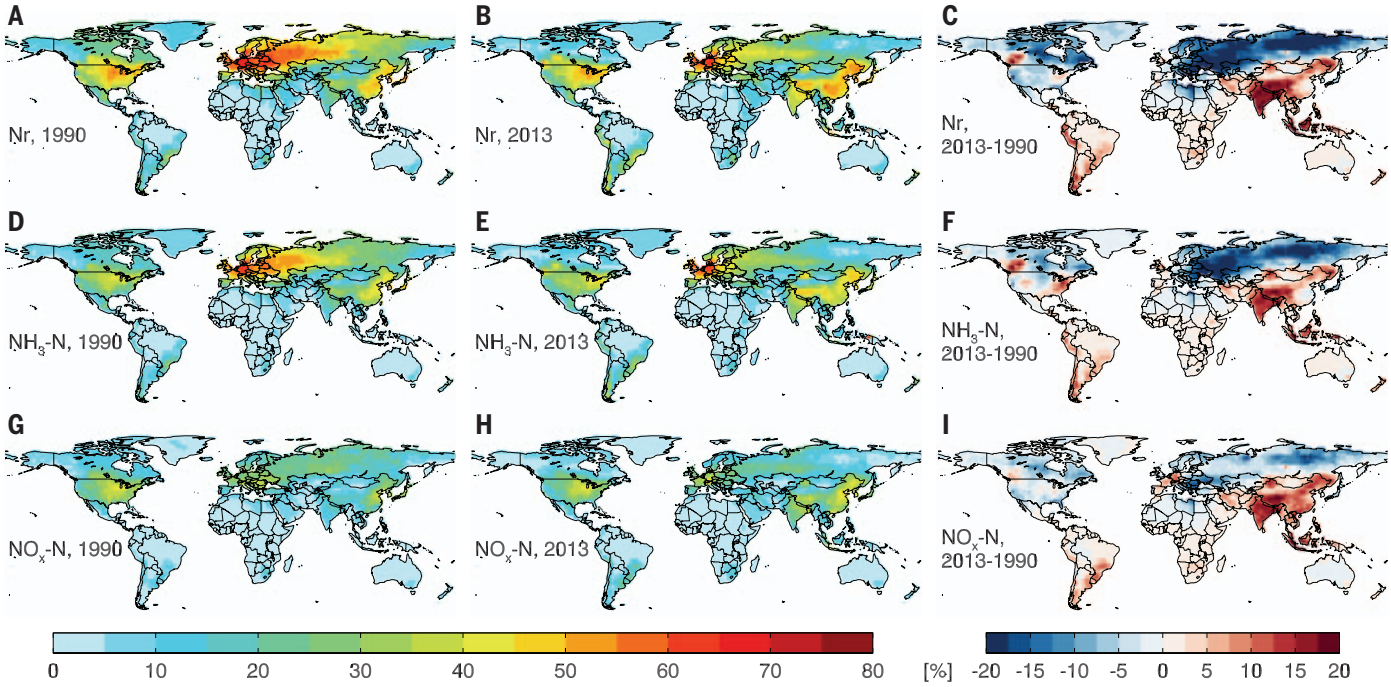



**Fig. 1. N-shares of $PM_{2.5}$ pollution and their changes between 1990 and 2013.** Values represent the percentage contribution of each form of $N_r$ emission to $PM_{2.5}$ pollution. (**A**) Total $N_r$-share in 1990. (**B**) Total $N_r$-share in 2013. (**C**) Change in total $N_r$-share between 2013 and 1990. (**D**) $NH_3$-N share in 1990. (**E**) $NH_3$-N share in 2013. (**F**) Change in $NH_3$-N share between 2013 and 1990. (**G**) $NO_x$-N share in 1990. (**H**) $NO_x$-N share in 2013. (**I**) Change in $NO_x$-N share between 2013 and 1990. (**J**) Change in $NO_x$-N share between 2013 and 1990. Results are based on the average value of simulations from GEOS-Chem and EMEP-WRF models using 100% $N_r$ emission reduction with uncertainty analysis from figs. S2 to S5. Base map is applied without endorsement from Natural Earth (https://www.naturalearthdata.com/).

**Table 1. Estimated mortality cost caused by $PM_{2.5}$ air pollution from total $N_r$ emissions**

| | $N_r$ shares in $PM_{2.5}$ formation (%) | | YLL attributable to $PM_{2.5}$ from $N_r$ emission (million years) | | Total mortality cost of $N_r$ in $PM_{2.5}$ pollution (billion USD) | | Marginal mortality cost (USD/kg N) | |
|---|---|---|---|---|---|---|---|---|
| | 1990 | 2013 | 1990 | 2013 | 1990 | 2013 | 1990 | 2013 |
| Africa | 5.9 | 7.3 | 0.9 | 0.9 | 6.3 | 9.3 | 1.3 | 1.1 |
| Asia | 31.0 | 42.3 | 13.4 | 18.1 | 49.4 | 212.4 | 1.6 | 4.1 |
| Europe | 57.9 | 47.4 | 3.8 | 2.9 | 148.6 | 124.6 | 8.4 | 13.1 |
| Latin America | 20.9 | 23.5 | 0.6 | 0.7 | 8.6 | 19.0 | 1.5 | 2.1 |
| North America | 46.3 | 44.2 | 0.8 | 0.8 | 47.6 | 53.6 | 4.3 | 6.8 |
| Oceania | 7.8 | 11.2 | 0.01 | 0.01 | 0.5 | 0.6 | 0.6 | 0.5 |
| World | 30.2 | 38.6 | 19.5 | 23.3 | 261.1 | 419.5 | 3.6 | 4.8 |

America, western Europe, and Oceania (fig. S6). In these areas, populations are generally exposed to lower levels of $PM_{2.5}$ pollution compared with Asia, and higher GDP per capita is associated with access to better medical care (table S2).

Quantifying the global welfare loss per YLL [in 2011 U.S. dollars (USD) converted at purchasing power parity rates], we found that the willingness to pay to reduce the risk of a YLL caused by $N_r$ emission represented a total cost of 261 (range, 156 to 398) billion USD in 1990, increasing by 60% to 420 (247 to 640) billion USD in 2013 (Table 1). Of this, by far the largest regional change was estimated for Asia (+330%, mainly linked to increased $N_r$ emissions), which may be contrasted with reduction in Europe (–16%, mainly linked to $NO_x$ control). Ammonia contributes most to the estimated total cost of mortality from $N_r$ emissions both because it is found in large amounts (table S1) and because it is the limiting (alkaline) compound for aerosol formation (*17*, *18*).

Globally, the average marginal cost of premature mortality caused by $N_r$ emission in 2013 was estimated at 4.8 (range, 2.8 to 7.3) USD/kg $N_r$ emission, compared with 3.6 (2.2 to 5.5) USD/kg $N_r$ emission in 1990 (Table 1, Fig. 2, and fig. S7). This 31% increase can be mainly attributed to increase of $N_r$ air pollution and higher costs per YLL caused by income growth. The marginal cost from $NH_3$ emission was ~44% higher than that of total $N_r$ emission (table S1), affecting the overall pattern of $N_r$ values (Fig. 2).

We found that the total social benefit of reduced mortality from the abatement of $NH_3$ emission was substantially larger than the

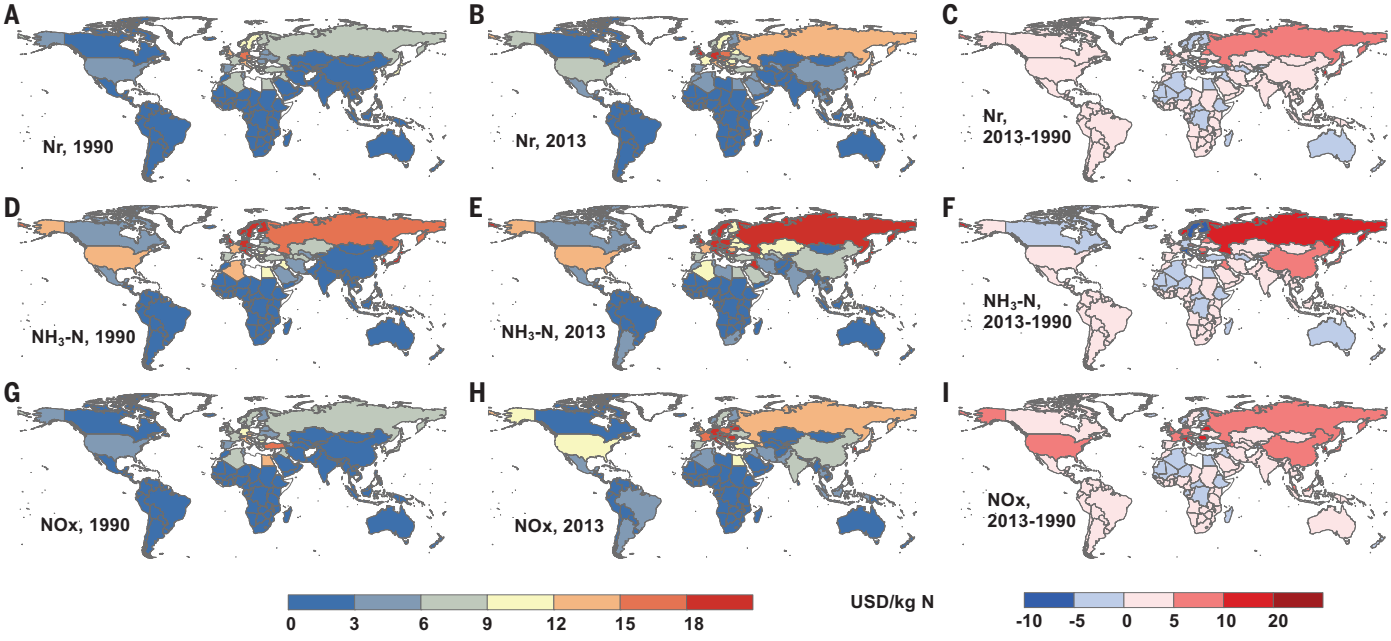



**Fig. 2. Changes in mortality cost per kilogram $N_r$ emission between 1990 and 2013 caused by $PM_{2.5}$ pollution.** Values are based on the control of different $N_r$ components. Shown are control of total $N_r$ emission for 1990 (**A**), 2013 (**B**), and differences between 2013 and 1990 (**C**); control of $NH_3$ emission for 1990 (**D**), 2013 (**E**), and differences between 2013 and 1990 (**F**); and control of $NO_x$ emission for 1990 (**G**), 2013 (**H**), and differences between 2013 and 1990 (**I**). Positive values for (C), (F), and (I) indicate an increase over time. Uncertainties for these figures can be found in fig. S7. White areas of some countries indicate a lack of data. Base map is applied without endorsement from Natural Earth (https://www.naturalearthdata.com/).

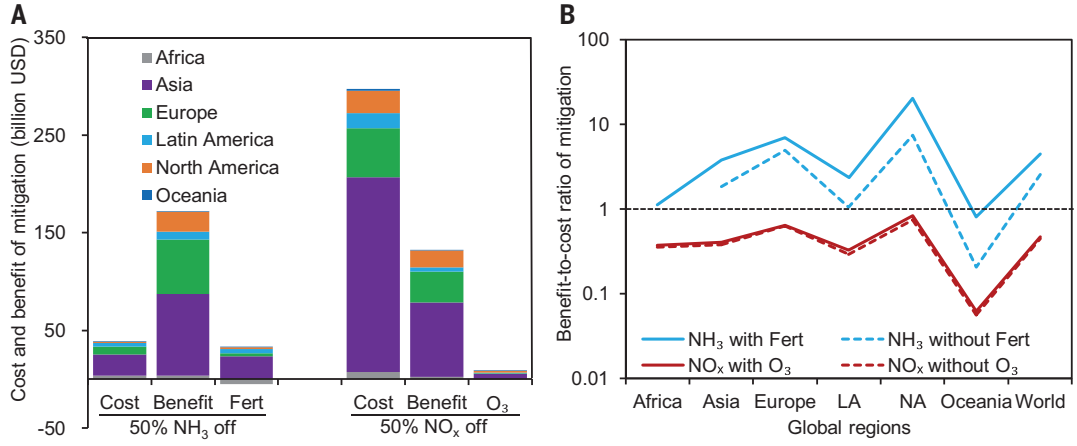



**Fig. 3. Cost and benefit to reduce 50% of $NH_3$ and $NO_x$ emission in 2013.** (**A**) Abatement cost, benefit of prevented mortality from reduced $PM_{2.5}$, fertilizer saving (Fert), and ground-level ozone pollution ($O_3$) mitigation. (**B**) Ratio of benefit of prevented mortality from $PM_{2.5}$ mitigation to cost, with values for $NH_3$ and $NO_x$ shown with and without the benefit from fertilizer saving and $O_3$ mitigation caused by $NO_x$ reduction, respectively. The benefit in (B) only refers to benefit of prevented mortality derived from $PM_{2.5}$ mitigation caused by $NH_3$ and $NO_x$ abatement. LA, Latin America, NA, North America. For Africa, fertilizer savings are negative because of too little $N_r$ input compared with crop need, and the best management practices would increase $N_r$ fertilizer use (see the supplementary materials).

abatement costs. We estimated the global average cost of reducing $NH_3$ emission at 1.5 USD/kg $NH_3$-N (weighted mean of measures) using the GAINS method (see the materials and methods), which is substantially less than the welfare benefit of the associated reduction of mortality at 6.9 (range, 3.8 to 10.9) USD/kg $NH_3$-N globally (Table 1 and fig. S8). By contrast, we estimated the abatement cost of $NO_x$ emission at 16.0 USD/kg $NO_x$-N, which is larger than the welfare benefits of reduced mortality at 7.3 (4.0 to 11.8) USD/kg $NO_x$-N (table S1).

North America has the largest benefit-to-cost ratio for $NH_3$ mitigation, followed by Europe and Asia, suggesting reduction of $NH_3$ emission as a favorable option to increase social benefit (Fig. 3B). The benefit-to-cost ratio for $NO_x$ mitigation is also the largest in North America compared with other world regions. However, globally, this ratio is <1 even when including both $PM_{2.5}$ and ozone effects, suggesting a negative net benefit to further reducing $NO_x$ emissions (Fig. 3B). The benefit-to-cost ratios to mitigate $NH_3$ and $NO_x$ emission are both <1 in Oceania (Fig. 3B). This indicates that the costs of $N_r$ mitigation are not justified by the expected regional benefits for $PM_{2.5}$, although other considerations may still justify $N_r$ mitigation in this region [e.g., ecosystem benefits (*25*)]. In Africa, food security is still a challenge, and improved management practices may need to increase nitrogen inputs (whether

by fertilizer or biological nitrogen fixation) and reduce wasteful nitrogen losses to avoid soil degradation (*26*).

The main opportunities for $NH_3$ abatement concern agricultural sources, for which abatement measures are relatively easy and inexpensive. For instance, optimizing nitrogen fertilization will not only abate $NH_3$ emission but will also reduce nitrogen fertilizer use, which can save the implementation cost while offering opportunities for net cost savings (*27*–*29*). For nonagricultural sources of $NH_3$, which account for ~25% of emissions based on the Community Emissions Data System (CEDS) inventory, the reduction is more related to fossil fuel combustion, biomass burning, and waste treatment, which have received much less attention in the past (*30*). If global $NH_3$ emissions (from all sources, including both agricultural and nonagricultural) were to be reduced by 50%, then the total implementation cost is estimated at 38 billion USD (Fig. 3A). This is smaller than the social benefit of prevented mortality (172 billion USD) derived from mitigation of $PM_{2.5}$ pollution for $NH_3$ emission (Fig. 3A). Additionally, this would save ~20% of global nitrogen fertilizer use with a net value of ~28 billion USD (see the materials and methods for the calculation). In practice, this means that many measures to control $NH_3$ emission can have a zero cost implementation or represent a net economic benefit for farmers. This is in addition to the substantial societal co-benefits for natural ecosystems from reducing $N_r$ air pollution (*8*, *31*, *32*), such as improved water quality, biodiversity conservation, and reduced nitrous oxide emissions (*33*).

