

Should race be a factor in
medical risk calculators? p. 380

Probing psychological understandings
of legal constructs p. 394

Cranial immune cells home in
on the brain pp. 396, 408, & 409

Science

\$15
23 JULY 2021
sciencemag.org

 AAAS



PLANETARY ANATOMY

Seismology on Mars reveals details
of planet's interior pp. 388 & 434-448



Frontiers
of Science

Gordon Research Conferences

Everyone Remembers Their First GRC! October-November 2021

“I remember my first GRC experience because I was a second-year PhD student and was introduced to an excellent community. It was an eye-opening experience to see that the leading figures in the field were approachable.”

DR. HOLGER WINKELS, Past Chair of the Atherosclerosis GRS

Alcohol-Induced End Organ Diseases

+ GRS

OCTOBER 24 – 29, 2021
FOUR POINTS SHERATON / HOLIDAY INN EXPRESS,
VENTURA, CA, US

CHAIRS: Gyongyi Szabo and Bernd Schnabl
VICE CHAIRS: Elizabeth Kovacs and Gavin Arteel

Biotherapeutics and Vaccines Development

+ GRS

NOVEMBER 14 – 19, 2021
VENTURA BEACH MARRIOTT, VENTURA, CA, US

CHAIRS: Michael Tarlov and Linda Narhi
VICE CHAIRS: Peter Tessier and Jessica Molek

Cannabinoid Function in the CNS

+ GRS

NOVEMBER 7 – 12, 2021
FOUR POINTS SHERATON / HOLIDAY INN EXPRESS,
VENTURA, CA, US

CHAIRS: Matthew Hill and Sachin Patel
VICE CHAIRS: Miriam Melis and Mario Stelt

Carbon Capture, Utilization and Storage

+ GRS

OCTOBER 24 – 29, 2021
JORDAN HOTEL AT SUNDAY RIVER, NEWRY, ME, US

CHAIR: Roger Aines
VICE CHAIR: Andre Bardow

Cardiac Arrhythmia Mechanisms

+ GRS

OCTOBER 17 – 22, 2021
FOUR POINTS SHERATON / HOLIDAY INN EXPRESS,
VENTURA, CA, US

CHAIRS: David Christini and Ursula Ravens
VICE CHAIRS: Crystal Ripplinger and Edward Vigmond

Cell Biology of Metals

+ GRS

OCTOBER 17 – 22, 2021
MOUNT SNOW, WEST DOVER, VT, US
CHAIRS: Caryn Outten and Mitchell Knutson
VICE CHAIRS: Angela Wilks and Kerry Kornfeld

Central Nervous System Injury and Repair

+ GRS

OCTOBER 17 – 22, 2021
JORDAN HOTEL AT SUNDAY RIVER, NEWRY, ME, US
CHAIRS: Yimin Zou and Simone Di Giovanni
VICE CHAIRS: Elizabeth Bradbury and Yishi Jin

Chronobiology

+ GRS

OCTOBER 10 – 15, 2021
MOUNT SNOW, WEST DOVER, VT, US
CHAIR: Samer Hattar
VICE CHAIR: Charlotte Helfrich-Forster

Computer Aided Drug Design

+ GRS

NOVEMBER 7 – 12, 2021
MOUNT SNOW, WEST DOVER, VT, US
CHAIR: Georgia McGaughey
VICE CHAIR: Matthias Rarey

Lung Development, Injury and Repair

+ GRS

NOVEMBER 7 – 12, 2021
WATERVILLE VALLEY, WATERVILLE VALLEY, NH, US
CHAIRS: Anne-Karina Perl and Daniel Tschumperlin
VICE CHAIRS: Rory Morty and Xin

Mammalian DNA Repair

+ GRS

OCTOBER 31 – NOVEMBER 5, 2021
FOUR POINTS SHERATON / HOLIDAY INN EXPRESS,
VENTURA, CA, US
CHAIR: Lee Zou
VICE CHAIR: Patricia Opresto

Medicinal Chemistry

+ GRS

OCTOBER 24 – 29, 2021
MOUNT SNOW, WEST DOVER, VT, US
CHAIR: Janeta Popovici-Muller
VICE CHAIR: Cynthia Parrish

Phagocytes

+ GRS

OCTOBER 24 – 29, 2021
WATERVILLE VALLEY, WATERVILLE VALLEY, NH, US
CHAIR: Michelle Lennartz
VICE CHAIR: Sergio Catz

Physical Science of Cancer

+ GRS

NOVEMBER 14 – 19, 2021
FOUR POINTS SHERATON / HOLIDAY INN EXPRESS,
VENTURA, CA, US
CHAIRS: Josef Käs and Sharon Gerecht
VICE CHAIRS: David Odde and Jennifer Munson

Stochastic Physics in Biology

+ GRS

OCTOBER 10 – 15, 2021
FOUR POINTS SHERATON / HOLIDAY INN EXPRESS,
VENTURA, CA, US
CHAIR: Yuhai Tu
VICE CHAIR: Jie Xiao

Stress Proteins in Growth, Development and Disease

+ GRS

NOVEMBER 7 – 12, 2021
JORDAN HOTEL AT SUNDAY RIVER, NEWRY, ME, US
CHAIR: Brian Freeman
VICE CHAIR: Harm Kampinga

Three Dimensional Electron Microscopy

+ GRS

OCTOBER 31 – NOVEMBER 5, 2021
WATERVILLE VALLEY, WATERVILLE VALLEY, NH, US
CHAIR: Sharon Wolf
VICE CHAIR: Daniela Nicastro

Apply now at www.grc.org/

Making new connections through spatial biology

The rapid development of analytical cellular tools has had a groundbreaking effect on scientific discovery in recent decades. Using newly available technologies such as next-generation sequencing (NGS), scientists have been racing to work out how genetic information can help to create treatments for major diseases, such as cancer. The advent of single-cell RNA sequencing has made it possible to analyze the transcriptomes of individual cells.

Yet single-cell technologies can only tell scientists so much, says immunologist Charlotte Scott of the VIB-UGent Center for Inflammation Research in Ghent, Belgium. Newly developed spatial biology technologies are the missing piece of the puzzle. Allowing specific cell types to be mapped locally within the tissue enables the investigation of biologically relevant cell-cell interactions based on spatial proximity, she says.

Scott's work focuses on understanding the specific functions of macrophages, which are large white blood cells in the liver. "To fully understand these cells, we need a thorough understanding of both their heterogeneity and the local microenvironments of each subtype," she says.

The advent of spatial biology was enabled by iterative rounds of technology innovation and scientific discovery, according to a recent white paper (resolve-biosciences.com/science) published by Resolve Biosciences. Affordable predecessor technologies such as microscopy and NGS showed the value of capturing the heterogeneity of cell populations, but did not provide in situ context. Spatial biology was the natural next step. This technology pairs the benefits of single-cell analysis with a view of natural tissue architecture based on recent advances in imaging technology, according to the report.

Spatial biology in itself is not a new concept, says Scott. "Researchers have been using confocal microscopy and immunohistochemistry to locate their cells of interest for decades. However, this has always been limited in terms of the number of parameters that could be assessed simultaneously. This has prevented the study of their interactions with other cells in the same microenvironment," she explains.

Scott and her colleague Martin Guillems at VIB-UGent have been working together on the Liver Cell Atlas project to generate the first spatial proteogenomics atlas of healthy and diseased livers in mice and humans.

He believes their project will facilitate the high-throughput screening of a whole range of transgenic mouse models and large patient cohorts, which will lead to a better understanding of how these cells are affected in disease without the need for expensive single-cell RNA sequencing.

Different spatial technologies have been benchmarked at VIB thanks to the institute's initiatives, such as Tech Watch and the Single-Cell Accelerator program. These advancements allow the Scott and Guillems teams to select the best techniques and combine complementary spatial transcriptomics approaches to unravel the location of different macrophage subsets in specific microenvironments. First, the team determines cellular distribution within the tissue and narrows down the cells of interest in a specific zone. They then use Resolve Biosciences' Molecular Cartography platform to study the most relevant cell-cell interactions within that zone at single-cell resolution.

Scott and Guillems say that in recent years it has become clear that the local environment in which a cell resides greatly influences its phenotype and function. Thus, by understanding the spatial arrangement of cells in a tissue, scientists stand to gain significant insight into the factors regulating the functional specialization of cells and may ultimately identify the local cues that control cell activation.

They predict that further advances in spatial biology techniques, such as the combination of protein and RNA analyses, could help scientists find ways to control the development, activation, or inhibition of cells that play fundamental roles in human diseases.

Sponsored by



CONTENTS

23 JULY 2021 • VOLUME 373
ISSUE 6553

380

NEWS

IN BRIEF

370 News at a glance

IN DEPTH

372 Europe's deadly floods leave scientists stunned

Despite improvements, flood forecasts sometimes failed to flag risks along smaller streams *By W. Cornwall*

373 Alzheimer's drug approval spotlights blood tests

Physicians hope technology will help screen patients eligible for Biogen's controversial antibody, aducanumab *By K. Servick*

PODCAST



372

374 A few lucky researchers return to the field

The pandemic has transformed fieldwork and upended careers
By E. Pennisi

376 DNA plucked from air identifies nearby animals

Analysis of airborne cells could survey biodiversity, two studies of zoo animals suggest
By E. Stokstad

377 Event Horizon Telescope images second black hole's jets

Globe-spanning instrument's snapshot of Centaurus A galaxy suggests all black holes operate similarly *By D. Clery*

378 WHO chief pressures China on pandemic origin

Agency director calls for lab audits and more studies of how SARS-CoV-2 emerged *By J. Cohen*

FEATURES

380 A troubled calculus

Researchers use race to build disease risk assessment tools. Can removing it help resolve medicine's race crisis?
By J. Madhusoodanan

INSIGHTS

POLICY FORUM

384 Global implications of the EU battery regulation

A much-needed law may have unintended global consequences
By H. E. Melin et al.

PERSPECTIVES

388 The interior of Mars revealed

Direct seismic observations provide clues to the red planet's structure and evolution
By S. Cottaar and P. Koelemeijer
REPORTS pp. 434, 438, & 443

390 Coherent manipulation of a spin qubit

A new device opens new avenues for spin qubits
By G. Wendin and V. Shumeiko
REPORT p. 430

391 The inner workings of an enzyme

A high-throughput mutation screen dissects the mechanistic basis of enzyme activity
By Z. T. Baumer and T. A. Whitehead
RESEARCH ARTICLE p. 411

392 Proximity and single-molecule energetics

Scanning probes measure how nearby oxygen molecules affect triplet lifetimes of pentacene *By L. Li and N. Jiang*
REPORT p. 452

394 Experimental jurisprudence

Psychologists probe lay understandings of legal constructs
By R. Sommers

396 Bespoke brain immunity

The brain and spinal cord maintain a distinct cache of immune cells
By R. H. Nguyen and P. Kubers
RESEARCH ARTICLES pp. 408 & 409

397 Scent of a vaccine

Intranasal vaccination should block SARS-CoV-2 transmission at the source
By F. E. Lund and T. D. Randall

400 Ei-ichi Negishi (1935–2021)

Pioneering chemist who used transition metals to make and break bonds *By S. Ma*

BOOKS ET AL.

401 The elusive quest to make consistent calls

Variability in high-stakes decision-making is more prevalent than we might like to believe *By D. Frey*

402 An atomic warning

Early efforts to expand nuclear energy were rife with racism and peril, reminds a historian
By M. V. Ramana

LETTERS

403 Include all fungi in biodiversity goals

By S. C. Gonçalves et al.

403 Climate lawsuits could protect Brazilian Amazon

By A. C. Haliuc Bragança et al.

404 The hidden Olympic spectator

By S.-J. Luo

RESEARCH

IN BRIEF

405 From Science and other journals

RESEARCH ARTICLES

Neuroimmunology

408 Heterogeneity of meningeal B cells reveals a lymphopoietic niche at the CNS borders *S. Brioschi et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABF9277

409 Skull and vertebral bone marrow are myeloid cell reservoirs for the meninges and CNS parenchyma *A. Cugurra et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABF7844

PERSPECTIVE p. 396

410 Physiology

Enterically derived high-density lipoprotein restrains liver injury through the portal vein
Y.-H. Han et al.

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABE6729



411 Enzymology

Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics *C. J. Markin et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABF8761

PERSPECTIVE p. 391

412 Neurodevelopment

Retinal waves prime visual motion detection by simulating future optic flow *X. Ge et al.*

RESEARCH ARTICLE SUMMARY; FOR FULL TEXT: DOI.ORG/10.1126/SCIENCE.ABD0830

413 Structural biology

Structure of an AMPK complex in an inactive, ATP-bound state *Y. Yan et al.*

420 Plant science

Plant “helper” immune receptors are Ca²⁺-permeable nonselective cation channels
P. Jacob et al.

425 Astroparticle physics

Peta-electron volt gamma-ray emission from the Crab Nebula
The LHAASO Collaboration

REPORTS

430 Quantum devices

Coherent manipulation of an Andreev spin qubit *M. Hays et al.*
PERSPECTIVE p. 390

Planetary science

434 Upper mantle structure of Mars from InSight seismic data *A. Khan et al.*

438 Thickness and structure of the martian crust from InSight seismic data
B. Knapmeyer-Endrun et al.

443 Seismic detection of the Martian core
S. C. Stähler et al.

PERSPECTIVE p. 388; PODCAST

448 Photophysics

Inhibited nonradiative decay at all exciton densities in monolayer semiconductors
H. Kim et al.

452 Spectroscopy

Atomically resolved single-molecule triplet quenching *J. Peng et al.*
PERSPECTIVE p. 392

456 Animal culture

Innovation and geographic spread of a complex foraging culture in an urban parrot
B. C. Klump et al.

DEPARTMENTS

369 Editorial

Colleges need vaccine mandates
By H. Holden Thorp

462 Working Life

For those we've lost *By Anna Moyer*

ON THE COVER

Vital clues to a planet's geologic history are contained in its interior. For 2 years on Mars, the InSight lander has recorded seismic data to constrain the planet's interior structure. The data reveal a layered crust, a mantle with a thick lithosphere, and a liquid iron-nickel



core enriched in light elements, suggesting that Mars formed and evolved differently from Earth. See pages 388, 434, 438, and 443. *Illustration: C. Bickel/Science; Data: InSight Mars SEIS Data Service (2019)*

Science Staff 366
Science Careers 461

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.sciencemag.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.