Agricultural $NH_3$ is distributed over many individual facilities, so control requires action by farmers in many different contexts (*34*). In the United States, agricultural $NH_3$ emission is not well controlled, and it is a cause of damage to both the environment and human health (*18*). For larger farms with access to advanced technologies, it is technically feasible to reduce $NH_3$ emission from agriculture without risk of production loss through measures such as adoption of enhanced-efficiency fertilizers and improved manure management practices (*35*, *36*). With the revised EU National Emission Ceilings directive (adopted in 2016), reduction of $NH_3$ emission will be considered across sectors for the target year 2030 (*37*). Policies in the Netherlands and Denmark have required $NH_3$ abatement, leading to reported emission reductions of 35 to 66% between 1990 and 2011 (*38*). China started to address reduction of $NH_3$ emission from agricultural sources during its 13th Five Year Plan (2016–2020) (*39*), and it has been estimated that agricultural $NH_3$ emission could be reduced by one-third and at a low cost through appropriate policies, such as subsidies for enhanced-efficiency fertilizers and fertilizer application machinery (*40*, *41*). Taking reduction of $NH_3$ emission into consideration for $PM_{2.5}$ pollution control is therefore critical, presenting the opportunity for future legislation to mitigate $NH_3$ emissions at national to global scales (*34*).

The estimated implementation cost (297 billion USD) for reducing $NO_x$ emissions is larger than the benefit of reduced mortality (132 billion USD) derived from reduced $PM_{2.5}$ pollution (fig. S8). However, reduction of $NO_x$ emission can also have other benefits, such as alleviation of ground-level ozone (9 billion USD, estimated based on the similar method with $PM_{2.5}$; see the supplementary materials). The modest 7% additional benefit shows how the $PM_{2.5}$ costs dominate welfare loss derived from $NO_x$ emission (Fig. 3B), whereas reducing both $NO_x$ and $NH_3$ emissions will also benefit terrestrial and aquatic ecosystems (*9*). $NO_x$ emissions have already been reduced in high-income regions such as North America and Europe between 1990 and 2013 (fig. S5), which means that the marginal implementation cost to reduce $NO_x$ emissions further is much higher in these countries because the most cost-effective measures are in place already. For other countries where $NO_x$ emissions are still increasing, implementation costs are smaller than those for Europe and North America, but even these costs are much higher than those of $NH_3$ mitigation (fig. S8 and Table 1). Innovation to recapture $NO_x$ as value-added $N_r$ products rather than the present focus on wasteful destruction to form di-nitrogen ($N_2$) may have considerable future potential to reduce costs (*33*) but is still far from commercial availability. Considering the overall costs and benefits, our analysis highlights the priority for air pollution policies to give increased attention to controlling $NH_3$ emissions, complementing successful policies on $NO_x$ and $SO_2$ (Fig. 3B) as part of implementing the ambition of the Colombo Declaration to "halve nitrogen waste" from all sources globally.

## REFERENCES AND NOTES


1. World Bank Institute for Health Metrics and Evaluation, "The cost of air pollution: Strengthening the economic case for action" (World Bank, 2016).
2. J. Lelieveld, J. S. Evans, M. Fnais, D. Giannadaki, A. Pozzer, *Nature* **525**, 367–371 (2015).
3. M. Zheng, C. Yan, T. Zhu, *Philos. Trans. R. Soc. London Ser. A* **378**, 20190325 (2020).
4. G. W. Fuller, A. Font, *Science* **365**, 322–323 (2019).
5. F. Paulot, D. J. Jacob, *Environ. Sci. Technol.* **48**, 903–908 (2014).
6. P. Thunis *et al.*, "Urban $PM_{2.5}$ atlas: Air quality in European cities" (European Union, 2017).
7. D. J. Sobota, J. E. Compton, M. L. McCrackin, S. Singh, *Environ. Res. Lett.* **10**, 025006 (2015).
8. M. A. Sutton *et al.*, *Nature* **472**, 159–161 (2011).
9. H. J. M. Van Grinsven *et al.*, *Environ. Sci. Technol.* **47**, 3571–3579 (2013).
10. B. Gu *et al.*, *Environ. Sci. Technol.* **46**, 9420–9427 (2012).
11. J. E. Compton *et al.*, *Ecol. Lett.* **14**, 804–815 (2011).
12. D. Simpson *et al.*, *Atmos. Chem. Phys.* **12**, 7825–7865 (2012).
13. R. Van Dingenen *et al.*, *Atmos. Chem. Phys.* **18**, 16173–16211 (2018).
14. I. Bey *et al.*, *J. Geophys. Res. Atmos.* **106** (D19), 23073–23095 (2001).
15. O. Hertel *et al.*, *Biogeosciences* **9**, 4921–4954 (2012).
16. M. Amann *et al.*, *Environ. Model. Softw.* **26**, 1489–1501 (2011).
17. J. Kirkby *et al.*, *Nature* **476**, 429–433 (2011).
18. J. Plautz, *Science* **361**, 1060–1063 (2018).
19. European Commission. Emissions Database for Global Atmospheric Research (EDGAR), version 5, Global air pollutant emissions (European Commission, 2020); https://edgar.jrc.ec.europa.eu/.
20. Y. Wu *et al.*, *Environ. Pollut.* **218**, 86–94 (2016).
21. J. Chen, G. Hoek, *Environ. Int.* **143**, 105974 (2020).
22. J. D. Sacks *et al.*, *Atmosphere* **11**, 1–15 (2020).
23. R. Martinez, P. Soliz, R. Caixeta, P. Ordunez, *Int. J. Epidemiol.* **48**, 1367–1376 (2019).
24. N. Mensah Abrampah *et al.*, *Int. J. Qual. Health Care* **30**, 5–9 (2018).
25. Y. Sun *et al.*, *Research* **2021**, 9804583 (2021).
26. M. A. Sutton *et al.*, "The Kampala statement-for-action on reactive nitrogen in Africa and globally," in *Just Enough Nitrogen: Perspectives on How to Get There for Regions With Too Much and Too Little Nitrogen, Proceedings of the 6th International Nitrogen Conference*, M. A. Sutton *et al.* Eds. (Springer, 2020).
27. X. Zhang *et al.*, *Nature* **528**, 51–59 (2015).
28. S. Reis, C. Howard, M. A. Sutton, *Costs of Ammonia Abatement and the Climate Co-Benefits* (Springer, Dordrecht, 2015).
29. M. A. Sutton *et al.*, *Our Nutrient World: The Challenge to Produce More Food and Energy with Less Pollution* (Centre for Ecology and Hydrology, 2013).
30. R. M. Hoesly *et al.*, *Geosci. Model Dev.* **11**, 369–408 (2018).
31. C. J. Stevens *et al.*, *Philos. Trans. A Math. Phys. Eng. Sci.* **378**, 20190317 (2020).
32. N. B. Dise *et al.*, "Nitrogen as a threat to European terrestrial biodiversity," in The European Nitrogen Assessment: Sources, Effects and policy Perspectives, M. A. Sutton *et al.* Eds. (Cambridge Univ. Press, 2011), pp. 463–494.
33. M. Sutton *et al.*, "The nitrogen fix: From nitrogen cycle pollution to nitrogen circular economy." in Frontiers 2018/19 Emerging Issues of Environmental Concern (United Nations Environment Programme, 2019), pp. 52–65.
34. B. Gu *et al.*, *The Innovation* **2**, 100079 (2021).
35. D. R. Kanter, T. D. Searchinger, *Nat. Sustain.* **1**, 544–552 (2018).
36. United Nations Economic Commission for Europe, "Guidance document on integrated sustainable nitrogen management" (UNECE, 2021); https://unece.org/sites/default/files/2021-04/Advance%20version_ECE_EB.AIR_149.pdf.
37. M. Amann *et al.*, "The final policy scenarios of the EU Clean Air Policy Package, TSAP report #11, version 1.1a" (International Institute for Applied Systems Analysis, 2014); https://ec.europa.eu/environment/air/pdf/TSAP.pdf.
38. M. Shulz, M. Gauss, A. Benedictow, J. E. Jonson, "Transboundary acidification, eutrophication and ground level ozone in Europe in 2011," in *EMEP Status Report 2013* ( Norwegian Meteorological Institute, 2013); https://doi.org/10.13140/RG.2.2.25401.77929.
39. B. Gu, Y. Song, C. Yu, X. Ju, *Environ. Sci. Pollut. Res. Int.* **27**, 25813–25817 (2020).
40. X. Zhang *et al.*, *Nat. Commun.* **11**, 4357 (2020).
41. Y. Guo *et al.*, *Nat. Food* **1**, 648–658 (2020).


## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grants 41822701, 42061124001, 41773068, 41922037, and 41721001). This work is a contribution from Activity 1.4 to the "Toward the International Nitrogen Management System" project (INMS, www.inms.international/) funded by the Global Environment Facility (GEF) through the United Nations Environment Programme (UNEP) and the UK Natural Environment Research Council (grants NE/S009019/1, NE/R016429/1, and NE/R000131/1 as part of the GCRF South Asian Nitrogen Hub, UK-SCAPE and SUNRISE programs). L.Z. and W.W. acknowledge support received from UNCNET, a project funded under the JPI Urban Europe/China collaboration, project numbers UMO-2018/

29/Z/ST10/02986 (NCN, Poland), 71961137011 (NSFC, China), and 870234 (FFG, Austria). D.C. acknowledges support received from the Australian Research Council through the Industrial Transformation Research Hub (IH200100023). **Author contributions:** B.G. and H.JM.V.G. designed the study as part of the wider INMS framework designed by M.A.S. R.V.D., M.V., and L.Z. applied the global atmospheric models. B.G. conducted the health impact calculation and analysis with help from R.V.D. L.Z. conducted the N-share analysis with help from R.V.D. and M.V. S.Z. and X.Z. conducted the abatement cost analysis using the GAINS model with help from W.W. S.W. and C.R. conducted the spatial and statistical analyses. B.G. analyzed all the data, interpreted the results, and wrote the first draft of the paper. All authors contributed to the discussion and revision of the paper, which was finalized under the lead of B.G. and M.A.S. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data supporting the main findings can be found in the supplementary materials. Further data that support the findings of this study are collated from online open databases or literature sources as cited in the reference list.

**SUPPLEMENTARY MATERIALS**

science.org/doi/10.1126/science.abf8623
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S6
References (*42*–*64*)

24 November 2020; accepted 31 August 2021
10.1126/science.abf8623

NEUROSCIENCE

# The glial framework reveals white matter fiber architecture in human and primate brains

**Roey Schurr*** **and Aviv A. Mezer**

**Uncovering the architecture of white matter axons is fundamental to the study of brain networks. We developed a method for quantifying axonal orientations at a resolution of ~15 micrometers. This method is based on the common Nissl staining technique for postmortem histological slices. Nissl staining reveals the spatial organization of glial cells along axons. Using structure tensor analysis, we leveraged this patterned organization to uncover local axonal orientation. We used Nissl-based structure tensor analysis to extract fine details of axonal architecture and demonstrated its applicability in multiple datasets of humans and nonhuman primates. Nissl-based structure tensor analysis can be used to compare fine-grained features of axonal architecture across species and is widely applicable to existing datasets.**

One of the long-standing goals of neuroscience is mapping the underlying neural circuitry of cortical and subcortical computations (*1*, *2*). This entails mapping the axons that communicate information between brain regions. It is therefore essential to develop tools for fine-grained measurement of white matter architecture at a cellular resolution. Existing methods are limited to animal studies (*3*), require highly specialized equipment for data acquisition and processing (*4*–*8*), or require specialized stains for myelin, the staining quality of which is highly sensitive to specific staining procedures (*9*, *10*).

One of the most common tools for studying brain tissue postmortem is Nissl staining, which targets cell nuclei. This technique has revolutionized our understanding of cortical gray matter and has been used extensively to inform cortical parcellations on the basis of cytoarchitectonic features (*11*). However, it has never been used to study the architecture of white matter, which mostly consists of axons and glial cells. As a result, white matter is "terra incognita" in Nissl-based histological atlases.

We developed a method for visualizing and quantifying in-plane fiber orientations at a resolution of ~15 μm, on the basis of postmortem histological slices stained for Nissl. To extract fiber orientations, we used the spatial organization of glial cells within white matter. Studies of specific white matter tracts have shown that astrocytes and myelinating oligodendrocytes cluster in short rows aligned with the axons that they support (*12*, *13*). Such studies refer to this organization as the "glial framework" of white matter.

We hypothesized that by measuring the local orientation of glial cells across white matter, we could infer the underlying axonal architecture. To measure the in-plane glial orientation, we used structure tensor analysis, a technique from computational image processing often used for quantifying local orientations in textured images (*14*, *15*). The structure tensor is a matrix derived from partial derivatives of image intensity along the $x$ and $y$ axes (see methods for mathematical formulation) (*16*). The second eigenvector of this matrix points in the direction of minimal changes in intensity values—i.e., along oriented structures such as glial rows in Nissl-stained white matter. We applied structure tensor analysis to quantify the orientation of the glial rows and visualize the underlying fiber architecture. We termed this technique "Nissl-based structure tensor" (Nissl-ST) and applied it to six independent datasets of humans and nonhuman primates.

First, we demonstrated the applicability of Nissl-ST in the human corpus callosum, the major white matter tract that connects the two hemispheres. Figure 1 shows that the stained cells tend to cluster in short rows, as predicted by the glial framework assumption (all figures are based on dataset 1 unless stated otherwise). The orientation of these short rows is not arbitrary; rather, nearby rows are similarly oriented, and their orientation agrees with that of neighboring axons (*17*). Local orientations generally agree with the macroscopic orientation of the corpus callosum.

To increase signal-to-noise ratio in the pixel-wise structure tensor calculation, we convolved each tensor element with a Gaussian smoothing kernel (standard deviation $\rho$ = 15 μm) (*18*) (fig. S1). Calculating the ST at every pixel (Fig. 1C) allowed us to construct a polar histogram summarizing the glial row orientation density function (gODF) of every image tile with dimensions 200 μm by 200 μm (Fig. 1D). Hence, the measurement's effective spatial resolution is determined by the Gaussian kernel (15 μm) and the resolution of the tiles used for extracting peak orientations (we used tiles of 50 μm by 50 μm or 200 μm by 200 μm). As expected in the corpus callosum, where axons are coherently oriented, the gODFs are narrow and present a sharp peak (*17*, *19*). Throughout this work, the peak in-plane orientations extracted from the gODF are presented in color-coded orientation maps.