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. Hurlley (UK), Phillip D. Szurmi, Sacha Vignieri **SR. EDITORIAL FELLOW** Andrew M. Sugden (UK) **SR. EDITORS** Gemma Alderton (UK), Caroline Ash (UK), Brent Grocholski, Pamela J. Hines, Di Jiang, 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, Laura M. Zahn **ASSOCIATE EDITORS** Michael A. Funk, Priscilla N. Kelly, Tage S. Rai, Seth Thomas Scanlon (UK), Yury V. Suleymanov **LETTERS EDITOR** Jennifer Sills **LEAD CONTENT PRODUCTION EDITORS** Harry Jach, Lauren Kmeck **CONTENT PRODUCTION EDITORS** Amelia Beyna, Jeffrey E. Cook, Chris Filiatreau, Julia 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), Alana Warnke, Alice Whaley (UK), Anita Wynn **PUBLICATIONS ASSISTANTS** Jeremy Dow, Alexander Kief, Ronnel Navas, Brian White **EXECUTIVE ASSISTANT** Jessica Slater **ASI DIRECTOR, OPERATIONS** Janet Clements (UK) **ASI SR. OFFICE ADMINISTRATOR** Jessica Waldoock (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 Brainerd, Kelly Servick, Catherine Maticic **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 **GRAPHICS MANAGING EDITOR** Alberto Cuadra **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 ILLUSTRATORS** Valerie Altounian, Chris Bickel **SCIENTIFIC ILLUSTRATOR** Ashley Mastin **SENIOR PHOTO EDITOR** Emily Petersen **PHOTO EDITOR** Kaitlyn Dolan **SENIOR MEDIA STRATEGIST** Jessica Hubbard **SOCIAL MEDIA PRODUCER** Sabrina Jenkins **WEB DESIGNER** Jennie Pajeroski **SENIOR PODCAST PRODUCER** Sarah Crespi **VIDEO PRODUCER** Meagan Cantwell **INTERN** Claire Hogan

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 Spiegel **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 Wovk **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 Dhene

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 MANAGEMENT** Kris Bishop **PRODUCT DEVELOPMENT MANAGER** Scott Chernoff **PUBLISHING TECHNOLOGY MANAGER** Michael Di Natale **SR. PRODUCT ASSOCIATE** Robert Koepke **PRODUCT ASSOCIATE** Anne Mason **SPI ASSOCIATE MANAGER** Samantha Bruno Fuller **SPI 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 Castellani, 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 Gonçalves **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

sciencemag.org/authors/
science-information-authors

REPRINTS AND PERMISSIONS

sciencemag.org/help/
reprints-and-permissions

MEDIA CONTACTS

scipak@aaas.org

MULTIMEDIA CONTACTS

SciencePodcast@aaas.org
ScienceVideo@aaas.org

INSTITUTIONAL SALES

AND SITE LICENSES

sciencemag.org/librarian

PRODUCT ADVERTISING

& CUSTOM PUBLISHING
advertising.sciencemag.org/
products-services
science_advertising@aaas.org

CLASSIFIED ADVERTISING

advertising.sciencemag.org/
science-careers
advertise@sciencecareers.org

JOB POSTING CUSTOMER SERVICE

employers.sciencemag.org
support@sciencecareers.org

MEMBERSHIP AND INDIVIDUAL

SUBSCRIPTIONS
sciencemag.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

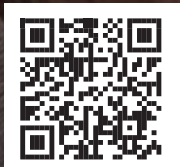
BOARD OF REVIEWING EDITORS (Statistics board members indicated with \$)

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, *NIH, NIH*
Paola Ariotta, *Harvard U.*
David Awschalom, *U. of Chicago*
Clare Baker, *U. of Cambridge*
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, *NIH, 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 (\$)*
Christian Büchel, *UCLA Hamburg*
Dennis Burton, *Scripta Res.*
Carter Tribble Butts, *UC Irvine*
György Buzsáki,
NYU School of Med.
Mariana Byndloss,
Vanderbilt U. Med. Ctr.
Annmarié 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. Dangel, *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 Hellariutta, *U. of Cambridge*
Janet G. Hering, *Eawag*
Heather Hickman, *NIH, 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 Kaeblerlein, *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, *Scripta Res. (\$)*
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*
Fabiene Mackay,
QIMR Berghofer
Anne Magurran, *U. of St. Andrews*
Asifa Majid, *U. of York*
Oscar Marin, *King's Coll. London*
Charles Marshall, *UC Berkeley*
Christopher Marx, *U. of Idaho*
David Masopust, *U. of Minnesota*
Geraldine Masson, *CHRS*
Jason Matheny, *Georgetown U.*
Heidi McBride, *McGill U.*
C. Robertson McClung,
Dartmouth
Rodrigo Medellín,
U. Nacional Autónoma de México
Jane Memmott, *U. of Bristol*
C. Jessica Metcalf, *Princeton U.*
Baotia Mi, *UC Berkeley*
Tom Misteli, *NCI, NIH*
Alison Motsinger-Reif,
NIH, NIH (\$)
Suresh Naidu, *Columbia U.*
Danielle Navarro,
U. of New South Wales
Daniel Nettle, *Newcastle U.*
Karen Neumark, *UC Berkeley*
Beatriz Noheida, *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. (\$)
Daniel Pauly, *U. of British Columbia*
Ana Pêgo, *U. do Porto*
Samuel Pfaff, *Salk Inst.*
Julie Pfeiffer,
Sharon Hammes-Schiffer, Yale U.
Philip Phillips, *UIUC*
Louise Harra, *U. Coll. London*
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 Silotti, *U. Coll. London*
Alison Smith, *Johns Hopkins U.*
Richard Smith, *UNC (\$)*
Mark Smyth, *QIMR Berghofer*
John Speakman, *U. of Aberdeen*
Tara Spire-Jones,
U. of Edinburgh
Allan C. Spradling,
Carnegie Institution for Sci.
V. S. Subrahmanian,
Dartmouth
Ira Tabas, *Columbia U.*
Eriko Takano, *U. of Manchester*
Patrick Tan,
Duke-NUS Med. School
Sarah Teichmann,
Wellcome Sanger Inst.
Rocio Titaniuk, *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*
Henrique Veiga-Fernandes,
Champalimaud Fdn.
Reinhold Veuglers, *KU Leuven*
Beit Vogelstein, *Johns Hopkins U.*
David Wallach, *Weizmann Inst.*
Jane-Ling Wang, *UC Davis (\$)*
Jessica Ware,
Amer. Mus. of Natural Hist.
David Waxman, *Fudan U.*
Chris Wickle, *U. of Missouri (\$)*
Terrie Williams, *UC Santa Cruz*
Ian A. Wilson, *Scripta Res. (\$)*
Hao Wu, *Harvard U.*
Wei Xie, *Tsinghua U.*
Yu Xie, *Princeton U.*
Jan Zaenen, *Leiden U.*
Kenneth Zaret,
UPenn School of Med.
Bing Zhu, *Inst. of Biophysics, CAS*
Xiaowei Zhuang, *Harvard U.*
Maria Zuber, *MIT*

NEWS FROM Science

**Up-to-the-minute research and
policy news you won't find in print**

Visit us online to read all the news coverage that
there just wasn't enough room to print in this issue.



ScienceMag.org/news

CALL FOR PAPERS



Ultrafast Science

 OPEN ACCESS

Ultrafast Science is an online-only, Open Access journal published in affiliation with **Xi'an Institute of Optics and Precision Mechanics (XIOPM)** of CAS and distributed by the **American Association for the Advancement of Science (AAAS)**. The mission of the journal is to build a platform for ultrafast scientific research; present new theories, concepts, ideas, technologies and progress in this field; publish the latest research representing the forefront of the discipline and that of international concern; and promote academic exchange and development worldwide. *Ultrafast Science* will publish high-quality original research articles, comprehensive reviews, editorials, and perspectives which feature high novelty, significance and technical quality in cutting-edge and emerging topics in ultrafast science with broad interest from scientific communities.

Submit your research to *Ultrafast Science* today!

Learn more at spj.sciencemag.org/ultrafastscience

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



ARTICLE PROCESSING CHARGES WAIVED UNTIL OCTOBER 2024

Colleges need vaccine mandates

Finally, in the United States, colleges and universities are ramping up for a relatively normal school year. Most pandemic restrictions have been lifted and—barring any unforeseen new coronavirus variants—parents and students will soon be pulling up to residence halls and unloading their belongings. At many schools in “blue” states, there will be a great deal of confidence that good times are on the horizon as everyone there will have shown proof of vaccination against COVID-19. But for some public universities in “red” states, where vaccine mandates are not permitted, an aura of uncertainty will hang over the campus and the local community, a foreboding sense that another outbreak could be at hand.

One risk is that unvaccinated students will contract and spread COVID-19, just as they did last fall. College presidents were quick to blame student partying and socializing for the outbreaks last year, without acknowledging their own culpability. No experienced college administrator can have truly believed that students would return to campus after months in lockdown without cutting loose and socializing. This year, some students will have been vaccinated or gained some immunity from contracting and then recovering from COVID-19. But many first-year students will be immunologically naïve. This raises the specter of continued testing, contact tracing, and potential lockdowns, provided schools will be allowed to enforce them.

Another risk will come from college sports. Last month, the baseball team at North Carolina State University (NCSU) was on an unexpected run into the final games of the College World Series in Omaha, Nebraska, when its season was suspended because of COVID-19 protocols. Some players fell ill with COVID-19, and it was later disclosed that many of them were among the unvaccinated. Eliot Avent, the NCSU baseball coach, said after the disappointing suspension that “My job is to teach them baseball, make sure they get an education and keep them on the right track forward, but I don’t try to indoctrinate my kids with my values or my opinions.” It’s hard to put all of the blame on Avent—the North Carolina legislature was already working on a bill prohibiting vaccine mandates when the University of North Carolina system stated that it would not be re-

quiring students to get the COVID-19 vaccines.

What the NCSU episode shows is that in the absence of a vaccine mandate, a testing program just isn’t enough. Beyond outbreaks like the one at NCSU, there are other worries. As Helen Branswell reported in *STAT*, many experts expect a powerful cold and flu season this year as masks are removed and viruses come out of hibernation. So, imagine a college campus where large numbers of students are coughing and sneezing, and even bedridden with normal colds and flu. In the absence of a vaccine mandate, it will be impossible for the college to reassure staff, faculty, and local residents that there is not a major outbreak of COVID-19. Further, many of these students who are unvaccinated could very well have severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in addition to whatever else ails them. This could lead to closures or substantial restrictions that could have easily been avoided if a vaccine mandate were in place.

I wrote to several higher education administrators and local government officials to ask their opinion about vaccine mandates. Although they privately agreed that every college should require mandates, no one was willing to say so on the record. They were all worried that Republican legislators will punish the universities if they come out swinging. Like other measures to limit academic

freedom and circumvent faculty expertise, this is another example of politicians and politically appointed trustees overreaching their appropriate roles in higher education policy.

When Arizona State University (ASU) tried to require that unvaccinated students would have to wear masks and be tested regularly, the governor issued an executive order banning any such expectation. The order says that ASU and other public colleges in Arizona cannot “place any conditions on attendance or participation in classes or academic activities, including but not limited to mandatory testing and mandatory mask usage.” ASU has about 75,000 students across its campuses. What could go wrong?

Officials at universities and in government need to take a stand regardless of the political consequences for the institutions. Lives are at stake.

—H. Holden Thorp



H. Holden Thorp
Editor-in-Chief,
Science journals.
hthorp@aaas.org;
@hholdenthorp

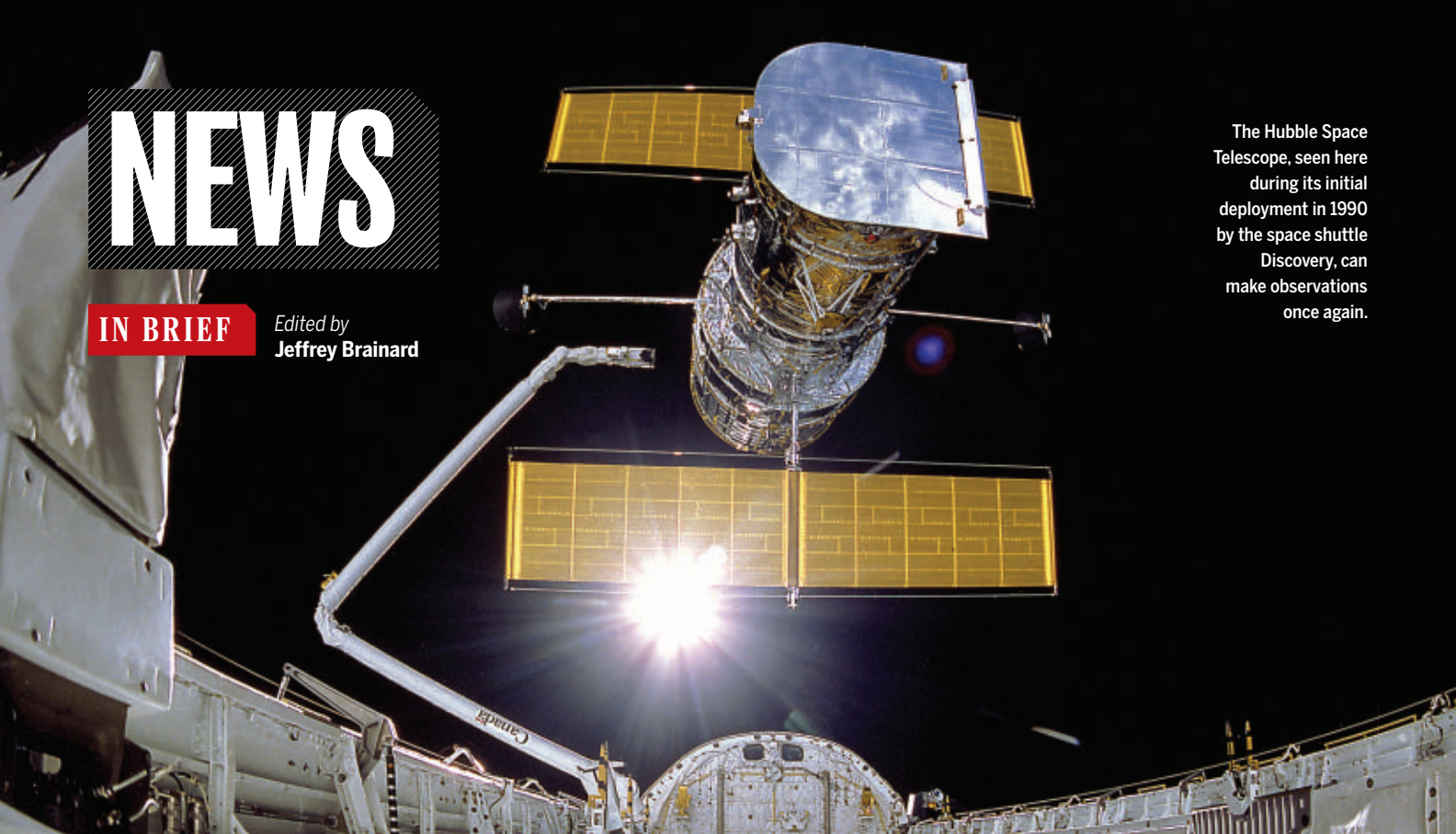
“... for some public universities... another outbreak could be at hand.”