In Fig. 2 we compared the orientation map derived from Nissl-ST with a published image of postmortem polarized light imaging (PLI) (*20*) and with the orientation map derived from in vivo diffusion magnetic resonance imaging (MRI) (*21*, *22*). For the in vivo MRI data, the in-plane peak orientation was extracted from the spherical harmonics representation of the fiber orientation distribution function (ODF; see methods and fig. S2). At the macroscopic scale, all methods yield similar orientation maps, both in deep white matter and near the cortex, where many axons that enter or leave the cortex are oriented toward major gyral crowns. A closer look at several regions of interest revealed the fine details obtained with Nissl-ST and PLI. First, Nissl-ST captures the local incoherence of fibers in the corpus callosum (Fig. 2D, red and yellow tiles) (*17*). It also reveals the fiber architecture known as Edinger's comb (Fig. 2E), which is the crossing of the lenticular fasciculus (red) through the internal capsule (green) (*23*, *24*). Finally, Nissl-ST reveals fine fiber bundles such as the angular

Edmond and Lily Safra Center for Brain Sciences, Hebrew University of Jerusalem, Jerusalem, Israel.
*Corresponding author. Email: roey.schurr@mail.huji.ac.il

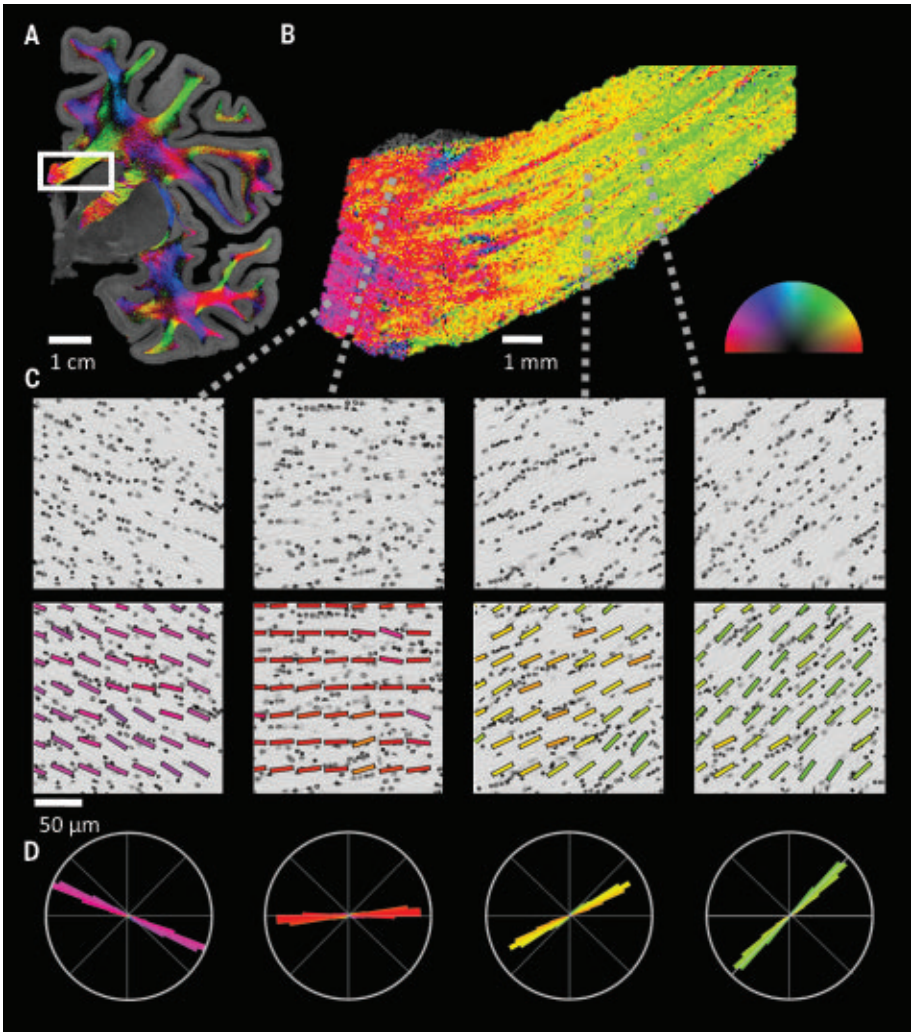



**Fig. 1. Nissl-based structure tensor analysis of the human corpus callosum.** (**A**) In-plane orientation maps in a coronal slice of the right hemisphere (in tiles of 200 µm by 200 µm smoothing kernel of 15 µm), color coded according to the semicircle [in (B) at right]. The in-plane orientation is calculated as the peak orientation of the structure tensors in each tile. (**B**) Magnified view of the corpus callosum (CC) region indicated by the rectangle in (A), calculated in 50 µm–by–50 µm tiles. (**C**) Example tiles from different locations along the CC. (Top) Glial cells are organized in short rows oriented along nearby axons. (Bottom) Subset of the pixel-wise orientations overlaid on the grayscale tiles. (**D**) Polar histograms of the gODF in each tile.

bundle (yellow), which connects the hippocampus and entorhinal cortex (Fig. 2F) (*25*). These fine details of fiber architecture are captured with Nissl-ST and PLI but are unresolved in the in vivo diffusion MRI, which is measured at a much lower resolution (1.25 mm isotropic).

Next, we tested whether the glial framework persists in regions of axon crossing. Complex architectures such as axon crossing are ubiquitous in the brain, estimated to occur in 60 to 90% of white matter voxels in a typical acquisition of diffusion MRI (*26*). To date, the glial framework has been explored in few studies, all of which focus on specific tracts in regions with a single fiber population (*13*, *27*–*29*). It is unknown how glial cells are distributed in regions of axon crossing—i.e., whether they are arranged in arrays of intersecting glial rows or randomly distributed. This question has notable biological implications (*28*): If the glial framework persists in regions of fiber crossing, this strengthens the assumption that the spatial organization of glial cells has functional implications, providing evidence that glial cells are tract-specific and are not shared across tracts. We found evidence supporting the persistence of the glial framework in regions of fiber crossing: First, our ability to obtain maps similar to those derived from PLI with only Nissl-stained images suggested that the glial framework is a brainwide feature. Second, a strong test case for the hypothesis that the glial framework persists in regions of fiber crossing is shown in regions in which the predominant orientation is of fibers that pass through the imaging plane. For example, the inferior longitudinal fasciculus (ILF) travels through the coronal plane and crosses fibers that enter the temporal lobe (Fig. 2, dashed ellipse; see fig. S2 for the visualization of ILF orientation). If the glial framework broke in this fiber crossing, we would not expect to find a clear orientation in this region. However, we found a clear in-plane orientation originating from axons that cross the ILF (magenta). This in-plane orientation was also captured by PLI and in vivo diffusion MRI data. Hence, it seems that the glial cells that support in-plane axons that cross the ILF retain the spatial organization of short glial rows. These results suggest that the glial framework persists even in regions of axonal crossing.

To further study the spatial organization of glial cells across the brain, we calculated how the coherence of in-plane orientation varies in space (Fig. 3A). The in-plane coherence quantifies the similarity between pixel-wise orientations across a region of interest. It is defined as the norm of the vector sum of all eigenvectors within the tile and ranges between 0 (incoherent orientations) and 1 (coherent orientations) (*30*). The coherence map in Fig. 3A reveals coherent regions both in areas with known single-bundle populations (e.g., the corpus callosum) as well as in areas of in-plane axons that cross through-plane axons (e.g., the ILF). By contrast, the centrum semiovale, which is a three-way crossing region (*31*), shows low coherence values. To test whether Nissl-ST detects any prominent glial orientations in such regions of low coherence, we focused on the centrum semiovale and Edinger's comb. In the centrum semiovale we found that, compared with other regions, the spatial arrangement of glial cells was less organized, as reflected by the more isotropic gODF (Fig. 3B). This may be partly a result of the limitation of Nissl-ST in recovering the orientation of through-plane axons. Although Nissl-ST extracts a single orientation per pixel, the in-plane fiber crossing in this region can be evaluated over defined regions of interest (fig. S4). The orientation of the through-plane crossing fibers remains unresolved as a result of the two-dimensional nature of Nissl-ST. In the region of Edinger's comb, two in-plane fiber bundles are intertwined, giving rise to a pattern of interchanging orientations (Fig. 3C). Close inspection of this region reveals a clear border between horizontal and vertical glial rows, which is also evident in the resulting two-peak gODF.

Regions of multiple fiber crossing may result in greater uncertainty in peak orientation. We quantified the uncertainty in peak

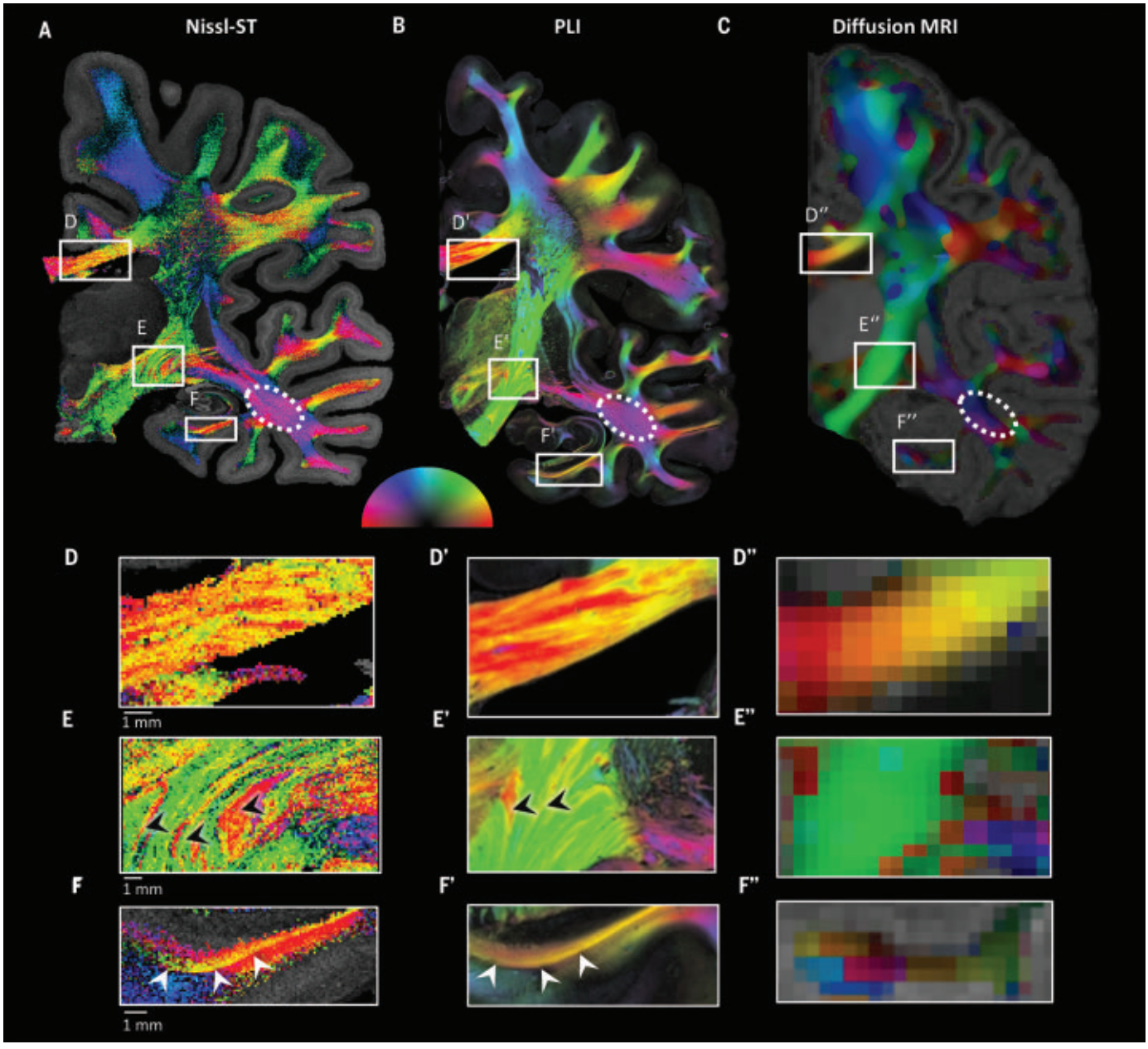



**Fig. 2. Comparison of Nissl-based structure tensor with other methods.** (**A** to **C**) In-plane orientation maps in coronal slices of the right hemisphere from three different human datasets. (A) Nissl-ST (effective resolution ~15 μm, visualized with 200 μm–by–200 μm tiles). (B) PLI (in-plane resolution 1.3 μm by 1.3 μm). [Image reproduced with permission from Axer, Amunts, and colleagues at Forschungszentrum Jülich GmbH]. (C) In vivo diffusion MRI (resolution of 1.25 mm isotropic), showing peak orientations projected onto the plane. To minimize sharp transitions in the more granular diffusion data, we smoothed the image using a median filter (see unsmoothed image in fig. S3). All methods estimate similar orientations for the in-plane axons, even in regions where the predominant fiber bundle crosses perpendicular to the plane, such as the ILF (dashed ellipse; see fig. S2) (**D** to **F**) Magnified views of the regions indicated in (A) to (C), visualized with 50 μm–by–50 μm tiles. (D) Local inhomogeneity in the corpus callosum. (E) Axons of the lenticular fasciculus (red, indicated by black arrowheads) that connect the subthalamic nucleus and the internal globus pallidus and pass through the internal capsule (green), giving rise to the interchanging pattern of crossing fibers known as Edinger's comb. (F) The angular bundle (yellow, indicated by white arrowheads), which connects the hippocampus and entorhinal cortex.

orientation using a bootstrap approach. Uncertainty was typically low, with a 95% confidence interval ≤3° in 90% of the white matter pixels (fig. S5). Higher values of uncertainty were localized mostly to the centrum semiovale. Besides fiber crossing, we found that other potential sources of signal typically did not affect the peak orientation extracted with Nissl-ST. These other sources include blood vessels (figs. S6 and S7) as well as background fibrous structures (figs. S8 and S9). Although the blood volume in white matter is estimated to be only 2.57% (*32*), the estimated peak orientation could be affected in atypical regions with very high blood vessel volume (fig. S7). Rare

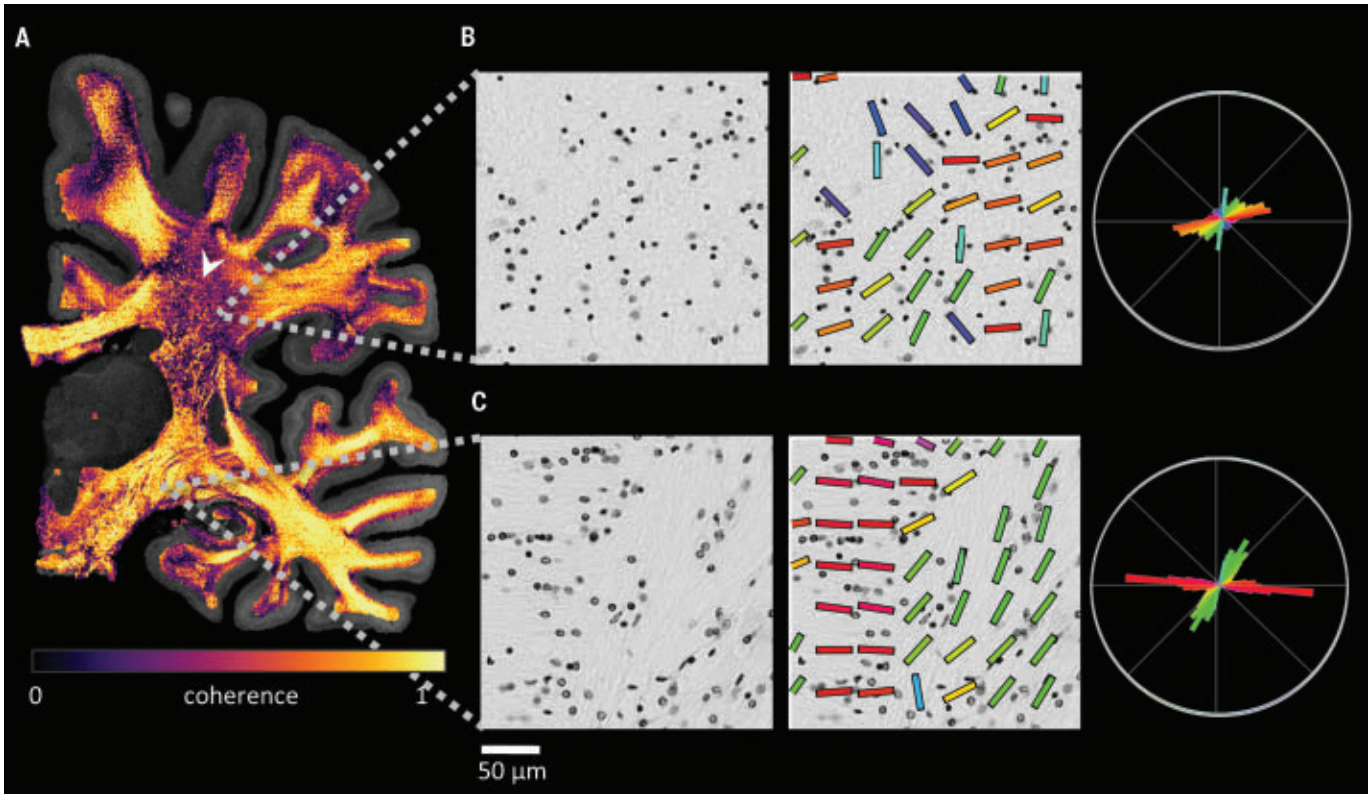



**Fig. 3. Coherence of Nissl-based structure tensor across the brain.** (**A**) Coherence map of the in-plane orientation of the same slice as in Fig. 2A. Low values indicate tiles of incoherent orientations, such as in the fiber-crossing region of the centrum semiovale (arrowhead). (**B**) Magnified view of a tile in a crossing region in the centrum semiovale. The incoherence is reflected in the isotropic gODF (right). (**C**) Magnified view of a tile in the region of Edinger's comb, where axons along the medial-lateral axis (red) cross axons along the inferior-superior axis (green). Crossing in this region manifests as a sharp change in the orientation of short glial rows.

white matter neurons likely do not affect the results of Nissl-ST, as Nissl-ST provided useful orientation information even near the gray matter (fig. S10), where the fraction of white matter neurons is higher than in deep white matter (*33*). See supplementary text for a detailed account of these analyses.