NEWS

IN BRIEF

Edited by
Jeffrey Brainard

The Hubble Space Telescope, seen here during its initial deployment in 1990 by the space shuttle *Discovery*, can make observations once again.



ASTRONOMY

Hubble returns to duty following fixed computer glitch

The Hubble Space Telescope ended a monthlong hiatus on 16 July when operators successfully switched a failed control system to backup devices. The trouble started on 13 June when Hubble's payload computer, which controls its instruments, halted, and the main spacecraft computer put all the astronomical instruments in safe mode. Operators were unable to restart the payload computer, and switching memory modules—which they initially thought were at fault—didn't wake the telescope. They tested and ruled

out problems in other devices before zeroing in on a power control unit. NASA called in retired staff to help devise a fix for the 31-year-old telescope, which involved remotely switching to a spare power control unit and other backup hardware for managing the instruments and their data. The agency practiced and checked the repair on the ground for 2 weeks before executing it. After powering up all the hardware, Hubble returned to work on 17 July, and has already beamed back new images. NASA says it expects Hubble to continue for many years.

Millions of kids go unvaccinated

PUBLIC HEALTH | Reflecting another toll of the coronavirus pandemic, 23 million children missed routine vaccinations in 2020, the most since 2009 and 19% more than in 2019, the World Health Organization (WHO) and UNICEF said last week. As many as 17 million didn't receive any childhood vaccine at all. The pandemic led to closures or cutbacks at vaccination clinics and lockdowns that prevented parents and their children from reaching them, the groups reported. In addition, 57 mass vaccination campaigns for non-COVID-19 diseases in 66 countries were postponed. Childhood vaccination rates decreased

across all WHO regions, with the Southeast Asian and eastern Mediterranean regions particularly affected. In India, more than 3 million children missed a first dose of the diphtheria, tetanus, and pertussis vaccine, more than double the number in 2019. "We [are] leaving children at risk from devastating but preventable diseases like measles, polio, or meningitis," says WHO Director-General Tedros Adhanom Ghebreyesus.

EU, China boost emissions curbs

CLIMATE POLICY | As part of the run-up to the U.N. climate summit in November, the European Union and China announced last week plans to follow through on

commitments to curb their carbon emissions. The European proposal, which must be approved by the bloc's member states, would steeply increase the price of carbon dioxide (CO₂) emissions; eliminate new gas-powered cars by 2035; require 38% of all energy to come from renewables by 2030, up from a previous goal of 32%; and impose tariffs on goods from countries that have not acted on climate change. (Democratic lawmakers in the United States proposed a similar tariff this week.) Meanwhile, China on 16 July launched a carbon trading scheme for power plants that instantly created the world's largest carbon market, triple the European Union's in size. China's plan incentivizes plants

to lower CO₂ emissions by allowing more efficient facilities to sell some of their reductions to less efficient ones. Although some observers call the plan weak because it covers a relatively small portion of China's emissions, it could be expanded to eventually incorporate three-fourths of the country's emissions from all sources.

Heat increases workplace injuries

PUBLIC HEALTH | When temperatures soar, workers and their employers need to take heed: Hot weather led to 20,000 more injuries annually in California between 2001 and 2018, according to a novel analysis of 11 million workers compensation claims. Economist Jisung Park at the University of California, Los Angeles, and colleagues classified work-related injuries by ZIP code and looked up local temperatures on the day each was recorded. They found increases of between 5% and 15% in claims, depending on the temperature and occupation, compared with those filed on a typical cooler day, defined as a temperature of 16°C. Few were attributed directly to heat, but the injuries connected to higher temperatures—such as falls and mishandling equipment—may have resulted because the heat made workers woozy, the researchers reported to Congress last week and in a preprint on the SSRN server. But mitigation may be possible: Heat-related injury claims declined after 2005, when California began to require shade, water, and breaks for outdoor workers—in industries such as construction, utilities, and farming—whenever temperature exceeded 35°C.

U.K., U.S. mull misconduct steps

RESEARCH INTEGRITY | Both the United Kingdom and the United States last week announced new high-level bodies to provide guidance on research integrity—but both lack the powers that many whistleblowers say are critical, such as independently investigating complaints of wrongdoing and pulling grant funding from institutions that fail to conduct misconduct probes properly. The umbrella funding body UK Research and Innovation (UKRI) launched the Committee on Research Integrity, which plans to operate for 3 years and accelerate existing projects in this area. The U.S. National Academies of Sciences, Engineering, and Medicine (NASEM) unveiled the Strategic Council for Research Excellence, Integrity, and Trust, which will have members from the U.S. National Institutes of Health and National Science Foundation. Unlike UKRI, NASEM does not fund researchers, so it cannot set



CONSERVATION AI flags wildlife diseases

A new automated alert system can help veterinarians get a jump on investigating disease outbreaks and disasters afflicting wildlife. Researchers at the University of California, Davis, and colleagues used a machine learning algorithm to scan case reports of sick and dead wildlife submitted to a database by wildlife clinics and rehabilitation centers in the United States and other countries. The researchers used data from 3081 reports filed from California to train the algorithm to detect patterns of species suffering common symptoms. The software is designed to identify unusual events in one of 12 clinical categories, such as mass starvation or an oil spill. The algorithm assigned the correct category to 83% of cases examined, including ones from an outbreak of neurological disease in California brown pelicans (above) and red-throated loons, the research team reported last week in the *Proceedings of the Royal Society B*. The system could help wildlife officials more quickly detect developing problems and confirm specific causes.

policies on how to handle misconduct allegations. But it could promote integrity in other ways—for instance by pushing for a central repository for researchers to report their financial conflicts of interest, says Marcia McNutt, president of the National Academy of Sciences and an ex officio member of the new panel.

'Borg' DNA carries microbe mix

MICROBIOLOGY | Sifting through DNA in the mud of her backyard, a geomicrobiologist discovered what may be the longest known extrachromosomal sequence, which includes genes from a variety of microbes—prompting her son to propose naming it after *Star Trek*'s Borg, cybernetic aliens that assimilate humans. Jill Banfield of

the University of California, Berkeley, was searching for viruses that infect archaea, a type of microbe often found in places devoid of oxygen. The 1-million-base-pair strand of DNA contains genes known to help archaea metabolize methane, suggesting the fragment might exist inside the microbes but outside their normal chromosome, the research team wrote in a preprint posted on 10 July on the bioRxiv server. Scanning a public microbial DNA database, the authors identified 23 possible Borks, with many of the same characteristics, in other U.S. locations. The Borks' role remains murky, but they may provide another example of DNA that can hop between an organism's chromosomes or between organisms, helping species adapt to changes in their environment.



IN DEPTH

Flash floods caught many communities in Germany, including the town of Insul, by surprise, leading to at least 165 deaths.

NATURAL DISASTERS

Europe's deadly floods leave scientists stunned

Despite improvements, flood forecasts sometimes failed to flag risks along smaller streams

By **Warren Cornwall**

Four days before deadly floods swept through western Germany and parts of Belgium last week, Hannah Cloke saw a forecast of extreme rain on a Europe-wide flood alert system to which she belongs. Researchers “were stupidly congratulating ourselves that we were forecasting something so early. ... The assumption was that would be really helpful,” says the hydrologist and flood forecaster at the University of Reading. Instead, she was stunned to see scenes of devastation and death despite the ample warnings. “We should not be seeing this number of people dying in 2021 from floods. It just should not be happening.”

As the magnitude of the destruction becomes clear, European scientists are wrestling with how such damage could happen in some of the world's wealthiest and most technologically advanced countries, despite major investments in flood forecasting and preparation catalyzed by previous inundations. And they are examining whether climate change helped fuel the disaster—and what that might mean for the future.

Beginning on 13 July, intense storms dropped as much as 15 centimeters of rain in

24 hours, swelling streams that then washed away houses and cars and triggered massive landslides. At least 196 people had died as *Science* went to press—165 in Germany and 31 in Belgium—and the number is expected to rise. On 18 July, German Chancellor Angela Merkel visited the stricken town of Adenau. The scene was “terrifying,” she said. “The German language can barely describe the devastation.” That same day more flash floods struck Bavaria, in southern Germany.

Researchers are just beginning to unravel the complex web of climatic, hydrological, and social factors that contributed to the catastrophe. But they already have some suspects in mind, including a warming climate that can supercharge rainstorms and European disaster plans that focused on major rivers, rather than the lower volume tributaries hit hardest by the storms.

“We’ve been concentrating a lot on the big rivers,” says William Veerbeek, an urban flood management expert at the IHE Delft Institute for Water Education. “There’s still a lot more to do in the smaller streams.”

For years, scientists have warned climate change will mean more flooding in Europe and elsewhere. Warmer air holds more moisture, which can translate into heavier rain-

fall. By 2100, flood damage on the continent could cost as much as €48 billion per year—up from €7.8 billion now—if nothing more is done to prepare, and the number of people affected could more than double to some 350,000, according to the European Commission's Joint Research Centre.

Some European rivers are already exhibiting climate-related changes, says Fred Hattermann, a hydrologist and flood expert at the Potsdam Institute for Climate Impact Research. Along the Danube River, for example, floods that once came every 50 years now happen roughly twice as often, according to recent unpublished work the institute conducted with the German insurance industry.

New research suggests such risks could grow if climate change slows the jet stream—the high-altitude winds circling the Northern Hemisphere—causing drenching rainstorms to linger longer over flood-prone landscapes. Storms that stall over Europe were once exceedingly rare. But according to a study published last month in *Geophysical Research Letters*, in a worst case scenario such storms could become as much as 14 times more common in 2100 than they were at the start of this century.

The exact weather pattern driving the current storm is different, says the paper's lead author, hydroclimatologist Hayley Fowler of Newcastle University. But the overall dynamic is similar, she adds: A slow-moving storm overwhelms a region's river system.

Floods already rank as the most destructive natural hazard in Northern Europe. Past events—including deadly floods in 2002—led the European Commission to launch the European Flood Awareness System, meant to provide emergency managers with early warnings. But last week the speed and intensity of the flooding in Germany—especially in towns perched next to smaller creeks—took most people by surprise. “There was simply no time,” Hattermann says. “Then, of course, people run to save their cars and whatever and bad things happen.”

In the Netherlands, decades of preparation appear to have helped. Even though the Meuse River, which runs through eastern Belgium and into the Netherlands, broke a high flow record set in 1993, it caused far less damage. A tributary of the Meuse flooded the city center of Valkenburg in the south, but no deaths were reported.

Dutch policies aimed at making “room for the river” have widened and deepened river channels, and set aside land where floodwaters can spread out. Those measures were projected to reduce flood peaks on the Meuse by as much as 1 meter, says Nathalie Asselman, a flood management expert and adviser to the Dutch government at Deltares, a research institute. Those predictions appear to be coming true. “When we look at how bad the flooding is, it's much less than what happened in the '90s,” she says.

Some researchers are examining how people responded to flood warnings. After the 2002 floods in eastern Germany, Bruno Merz, a hydrologist and civil engineer at the Helmholtz Centre Potsdam of the German Research Centre for Geosciences, surveyed residents. One-third of the people who received flood warnings reported having “no clue” what to do next. Now, Merz wants to know whether such confusion contributed to the current death toll. “All the data shows ... that when people get a warning and they take [it] seriously, they can rescue themselves,” he says. “They need 1 hour, not more.”

Merz also has some immediate and personal concerns about the floods. Last week, he learned that a swollen creek had swamped the ground floor of a vacation cottage—named “Brook view”—that his family had rented in a German village. The trip is still on; the landlord pleaded with Merz not to cancel because he needs the money. But, “If there are any problems,” Merz says, “we will just make a detour.” ■

BIOMEDICINE

Alzheimer's drug approval spotlights blood tests

Physicians hope technology will help screen patients eligible for Biogen's controversial antibody, aducanumab

By Kelly Servick

When the U.S. government approved the Alzheimer's disease drug aducanumab last month despite shaky evidence of clinical benefits, Suzanne Schindler saw an immediate consequence: “We're going to have to do a lot more biomarker testing.” Schindler, a neurologist at the Washington University School of Medicine in St. Louis, expects many patients with memory problems will want to know whether they are eligible for the drug, the first meant to address the underlying disease process and slow cognitive decline. But a definitive diagnosis is expensive and time-consuming, involving a hunt for indicators of disease, or biomarkers, in patients' spinal fluid or in positron emission tomography (PET) scans. “We simply don't have the capacity” to do the tests on everyone potentially eligible for aducanumab, Schindler says.

An innovation could help manage the deluge: simple blood tests for molecules that may indicate signs of disease in the brain. “A lot of data now [suggest] that what we find in the blood actually can reflect what's going on in the brain,” says Nancy Ip, a neuroscientist at the Hong Kong Uni-

versity of Science and Technology. Her team developed a test to identify Alzheimer's and determine a patient's stage of disease from levels of 19 blood proteins.

One blood test is already commercially available in the United States, though some researchers doubt it can meet demand. Companies are now working to make other tests fit for widespread clinical use—and doctors are scrambling to decide when and how to use them. Before the FDA decision, many physicians “had a somewhat leisurely approach” to determining how Alzheimer's blood tests might fit into their clinical practice, says Douglas Galasko, a neurologist at the University of California (UC), San Diego. “And now life is much more complicated.”

Aducanumab, approved to treat early-stage disease, reduces buildup in the brain of beta amyloid, a sticky protein thought to drive neuronal damage in Alzheimer's. But its evidence of clinical benefit is thin (*Science*, 11 June, p. 1141), the drug requires monthly infusions with a list price of \$56,000 per year, and it can cause swelling and bleeding in the brain. Deciding to prescribe it “is a really big deal,” Schindler says. “You want to be absolutely certain [patients] have Alzheimer's disease.”

Either a PET scan that detects beta amy-



A scientist at C₂N Diagnostics prepares to analyze blood samples with mass spectrometry.

loid in the brain or a spinal fluid test that measures amyloid and another Alzheimer's-associated protein, tau, can confirm that diagnosis. However, PET scans can cost thousands of dollars, and many patients dread a lumbar puncture to collect spinal fluid. And if every older person with cognitive impairment required evaluation by a dementia specialist and a PET scan or spinal fluid test, they would quickly overwhelm the medical system, says Soeren Mattke, a health services researcher at the University of Southern California's Center for Economic and Social Research.

In a study published last year, he and colleagues predicted that the first disease-modifying Alzheimer's drug available in the United States would lead to wait times of up to 45 months for a specialist evaluation and confirmatory tests. But blood tests that could screen out patients who don't have Alzheimer's could reduce wait times to about 10 months, the researchers found.

Reliably capturing the state of the brain from proteins that slip into blood has been a challenge. Blood contains so many different proteins that "you're looking in what is essentially very dirty pond water," says Ashvini Keshavan, a neurologist at University College London. Beta amyloid hides from detection by clinging to surfaces or to other proteins—and it can quickly degrade in a blood sample, adds Henrik Zetterberg, a neurochemist at the University of Gothenburg.

One type of blood test uses an instrument called a mass spectrometer to measure the ratio of two types of beta amyloid in blood. Decreased levels in blood of certain forms of the protein indicate it is instead accumulating in the brain as plaques. In 2019, researchers reported that a combination of this ratio, a person's age, and a measure of genetic Alzheimer's risk agreed with results of a PET scan in 94% of cases.

That test, marketed by C₂N Diagnostics, has been available to U.S. clinicians since October 2020 under a Centers for Medicare & Medicaid Services (CMS) certification process for laboratories that test patient samples. The test costs \$1250; a financial assistance program can bring that to \$25 to \$400, says C₂N CEO Joel Braunstein. The test must be done in the company's lab, which can run 250,000 to 300,000 tests per year, Braunstein says.

Other types of blood tests, so-called immunoassays, use antibodies to sandwich and isolate a protein of interest. Some might run on equipment common in clinical labs. Many recent studies of these tests have focused not on beta amyloid,

but on certain forms of tau. Beta-amyloid plaques can appear decades before patients develop dementia—if they ever do. But beta-amyloid buildup is thought to somehow prompt tau to form toxic tangles that damage neurons, explains William Jagust, a neuroscientist at UC Berkeley. As a result, he says, "It may well be that some of these tau biomarkers are better predictors" of symptom onset.

In May, Lund University neurologist Oskar Hansson and colleagues reported in *Nature Medicine* that combining a tau blood test with a person's Alzheimer's genetic status and two brief cognitive tests could predict with about 90% accuracy which people with mild cognitive impairment would develop Alzheimer's dementia within 4 years. "That was really groundbreaking," Keshavan says.

Several companies have developed blood tests sensitive to tau, including Roche and Eli Lilly and Company. A large clinical trial of Eli Lilly's antiamyloid antibody donanemab will use a tau blood test to help select patients. But to be used as clinical decision tools, the tests need to show consistent results under variable conditions and in more diverse patient populations, Hansson says. Most studies have used blood

samples from patients at specialized memory clinics. Many of these groups are disproportionately white or of high socioeconomic status; and they generally don't have other conditions, such as heart and kidney failure, which could influence levels of certain biomarkers, he says.

Some researchers caution that blood tests aren't ready to replace a PET scan or spinal fluid test. And confirming Alzheimer's will still require careful clinical examination, Zetterberg says, because many factors can cause memory problems in aging patients, including depression and sleep disruptions. Prescribing the new drug to anyone with a positive blood test is "the nightmare scenario of starting to use the biomarkers in a careless way," he says.

But blood tests could act as an initial screen. A test in primary care offices would offer peace of mind to those who test negative—and could reduce the number of people seeking a PET scan or spinal fluid test such that CMS might cover these costs, says Sid O'Bryant, a clinical neuropsychologist at the University of North Texas Health Science Center. "Without a blood test, access to the drug is going to be so incredibly limited," he says. He expects such tests to be available within the next 2 years. "The data and the science are so close." ■

"We're going to have to do a lot more biomarker testing."

Suzanne Schindler,
Washington University School
of Medicine in St. Louis

THE NEW NORMAL

A few lucky researchers return to the field

The pandemic has transformed fieldwork and upended careers

By Elizabeth Pennisi

Just before dawn in the Jama-Coaque Ecological Reserve, a patch of Ecuador's lush coastal forest, Abhimanyu Lele unfurls a tall net between two poles, then retreats out of sight. A half-hour later, he and a local assistant reappear and smile: Their catch—10 birds that collided with the net and tumbled into a pocket along its length—was a good one. The pair records species, measures and photographs the captives, and pricks wings for blood that can yield DNA before releasing the birds back into the forest. The data, Lele hopes, will shed light on how Ecuadorean songbirds adapt to different altitudes and other conditions.

The third-year graduate student at the University of Chicago (UC), who returns next week from a 10-week field season, was delighted to have made it to his destination. In a typical year, thousands of graduate students and faculty fan out across the world to tackle important research in climate change, fragile ecosystems, animal populations, and more. But the pandemic shut down travel, and fieldwork can't be done via Zoom, depriving young scientists like Lele of the data and publications they need to climb the academic ladder and help advance science. Now, he and a few others are venturing out—into a very different world.

They are the exceptions. "Most folks have never been able to get back out there," because COVID-19 continues to spread in much of the world, says Benjamin Halpern, an ecologist with the National Center for Ecological Analysis and Synthesis at the University of California, Santa Barbara. "They are just waiting."

At the American Museum of Natural History, which mounts about 100 international expeditions a year, "Travel to countries still having trouble [is] just not

going to happen,” says Frank Burbrink, a herpetologist there. “This is the longest I’ve ever gone without catching snakes since I was 12 years old.” The Smithsonian Institution’s National Museum of Natural History likewise “is not putting people overseas,” says Director Kirk Johnson.

Some institutions are allowing international travel on a case-by-case basis, but the process can be a frustrating one. “I’ve just been given clearance to go to Ecuador, but not to one province where two-thirds of my field sites are located,” says Michael Ellis, a graduate student at Tulane University who is studying environmental and human factors that affect where Ecuador’s birds live. The loss of fieldwork is pushing Ellis to reconsider his research focus.

Researchers are quick to say their frustration pales in comparison with the losses many worldwide are enduring. “The personal disappointment was completely dwarfed by the scale of the tragedy,” says Kristina Fialko, another UC graduate student who was 3 days from planned fieldwork in India to study sunlight’s effects on visual communication in warblers when her university deemed the trip too risky and pulled the plug in May. She will make do with a literature review and local fieldwork to stay on track for getting her Ph.D. next year.

Despite their efforts to adapt, for untenured researchers and graduate students, delays can be career breaking. Two years “is an eternity,” says Shannon Hackett, an ornithologist at the Field Museum and one of Lele’s unofficial advisers. Because fieldwork must often happen in a narrow time window—during a mating season, say, or a seasonal migration—a few months’ delay can mean a lost year of work.

Lele was set to start his first season in March 2020 when the pandemic locked everyone in place. So instead, he helped teach undergraduates remotely and wrapped up papers about earlier research. It was “a difficult time to be productive,” he recalls.

Ever the optimist, Lele began to lobby his advisers last fall about heading to South America again. “This conversation did not go anywhere,” he says, but then came the vaccines for COVID-19. He and many in the biology department lucked out and got shots in January: The university had leftover doses from its campaign to protect its health care workers. After that, “My advisers did not need convincing,” Lele says. He still had to detail the precautions he would take to ensure his safety and that of local collaborators. But Ecuador and the university agreed, and he landed in Quito in late May. “I have been feeling a deep sense of relief and satisfaction to be working on the substance of my dissertation at long last,” he says.



Still, he couldn’t escape the pandemic’s shadow. In the remote Ecuadorean forest, Lele could easily limit his exposure to other people. But none of the local staff at the reserve’s field station had been vaccinated and all wore masks and kept distanced from each other. For the first month, Lele did all the shopping for the group and later had to cope with local colleagues who were casual about COVID-19 precautions.

Thousands of kilometers away, “I have a twinge of worry every day,” Hackett says. With the pandemic barely under control in Ecuador, Lele might still get sick, she says, or face antifoiegn sentiment. Hackett thinks her heightened concern for students doing distant fieldwork may persist even after the global pandemic subsides. The crisis has reminded her of the instability in many countries and the immense stress on her mentees, she says.

The pandemic has also created a disparity that may be slow to abate: Vaccinated U.S. scientists working domestically can now easily pursue the project of their dreams, whereas those seeking to venture farther from home often cannot. “We are having a very successful field season this year,” says Robin Hopkins, a plant evolutionary biologist at Harvard University who conducts fieldwork in rural Texas. Already, two of her students have spent 1 month measuring plants and collecting seeds and other material to grow in the lab, without leaving the United States.

Kevin Langergraber, a primatologist at Arizona State University, Tempe, has con-

tinued his work on chimps in Uganda, but with a change that could prove lasting: He and colleagues have set up a quarantine “camp” 2 kilometers away from their main camp. Each newcomer spends 1 week there before starting fieldwork, to reduce the risk of transmitting disease to the chimps. They expect to continue the practice when COVID-19 finally recedes, to guard against other infections.

Others are seeking new ways to do fieldwork remotely. Harvard deep-sea biologist Peter Girguis, for example, couldn’t get to sea to test an idea for generating electricity by harnessing methane bubbling up from the sea floor, as the ocean research fleet sponsored by the National Science Foundation was still grounded. So, he hired commercial remotely operated vehicle operators to help retrieve a seafloor instrument platform, which has a mass spectrometer and other sensors that measure methane flux.

As a generation of young field researchers considers whether and how to change course, Lele is grateful to not be among them. He will soon be settling back into Chicago, reams of samples and data in hand. And he’s eager to make the most of what he has gathered. “Making this trip happen required putting a lot of people to considerable trouble, in Chicago and especially in Ecuador,” he explains. “I really do not want their efforts on my behalf to be for nothing.” ■

With reporting by David Grimm.

CONSERVATION BIOLOGY

DNA plucked from air identifies nearby animals

Analysis of airborne cells could survey biodiversity, two studies of zoo animals suggest

By Erik Stokstad

DNA is everywhere, even in the air. That's no surprise to anyone who suffers allergies from pollen or cat dander. But two research groups have now independently shown the atmosphere can contain detectable amounts of DNA from many kinds of animals. Their preprints, posted on bioRxiv last week, suggest sampling air may enable a faster, cheaper way to survey creatures in ecosystems.

The work has impressed other scientists. "The ability to detect so many species in air samples using DNA is a huge leap," says Matthew Barnes, an ecologist at Texas Tech University. "It represents an exciting potential addition to the toolbox."

"The surprising part is that you're able to get birds and mammals—wow," says Julie Lockwood, a molecular ecologist at Rutgers University, New Brunswick. The new studies suggest "there's more than just spores; there's cells and hair and all kinds of interesting things that float through the air."

For more than a decade, researchers have analyzed those disparate sources of DNA in water to identify elusive organisms. Researchers' sampling of environmental DNA (eDNA) in lakes, streams, and coastal waters has let them identify invasive species like lionfish as well as rare organisms such as the great crested newt. More recently, some scientists have tracked insects by eDNA on leaves, and also found soil eDNA apparently left by mammals loping along a trail.

Far fewer studies have been done on animal eDNA in air. It's not obvious how much tissue wafts off animals or how long the genetic contents of those cells persist in air. Some earlier studies used metagenomic sequencing—an approach to identify mixtures of DNA—to detect microorganisms including bacteria and fungi that are abundant in air. And a 2015 study of air monitors for pathogens in the Washington, D.C., area found traces of eDNA from many kinds of vertebrates and arthropods. But it wasn't obvious how useful the technique would be, and it's not clear how terrestrial animals shed cells that float away.

Earlier this year, Elizabeth Clare, a molecular ecologist now at York University, reported in *PeerJ* that eDNA from naked mole rats could be detected in air samples taken in the laboratory. To find out whether animal eDNA could be detected outdoors, she and colleagues from Queen Mary University of London went to a zoo: There, the species are known and absent from the surrounding landscape, so the team could determine the source of airborne eDNA they found. In December 2020, Clare set up vacuum pumps with filters in 20 locations in Hamerton Zoo Park and let each run for 30 minutes.

molecular ecologist at the University of Copenhagen, recalls inspiration struck while brainstorming proposals for a high-risk grant program. "I remember saying, it has to be crazier—like vacuuming DNA from air, that would be insane." They won the grant and sucked up air from three locations in the Copenhagen Zoo with vacuums and fans in three types of samplers. They consistently detected animals—a total of 49 species of vertebrate.

"These preprints are exciting and show some great data," says Kristy Deiner, a conservation ecologist at ETH Zürich. She leads an XPRIZE Rainforest team to develop airborne DNA technology for monitoring biodiversity.

Airborne DNA may help reveal the presence of otherwise hard to detect animals, such as those in dry environments, burrows, or caves, and those that fly out of sight of wildlife cameras, like some birds, Lockwood says.

She cautions that many questions remain about the approach, including the key issue of how far eDNA travels on air, which will influence how well the method can pinpoint the recent location of animals. That distance will depend on many factors, including the environment; eDNA will probably waft farther in a grassland than in a forest. Another question is how exactly animals shed the DNA. It

could be when cells are freed as they scratch or rub their skin, sneeze, or do any vigorous activity like fighting or subduing prey. But even sloth eDNA turned up, says molecular ecologist Christina Lynggaard, a postdoc at the University of Copenhagen who did the sampling at the zoo.

Preventing contamination—always an issue with eDNA studies—is particularly thorny. Sampling eDNA in air, Barnes says, is like "pipetting underwater." One problem, Clare says, is how to find a negative control, or a test sample with no DNA in it. "I don't know where to buy a balloon of sterile air."

Despite the unknowns, Barnes and others have high hopes. Lockwood, who studies forest pests and has identified eDNA traces on bark and leaves, is already hoping to identify insect pests from air. "I can't wait to try it," she says. ■



This kingfisher stunning its prey may spread DNA of both species into the air.