We further explored the effects of three nonbiological sources that could also affect the results of Nissl-ST: image resolution, imaging noise, and staining variability. First, to test the effect of the spatial resolution of the raw data on Nissl-ST, we compared the estimated orientations in downsampled versions of the same data. We found that downsampling increased the angular difference with respect to the native-resolution image (fig. S11 and table S1). Furthermore, an in-plane resolution lower than 5 μm by 5 μm limited the accuracy of the recovered orientations, with a median angular difference of 15°. Second, because Nissl-ST is based on computing image gradients, imaging noise can negatively affect the orientation estimation (*16*). Qualitatively, visual inspection of the high-quality data used here suggests that imaging noise is very low and unlikely to affect the results of Nissl-ST. Quantitatively, we found an extremely low level of imaging noise in these data, with a normalized variance $<10^{-4}$ (see methods). We performed simulations in which we added higher levels of noise and found that this increased the angular error of Nissl-ST (though denoising methods can reduce the effects of severe noise; table S2 and figs. S12 to S14). Finally, some nuclei may appear lighter than others in Nissl-stained slices, as a result of either actual variability in staining or their different positions along the depth of the slice. We simulated variable levels of cell staining and found that Nissl-ST is robust in terms of staining variability (fig. S15). See supplementary text for more details on these analyses.

One application of the orientation maps derived from Nissl-ST is the digital reconstruction of white matter pathways using existing histological datasets. We used deterministic tractography (*34*) over the peak in-plane orientations derived from Nissl-ST to reconstruct white matter pathways in the human brain. Fig. S16 shows the reconstruction of the corpus callosum. In addition, we tested the ability of Nissl-ST to resolve short association fibers (U-fibers). We found that the high resolution of Nissl-ST allows for reconstruction of U-fibers, and we provided evidence for a U-fiber system around the occipitotemporal sulcus. Although the exact cortical endpoints of reconstructed fibers can be measured only in axonal tracing studies, the ultrahigh resolution of Nissl-ST can provide complementary information to existing techniques for U-fiber mapping (*5*, *35*) (see supplementary text).

To establish the wide applicability of Nissl-ST to multiple datasets and species, we applied it to histological sections of four independent datasets: two human specimens (*36*, *37*) and two nonhuman primates (a rhesus macaque and a vervet monkey). Fig. 4A shows the orientation maps derived from each dataset. Nissl-ST successfully extracted meaningful orientation maps with similar features across datasets. Furthermore, we found that the region of Edinger's comb (white arrowhead) is conserved across species, showing a similar characteristic crossing of glial rows in each specimen (Fig. 4B).

Nissl-ST has potential for use in future studies of white matter in normal development and aging (*38*), as well as pathological states that affect white matter, such as schizophrenia (*39*). Such applications of Nissl-ST will need to

account for the potential challenges that the histological data may present. For example, pathological conditions accompanied by microglial response may also be accompanied by an alternation to the spatial distribution of glial cells (*40*). Development, especially at prenatal stages, poses an additional challenge for Nissl-ST: On one hand, Nissl-ST provides a potential means to study the initial development of myelin, as some studies in rats found that at birth, oligodendroglial progenitor cells were already assembled in short rows resembling

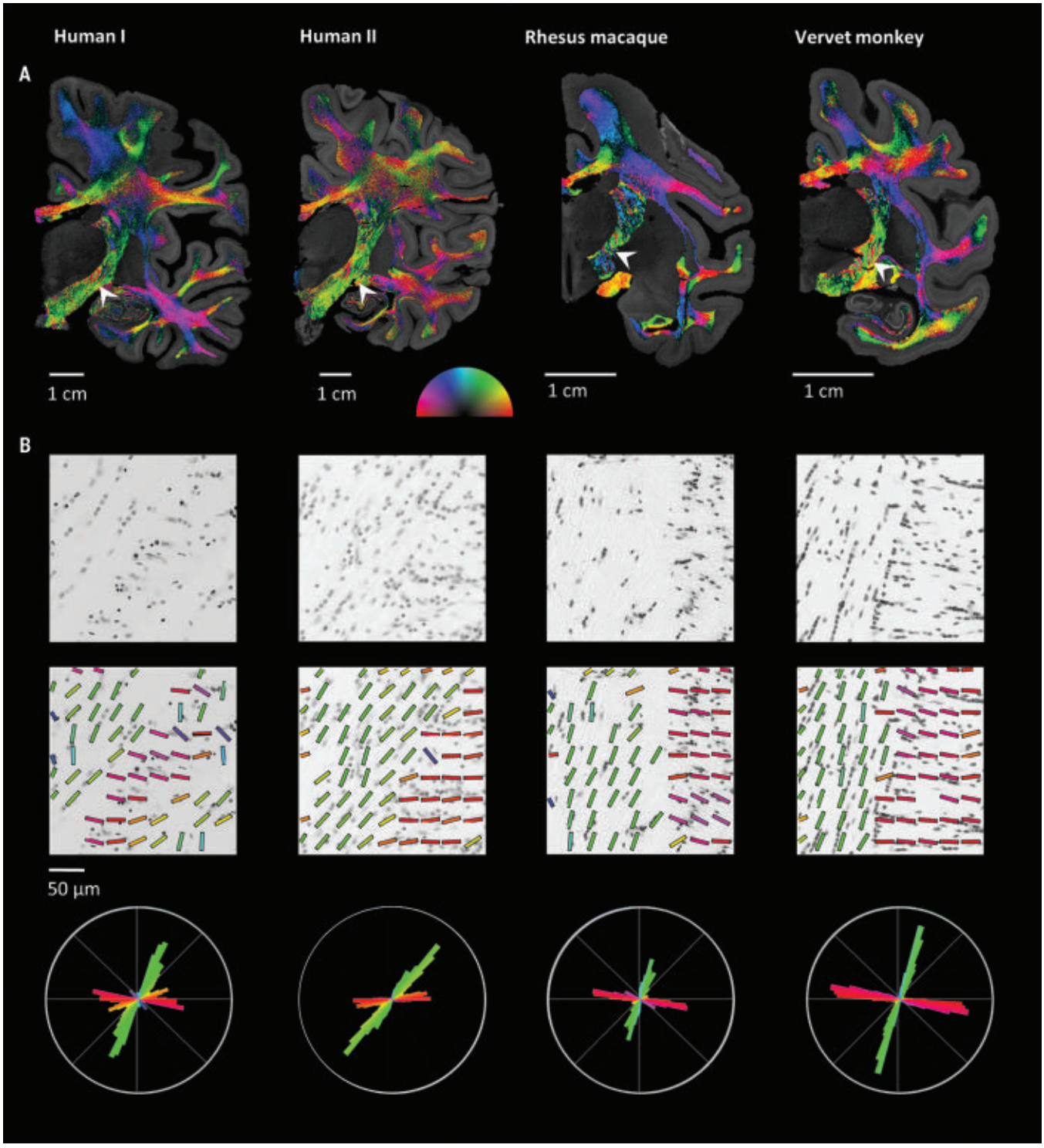



**Fig. 4. Applicability of Nissl-based structure tensor across datasets.** (**A**) In-plane orientation maps in coronal slices of four independent datasets, demonstrating the applicability of Nissl-ST to existing histological datasets of humans and nonhuman primates (rhesus macaque and vervet monkey), establishing the existence of the whole-brain glial framework across species. (**B**) Magnified view of tile from Edinger's comb in each specimen [white arrowheads in (A)]. The characteristic axonal crossing in this region is preserved across species.

those of the adult brain (*27*). On the other hand, because there are changes in the relative proportion of different cell types during development (*41*, *42*), it is possible that in early developmental stages, the main source of the signal in Nissl-ST would be other cell types (e.g., endothelial cells), thereby masking the effect of the short glial rows. To demonstrate the applicability of Nissl-ST to developmental data, we compared the dataset of the adult rhesus macaque (Fig. 4) with data from both the embryonic stage at 120 days and the postnatal stage at 14 days. These data suggest that at 120 embryonic days, the density of glial cells (or their progenitors) is very low compared with that of the postnatal brain at 14 days (fig. S17), with the latter having a much higher similarity to the brain of an adult. Furthermore, at 120 embryonic days, the cells did not seem to be organized in short rows. By contrast, at postnatal day 14, glial cells were already organized in short rows along the axons' expected orientation. Therefore, future research focusing on the intermediate stages of development could shed light on the process and ordering of myelination across white matter pathways.

A common feature of Nissl-ST, PLI, and structure tensor analysis on the basis of myelin staining is their sensitivity to myelinated (rather than nonmyelinated) axons. In addition, the proposed implementation of Nissl-ST is two dimensional in nature. Hence, it is unable to resolve the orientation of through-plane fibers. Because glial cells associated with out-of-plane axons cannot be easily identified as such, they might introduce uncertainty into the method's orientation estimates. Future studies directly comparing Nissl-ST and PLI in the same slice could shed more light on their differential sensitivity to both in-plane and through-plane orientations. As an extension of the proposed technique, the Nissl-ST method could be generalized to three-dimensional datasets, such as the future high-resolution version of BigBrain (which used silver staining for cell bodies rather than Nissl staining) (*43*). Notably, this would require specialized tools for correcting staining variability between sections (*43*). Such three-dimensional datasets could be integrated with neuroimaging data using specially tailored toolboxes (*44*).

The glial framework—the patterned spatial organization of glial cells in the white matter—has been described in only a handful of pathways, namely the rat fimbria (*13*), the mouse corpus callosum (*28*), and the vervet corpus callosum (*29*). Additionally, Pandya and Schmahmann (*3*) briefly mentioned the glial framework (i.e., the "glial matrix") in the context of the ILF and nearby tracts. It is a long-standing question whether the glial framework extends to all myelinated tracts, especially phylogenetically younger tracts. We found direct evidence that this is the case for the corpus callosum, the lenticular fasciculus, and the internal capsule in humans, rhesus macaques, and vervet monkeys, as well as a U-fiber around the occipitotemporal sulcus in the human brain. The overall similarity of in-plane orientation maps derived from Nissl-ST and from PLI suggests that the structured organization of glial cells along axons extends to other white matter tracts.

The model-free approach of Nissl-ST reveals a rich layer of information in Nissl-stained brain slices; this information has been hitherto unused. An advantage of Nissl-ST compared with existing methods is the inherent co-registration of the characteristic maps that can be derived from white matter, as well as the maps derived from cortical and subcortical structures. Such coregistration would allow for easy integration of Nissl-based atlases of gray matter structures alongside maps of white matter architecture. The abundance of Nissl-stained resources in labs worldwide, as well as the prevalence of digitized datasets and open-source atlases including high-resolution Nissl-stained slices (*36*, *37*, *45*, *46*), makes the proposed technique readily applicable to numerous datasets. Notably, such datasets would allow for the comparison of fine-grained features of fiber architecture across species as well as between healthy and diseased brains, and would provide a simple way to obtain a histological reference for in vivo white matter mapping.

## REFERENCES AND NOTES


1. B. A. Wandell, *Annu. Rev. Neurosci.* **39**, 103 (2016).
2. H. Takemura, M. Thiebaut de Schotten, *Brain Struct. Funct.* **225**, 1189 (2020).
3. J. D. Schmahmann, D. N. Pandya, *Fiber Pathways of the Brain* (Oxford Univ. Press, 2006).
4. M. Axer *et al*., *Front. Neuroinform.* **5**, 34 (2011).
5. H. Takemura *et al*., *eLife* **9**, e55444 (2020).
6. J. Men *et al*., *IEEE J. Sel. Top. Quantum Electron.* **22**, 120 (2016).
7. M. Menzel *et al*., *Neuroimage* **233**, 117952 (2021).
8. H. R. Ueda *et al*., *Nat. Rev. Neurosci.* **21**, 61 (2020).
9. C. Stüber *et al*., *Neuroimage* **93**, 95 (2014).
10. A. Yendiki *et al*., *bioRxiv* (2021.04.16.440223) [Preprint] (2021).
11. K. Amunts, K. Zilles, *Neuron* **88**, 1086 (2015).
12. N. Baumann, D. Pham-Dinh, *Physiol. Rev.* **81**, 871 (2001).
13. M. Suzuki, G. Raisman, *Glia* **6**, 222 (1992).
14. J. Bigun, G. H. Granlund, J. Wiklund, *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 775 (1991).
15. M. D. Budde, J. A. Frank, *Neuroimage* **63**, 1 (2012).
16. W. Zhang *et al*., Structure Tensor Based Analysis of Cells and Nuclei Organization in Tissues, *IEEE Trans. Med. Imaging* **35**, 294 (2016).
17. J. Mollink *et al*., *Neuroimage* **157**, 561 (2017).
18. C. Bailey *et al*., *NMR Biomed.* **32**, e4073 (2019).
19. M. D. Budde, J. Annese, *Front. Integr. Nuerosci.* **7**, 3 (2013).
20. K. Amunts, A. Lindner, K. Zilles, The human brain project: neuroscience perspectives and German contributions. *e-Neuroforum* **5**, 43 (2014).
21. P. J. Basser, J. Mattiello, D. LeBihan, *J. Magn. Reson. B.* **103**, 247 (1994).
22. J. D. Tournier, F. Calamante, D. G. Gadian, A. Connelly, *Neuroimage* **23**, 1176 (2004).
23. A. Horn *et al*., *Neurology* **92**, e1663 (2019).
24. E. J. L. Alho *et al*., *Mov. Disord.* **35**, 75 (2020).
25. M. M. Zeineh *et al*., *Cereb. Cortex* **27**, 1779 (2017).
26. B. Jeurissen, A. Leemans, J.-D. Tournier, D. K. Jones, J. Sijbers, *Hum. Brain Mapp.* **34**, 2747 (2013).
27. R. Reynolds, R. Hardy, *J. Neurosci. Res.* **47**, 455 (1997).
28. T. Tanaka *et al*., *Glia* **69**, 2488 (2021).
29. M. Andersson *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 33649 (2020).
30. J.-P. Lachaux, E. Rodriguez, J. Martinerie, F. J. Varela, *Hum. Brain Mapp.* **8**, 194 (1999).
31. K. Schilling *et al*., *NMR Biomed.* **30**, e3787 (2017).
32. J. Doucette, C. Kames, E. Hernández-Torres, A. Traboulsee, A. Rauscher, *bioRxiv* (614008) [Preprint] (2019).
33. V. García-Marín, L. Blazquez-Llorca, J. R. Rodriguez, J. Gonzalez-Soriano, J. DeFelipe, *J. Comp. Neurol.* **518**, 4740 (2010).
34. S. Mori, B. J. Crain, V. P. Chacko, P. C. M. van Zijl, *Ann. Neurol.* **45**, 265 (1999).
35. M. Guevara, P. Guevara, C. Román, J. F. Mangin, *Neuroimage* **212**, 116673 (2020).
36. J. K. Mai, M. Majtanik, G. Paxinos, *Atlas of the Human Brain* (Academic Press, ed. 4, 2015).
37. J. A. Miller *et al*., *Nature* **508**, 199 (2014).
38. J. D. Yeatman, B. A. Wandell, A. A. Mezer, *Nat. Commun.* **5**, 4932 (2014).
39. J. Fitzsimmons, M. Kubicki, M. E. Shenton, *Curr. Opin. Psychiatry* **26**, 172 (2013).
40. S. Hickman, S. Izzy, P. Sen, L. Morsett, J. El Khoury, *Nat. Neurosci.* **21**, 1359 (2018).
41. R. R. Sturrock, *J. Anat.* **122**, 521 (1976).
42. R. K. Sigaard, M. Kjær, B. Pakkenberg, *Cereb. Cortex* **26**, 89 (2016).
43. K. Amunts *et al*., *Science* **340**, 1472 (2013).
44. C. Paquola *et al*., *bioRxiv* (2021.05.04.442563) [Preprint] (2021).
45. S. Mikula, I. Trotts, J. M. Stone, E. G. Jones, *Neuroimage* **35**, 9 (2007).
46. G. Striedter, The Zoomable Human Brain Atlas, https://zoomablebrain.bio.uci.edu/.
47. R. Schurr, A. Mezer, Nissl-ST, Version 1.0, Zenodo (2021); https://doi.org/10.5281/zenodo.5245155.