Clare collected 72 air samples from both outside and inside zoo buildings. She used polymerase chain reaction to amplify the scant genetic fragments left on the filters into enough DNA for sequencing. "We had to take a leap of faith that it was there because it wasn't something you can measure," she says. After sequencing the eDNA, she matched the snippets to known sequences in a database. The team identified 17 species kept at the zoo and others living near and around it, such as hedgehogs and deer. Some zoo animal DNA was found nearly 300 meters from the animals' enclosures. She also detected airborne DNA likely from the meat of chicken, pig, cow, and horse fed to captive predators indoors. All told, the team detected 25 species of mammals and birds.

Meanwhile, researchers in Denmark had pursued the same idea. Kristine Bohmann, a



Massive jets of material shooting out of Centaurus A's center are powered by a matter-guzzling black hole.

ASTRONOMY

Event Horizon Telescope images second black hole's jets

Globe-spanning instrument's snapshot of Centaurus A galaxy suggests all black holes operate similarly

By Daniel Clery

The astronomy team that 2 years ago captured the first close-up of a giant black hole, lurking at the center of the galaxy Messier 87 (M87), has now zoomed in on a second, somewhat smaller giant in the nearby active galaxy Centaurus A. The Event Horizon Telescope's (EHT's) latest image should help resolve questions about how such galactic centers funnel huge amounts of matter into powerful beams and fire them thousands of light-years into space. Together the two galaxy images also support theorists' belief that all black holes operate the same way, despite huge variations in their masses.

"This is really nice," astronomer Philip Best of the University of Edinburgh says of the new EHT image. "The angular resolution is astonishing compared to previous images of these jets."

The EHT merges dozens of widely dispersed radio dishes, from Hawaii to France and from Greenland to the South Pole, into a huge virtual telescope. By pointing a large number of dishes at a celestial object at the same time and carefully time stamping observations from each with an atomic clock, researchers can later reassemble the data with massive computing clusters—a process that takes years—to produce an image with a resolution as sharp as that of a single Earth-

size dish. One challenge is getting observing time at 11 different observatories simultaneously, so the EHT only operates for a few weeks each year; poor weather and technical glitches often further narrow that window.

The virtual telescope probed Centaurus A during the same 2017 observing campaign that produced the now-famous image of the supermassive black hole in M87—*Science's* Breakthrough of the Year for 2019. Centaurus A, about 13 million light-years away, is one of the closest galaxies to Earth that is bright at radio wavelengths. It also has obvious jets spewing matter above and below the galactic disk, a hallmark of a central, active giant black hole. "We wanted to see what the

jets looked like at the resolution" the EHT could offer, says team leader Michael Janssen of the Max Planck Institute for Radio Astronomy. "We didn't know what to expect."

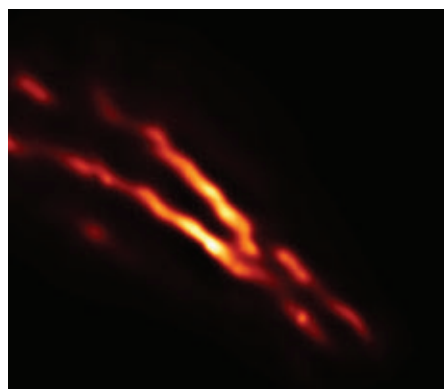
The result, which he and colleagues report this week in *Nature Astronomy*, was a detailed image of how the jets emerge from the region around Centaurus A's supermassive black hole, showing a remarkable similarity to the EHT's images of M87's jets on a much smaller scale. Images of Centaurus A's jets taken by other telescopes at different wavelengths revealed little detail, but the EHT image shows the jets with dark centers flanked by bright stripes; Best suggests the jets may appear bright around their edges because outer regions rub against surrounding gas and dust, causing them to glow.

Astrophysicists don't fully understand how galactic nuclei drive these fantastically powerful fountains. One theory holds that an accretion disk, the swirling whirlpool of matter spiraling into the black hole, generates a magnetic field that funnels some of the matter into a jet. Others argue this magnetic field must tap into the rotational energy of the black hole itself to be able to achieve such colossal power.

The new observation of Centaurus A doesn't resolve that question, but it holds clues. Janssen says the image shows that the remarkably parallel edges of the jets narrow into cones close to the black hole. The bases of the cones remain wide, which might suggest they are coming from the accretion disk. "It remains to be seen," he says.

Theoretical astrophysicist Jim Beall of St. John's College says there may be no single answer: The spin of the black hole drags on the innermost stable orbit of the infalling matter, which in turn affects how the accretion disk shapes and powers the jet. "It's a symbiotic relationship," he says. "The EHT takes us down close to the accretion disk. The results are really quite beautiful."

The picture of Centaurus A also fills in a size gap in black hole observations. Observers have studied the workings of jets coming from the very largest of black holes—including M87—weighing billions of times the mass of our Sun. They've also seen jets from much smaller black holes, with masses of a few tens of Suns. The new view of Centaurus A's black hole, at 55 million times the Sun's mass, looks just like its bigger and smaller relatives. This confirms the idea that black holes are essentially simple objects, defined only by their mass, charge, and spin, so those with the mass of a large star should behave no differently from one with the mass of a galaxy. ■



The EHT's image of Centaurus A reveals jets of material with dark centers paralleled by glowing edges.

COVID-19

WHO chief pressures China on pandemic origin

Agency director calls for lab audits and more studies of how SARS-CoV-2 emerged

By Jon Cohen

In a sharp tightening of the diplomatic screws, the director-general of the World Health Organization (WHO), Tedros Adhanom Ghebreyesus, is urging China to increase its transparency about the early days of the COVID-19 pandemic to help resolve the origin of SARS-CoV-2. Tedros also says WHO will create a new body to conduct the next phase of studies into the emergence of the virus, an unexpected move that concerns some scientists, including at least one member of an existing mission the agency organized to study COVID-19's origin. "I'm worried about delays and of course it's a bit strange," says virologist and veterinarian Marion Koopmans of Erasmus University Medical Center. "We're losing valuable time."

At a press conference on 15 July and in a statement the next day at an information session on the pandemic's origin for WHO members, Tedros called for more aggressively probing the two leading theories of how SARS-CoV-2 first infected humans and then emerged in Wuhan, China, in December 2019: that the virus made a natural "zoonotic" jump from an unknown animal species into humans or, more controversially, that it first infected a human during laboratory or field studies of coronaviruses found in animals. (An even more contentious theory suggests the virus was genetically engineered in a Wuhan lab.)

Tedros, who has been accused of being too deferential to Chinese President Xi Jinping, said China has not shared "raw data" from the pandemic's early days and called for "audits of relevant laboratories and research institutions operating in the area of the initial human cases identified in December 2019." The Wuhan Institute of Virology is world famous for its study of bat coronaviruses, and an outpost of China's Center for Disease Control and Prevention also has a lab in the city that does similar work.

Researchers who have been critical of WHO's handling of the origin issue welcome Tedros's tougher tone. "It's a sign that the WHO might be able to do more cred-

ible or balanced investigation," says Alina Chan, a gene therapy researcher at the Broad Institute, who with 17 other scientists co-authored a 14 May letter in *Science* that argued the lab theory deserves a more balanced assessment. But Chan doubts China will agree to audits of its labs. "Right now, the lack of clarity is in China's interest."

Another author of the *Science* letter, microbiome researcher David Relman of Stanford University, wished Tedros had owned up to past WHO "missteps." "I don't think he can simply just take the next step and not worry about what's happened so far."

But other researchers think Tedros has been caught up in what Gerald Keusch, as-

issued in March by the joint mission, which had just completed the first of two planned phases of studies, declared the lab origin hypothesis "extremely unlikely" and favored the zoonotic theory.

That sparked controversy, and even Tedros was chagrined. Last week, he said it was "premature" to discount the lab theory. The next day, China's Ministry of Foreign Affairs Spokesperson Zhao Lijian pushed back on Tedros's remarks at a press conference, stressing that the joint mission report reached "important conclusions" and suggesting the WHO director was "politicizing the issue."

A "circular" Tedros presented to member states last week spelled out other data the phase 2 studies should attempt to gather, such as testing of captive and wild animals, particularly in regions where SARS-CoV-2 first circulated, and of humans who came in contact with them. Tedros also wants more "studies of animal markets in and around Wuhan"; the joint mission report said it could not verify that live mammals were available at these markets, but a study later showed thousands were sold there.

Whether Koopmans and other members of the existing joint mission will help conduct those studies is murky. Tedros said a new WHO International Scientific Advisory Group for Origins of Novel Pathogens "will play a vital role in the next phase of studies into the origins of SARS-CoV-2, as well as the origins of future new pathogens." WHO will soon make an open call for "highly qualified experts" to apply.

Keusch says the current group has highly qualified, diverse experts who worked "diligently" and established important ties to their Chinese colleagues. "I'm very suspicious about dismissing the initial task force and now allowing individuals and governments to nominate themselves."

Relman wonders whether WHO is the best organization to oversee an independent probe into the SARS-CoV-2 origin, given its ongoing need to please donor nations. But he is pleased at its new push for answers. "I really do hope that good science can rule the day," he says. ■



World Health Organization chief Tedros Adhanom Ghebreyesus wants a new scientific body to conduct the next phase of studies into the origin of COVID-19.

sociate director of the National Emerging Infectious Diseases Laboratory Institute at Boston University, calls "the barrage of media and political commentary"—particularly sharp in the United States, WHO's largest funder—about a potential lab leak. "I think he's under enormous pressure, and he's capitulated," says Keusch, who co-authored two letters in *The Lancet* that favor the natural origin theory and criticize the "conspiracy theories" and speculation that fuel some lab-origin arguments. "It's sad." (Tedros declined an interview request.)

Earlier this year, WHO sent a team of international scientists to China to work with colleagues there on a joint mission to study the origin of SARS-CoV-2. The report

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](https://www.scienceadvances.org)

Science
Advances
AAAS

GOLD OPEN ACCESS, DIGITAL, AND FREE TO ALL READERS



FEATURES

A TROUBLED CALCULUS

Researchers use race to build disease risk assessment tools.
Can removing it help resolve medicine's race crisis?

To pediatrician Nader Shaikh, the rhythm of treating babies running high fevers is familiar. After ruling out the obvious colds and other common viruses, he must often thread a catheter into a months-old baby to draw a urine sample and check for a urinary tract infection (UTI). “You have to hold the baby down, the baby’s crying, the mother is usually crying too,” says Shaikh, who works at the University of Pittsburgh. “It’s traumatic.”

By **Jyoti Madhusoodanan**

UTIs, although relatively rare in children under age 2, carry a high risk of kidney damage in this group if left untreated. Often, the only symptom is a high fever. But high fevers can also signal a brain or blood infection, or a dozen other illnesses that can be diagnosed without a urine sample. To help clinicians avoid the unnecessary pain and expense of catheterizing a shrieking infant, Shaikh and his colleagues developed an equation that

gauges a child’s risk of a UTI based on age, fever, circumcision status, gender, and other factors—including whether the child is Black or white. Race is part of the equation because previous studies found that—for reasons that aren’t clear—UTIs are far less common in Black children than in white ones.

The UTI algorithm is only one of several risk calculators that factor in race, which doctors routinely use to make decisions about patients’ care. Some help them decide what tests to perform next or which

ILLUSTRATION: ANUJ SHRESTHA



patients to refer to a specialist. Others help gauge a patient's lung health, their ability to donate a liver or kidney, or which diabetes medicines they need.

In the past few years, however, U.S. doctors and students reckoning with racism in medicine have questioned the use of algorithms that include race as a variable. Their efforts gained momentum thanks to the Black Lives Matter movement. In August 2020, a commentary published in *The New England Journal of Medicine (NEJM)* highlighted the use of race in calculators as a problem “hidden in plain sight.” It's widely agreed that race is a classification system designed by humans that lacks a genetic basis, says Darshali Vyas, a medical resident at Massachusetts General Hospital and co-author on the paper. “There's a tension between that [understanding] and how we see race being used ... as an input variable in these equations,” Vyas says. “Many times, there's an assumption

that race is relevant in a biological sense.”

Vyas and others warn that using race to adjust risk calculators may also widen existing health disparities. Black Americans are generally diagnosed with kidney disease later than white Americans, which delays treatment and puts them at greater risk of developing kidney failure—yet an equation widely used to measure kidney function tends to estimate better function for Black patients relative to non-Black patients. Osteoporosis is underdiagnosed and undertreated in Black women, but a common bone fracture risk calculator places them, along with Asian and Hispanic women, at lower risk than white women. “We know these disparities exist, yet the calculators tell us that we don't need to worry about this population,” says epidemiologist Anjum Hajat of the University of Washington, Seattle.

Some of these calculations are rooted in racist assumptions. Others emerged out of an effort to improve predictions across racial groups. The challenge of defining “normal” versus “diseased” and capturing these qualities accurately in a simple test led scientists to grasp whatever data they could to make their tools more accurate. And at a population scale, race often does correlate with medical outcomes, in part because it acts as a proxy for the influence of other socioeconomic factors on health.

But even if racial trends sharpen predictions, using them to make decisions about an individual's treatment is problematic, Hajat says. “Even if a calculator is not causing disparities, it is maintaining and perpetuating them,” she says. For some, applying a different standard to Black patients than to white ones recalls a long history of neglect and discrimination in medicine. “I don't think people had bad intentions when they were creating these calculators,” Hajat says. “But we have to be aware that biomedical research has really contributed to upholding white supremacy, which is why we're reexamining the calculators now.”

The questions are already spurring change. In March, a task force from the American Society of Nephrology and the National Kidney Foundation recommended removing race as a variable in the kidney function calculator, known as the estimated glomerular filtration rate (eGFR) equation. The University of Washington, Beth Israel Deaconess Medical Center, and others have already dropped race from their eGFR calculations.

But similar efforts met resistance at other institutions. To some researchers and clinicians, the use of calculators that incorporate race seems not just appropriate, but a crucial measure to avoid unnecessary medication or invasive treatments, such as a catheter in a

6-month-old baby. Shaikh sees the UTI equation's use of race as an effort to achieve equity, not worsen disparities. “It sounds weird to use race to pick patients, and it doesn't look good on the surface,” he says. “But which one is worse: catheterizing kids who don't need it or using race in an algorithm? It's more complicated than it seems.”

THE HISTORY of racism in U.S. medicine dates back to the nation's earliest medical schools. Benjamin Rush, one of the physicians who signed the Declaration of Independence, once described Blackness as a form of leprosy that could be cured to restore the “natural white flesh color.”

At least two modern-day risk calculators have been accused of having similarly racist logic: One, which estimates a woman's odds of successful vaginal birth after cesarean section (VBAC), falsely assumes that women's pelvis shapes differ based on race, making this form of childbirth riskier for Black and Hispanic women compared with white women. Another equation estimates lung function by gauging the maximum amount of air a person can exhale forcefully into an instrument called a spirometer. Lower measurements are considered normal for Black and Asian people, based on the disputed assumption that their lung capacity is lower. “The spirometer was built on anti-Black racism,” says Lundy Braun of Brown University, who studies the history of racial health disparities. The VBAC calculator was updated to remove race in May, but spirometers still include a race adjustment. The American Thoracic Society (ATS) has begun to examine its use, Braun says.

In other calculators, race has been added to bring measurements in line with the best available data. The eGFR equation, developed in 1999, estimates how well a person's kidneys function based on urinary levels of a compound called creatinine, which builds up in blood when kidney filtration declines. Because the equation doesn't test kidney function directly, its developers compared its results with kidney filtration rates measured using a more definitive test, based on a radioactive tracer, that is too complex to perform routinely. They found the eGFR equation consistently underestimated kidney function in Black patients, so they used a common statistical method called curve fitting to adjust the estimates according to race.

Other risk calculators have added race in an effort to better match epidemiological data. In 1992, the World Health Organization recognized an epidemic of osteoporosis and funded research to develop a tool that could assess a person's risk for fractures based on the brittleness of their bones.

Researchers developed several country-specific versions of the tool for the United States, Canada, South Africa, and others, which incorporated race-specific prevalence where data were available.

When adapting the equation to U.S. populations, the researchers included a race correction to account for the lower reported occurrence of osteoporosis in Black women. The goal was to avoid medicating

from research studies of the condition.

Despite such uncertainties, it has been hard to resist including race in the calculators. “When you plug variables into a model and see a large effect, it seems like race is representing something and it should be in the calculator,” Hajat says. It’s convenient to see race as a variable like age, which is “predictive of everything related to health,” she says. “The problem is that race and

medical risk tool. The eGFR is especially controversial because “you can consider the trade-offs of accuracy versus the harms of the equation,” says medical student James Diao of Harvard Medical School, who is studying alternative, race-free equations. Adding race to the eGFR equation may have made it technically more accurate for Black patients, but it also results in fewer of them being diagnosed with chronic kidney disease.

Deleting race from the tool could counteract long-standing disparities in care, because more Black patients would receive earlier referrals to specialists and get placed on transplant lists sooner. (On the other hand, it might also mean fewer receive certain life-saving medications for blood pressure, diabetes, and other conditions because of a risk of renal side effects.) Similarly, an analysis presented at the ATS International Conference in May suggests nearly 21% more Black patients would be diagnosed with more severe pulmonary disease—and receive earlier care—if race were removed from the lung function calculator.

But removing race from medical equations is not a matter of simple math. The eGFR equation for kidney function is embedded in electronic medical systems used in hospitals and commercial laboratories. A technician using a spirometer to test lung function must begin by entering a patient’s race along with age, height, and other details. And the fracture risk calculator is built into scanners that test bone density.

Aside from the technical challenges of updating instruments, testing labs are obliged to follow current regulations and standards of care endorsed by professional societies and used in clinics. Without a formal change in guidelines, individual test providers or clinics may find it tough to discard risk models that include race.

Legal constraints might also make it difficult for physicians to phase out race-based calculations, such as those used to gauge the risks of a surgery. “From a liability standpoint, surgeons might be compelled to use what’s validated as the most accurate equation,” Diao says. “Otherwise, it might be seen as a failure to adequately inform a patient about a procedure’s risks prior to consent.”

WHEN VYAS and her colleagues published a list of problematic algorithms in their *NEJM* commentary, it raised concerns—and ire—across clinical specialties. In September 2020, osteoporosis researcher John Kanis of the University of Sheffield, who developed the bone fracture risk tool, published a commentary in the journal *Osteoporosis International* arguing that using country-specific incidence trends, including by race, is important to the tool’s accu-

Deeper than skin color

Doctors use risk calculators to help decide a person’s prescriptions, their risks during surgery or childbirth, or when to refer them to a specialist. Race corrections sway this math in varied ways.

Lung

Instruments used to measure lung function set lower values as the “normal” range for Black patients. The adjustment could delay or prevent diagnoses of lung disease.

Heart

One equation assessing the risk of death after heart failure gives lower estimates for Black patients than white patients. Doctors use the tool to prioritize high-risk patients for treatments.

Breast

A breast cancer risk equation places Black, Hispanic, and Asian American women at lower risk of disease. Low-risk patients may miss out on screening and preventive care.

Kidney

A race adjustment makes results of a kidney function test appear better in Black people. That could delay medication, referrals, or other supportive care.

Uterus*

A childbirth risk calculator gave Black and Hispanic women lower odds of a successful vaginal childbirth after a previous cesarean, making doctors more likely to recommend a second surgical delivery.

Bladder*

A urinary tract infection risk calculator estimated Black children have lower odds of UTIs, potentially reducing the likelihood that doctors would run additional tests.

Bone

A fracture risk tool places Black women at lower risk of osteoporosis. High-risk patients receive preventive drugs to minimize fractures.

*Indicates tests that have been updated to remove race.

people who didn’t need it, and the correction brought fracture predictions in line with official rates of disease.

And in this case, the differences may have a physiological underpinning, says epidemiologist Nicole Wright of the University of Alabama, Birmingham, who studies disparities in bone health. “Genetically, people of African descent have higher bone mass than others, so you need to account for that,” she says. “If you don’t have osteoporosis and you’re taking these medications, they do come with some risks.”

It’s unclear, however, whether the lower incidence of osteoporosis in Black women is also influenced by missed diagnoses due to lack of access to care, delayed screening, or the exclusion of these women

age are fundamentally different things.”

If race doesn’t represent a biological difference between patients, why would including it in risk equations improve predictions at all? Even researchers who develop such equations often don’t know exactly why race matters. “We sometimes use surrogate measures that can identify people at different levels of risk, even if we don’t understand the exact factors driving their risk,” says epidemiologist Montserrat García-Closas of the National Cancer Institute, who has worked on various cancer risk calculators. “It doesn’t really matter if these are true [disease-causing] factors, which are much harder to establish.”

Also hard to establish are the public health consequences of including race in a given

racy and avoids overdiagnosis. He and his co-authors note, for example, that Black people in the United States have lower fracture risk than white Americans, but their risk is far higher than that of Black people in African countries.

Although the underdiagnosis of osteoporosis among Black people in the United States is a problem, “I don’t think removing race from the calculator would do anything to reduce disparities,” says osteoporosis researcher Michael Lewiecki of the University of New Mexico Health Sciences Center. “These are important problems, but the calculator is not their cause.”

At a faculty meeting last summer, Shaikh says his colleagues discussed whether to respond to the *NEJM* commentary to refute the suggestion that the use of race in algorithms is always problematic. But Shaikh didn’t see the commentary as a call to change the UTI risk calculator. “To me, it reflected a desire to bring the issue [of race] to the surface, which is laudable,” he says. “I don’t know who is right or wrong here. My view is, these tools have to be based on data.”

It’s tough to know whether the UTI calculator leads to disparities, and there’s little evidence to suggest UTIs are underdiagnosed or undertreated in Black children. But social justice movements in recent years prompted Shaikh to re-examine the data he used to create the UTI calculator, including the information in previous studies that suggested the occurrence of infections varied by race. “It’s true what the criticisms say that using race is not free of problems,” he says. “If we test people based on race, we are creating a difference. The question is if that difference is to the patient’s benefit—I don’t think the difference itself bothers me as much as the idea that we might cause harm.”

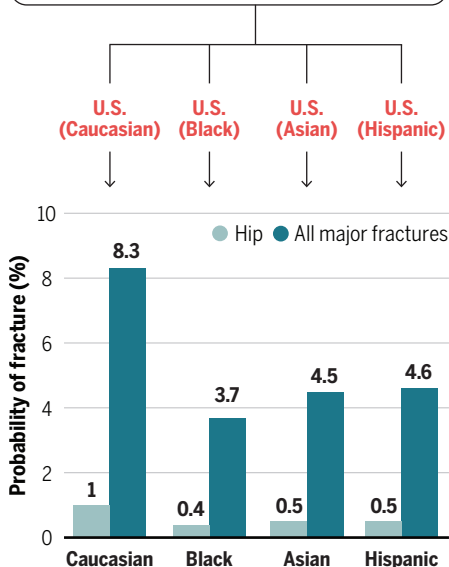
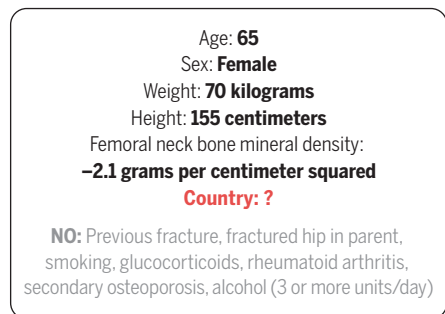
Going forward, any large study that relies on racial differences to develop models of disease risk should undergo additional review before publication to consider potential consequences of applying the research to medicine, says nephrologist Nwamaka Eneanya of the University of Pennsylvania. Researchers also need to go beyond correlating race with health outcomes to pinpoint the actual drivers of health disparities, such as income, education, or neighborhood environmental exposures, she adds. “That’s not a standard that is expected of scientists in this day and age, and it needs to be,” Eneanya says. “This is a wake-up call for the scientific community.”

Replacing race with a different metric is not always easy. Recent studies have attempted to use ZIP codes, income or education levels, or a measure of socioeconomic

Fractured risks

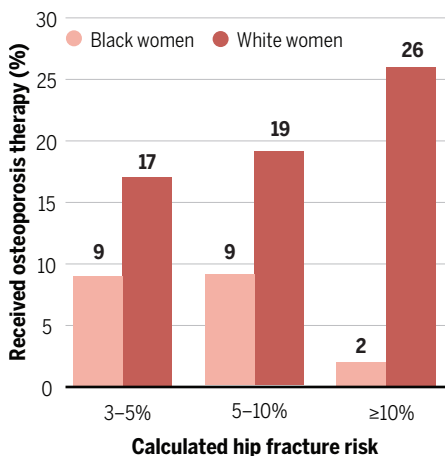
A bone fracture risk assessment tool includes individual patient details and history. But it also accounts for race, placing Black, Asian, and Hispanic people at lower risk of osteoporosis than white people, based on the lower incidence of the condition in these groups.

Fracture risk assessment tool inputs



Delayed doses

Lifestyle changes or treatment are recommended when fracture risk exceeds 3%. But in a recent study, Black women were consistently less likely to receive prescription osteoporosis therapies—even when they were at high risk.



status called the area deprivation index instead of race to capture conditions that influence health. Precisely how they’d be implemented isn’t clear, and few have been put to work in clinics or endorsed by professional societies of clinicians.

But once researchers and clinicians commit to equity, they often find good alternatives to the use of race, Diao says. In addition to recommending the removal of race from the eGFR equation, nephrology researchers are evaluating a handful of different race-free equations that combine creatinine with other biomarkers such as the protein cystatin C. The tests seem to perform just as well, though they will need further validation and can be more expensive.

Researchers who developed the child-birth risk calculator, which assumes pelvis shapes vary by race, published a new version in May that differs from the original only in the removal of race—and performs just as well. “Presenting this as a choice of using a more accurate equation with race or a less accurate one without it is a false dilemma,” Diao says. “When there’s enough pressure to create an equation that is both accurate and does not use race, researchers rise to the challenge.”

Last year, Shaikh began to work with community organizers to gather parents’ perspectives on the infant UTI calculator. In virtual meetings with these participants, he reenacted the familiar emergency room scenario: A research assistant played the worried parent with a feverish child, while Shaikh played the doctor and explained the use of the calculator.

One participant, a Black father, understood the need to keep race in the equation and drew a parallel to affirmative action. Still, he wondered, was there an alternative? “He didn’t have a problem with it as long as it improved outcomes,” Shaikh recalls. “But given the history of racism in the U.S., it’s a lot to expect people to just trust that it’s going to improve outcomes.” Shaikh eventually homed in on two replacement variables for race in the calculator: the duration of a child’s fever and a prior history of UTIs. He validated the tool and updated the online calculator to a race-free version this week.

Replacing race doesn’t imply researchers shouldn’t continue to seek causes for disparities, Shaikh warns: A child’s history of UTIs is useful in the calculator in part because it captures differences between Black and white children. “We didn’t solve the problem: The data still show a link between race and UTI,” he says. “It’s important to understand that, not bury it.” ■

Jyoti Madhusoodanan is a journalist in Portland, Oregon.

INSIGHTS

POLICY FORUM

RESOURCE POLICY

Global implications of the EU battery regulation

A much-needed law may have unintended global consequences

By **Hans Eric Melin**¹, **Mohammad Ali Rajaeifar**^{2,3}, **Anthony Y. Ku**^{4,5}, **Alissa Kendall**⁶, **Gavin Harper**^{3,7}, **Oliver Heidrich**^{2,3,8}

Transport electrification is a key element of decarbonization strategies; thus, the design, production, manufacture, use, and disposal of lithium-ion batteries (LIBs) are taking center stage. The environmental, economic, and social consequences of the battery life cycle are high on political agendas, owing to exponential growth in metals extraction; the climate impacts of battery production; and uncertainties in battery end-of-life (EOL) safety, recyclability, and environmental consequences (1) [see figs. S1 to S3 in the sup-

plementary materials]. The European Union (EU) has proposed a new Battery Regulation (2) that intends to ensure sustainability for batteries placed on the EU market (see the figure), developing a robust European battery industry and value chain. The Regulation is very much needed, but, as discussed below, it will have global implications, with perhaps some unintended consequences. If left unaddressed, the Regulation, at worst, could hamper climate change mitigation targets and fall short of its intentions to promote a circular economy and establish a socially acceptable raw material supply chain.

The proposed Regulation will build upon and replace the 2006 EU Battery Directive (3) to address this era's challenges, stipulating

labeling and information provisions, setting out supply chain due diligence requirements, and enforcing the use of recycled materials for batteries over 2 kilowatt-hours (kWh), most of which are used in electric vehicles (EVs). Historical data show an almost 10-fold increase of LIBs placed on global markets in the past 10 years (see figs. S4 and S5), and a similar growth rate is expected during this decade (see fig. S1). Therefore, it is imperative to address the global climate change challenge and battery issues in tandem (4–6).