## ACKNOWLEDGMENTS


We thank B. Wandell, A. Rokem, S. Berman, S. Filo, E. Drori, H. Takemura, J. Yeatman, and K. Weiner for helpful comments on this manuscript. Dataset 1 was provided by J. Mai and M. Majtanik. Dataset 2 © 2010 Allen Institute for Brain Science. BrainSpan Atlas of the Developing Human Brain is available from human.brain-map.org. Dataset 3 (*Macaca mulatta*) was provided by E. G. Jones. Dataset 4 (*Chlorocebus aethiops*) was provided by NeuroScience Associates. Datasets 5 and 6 © 2010 Allen Institute for Brain Science. NIH Blueprint Non-Human Primate (NHP) Atlas is available from www.blueprintnhpatlas.org. **Funding:** This research was supported by the Israel Science Foundation (grant 1169/20) given to A.A.M., and by the Jerusalem Brain Community (JBC) award granted to R.S. **Author contributions:** Conceptualization, methodology, formal analysis, visualization, and writing: R.S. and A.A.M. Supervision: A.A.M. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All original datasets used in this work can be accessed on the websites listed in the methods section. The MATLAB code for calculating and visualizing the structure tensor is available along with the code for downloading example data on Zenodo (*47*). Fig. S19 shows the expected outputs of the example code.


## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj7960
Materials and Methods
Supplementary Text
Figs. S1 to S19
Tables S1 and S2
References (*48*–*71*)
MDAR Reproducibility Checklist


4 June 2021; accepted 14 September 2021
Published online 7 October 2021
10.1126/science.abj7960

HEMATOPOIESIS

# Resistance to inflammation underlies enhanced fitness in clonal hematopoiesis

S. Avagyan[1]†, J. E. Henninger[2]†, W. P. Mannherz[3], M. Mistry[4], J. Yoon[4], S. Yang[5], M. C. Weber[5], J. L. Moore[5], L. I. Zon[5,6]*

Clonal hematopoiesis results from enhanced fitness of a mutant hematopoietic stem and progenitor cell (HSPC), but how such clones expand is unclear. We developed a technique that combines mosaic mutagenesis with color labeling of HSPCs to study how acquired mutations affect clonal fitness in a native environment. Mutations in clonal hematopoiesis–associated genes such as *asxl1* promoted clonal dominance. Single-cell transcriptional analysis revealed that mutations stimulated expression of proinflammatory genes in mature myeloid cells and anti-inflammatory genes in progenitor cells of the mutant clone. Biallelic loss of one such immunomodulator, *nr4a1*, abrogated the ability of *asxl1*-mutant clones to establish clonal dominance. These results support a model where clonal fitness of mutant clones is driven by enhanced resistance to inflammatory signals from their mutant mature cell progeny.

Clonal hematopoiesis reflects an imbalance of hematopoietic stem and progenitor cell (HSPC) output. Imbalances may result from mutations that enable competitive outgrowth during aging, selective resistance to chemotherapy, or founder effects in HSPC transplantation (*1*). Clonal hematopoiesis with recurrent mutations in genes such as *DNMT3A*, *TET2*, and *ASXL1* is associated with increased risk for developing hematologic malignancies (*2*, *3*). It is not known how mutations confer competitive advantage to stem cells. Although a number of gene mutations induce clonal hematopoiesis in preclinical models using transplantation assays (*4*, *5*), prospective clonal competition in a native environment has not been studied. Here, we used an endogenous color labeling system of HSPC clones in zebrafish, called Zebrabow, to study the effect of germline and acquired mutations on stem cell clonality.

We previously used Zebrabow to evaluate HSPC composition and found that cytotoxic irradiation and transplantation reduced clonal diversity (*6*). To determine whether clonal changes could result from genetic alterations, we evaluated the clonal composition in a zebrafish *runx1* germline mutant that generates fewer HSPCs during development (*7*). Embryos from an intercross of *runx1*$^{W84X/+}$;*Tg(ubi:Zebrabow-M)* and *runx1*$^{W84X/+}$;*Tg(draculin(drl):CreER*$^{T2}$ zebrafish were treated with 4-hydroxytamoxifen (4-OHT) at 24 hours post-fertilization (hpf) to induce color labeling of HSPCs, and clonal composition was assessed in adult marrow

[1]Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Boston, MA, USA. [2]Whitehead Institute for Biomedical Research, Cambridge, MA, USA. [3]Harvard Medical School, Boston, MA, USA. [4]Harvard Chan Bioinformatics Core, Boston, MA, USA. [5]Boston Children's Hospital, Boston, MA, USA. [6]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA.
*Corresponding author. Email: zon@enders.tch.harvard.edu
†These authors contributed equally to this work.

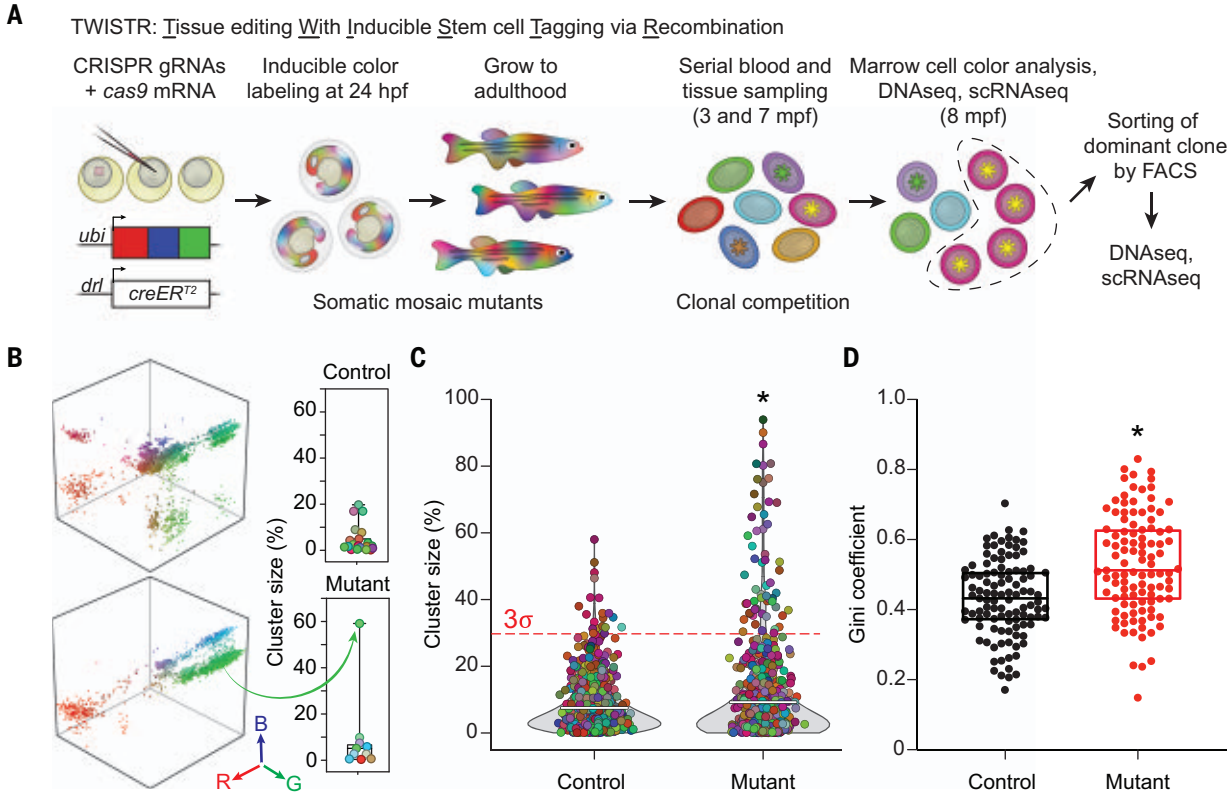



**Fig. 1. Mutations of clonal hematopoiesis induce clonal dominant states with clonal skewing.** (**A**) Schematic of TWISTR. gRNA, guide RNAs; hpf/mpf, hours/months post-fertilization; DNA-seq, DNA sequencing; scRNA-seq, single-cell RNA sequencing; FACS, fluorescence-activated cell sorting. (**B**) Example of Zebrabow analysis in control and mutant zebrafish. Dot plots show cluster size of each color as percent of marrow myelomonocytes. (**C**) Contribution of clusters in myelomonocytes of control ($n$ = 107) and mutant ($n$ = 105) zebrafish. Dashed line marks three standard deviations of mean cluster size of control zebrafish. *$P$ < 0.0001 (Mann-Whitney U test comparing the largest clusters per zebrafish). (**D**) Gini coefficient in control ($n$ = 107, median 0.4322) and mutant ($n$ = 105, median 0.5121) zebrafish. *$P$ < 0.0001 (Mann-Whitney U test).

myelomonocytes (fig. S1A). In wild-type zebrafish, hematopoiesis is polyclonal and appears multicolored (*6*) (fig. S1B). Hematopoiesis in *runx1* homozygous mutants was driven by fewer larger clones, consistent with an oligoclonal state resulting from a smaller HSPC pool (fig. S1, B and C). Thus, Zebrabow can detect an aberrant clonal composition in hematopoiesis caused by a germline genetic mutation.

We next sought to determine how acquired mutations affect clonal composition. We combined mosaic mutagenesis using CRISPR-Cas9 with Zebrabow labeling, a technique we call TWISTR (tissue editing with inducible stem cell tagging via recombination). Embryos of *Tg(ubi:Zebrabow-M)* and *Tg(drl:CreER$^{T2}$)* crosses were injected with a pool of guide RNAs (gRNAs) targeting zebrafish orthologs of genes implicated in clonal hematopoiesis and two control genes (Fig. 1A and table S1). HSPC color labeling was induced by 4-OHT at 24 hpf. We hypothesized that a clonal outgrowth due to enhanced fitness would be detectable by a dominant single color. Although zebrafish injected with control gRNAs had normal multicolored adult hematopoiesis, mutant zebrafish often exhibited a large single-colored cluster (Fig. 1, B and C), with increased average size of the largest cluster (fig. S2, A and B). Cluster size in mutant zebrafish was more variable than that in controls, consistent with abnormal clone size distribution (fig. S2C). We used the Gini coefficient as a measure of inequality (*8*) and found that individual mutant zebrafish had higher Gini coefficients than controls, indicating greater clonal skewing (Fig. 1D). There were no significant differences in hematopoietic cell types or morphology between cohorts (*9*) (fig. S2, D and E). These results show that TWISTR can induce and detect clonal dominance in the presence of acquired mutations without disturbance of hematopoiesis.

We determined the mutational landscape in TWISTR mutants to characterize genetic drivers of clonal fitness. Although embryos were injected at the one-cell stage, translation of *Cas9* mRNA delays mutagenesis, generating a mosaicism of wild-type and mutant alleles (*10*, *11*). We collected serial peripheral blood samples starting at 3 months and marrow samples at 8 to 9 months (fig. S3A). As a control for overall gRNA targeting, we collected fin tissue largely composed of nonhematopoietic cells. Sequencing showed mutations in all targeted genes, with variable numbers of mutations per zebrafish (fig. S3, B and C). Although all gRNAs were tested individually to confirm target editing and gRNA efficiency was concordant across tissues, the numbers of detected edits varied between genes (fig. S3, D and E). This could be due to gene-specific selection or may represent differences in pooled gRNA efficiencies. Comparing types of insertions/deletions (indels) for a given gene in marrow (hematopoietic) versus fin (nonhematopoietic) could show tissue-specific selection. We found that disruptive frameshift mutations were more likely to occur in marrow than in fin for *asxl1*, *tet2*, and *tp53* (fig. S4A). These results demonstrate multiplexed mutagenesis with TWISTR and reveal that frameshift alleles of specific genes undergo positive selection in marrow.

To examine temporal changes of variant allele fractions (VAFs) in mutant zebrafish, we sequenced and tracked unique variants in

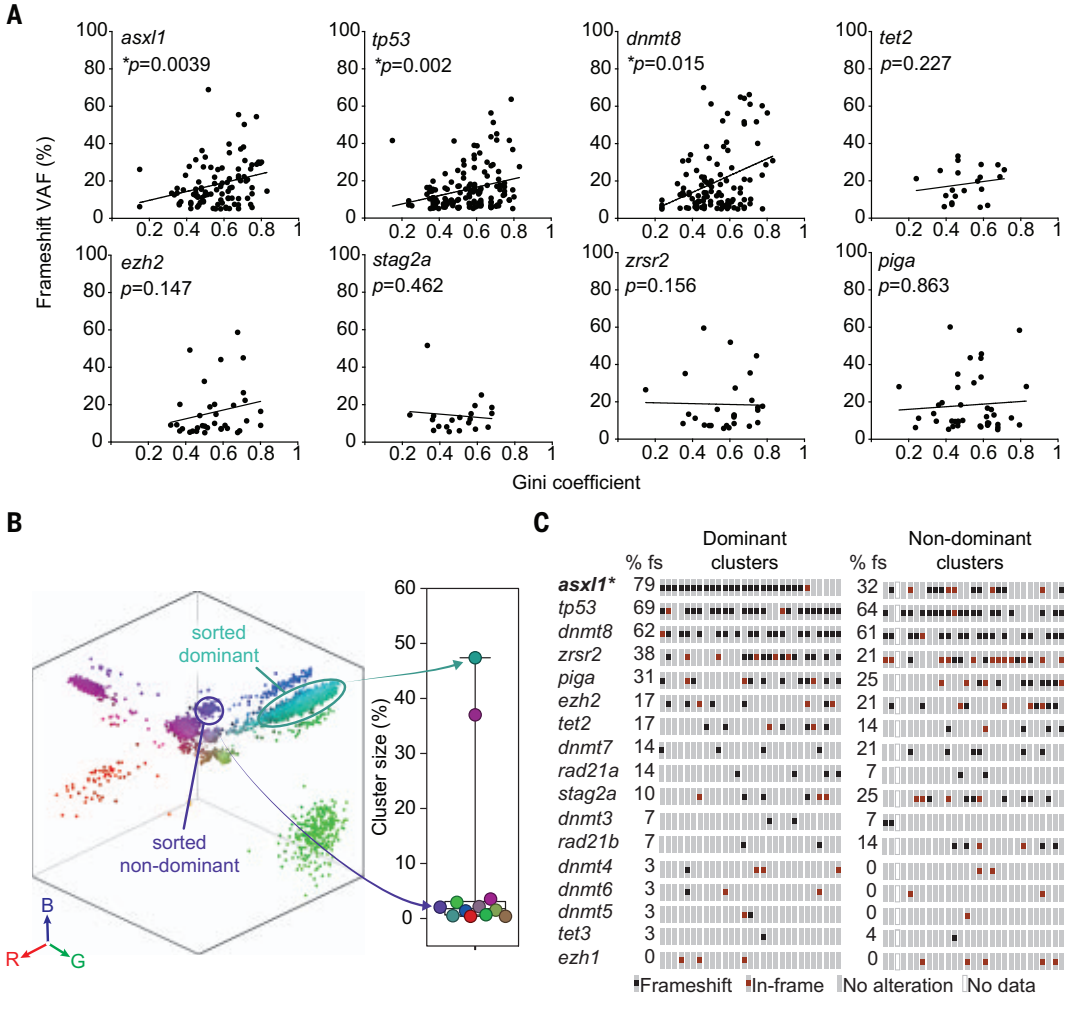



**Fig. 2. Multiplex mosaic mutagenesis and Zebrabow labeling reveal mutant allele dynamics and identify drivers of clonal hematopoiesis.** (**A**) Plots of frameshift indel variant allele fractions (VAFs) in marrow cells and Gini coefficient in myelomonocytes of mutant zebrafish. *$P$ < 0.02 (Spearman's rank correlation). (**B**) Sorting strategy of individual clusters. (**C**) Tiling plot of alleles in sorted clusters from 29 mutant zebrafish. Columns correspond to single zebrafish in the same order in each set. Percent of frameshift (fs) VAF containing clusters is shown. *$P$ < 0.001 ($\chi^2$ test).

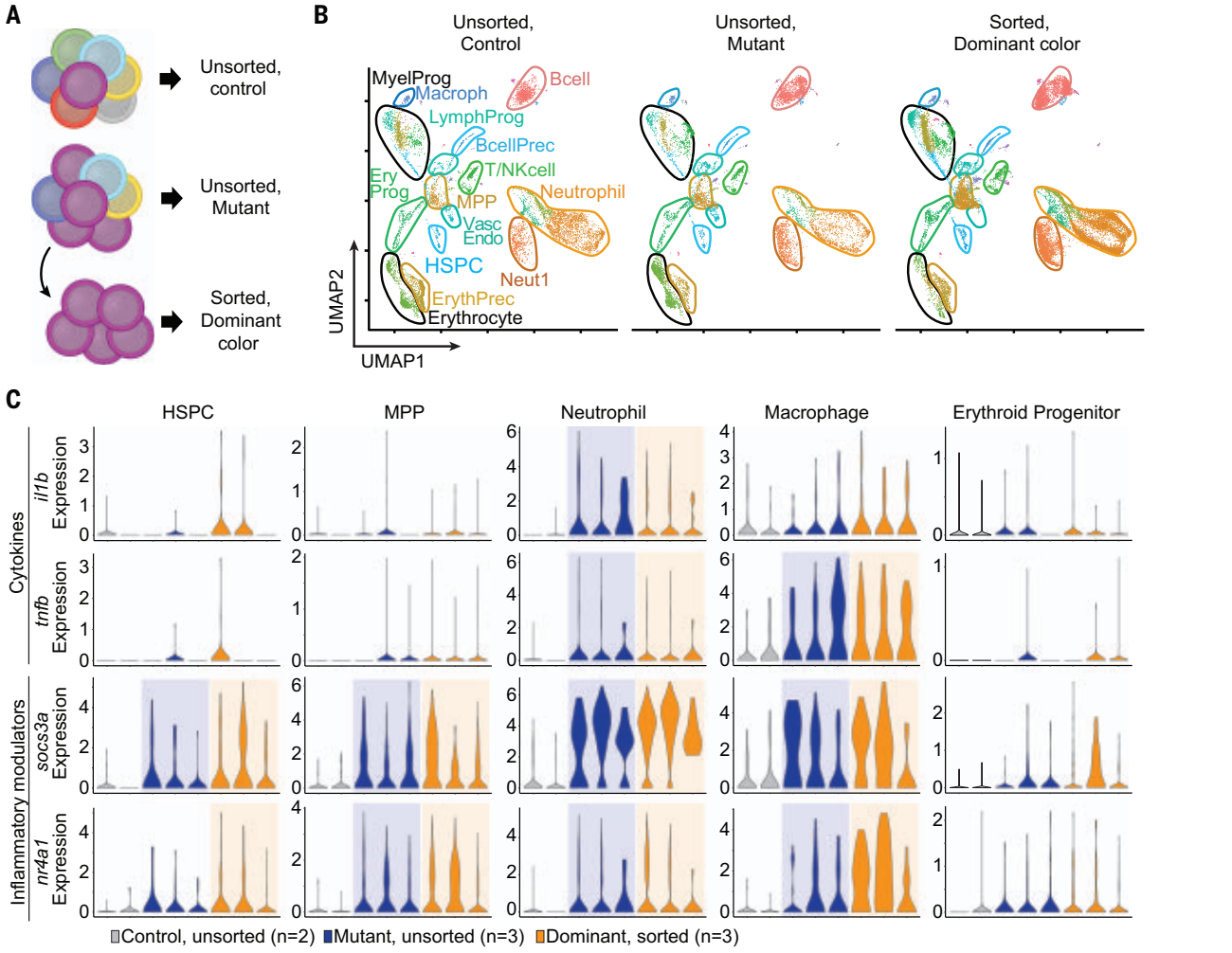



**Fig. 3. Single-cell RNA sequencing reveals an inflammatory state in hematopoiesis and differential response of mutant progenitors.** (**A**) Schema of scRNA-seq experiment. (**B**) Split uniform manifold approximation and projection (UMAP) of marrow cells in control zebrafish ($n$ = 2, unsorted), mutant zebrafish ($n$ = 3, unsorted), and dominant color from these mutant zebrafish ($n$ = 3, sorted). (**C**) Violin plots showing select gene expression (log-transformed) in the eight samples. Shaded blue and orange boxes indicate significant difference relative to controls in mutant and dominant groups, respectively; $P^{adj} < 5 \times 10^{-4}$. MPP, multipotent progenitors.

peripheral blood over time. VAFs of frameshift indels in *asxl1* increased, consistent with their positive selection, whereas other gene variants exhibited mixed dynamics or no significant enrichment for frameshift indels (fig. S4B). Three observations suggested that positive selection was gene- and mutation type–specific. First, nondisrupting in-frame indels of these genes did not change (fig. S4C). Second, VAFs of control genes in control zebrafish remained stable (fig. S4D). Third, we observed no change in VAFs of genes in fin tissue of mutant zebrafish (fig. S4, E and F). Expanding variants were invariably found in marrow dominant clones sorted by color (fig. S4G). We observed no correlation between clone size and number of frameshift mutations for the dominant clones (fig. S4H). These results show that frameshift mutations in some genes, such as *asxl1*, exhibit positive selection over the study time period, resulting in clonal expansion.

To assess the potential of these mutations to promote clonal dominance, we analyzed the correlation between VAFs and the Gini coefficient of clonal skewing. We found that frameshift indels in the common clonal hematopoiesis genes *asxl1*, *dnmt8* (*DNMT3A* ortholog), and *tp53* positively correlated with the Gini coefficient, indicating that disruption of these genes was associated with clonal dominance (Fig. 2A). There was no correlation between in-frame indels and the Gini coefficient (fig. S4I). These data show that frameshift mutations in select clonal hematopoiesis–associated genes lead to clonal dominance.

TWISTR enables the prospective isolation of live cells from dominant and nondominant clones from the same animal, allowing their direct comparison to identify drivers of clonal fitness. We sorted myelomonocytes from dominant clusters and compared their mutational landscape to that of myelomonocytes in smaller clusters from the same mutant zebrafish (Fig. 2B). The dominant clusters harbored frameshift indels in *asxl1* significantly more commonly than did nondominant clusters (Fig. 2C). Single targeting of *asxl1* resulted in clonal dominance associated with frameshift indels, which suggested that mutant *asxl1* alone can promote clonal dominant states (fig. S5). These results show that frameshift mutations in *asxl1* confer clonal fitness.

To identify pathways underlying clonal fitness with acquired mutations, we performed single-cell RNA sequencing (scRNA-seq) on (i) unsorted marrow cells from control zebrafish, (ii) unsorted marrow cells from mutant zebrafish with unique sets of mutations, and (iii) dominant clusters that contributed to >50% of marrow myelomonocytes in the mutant zebrafish (Fig. 3A). All cell types were present in mutant zebrafish, except for mildly decreased erythroid precursors (*12*) (Fig. 3B, fig. S6, and

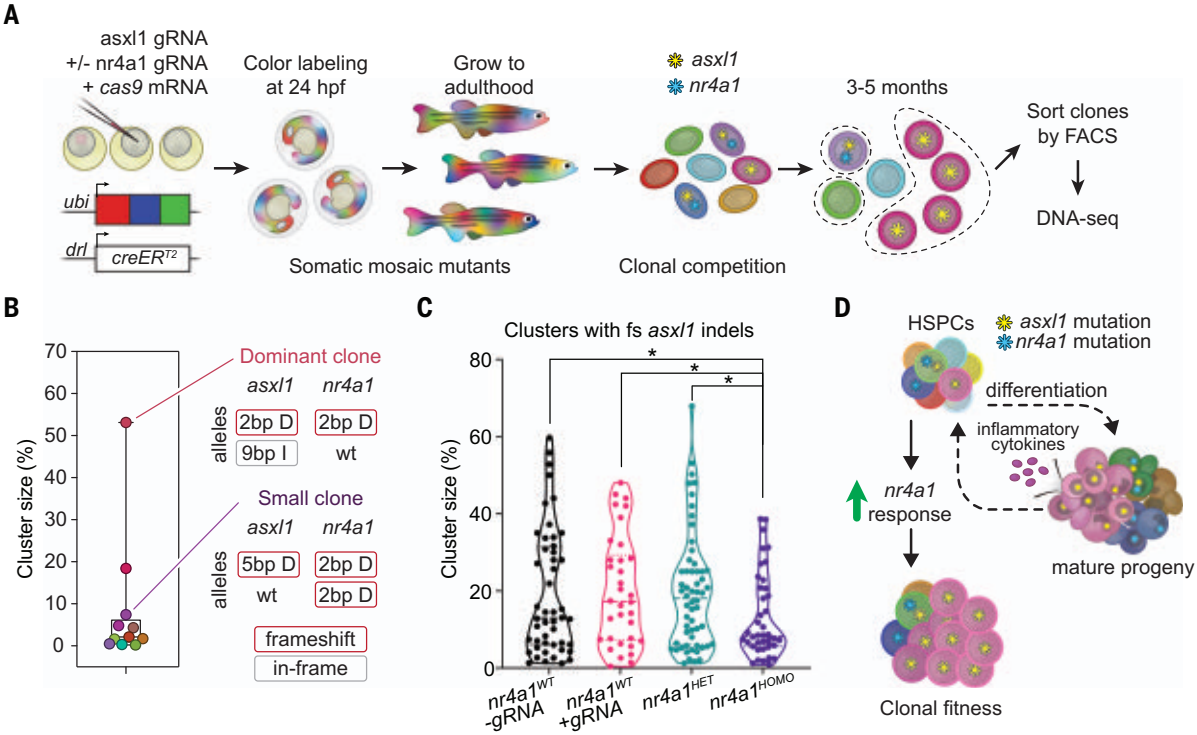



**Fig. 4. Biallelic *nr4a1* loss abrogates fitness of *asxl1*-mutant clones.** (**A**) Testing the effect of *nr4a1* with TWISTR. (**B**) Example DNA-seq analysis of sorted clusters from a single zebrafish. D, deletion; I, insertion; bp, base pair. (**C**) Size of sorted *asxl1*-mutant clusters in zebrafish injected with only *asxl1* gRNA ($nr4a1^{WT}$, $n$ = 48) or with *asxl1* and *nr4a1* gRNAs that resulted in no deleterious *nr4a1* indels ($nr4a1^{WT}$, $n$ = 34), heterozygous deleterious indels ($nr4a1^{HET}$, $n$ = 58), or homozygous deleterious indels ($nr4a1^{HOMO}$, $n$ = 46). *$P$ = 0.029 *asxl1*-only $nr4a1^{WT}$ versus $nr4a1^{HOMO}$, $P$ = 0.014 $nr4a1^{WT}$ versus $nr4a1^{HOMO}$, $P$ = 0.013 $nr4a1^{HET}$ versus $nr4a1^{HOMO}$ (unpaired $t$ test). (**D**) Model of inflammation-driven clonal fitness with acquired mutations.

table S2). The mutant marrow cells expressed a stimulated inflammatory response relative to controls, with increased expression of genes such as *fosab* and *chac1* (fig. S7, A to C). The source of inflammation appeared to be mutant macrophages and neutrophils, as niche cells were devoid of inflammatory cytokine expression (Fig. 3C and fig. S7, A and B). Elevated inflammatory cytokine levels have been described in clonal hematopoiesis, but their effect on clonal expansion is not known (*13*, *14*). Acute inflammation enhances HSPC cycling but can lead to exhaustion if it persists (*15*), and chronic infection can potentiate *Dnmt3a* mutant hematopoiesis (*16*). We hypothesized that mutant HSPCs moderate their response to chronic inflammation maintained by their mutant mature progeny, rendering them resistant to replicative exhaustion. We identified elevated gene expression of suppressors of inflammation, such as *nr4a1* and *socs3a*, in mutant HSPCs and progenitors (Fig. 3C and fig. S7, A and B). We compared the gene expression of sorted marrow cells of dominant clones from mutant zebrafish to that of total marrow cells from control zebrafish and found similarly dysregulated inflammatory gene signatures (Fig. 3C and fig. S7D), suggesting that mutant dominant clones express an anti-inflammatory gene signature in response to inflammation. *Runx1* deficiency has been associated with inflammation (*17*). A scRNA-seq comparison of germline homozygous *runx1* mutants with oligoclonal hematopoiesis to heterozygous mutants with polyclonal hematopoiesis revealed a similar pro-inflammatory/anti-inflammatory regulatory axis (fig. S8).

We reasoned that genetic abrogation of the anti-inflammatory response in mutant clones would reduce their fitness. We used TWISTR to generate a cohort with gRNAs targeting *asxl1* and *nr4a1*, a gene that encodes a nuclear receptor antagonizing immune activation and overexpressed in dominant mutant clones (*18*) (Fig. 4A). We observed normal hematopoiesis in these zebrafish (fig. S9, A and B). We sorted 306 clusters from 20 zebrafish injected with *asxl1*-targeting gRNA and 58 zebrafish injected with gRNAs targeting both *asxl1* and *nr4a1*, and assessed indels (Fig. 4B). Of these, 186 harbored deleterious *asxl1* indels: 48 from *asxl1*-injected and 138 from *asxl1*/*nr4a1*-injected zebrafish. Comparing cluster sizes, we observed that clusters with biallelic deleterious indels of *nr4a1* were significantly smaller relative to those with no *nr4a1* indels, in-frame/no indels, or heterozygous deleterious indels (Fig. 4C and fig. S9C). *Asxl1*-mutant clones with biallelic *nr4a1* indels diminished in peripheral blood over time, whereas *asxl1*-mutant clones with intact *nr4a1* expanded (fig. S9D). There was no difference between wild-type *asxl1* clusters across *nr4a1* genotypes, which suggests that *nr4a1* alone is not sufficient to modulate clonal fitness (fig. S9E). Within individual zebrafish, smaller *asxl1*-mutant clusters were more likely than larger *asxl1*-mutant clusters to harbor homozygous deleterious *nr4a1* indels (fig. S9F). These results suggest that a blunted anti-inflammatory response in *asxl1*-mutant clones due to *nr4a1* loss diminished their clonal fitness.