Despite being the second-largest market for EVs in the world, Europe does not dominate LIB supply chains, instead relying on global markets for raw material extraction, refining, and battery manufacturing (7).

Voltage of a used lithium-ion car battery is checked by an employee of the German recycling firm Accurec in Krefeld, Germany, 16 November 2017.

From the moment the Regulation enters into force, potentially in 2023, producers and importers of batteries will need to comply with the Regulation to sell or use batteries on the European market. EVs not only help achieve global carbon reduction targets; they also serve as a driver of growth and job creation in Europe (8), the United States (9), China (4), and a few other markets such as South Korea and Japan. As such, the Regulation must balance concerns about environmental stewardship with regional interests in economic competitiveness. The Regulation thus comes with explicit aims to obtain a global competitive advantage, build barriers to entry, and provide incentives to invest in production capacity for sustainable batteries.

The new Regulation (2) brings European legislation up-to-date, anticipating near-term developments (in 2035) (see fig. S6). We distill the Regulation's 79 Articles into four key elements that are core to improving the sustainability of LIBs. First, the Regulation intends to increase transparency and traceability across the battery life cycle (e.g., Articles, 10, 47, and 65; see the figure), mandating third party due diligence of the supply and value chains for batteries >2 kWh (e.g., Articles 8, 39, and 72). Second, it addresses climate impact throughout the battery life cycle (e.g., Article 7) by mandating carbon footprint declaration and later establishing maximum thresholds. These first two elements seek to ensure the quality and availability of underlying data needed to guide environmental responsibility and track compliance by industrial players. Third, concrete actions to promote the circularity of critical materials are emphasized (e.g., Articles 47, 55, and 57), targeting increased collection and recycling efficiency; improving recovery rates for lithium, cobalt, and nickel; and mandating the use of recycled materials in new batteries (e.g., Articles 8, 55, and 59). Fourth, requirements for longevity and performance management are proposed, including access for waste processors to the battery management system (BMS), which verifies the state of health of the battery in real time and can determine the potential for the battery to be reused or repurposed before being recycled (e.g., Articles 51, 59, and 65; see the figure). These last two elements provide a foundation for EU efforts to establish circularity as an economic opportunity aligned with its long-term sustainability objectives.

A GLOBAL SYSTEM

In addition to the proposed Regulation, the European Commission is supporting battery developments through a range of initiatives such as the creation of the European Battery Alliance and financial aid packages to support research and innovation along the entire battery value chain (4). Under the label "open strategic autonomy," the union seeks to combine free trade with an ability to manage its own destiny in key sectors. The LIB market is dominated by Chinese companies, which occupy more than two-thirds of the supply chain (7), including resources outside of China. The United States market has, through Tesla, been on the forefront of large-scale battery production but has, like Europe, failed to secure key elements of the supply and value chain such as refining, production, or recycling of battery materials (see fig. S7). The regulatory environment in each market will play a key role as electrification of the transportation sector intensifies.

In China, the LIB market has profited from rapid growth in the EV sector, fueled by strong government support in the form of both subsidies and investment stimulus. A whole-of-government approach has provided clear strategic guidance; and the Interim Measures for the Administration of the Recycling and Utilization of Power Batteries for New Energy Vehicles from 2018 reflect a number of the themes found in the EU Regulation (10). These included minimum standards for the reclassification of batteries for reuse applications, the recycling efficiency of plants treating batteries at the EOL, and requirements for product labeling to qualify for subsidies. In 2019, these regulations were tightened to levels stricter than those the EU Regulation plans to enforce a decade from now (2). In December 2020, China's State Council issued "Energy in China's New Era," a blueprint for energy sector development through 2030. It includes guidance on the development of battery supply chains, including recycling to support energy efficiency, and provisions for reducing the carbon intensity of electricity used to power EVs (9).

In the United States, electrification of transportation, and the supply of critical materials that enable it, has been declared a top priority of the Biden administration. This is reflected in an investment proposal of up to \$174 billion in EV programs focused on reestablishing primary metals refining and manufacturing in the United States, and creating new recycling capabilities as part of materials criticality efforts. In the policy arena, however, the United States lags both the EU

and China in mandating Extended Producers Responsibility or promoting circular economy principles other than basic legislation to keep batteries classified as hazardous waste out of landfills and incinerators.

CHALLENGES TO OVERCOME

Current battery legislation in all three major markets—China, the EU, and the United States—is geared toward the protection of local environment and human health. However, the scopes of China's and the EU's policies include measures that, directly or indirectly, give companies that manage to comply important competitive advantages both domestically and globally, the mandatory recovery rates for recyclers being a standout example (see table S1). An important distinction arises, however, in the manner by which Chinese and EU regulations affect the global value chain. Compared to the EU, China has more direct leverage owing to its dominance in materials refining, battery material and cell production, and a much more mature recycling infrastructure. Although tougher requirements for placing batteries on the European market force players around the world, including the EU, to comply, the domestic requirements imposed on the Chinese companies have already helped them do that.

This imbalance between new and mature markets is one of several factors that can lead to unintended consequences, such as described below, not just in Europe but globally, when a nascent market is shaped by regulation.

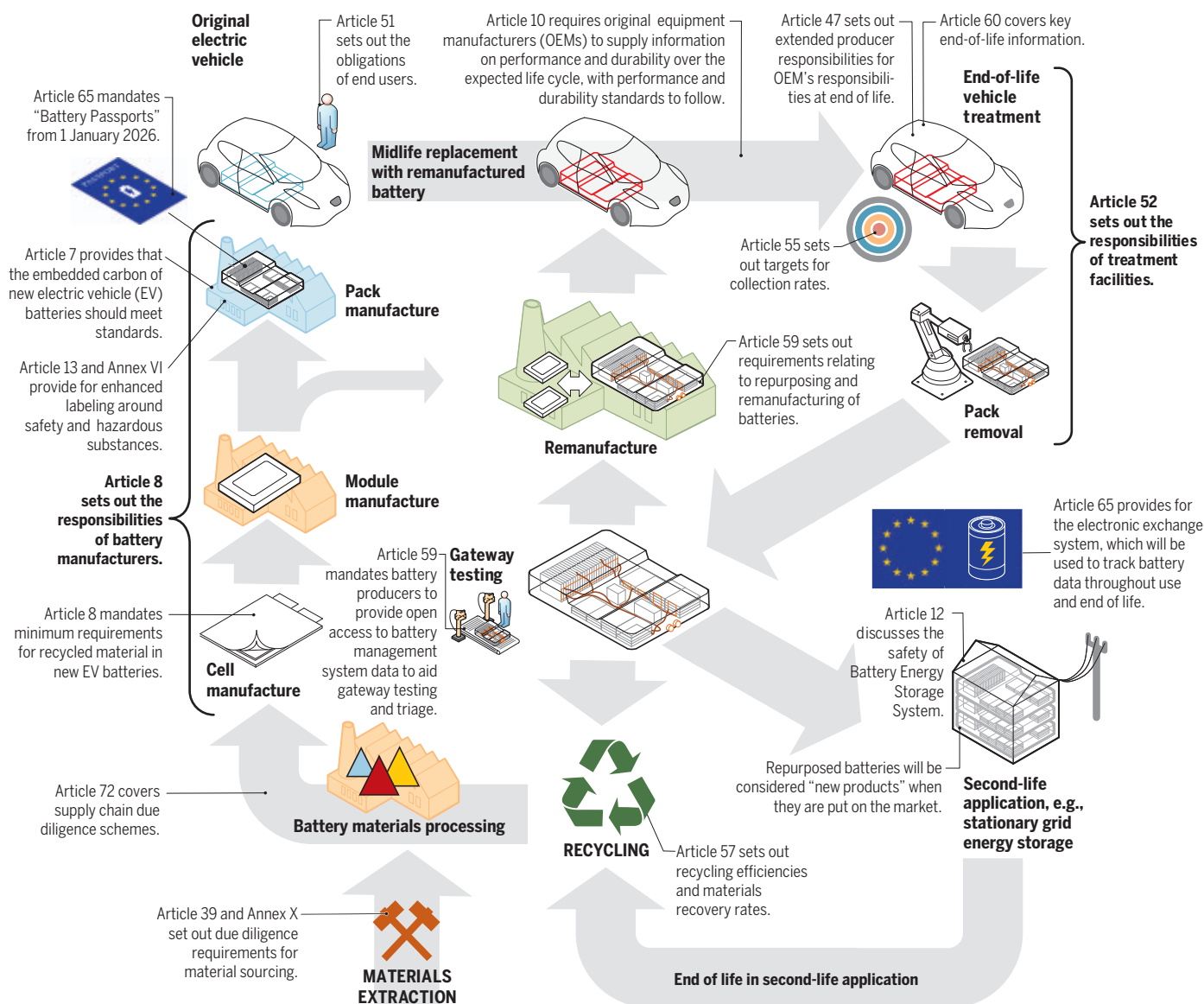
Distorted innovation

The LIB industry is experiencing both rapid growth and innovation. Together, these dynamics create uncertainty around how the markets will evolve, making it difficult for industry to properly invest for future growth. In this context, regulations offer considerable impetus by creating stable frameworks under which market players operate. Overly stringent regulations on batteries risk imposing compliance costs that hinder competitiveness, leading to reduced innovation and lower EV adoption rates. Few, if any, other technological products appear to be presently regulated in quite the same way—with manufacturers needing to provide details on product life cycle, material sourcing, third-party audit, and recycled content. This means that batteries, the product at the core of transport decarbonization, may in some respects face tougher requirements than the incumbent industry serving the internal combustion engine vehicle market. Even if

¹Circular Energy Storage Research and Consulting, London, UK. ²School of Engineering, Newcastle University, Newcastle upon Tyne, UK. ³Faraday Institution (ReLiB project), Quad One, Harwell Science and Innovation Campus, Didcot, UK. ⁴National Institute of Clean and Low Carbon Energy (NICE) America Research, Mountain View, CA, USA. ⁵NICE, Future Science City, Changping District, Beijing, China. ⁶Department of Civil and Environmental Engineering, University of California, Davis, CA, USA. ⁷Birmingham Centre for Strategic Elements and Critical Materials, University of Birmingham, Edgbaston, UK. ⁸Tyndall Centre for Climate Change Research, Newcastle University, Newcastle upon Tyne, UK. Email: oliver.heidrich@ncl.ac.uk; hanseric@circularenergystorage.com

A circular economy for electric vehicle batteries: Key articles from the proposed EU Battery Regulation

The proposed Regulation addresses the battery life cycle, from initial extraction of raw materials (bottom left) through end of life and recycling.



stricter emission standards are underway, there are no proposals to provide the same value chain transparency.

The new regulation also means that European EV manufacturers will be more constrained in their options to source batteries than manufacturers in the less regulated United States but also in a strongly regulated China as their share of the LIB market is simply so much bigger and their players much more mature. The fact that smaller producers are not exempted from the many potentially burdensome requirements presents a considerable risk that innovation in alternative transportation areas and technological niches may be disadvantaged.

A second mechanism for distorted inno-

vation involves situations where regulatory guidance lags the technical realities in the market. In an extreme case, specific regulatory targets might become obsolete even before they enter into force, owing to the introduction of new chemistries or battery designs. A second case concerns unexpected "work-arounds." Rather than direct compliance on specific materials, producers might opt for raw material substitutions; such decisions upstream in the value chain could unintentionally weaken future markets for recyclers expecting different levels of material availability. In addition, manufacturers might simply incorporate the threshold of 2 kWh as a design parameter to avoid the heavy burden of supply chain due diligence and the

use of recycled content, which could make the Regulation less effective. A third case of distorted innovation is that the requirements for recycled content could hamper the pace of the roll-out of new technologies as well as higher ambitions for the overall growth of the LIB market. This could be particularly relevant for individual manufacturers, as the availability of recyclable materials effectively establishes a limit on the amount of certain materials that can be used in the batteries (11). All three cases reinforce the importance of continual monitoring and adjustment of regulations to ensure that outcomes remain aligned with both their environmental stewardship and economic goals. This importance has been addressed in work preceding the

Regulation that motivated the use of “secondary legislation” for specific measures, including recycled content, which will be adaptable over time. However, there is a trade-off between adaptability and stability, which is one of the intentions of the regulation.

Material leakage, investment slowdown

One objective of the Regulation is to responsibly secure material supply chains for Europe’s EV industry. Recycling is here a core activity, with recycled content and recovery targets as main measures. However, the Regulation neither requires recycled material to be sourced in Europe nor restricts the source of recyclables to EOL batteries. With considerable experience in battery material production, including the use of recycled materials from both EOL batteries and production scrap, several companies in markets like China and South Korea are as well, and likely better, positioned to meet EU requirements for recycling than European companies. With China’s and South Korea’s dominance in battery production, these companies also have better access to feedstock of recyclables, which already has provided many of the companies with economies of scale and thus a competitive advantage in sourcing recyclables from both Asia and Europe. This extends to both Asian battery material producers, which are the users of the recycled materials, and battery manufacturers with close access to these materials. The net effect of this dynamic could be a higher barrier to entry for European material producers, and ultimately battery manufacturers, in their own market. This issue has been raised by battery manufacturers in consultations both before and after the adoption of the Regulation.

This potential outflow of recyclable materials comes on top of the fact that battery-containing products, as well as the batteries themselves, usually cascade through multiple owners and are traded on international markets (11). EOL batteries still have value and are often exported for reuse and refurbishment, or recycling (12). This outflow makes it more difficult for recyclers in Europe to reach the necessary economies of scale, making them less competitive and less attractive to invest in.

The Regulation’s clarification of producer responsibility for repurposed batteries, as well as its waste status, and the mandated access to BMS data might help keep batteries in the EU, facilitating the manufacturing of new battery products, which ultimately can reach end of life in Europe. However, it also provides a route for traders to declare the used battery “a resource or product,” which facilitates its export, adding to the leakage of recyclable materials. Access to EOL batteries and recyclable materials is primarily a concern

for the companies involved in the end-of-life value chain. However, with requirements for recycled content in the batteries, this dependency extends to the entire industry, including the manufacturing of electric vehicles.

CONCLUSIONS

The EU Battery Regulation has admirable intentions—from spurring the growth of domestic industries, to environmental protection at local and global scales, to achieving a truly circular economy. All these goals are intertwined, and one could reasonably argue that they need a unified policy. However, development of largely uncoordinated regional and domestic policies across globally important consumption and production regions makes the consequences of the proposed EU Regulation hard to predict. Although unintended consequences might first affect the EU, they will also be felt by manufacturers, recyclers, and other actors across the world. Ultimately, disruptions in the European battery value chain might limit automotive makers’ ability to produce EVs at the scale that is required by 2030, when several countries have placed a ban on the sale of new internal combustion engine vehicles.