Our results provide evidence that progeny of mutant clones generate an inflammatory environment to which they are resistant (Fig. 4D). Certain mutations in clonal hematopoiesis have an association with specific cytokines, including interleukin-6 with *TET2* mutations and tumor necrosis factor–α with *ASXL1* mutations (*13*, *14*). It is plausible that there is mutation-dependent cytokine expression in mutant mature cells, and an up-regulation of mutation-associated anti-inflammatory pathways in mutant progenitor cells. Thus, clonal fitness is achieved by a pro-survival network engaged by mutations, which offers an explanation for how gene mutations with diverse functions confer similar clonal fitness. Our study developed a genetic and live clonal tracking technology to dissect the progression to clonal dominance, identifying new

pathways that may enable therapeutic approaches to alter the fitness of mutant clones.

## REFERENCES AND NOTES


1. J. N. Cooper, N. S. Young, *Blood* **130**, 2363–2372 (2017).
2. G. Genovese *et al*., *N. Engl. J. Med.* **371**, 2477–2487 (2014).
3. S. Jaiswal *et al*., *N. Engl. J. Med.* **371**, 2488–2498 (2014).
4. Z. Tothova *et al*., *Cell Stem Cell* **21**, 547–555.e8 (2017).
5. X. Shi *et al*., *Exp. Hematol.* **64**, 33–44.e5 (2018).
6. J. Henninger *et al*., *Nat. Cell Biol.* **19**, 17–27 (2017).
7. R. Sood *et al*., *Blood* **115**, 2806–2809 (2010).
8. H. A. David, *Biometrika* **55**, 573–575 (1968).
9. D. Traver *et al*., *Nat. Immunol.* **4**, 1238–1246 (2003).
10. J. Ablain, E. M. Durand, S. Yang, Y. Zhou, L. I. Zon, *Dev. Cell* **32**, 756–764 (2015).
11. A. Alemany, M. Florescu, C. S. Baron, J. Peterson-Maduro, A. van Oudenaarden, *Nature* **556**, 108–112 (2018).
12. F. E. Moore *et al*., *J. Exp. Med.* **213**, 979–992 (2016).
13. S. Jaiswal *et al*., *N. Engl. J. Med.* **377**, 111–121 (2017).
14. A. G. Bick *et al*., *Nature* **586**, 763–768 (2020).
15. K. A. Matatall *et al*., *Cell Rep.* **17**, 2584–2595 (2016).
16. D. Hormaechea-Agulla *et al*., *Cell Stem Cell* **28**, 1428–1442.e6 (2021).
17. D. C. Bellissimo *et al*., *Blood Adv.* **4**, 1145–1158 (2020).
18. S. Safe *et al*., *J. Steroid Biochem. Mol. Biol.* **157**, 48–60 (2016).


## ACKNOWLEDGMENTS

We thank P. Liu for the *runx1* mutant zebrafish, J. Mito for assistance with planning the scRNA-seq experiment with TWISTR zebrafish, B. Gheller for assistance with GSEA, and colleagues for critical reading of the manuscript. **Funding:** NIH grants P01HL131477 (L.I.Z.) and 5F31HL126338-03 (J.E.H.), Pedal for Pediatrics (S.A.), and Alex's Lemonade Stand Foundation and Babich Family Foundation Familial RUNX1 grant (L.I.Z.) L.I.Z. is a Howard Hughes Medical Institute Investigator. **Author contributions:** Conceptualization: S.A., J.E.H., W.P.M., L.I.Z.; methodology: S.A., J.E.H., W.P.M.; investigation: S.A., J.E.H., W.P.M., M.C.W., J.L.M.; visualization: S.A., J.E.H.; funding acquisition: S.A., J.E.H., L.I.Z.; project administration: S.A., J.E.H.; resources: S.Y.; software: M.M., J.Y.; supervision: S.A., J.E.H., L.I.Z.; writing–original draft: S.A., J.E.H.; writing–review and editing: S.A., J.E.H., L.I.Z. All authors commented on the manuscript. **Competing interests:** L.I.Z. is a founder and stockholder of Fate Therapeutics, CAMP4Therapeutics, and Scholar Rock. All other authors declare that they have no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. The scRNA-seq data have been deposited in GEO under accession code GSE176037.

## SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.aba9304
Materials and Methods
Figs. S1 to S10
Tables S1 and S2
References (*19*–*33*)

16 January 2020; accepted 20 September 2021
10.1126/science.aba9304




IMAGE: © PX MEDIA/SHUTTERSTOCK.COM

# How to begin building a culture of diversity, equity, and inclusion in your research group

*A strategic plan for your team should include more than just the infrastructure to achieve your research agenda: It must include a thoughtful approach to crafting a culture that honors the principles of diversity, equity, and inclusion. In doing so, you create the means for science to advance—and everyone to triumph.* **By Alaina G. Levine**

A lawnmower served as the vehicle to illustrate to environmental engineer **Sharon Walker** (she/her) the transformative power of investing in diversity, equity, and inclusion (DEI) practices and culture. Walker, dean of the College of Engineering at Drexel University, Philadelphia, Pennsylvania, was mentoring a student, the daughter of Mexican immigrants and whose family members worked as gardeners and landscapers. She spotted a problem: Her relatives had developed lung problems as a result of breathing in the emissions from common lawnmowing equipment. "Her goal as an environmental engineer was to protect them," recalls Walker. For her senior design project, the first-generation college student invented a catalytic converter that used non-precious metals, significantly reducing emissions that are inhaled by the users of landscape equipment typically used for "the mow and blow." The student took the project further, applying for patent protection, and won an award from the U.S. Environmental Protection Agency. She ended up landing roles at Google and a municipal water district. "This is why you need people with different experiences in your lab: to make sure we answer questions for everyone who matters, not just faculty in elite institutions," says Walker. "We are missing innovation."

The experience of Walker and her student is a commanding example of why DEI is so important and relevant in STEM. But DEI initiatives are not limited to the overall field; it is something that individual faculty members can facilitate in their own spaces. While there are plentiful reference materials guiding new professors as to how they can establish research infrastructure and collaborations necessary to their endeavors, there are less resources as to the why and how of interweaving DEI principles into a research group's strategic plan. Experts argue this has to change, for the sake of science and scientists. "The best ideas come in a diverse and inclusive workspace," says **Arti Agrawal** (she/her), associate adjunct professor in the School of Electrical and Data Engineering at the University of Technology Sydney in Australia. Studies have shown that "we get improved innovation outcomes when we have diversity attained," says **Bree Gorman** (they/them), managing director of Bree Gorman Consulting, a Geelong, Australia-based firm which works with organizations to create DEI initiatives that are sustainable, realistic, and achievable. "It's a competitive advantage that you are missing out on if your team is not diverse."

That advantage is a serious concern. "It goes back to a core definition of who engineers are and what we do," says **Stephanie G. Adams** (she/her), dean of the Erik Jonsson School of Engineering and Computer Science, Lars Magnus Ericsson Chair in Electrical Engineering, and professor of systems engineering at the University of Texas at Dallas. "You have to design them for all people—not just white people or women or x. It can only be as effective as the data you pour into it, and if your data is one dimensional you have problems." We have seen this outcome, she notes, with AI systems, which have been shown to have significant errors when it comes to facial recognition of people of color.

An inclusive, equitable culture that engenders a diversity of perspectives is what can alleviate these pain points and cont.>

can improve how your lab's science is viewed. "By not creating more inclusive lab spaces, you risk becoming illegitimate in the eyes of those who will be seeking these opportunities in the future," says **Donald A. Outing** (he/him), a mathematician and vice president for equity and community at Lehigh University in Bethlehem, Pennsylvania, and the former inaugural chief diversity officer at the U.S. Military Academy, West Point, New York.

But there is another reason to enact a DEI strategy for your group: "Because it's the right thing," says **Omar A. Quintero-Carmona** (he/him), associate professor of biology at the University of Richmond in Virginia. "Hiring practices that work to broaden the applicant pool, giving consideration to as many people as possible, increases your ability to identify talent that you might have missed otherwise. Equally important is developing mentoring strategies to aid talent to transitioning into a new environment in a way that supports their success." **Asia Eaton** (she/her), associate professor of psychology at Florida International University in Miami, who studies how gender interacts with other identifiers such as race to affect experiences and access to power, agrees: "For anyone who has their own lab in the privileged position to mentor grad students, it is their personal responsibility to change the face of science," she argues.



Stephanie G. Adams (**left**), Bree Gorman (**right**)

## When best intentions fall short: Challenges

While many faculty have excellent intentions, they face challenges in orchestrating a sustainable DEI strategy. The first area for improvement is related to a cornerstone of DEI: putting yourself in others' shoes and looking for the barriers, gaps, and roadblocks that are negatively impacting their success. "It takes a real sense of self-awareness to see where your blind spots are located and that you are not setting everyone up for success," says Walker.

Part of that trepidation is tethered to how academia views DEI. While more and more universities publicly tweet that DEI is vital, there are not as many accountability, measurement, and reward systems in place for DEI practices. "DEI is still not part of the day-to-day reality of professors," says Agrawal. "The pressure on them is to produce papers and get research grants or to teach. There is no room or space or time or resources [for DEI] and there are very few people doing it." Without policies and procedures in place to check and assess DEI initiatives, they falter, and professors lose out on a means to fortify their own research agenda and team.

## Educating yourself: Strategies

Even if you don't have a doctorate in diversity, fortunately, there are a number of mechanisms and resources to school yourself on how to spearhead a strategy that fosters DEI in your team. "First, agree there is an issue," says **Jessica Harris**, she/her, a physicist and DEI consultant who helps organizations address cultural change through her business, Jessica A. Harris, LLC. "If the team doesn't recognize a problem, they won't be able to address it." "When building a lab, you're thinking about the equipment and supporting infrastructure to allow the people you recruit to push the science forward," says Quintero-Carmona. "But a different kind of infrastructure has to exist that supports lab members' success and growth. Sustained investment in your people is as important as having working equipment." Gorman sees it as a leadership exercise. "It's about the leader recognizing their responsibility to give everybody equal opportunity, to [help] every one of your team members to achieve their career aspirations no matter who they are," they say. "If you take that approach as a leader you will begin to see the barriers" that individuals face that can translate into bottlenecks for the team and project.

"It starts with educating yourself always," adds Outing. "Discuss with students how diversity can enhance an education, how examining the value of diversity of ideas and hypotheses and strategies is important to the world." He also suggests that faculty examine their implicit biases through assessments such as the Harvard Implicit Association Tests. "Take a reality check of where are you," says Adams, "identifying your own shortcomings, biases, myths, and realities." She describes her own experience when she needed some tables moved and announced that she needed some "guys" to help her. "A student called me on it, and it caused me to look at pronouns and differences on the spectrum. Until I can own [my mistakes] I can't begin to change."

You can often access training programs and resources within your institutions and professional societies. Biologist and diversity consultant **Lorenza D'Alessandro** (she/her) encourages faculty to take training in "intercultural communications and awareness on biases, their effects at work and consequences also in regards to harassment and discrimination," she notes. "By being better informed, [faculty] could install a cultural change in their groups by encouraging open and respectful communication with a deeper understanding of diversity."

The training can assist a professor in exploring exactly what they don't know and becoming comfortable with the uncomfortable questions and conversations. "One of the first steps is to be serious about understanding the institutional and structural barriers that have excluded and continue to exclude people from marginalized communities," says **Lyndele von Schill** (she/her), director of diversity and inclusion at the National Radio Astronomy Observatory in Charlottesville, Virginia. "If we don't explore that, we will run into barriers that exist for other people because we don't see them."

cont.>

PHOTOS: STEPHANIE ADAMS COURTESY OF UT DALLAS; BREE GORMAN BY 5-12 IMAGING

VAN ANDEL INSTITUTE®



## CANCER EPIGENETICS TRAINING PROGRAM

VAI's Cancer Epigenetics Training (CET) Program prepares postdoctoral fellows to become independent research leaders in the field of cancer epigenetics. The program offers:

- **Interdisciplinary** research training in cancer epigenetics
- Access to state-of-the-art Core Technologies and Services
- Extensive professional development
- Integration with established translational research programs and partnerships led by VAI investigators (such as the Van Andel Institute–Stand Up To Cancer Epigenetics Dream Team)

**Eligibility:** Applicants must be within 12 months of completing a Ph.D., M.D., or other appropriate terminal degree. Applicants must be U.S. citizens or permanent residents. Individuals who identify as members of an underrepresented racial and/or ethnic group (as defined by the National Institutes of Health) are strongly encouraged to apply.

**Learn more and apply at https://bit.ly/VAICETP**

*The Cancer Epigenetics Training Program is supported by a National Cancer Institute T32 training grant (no. T32CA251066). The content is solely the responsibility of VAI and does not necessarily represent the official views of the National Institutes of Health.*

## INSPIRE FELLOWSHIP

Diverse backgrounds, expertise and viewpoints are foundational to exceptional science and innovation. VAI's Inspire Fellowship is designed to increase representation at VAI and in the broader biomedical research community by supporting postdoctoral fellows who identify as members of underrepresented racial and ethnic groups (as defined by the National Institutes of Health). The fellowship offers:

- Extensive research training in the Institute's labs
- Access to state-of-the-art Core Technologies and Services
- Fellowship-specific professional development activities

**Eligibility:** Applicants must be within 12 months of completing a Ph.D., M.D., or other appropriate terminal degree. Applicants must be U.S. citizens or permanent residents who identify as members of an underrepresented racial and/or ethnic group (as defined by the National Institutes of Health).

**Learn more and apply at https://bit.ly/InspireVAI**

**Questions? Please contact Dr. Erica Gobrogge, program director of VAI's Office of Postdoctoral Affairs, at opa@vai.org**

## Featured participants

**Bree Gorman Consulting**
breegorman.com

**Drexel University**
drexel.edu

**Florida International University**
www.fiu.edu

**JA Harris, LLC**
www.jessicaaharris.com

**Lehigh University**
www1.lehigh.edu

**Miami University**
miamioh.edu

**National Radio Astronomy Observatory**
public.nrao.edu

**University of Richmond**
www.richmond.edu

**University of Technology Sydney**
www.uts.edu.au

**The University of Texas at Dallas**
www.utdallas.edu

Adams, who has been involved in DEI efforts for years, offers a list of how to structure your DEI strategy, beginning with mission, values, and goals, and then move into accountability and consequences. "A PI should start by defining a vision for their lab. They should decide what DEI looks like in that space. Do they want representation that looks like society?" she says. "If so, then devise a plan to build that vision. How are we going to get there and how will we know we are there?"

Scientists are well positioned to create their own DEI strategy because many of the steps mirror that of the scientific problem solving process, including hypothesis, data collection, and experimentation. "Create the code of conduct first—it's how are we agree to treat each other," says Harris. The code should include inclusive language, a task which presents its own challenges. "Trying to write them well in an inclusive way is hard," admits Walker. "It takes expertise to not say or canonize things you probably shouldn't." **Lesley Chow** (she/her), Frank Hook Assistant Professor of materials science and bioengineering at Lehigh University, intentionally plasters her lab ethos on her website. "Love is love, immigrants are welcome, Black Lives Matter, feminism for everyone," she says. Her intention in explicitly stating this is to create a safe space, where each individual feels heard and seen and respected. "I talk very openly with my lab. By talking about things, it validates it. I also tell my group if something happens [to you], tell me so I can stand up and fight for [you]." Adds Eaton: "You have to be adaptable, creative, and vulnerable; tell them about your own struggles. It helps build trust, cohesion, and authenticity."

A safe space can encompass many attributes. It should include a diversity of trainees and staff, and the means to support them, advises Harris. It can feature cultural elements such as reasonable work hours and mechanisms for feedback. "Part of the DEI plan should also include support for families for childcare facilities and people of trust to whom faculty members can openly report," says D'Alessandro. "Monitoring and transparent reporting of data on DEI is among the key elements to identify where measures need to be modified or newly introduced and therefore to make progress. Preferably, the DEI plan should be prepared by a diverse team and importantly it should reflect the real values of the faculty."

## Walking the walk: Tactics

There are many tactics to craft your strategic plan. Certainly, a major component is committing to using inclusive language. "Asking pronouns and accommodations is a baseline must, but it is more about getting to know them individually," says Eaton. Introduce your team to diverse sources in your reference material, syllabus, and networks. "If you look at our textbooks, it has always been a celebration of white, straight men," says **Beena Sukumaran**, she/her, professor of mechanical and manufacturing engineering and dean of the College of Engineering and Computing at Miami University in Oxford, Ohio. "Provide aspirational role models highlighting people from diverse backgrounds in the professions."

Inclusivity is aligned with accessibility, something that Agrawal recognizes. In fact, she coaches her proteges to be thoughtful when preparing presentations to ensure they are as accessible as possible. Can those who are visually impaired, dyslexic, or neurodiverse follow it? How about at the conference itself: Will there be a sign language interpreter at the meeting? Can someone in a wheelchair maneuver in the space? These are some of the queries she challenges her students and postdocs to consider.

There is also team and personal accountability. "I have discontinued faculty searches when I see the applicant pool is not diverse. Shame on them," says Walker. "We need to see good behaviors in all our practices and policies and when we don't see it, we have to call it out to not reward those behaviors."

## Do the just thing

It may not be easy to topple the status quo, or to stand up for righteousness, especially in the context of our careers. But "creating inclusive spaces is not a fad. The alternative is an injustice," says von Schill. "We have to be really willing to do some work and look at cherished ideas about STEM and our educational systems." Adds Chow: "If you are not purposeful and genuine about it, it's not going to work. But if you choose to not do it, you will be left behind."

Furthermore, DEI is more than a framework. It's a state of mind. "It's about what you do day-to-day," says Eaton, "and how you're showing support for DEI in your field more so than any proclamation you might make or contract you might have someone sign."

DEI is tough territory for the uninitiated but it is not outside the boundaries of possibilities for STEM professors to construct and execute a research strategy entwined with DEI principles that support and nurture trainee triumph. "As an inclusive leader, your role is to amplify their voice and provide them the platform to speak," advises Gorman. "This is difficult for people who hold privilege and power. They might think it is easy to jump in and fix all the problems but it is really more about listening and empowering."

And when you do listen, pay attention to the outcomes of your DEI measures, too. Notes Walker: "The benefits are amazing in the long term and your lab will flourish."

*Alaina G. Levine is a science writer, professor speaker, and author of* Networking for Nerds *(Wiley, 2005).*

## FOCUS ON DIVERSITY AND INCLUSION



UNIVERSITY OF MINNESOTA
Driven to Discover℠
Crookston Duluth Morris Rochester Twin Cities

### Assistant Professor in Cell and Development Biology

The Department of Genetics, Cell Biology and Development (GCD) at the University of Minnesota – Twin-Cities has two tenure-track openings at the Assistant Professor level. We seek applicants in the broad area of cellular & developmental biology. We are particularly interested in applicants whose research programs combine approaches from biology and engineering disciplines and utilize emerging technologies to study cellular and developmental processes and how they go awry in disease.

Research topics of interest include, but are not limited to, a study of the biophysical mechanisms that guide cell shape change and collective movements during development, metabolic control of developmental processes, matrix-cell interactions during gastrulation or organ formation, and the cell biology of signal transmission and reception in developing systems. Individuals studying systems principles such as self-organization, cellular decision making, or cellular cooperation, as applied to translational medicine (cell therapy, organ development, or understanding the role of tissue structure in disease etiology) are also encouraged to apply.

More information, required qualifications, and applications instructions can be found at **https://apptrkr.com/2590934**

## POSITIONS OPEN

Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences



Established in partnership between the Chinese Academy of Sciences and the Shenzhen Municipal Government, the Shenzhen Institute of Advanced Technology (SIAT) is a newly-created university with an objective to become the world's preeminent institute for emerging science and engineering programs. SIAT is equipped with state-of-art teaching and research facilities and is dedicated to cultivating international, visionary, and interdisciplinary talents while delivering research support to pursue innovation-driven development.

SIAT is located in Shenzhen, also known as the "Silicon Valley of China," a modern, clean, and green city, well-known for its stunning architecture, vibrant economy, and its status as a leading global technology hub. SIAT is seeking applications for faculty positions of all ranks in the following academic programs: Computer Science and Engineering, Bioinformatics, Robotics, Life Sciences, Material Science and Engineering, Biomedical Engineering, Pharmaceutical Sciences, Synthetic Biology, Neurosciences, etc. SIAT seeks individuals with a strong record of scholarship who possess the ability to develop and lead high-quality teaching and research programs. SIAT offers a comprehensive benefits package and is committed to faculty success throughout the academic career trajectory, providing support for ambitious and world-class research projects and innovative, interactive teaching methods.

Further information:
http://english.siat.cas.cn/DP/Va/

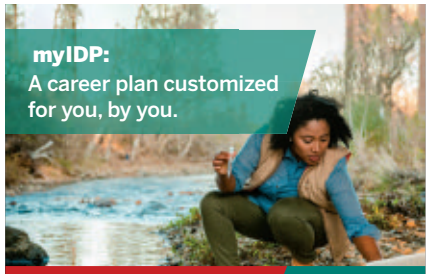
ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

Confused about your next career move?

Download Free Career Advice Booklets!
ScienceCareers.org/booklets








There's only one Science

**Features in myIDP include:**

- Exercises to help you examine your skills, interests, and values.
- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

myIDP
Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers AAAS In partnership with:

UCSF University of California San Francisco
UIC
BURROUGHS WELLCOME FUND
University of Massachusetts Medical School UMASS
FASEB Federation of American Societies for Experimental Biology

## POSITIONS OPEN

**Faculty Positions**
**Center for Cardiovascular Research**

The Abigail Wexner Research Institute at Nationwide Children's Hospital seeks an outstanding faculty candidate to lead a research program in the Center for Cardiovascular Research.

A successful candidate for any rank (tenure track assistant, associate or full professor):

- currently directs a research laboratory focused on pediatric cardiovascular disease and/or relevant basic science areas,
- has an MD, MD/PhD, PhD, or equivalent degree,
- has a strong track record of translational research and publication in quality journals and
- has current federal grant funding.

Principal investigators are also faculty members at The Ohio State University College of Medicine.

Inquires and referrals welcome.

To learn more about the opportunity and apply, visit the Careers tab on **NationwideChildrens.org/CCVR.**

NATIONWIDE CHILDREN'S
When your child needs a hospital, everything matters.
THE OHIO STATE UNIVERSITY
COLLEGE OF MEDICINE

# YOU'D SMILE TOO IF YOU JUST ADVANCED YOUR CAREER.



Find your next job at ScienceCareers.org

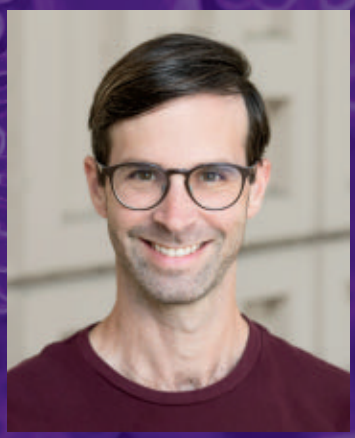
There's scientific proof that when you're happy with what you do, you're better at what you do. Access career opportunities, see who's hiring and take advantage of our proprietary career-search tools. Get tailored job alerts, post your resume and manage your applications all in one place: sciencecareers.org

ScienceCareers

FROM THE JOURNAL SCIENCE AAAS

# TRILLIONS OF MICROBES ONE ESSAY

**The NOSTER *Science* Microbiome Prize is an international prize that rewards innovative research by investigators, under the age of 35, who are working on the functional attributes of the microbiota. The research can include any organism that has potential to contribute to our understanding of human or veterinary health and disease, or to guide therapeutic interventions. The winner and finalists will be chosen by a committee of independent scientist, chaired by a senior editor of *Science*. The top prize includes a complimentary membership to AAAS, an online subscription to *Science*, and $25,000 (USD). Submit you research essay today.**



Eran Blacher, Ph.D.
2021 Winner

NOSTER | Science
MICROBIOME
PRIZE

**Apply by 1/24/22 at www.science.org/noster**

**Sponsored by Noster, Inc**

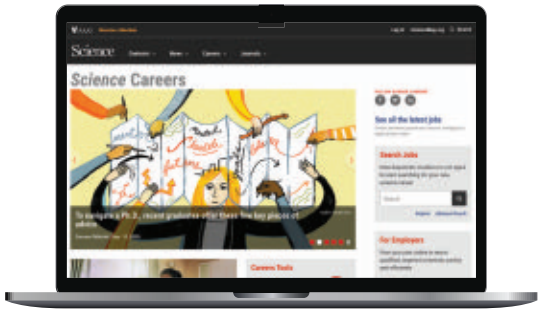
Science Careers online @sciencecareers.org




# *Science* Careers helps you advance your career. Learn how!

- Register for a free online account on **ScienceCareers.org**.
- Search hundreds of job postings and find your perfect job.
- Sign up to receive e-mail alerts about job postings that match your criteria.
- Upload your resume into our database and connect with employers.
- Watch one of our many webinars on different career topics such as job searching, networking, and more.
- Download our career booklets, including Career Basics, Careers Beyond the Bench, and Developing Your Skills.
- Complete an interactive, personalized career plan at "my IDP."
- Visit our Employer Profiles to learn more about prospective employers.
- Read relevant career advice articles from our library of thousands.

Visit **ScienceCareers.org** today — all resources are free




SCIENCECAREERS.ORG

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

# Who's the top employer for 2021?



Science 2021 TOP EMPLOYER

*Science* Careers' annual survey reveals the top companies in biotech & pharma voted on by *Science* readers.

Read the article and employer profiles at sciencecareers.org/topemployers

*By* **Julia Behnfeldt**

# Keeping science honest

I was sitting at my computer, poring over files my office had sequestered from a professor's lab. We'd received an anonymous tip alleging the lab had published fraudulent data. As a research integrity officer, I was tasked with reviewing the data files, so I clicked on an image depicting the results of a Western blot. Sure enough, it appeared to have been digitally manipulated: Parts had been cropped and moved around in a manner that obscured the true results. The file's metadata held a record of who generated the altered image and when, a key clue in what ended up being a massive case of research misconduct. These weren't the tasks I would have ever imagined doing when I started graduate school. But it's the right career for me, and I owe it to a chance assignment during a seminar class.

The professor teaching the course assigned us publications to present, and when I read mine, I noticed some of the Western blot images had been reused and relabeled in multiple figures. As a fourth year Ph.D. student, I had spent years producing those kinds of images and I knew something was wrong.

I Googled "what to do if I suspect research misconduct" and landed on a web page that suggested I reach out to a trusted mentor or to my institution's research integrity officer—a completely foreign job title to me at the time. I decided to start with the professor teaching the seminar. I stumbled over my words as I tried to explain my concerns. He reassured me that my suspicions were justified, adding that he'd report them to the journal. A few months later the journal posted a corrigendum, with new figures in place of the old, suspicious ones.



**"It was hard for me to believe the original figures in the paper were produced by accident."**

Before this episode, I'd had no experience with wrongdoing in science; my collaborators were all careful scientists, and I trusted that the papers I was reading by others were sound. But it was hard for me to believe the original figures in the paper were produced by accident. Science wasn't always honest, I realized—a sense that grew when I read a slew of articles about scientists found guilty of misconduct.

It was clear to me that investigating these incidents was important work; it helped strike fraudulent science from the literature and clear the names of innocent people—co-authors, grad students—whose careers were affected by the actions of others. I noticed many people investigating the incidents held Ph.D.s. The more I read and learned, the more appealing a career in research integrity became.

I had already been exploring nonresearch careers, as the grind of lab work and manuscript writing had left me deflated. Reading about science still invigorated me, but working nights and weekends did not. I wanted a 9-to-5 job that would keep me engaged in science and utilize skills I'd developed during grad school.

As I was nearing graduation, I applied for a fellowship that gave Ph.D. graduates an opportunity to work for the federal government. I hoped for a placement within the U.S. Office of Research Integrity—and that's exactly what I got.

It was a good fit. I liked that my colleagues were all scientists. I enjoyed learning about methods for detecting falsified data. And I found my mind being challenged by the volume of details included in investigation reports, which read like true crime books.

Since then, I've worked in research integrity offices at two universities. I spend my days looking into misconduct allegations, assisting investigative committees, and developing training materials designed to prevent misconduct. The job can be challenging at times. For instance, it's hard to tell scientists they're under investigation, and I find it disheartening to uncover evidence of misconduct and the various reasons behind it—including the pressure to publish and toxic lab environments.

But overall, I enjoy the path I'm on. The work is varied and intellectually stimulating. I appreciate having my evenings and weekends free. And I like seeing positive outcomes of my work. In the case of the altered Western blot images I analyzed, for example, the clues I uncovered helped exonerate a grad student and zero in on their adviser. To me, it's research—a different but still fulfilling kind. ■

Julia Behnfeldt is an assistant research integrity officer at the University of Michigan. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER

CALL FOR PAPERS

SPJ
SCIENCE PARTNER JOURNALS

PEKING UNIVERSITY 1898



# Health Data Science

*Health Data Science* is an online-only Open Access Science Partner Journal published in affiliation with **Peking University (PKU)** and distributed by the **American Association for the Advancement of Science (AAAS)**. This journal is committed to promoting the appropriate application of cutting-edge technologies and analytic approaches for the health field, translating the value of health data into practice for the benefit of patients' health and populations' well-being and advancing health data science. *Health Data Science* publishes innovative, scientifically-rigorous original research, comprehensive reviews, insightful perspectives, viewpoints and editorials. The topics of interest include but are not limited to: data management, governance and provenance; comparative effectiveness research; ethics issues; public health and health policies; and more.

**Submit your research to *Health Data Science* today!**

Learn more at **spj.sciencemag.org/hds**

The Science Partner Journal (SPJ) program was established by the American Association for the Advancement of Science (AAAS), the nonprofit publisher of the *Science* family of journals. The SPJ program features high-quality, online-only, Open Access publications produced in collaboration with international research institutions, foundations, funders and societies. Through these collaborations, AAAS furthers its mission to communicate science broadly and for the benefit of all people by providing top-tier international research organizations with the technology, visibility, and publishing expertise that AAAS is uniquely positioned to offer as the world's largest general science membership society.

Visit us at **spj.sciencemag.org**

**@SPJournals** **@SPJournals**

**OPEN ACCESS**

**ARTICLE PROCESSING CHARGES WAIVED THROUGH DECEMBER 2024**

# THE 2022 LASKER AWARDS

## Call for Nominations

Submit a nomination online at laskerfoundation.org

**Deadline: February 7, 2022**

**Basic Medical Research**
**Clinical Research**
**Public Service**

*Nominations of women and minorities are encouraged.*

## Congratulations to the 2021 Lasker Award Winners

**Basic Research Award**
Light-Sensitive Microbial Proteins and Optogenetics

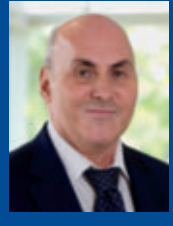

DIETER OESTERHELT


PETER HEGEMANN


KARL DEISSEROTH

**Clinical Research Award**
Modified mRNA Vaccines


KATALIN KARIKÓ


DREW WEISSMAN


DAVID BALTIMORE

**Special Achievement Award**
Fundamental Discoveries, Academic Leadership, and Public Advocacy

LASKER FOUNDATION

CELEBRATING 75 YEARS