Efforts to disentangle the goals of different measures and instruments, and better understand the effects on global supply and value chains, may be made more predictable and powerful if addressed by the global community in coordinated fashion. For example, shifting to open BMS architectures (namely, battery passports, a digital representation of battery environmental, social, governance, and life cycle-specific information in Article 65) represents a major departure from existing business models in which manufacturers maintain tight control over the data generated. It is unknown whether and how locally acquired battery data will be shared across global value chains. What is likely is that in a unified global market, the dominant market’s standard will drive compliance.

Having clear and stringent global standards in areas where regulations are limited in scope such as recycling, circularity, and cascaded use for batteries could provide EU firms with a first-mover advantage in achieving product differentiation in battery markets. Perhaps this can form a non-tariff barrier to cheaper imported products, affording some protection to European battery makers. Over the longer term, a more robust global supply chain with many players is good for the environment overall—and will benefit innovative players regardless of their origin.

Because standards on battery supply and value chains are uneven around the world, establishing global standards that all coun-

tries adhere to could help ensure a level (and sustainable) playing field and facilitate coordination on these global challenges and solutions. The EU Regulation is among the most advanced environmental standards influencing environmental stewardship and sustainability. Though embedded resistance from market players alongside the challenges inherent in international negotiations presents formidable hurdles to the establishment of global standards, the EU’s share of the EV market provides it with considerable leverage to set standards that might be accepted on a de facto basis. In similar fashion, there should be an ambition to align other product areas with the same standards, ensuring that the battery industry is not disadvantaged in relation to other energy storage and power technologies.

The EU Regulation places the EU on the forefront of regulating battery markets, providing much-needed policies and legislation that address environmental and social issues. However, these policies need to contemplate the unintended consequences, both in the near and distant future. ■

REFERENCES AND NOTES

1. T. L. Curtis, L. Smith, H. Buchanan, G. Heath, “A Circular Economy for Lithium-Ion Batteries Used in Mobile and Stationary Energy Storage: Drivers, Barriers, Enablers, and U.S. Policy Considerations” [National Renewable Energy Laboratory (NREL), 2021].
2. European Union, “Proposal for a regulation of the European parliament and of the council concerning batteries and waste batteries, repealing Directive 2006/66/EC and amending Regulation (EU) No 2019/1020” (European Commission, Brussels, 2020).
3. European Union, in *OJ L266/1* (Official Journal of the European Union, Brussels, 2006).
4. “State aid: Commission approves €2.9 billion public support by twelve Member States for a second pan-European research and innovation project along the entire battery value chain” (2021); https://ec.europa.eu/commission/presscorner/detail/en/IP_21_226.
5. G. Crabtree, *Science* **366**, 422 (2019).
6. B. K. Sovacool et al., *Science* **367**, 30 (2020).
7. J. Baars, T. Domenech, R. Bleischwitz, H. E. Melin, O. Heidrich, *Nat. Sustain.* **4**, 71 (2021).
8. “FACT SHEET: The American Jobs Plan” (2021); www.whitehouse.gov/briefing-room/statements-releases/2021/03/31/fact-sheet-the-american-jobs-plan/.
9. “Energy in China’s New Era” (2020); www.xinhuanet.com/english/2020-12/21/c_139607131.htm.
10. State Council, Ministry of Industry and Information Technology (Ministry of Industry and Information Technology, Beijing, China, 2018).
11. R. Sommerville et al., *Resour. Conserv. Recycling* **165**, 105219 (2021).
12. UNEP, “Used Vehicles and the Environment” [United Nations Environment Programme (UNEP), Nairobi, Kenya, 2020].

ACKNOWLEDGMENTS

This research was supported by the UK’s Engineering and Physical Sciences Research Council (EPSRC), the Faraday Institution (EP/S003053/1), and its Recycling of Li-Ion Batteries (ReLIB) project (FIRG005). None of the funders had input or a role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the study or this manuscript. We thank reviewers for comments.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/384/suppl/DC1

10.1126/science.abh1416

PERSPECTIVES

PLANETARY SCIENCE

The interior of Mars revealed

Direct seismic observations provide clues to the red planet's structure and evolution

By **Sanne Cottaar** and **Paula Koelemeijer**

The interior of a planet holds important clues to its origin and thermal and dynamic evolution. Exploring a planet's deeper layers can reveal how a planet accreted and differentiated into layers, whether its core sustains a geodynamo that creates a magnetic field, and what the origin is of any tectonic and volcanic activity. New clues are now uncovered for Mars. On pages 438, 434, and 443 of this issue, Knapmeyer-Endrun *et al.* (1), Khan *et al.* (2), and Stähler *et al.* (3), respectively, present the first findings of the interior structure of Mars based on data from the NASA lander InSight (Interior Exploration using Seismic Investigations, Geodesy and Heat Transport). These studies provide the first direct observations of the crust, mantle, and core structure on another rocky planet, for which the results and implications can be compared and contrasted with the characteristics of Earth.

A planetary subsurface can be probed with recordings of seismic waves that propagate through the planet after a quake. For Earth, the first measurement of waves diving deep through the mantle was made in 1889 in Germany from an earthquake in Japan (4). Global seismology is thus a relatively young science. In the first decades of the 20th century, seismologists discovered the main layers of Earth—that is, the boundaries between the crust and the mantle 30 km below Earth's surface (5), the mantle and core roughly halfway to the center (6), and the small inner core (7).

In a great feat of engineering, InSight landed on Mars in November 2018. Its seismometer, SEIS (Seismic Experiment for Interior Structure), started recording marsquakes in February 2019. The measurement of ground movements on Mars comes with a range of challenges. Although the sensitive sensors in SEIS measure ground movement along three axes and across frequencies, its recordings are only obtained at a single location. This limits how well it can constrain marsquake

epicenters and timings. And although a protective shield surrounds SEIS, its sensors still record perturbations caused by weather variations, including atmospheric pressure waves and dust storms. These signals give interesting clues about the martian atmosphere and its daily and seasonal variability (8), but for those studying marsquakes and Mars' interior, these signals are generally considered “noise.” The geology of Mars also creates challenges. Its rocky soil surface disperses seismic energy near the seismometer, complicating the identification of waves.

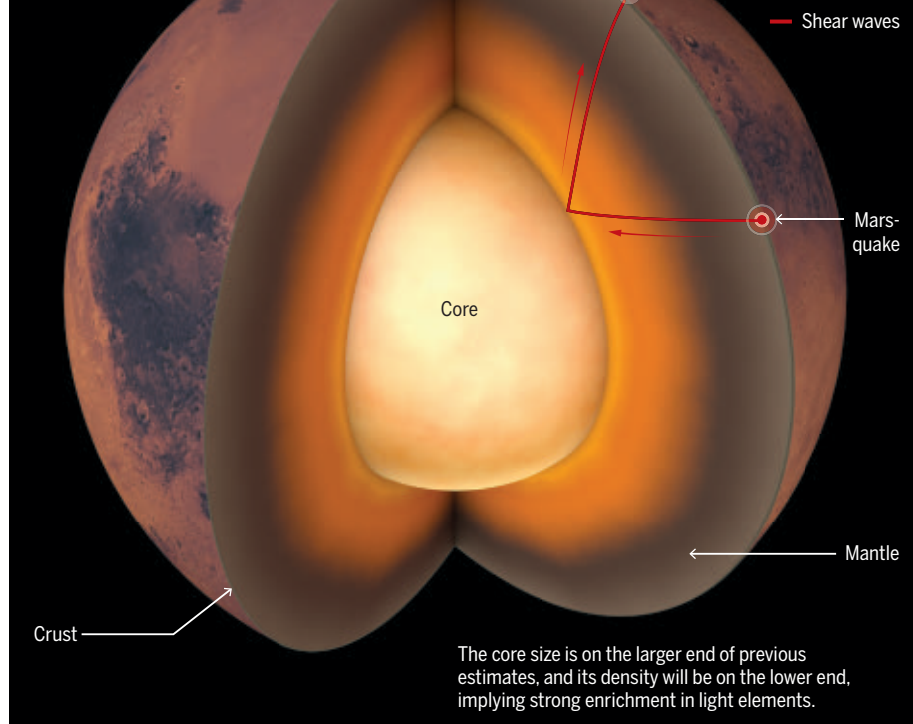
Results from the InSight mission have already demonstrated that Mars is seismically active. Marsquakes are plentiful, albeit small in magnitude (9). All marsquakes observed in the 2 years of recording are estimated to have magnitudes below 4.0, which humans would only notice within several kilometers of the epicenter. The vast majority of marsquakes originate in the crust and create strong reverberations within, making individual waves difficult to identify (9). A smaller number are generated below the crust, and their appearance resembles tectonic events on

Martian core measurements

Seismic waves tell us about the properties and boundaries of a planet's interior. Shear waves that travel from a marsquake and reflect off of the iron-nickel core are detected by the InSight seismometer and give us an estimate of the core size. The strength of the reflected waves shows that the core is in a liquid state, which shear waves cannot propagate through.

The InSight robotic lander reached the surface of Mars in November 2018 with a suite of geophysical instruments.

The seismometer was deployed on the surface in December 2018 and became operational in February 2019.



Earth. The studies of Knapmeyer-Endrun *et al.*, Khan *et al.*, and Stähler *et al.* used data from about 10 of these subcrustal marsquakes to reveal the inner structure of Mars (see the figure).

To overcome the challenges of working with noisy and sparse seismic data, the studies prepared and benchmarked a range of data analysis techniques, exploring signals in different frequency ranges and combining data from multiple marsquakes. For all three studies, independent research groups applied their methods of expertise to come to robust results or provide complementary constraints. Notably, seismologists worked together with a broad team of planetary scientists specializing in petrology, geochemistry, mineral physics, and geodynamics to understand the results and the implied structure and evolution of Mars. Seismic data were combined with compositional constraints from martian meteorites as well as with other geophysical observables such as surface topography and heat flow, gravity and geodetic data, and observations of crustal magnetization.

Although the detection of waves from small marsquakes is far from trivial, waves bouncing off a martian core will be even weaker and are thus more challenging to observe. Stähler *et al.* report the first observations of faint signals bouncing off the boundary to the martian core. They determined that the core starts about 1560 km deep, nearly halfway to the center. This is on the large end of previous estimates. Given the known mass and moment of inertia of Mars, this implies that the core is less dense than previously thought and that its iron-nickel alloy must be strongly enriched in sulfur and other light elements. The strength of the bouncing waves confirms that the core is still in a liquid state, as has been suspected from Mars' tidal response (10) and for a sulfur-enriched alloy at estimated martian core temperatures above 1800 K (11).

The observation of a relatively thin mantle means that Mars lacks the dense, insulating layer of the bridgmanite mineral that becomes stable under large pressures in Earth's mantle. The absence of this mineral would have led to a more rapid cooling of the early martian core, potentially driving a geodynamo to create a magnetic field. This geodynamo has since ceased and is now only evidenced by magnetized older crustal rocks (12). The magnetometer on InSight found that magnetization in the

crust observed at the surface is 10 times stronger than that modeled on the basis of satellite data. These new observations imply that Mars' early geodynamo must have been similar in strength to Earth's present-day geodynamo (13).

Instead of using seismic waves that bounce off the core, Khan *et al.* used waves that either travel directly from the marsquake to the seismometer or bounce off Mars' surface to reveal the shallow martian mantle structure. The travel times and amplitudes of these waves show that the seismic shear-wave speed decreases gradually in the martian mantle down to depths of 400 to 600 km. This reduction in seismic wave speed could be caused by the thermal structure in a static, thick outer shell (the lithosphere) on top of a convecting mantle. Closer to the surface, Knapmeyer-Endrun *et al.* imaged the local martian crustal structure by identifying energy conversions from shallow layers through a range of methods. The study was agnostic as to whether there is a 20-km-thick two-layer crust or a 39-km-thick three-layer crust at the InSight landing spot. Both crustal thickness models point to the subsurface crust as less dense than the surface materials to different degrees, indicating that the material has been highly altered over time.

These three studies provide important constraints on the present-day structure of Mars and are also key for improving our understanding of how the planet formed billions of years ago and evolved through time. Knapmeyer-Endrun *et al.* and Khan *et al.* both model the cooling and differentiation history of Mars and test which parameters result in the proposed crustal and thick lithospheric structure. They find that the crust must be 13 to 21 times more enriched in radioactive heat-producing elements compared with the mantle. This is greater than estimates based on measurements of surface materials and puts new bounds on Mars' crustal composition and formation. The models also find that the mantle beneath the thick stagnant lithosphere convects sluggishly.

The observations of a highly enriched crust, a thick thermal lithosphere, a sluggish mantle, and the lack of an insulating lower mantle will now have to be investigated further in dynamical mantle models. Such models will test whether internal dynamics, rather than a giant impact, could have caused the strong topographic dichotomy of Mars—the heavily cratered southern highlands and the smooth plains of the northern lowlands (14)—or whether a single mantle plume could have produced the volcanism beneath the broad Tharsis Rise, the most extensive topographic feature on the planet (15). These kinds of dynamic processes control the rate of volcanism, volatile outgassing, and early habitability.

The decrease in seismic velocities across the shallow mantle and the presence of a relatively large core both contribute to more bending of seismic energy from marsquakes deeper into the planet. This predicts the existence of so-called seismic shadow zones—less direct or no seismic energy would arrive at greater distances from a marsquake. SEIS would thus not observe marsquakes at certain distances, thereby underestimating the seismic activity on Mars. Crucially, because of the larger core, SEIS lies in the seismic shadow zone for seismicity in the tectonically and volcanically active Tharsis Rise. On the positive side, the preliminary models of the martian mantle presented by these studies will help locate more subcrustal marsquakes and identify more core-bouncing waves, possibly even core-traversing waves. With the InSight mission currently extended until the end of 2022, the number of high-quality observations is expected to double, leaving plenty of opportunity for adding detail and improving models of Mars.

Direct seismic observations on Mars represent a major leap forward in planetary seismology. The size of the martian core, the crustal layering, and the thick lithosphere provide important insights into the thermal and dynamic evolution of Mars. Over the coming years, as more marsquakes are measured, scientists will refine these models of the red planet and reveal more of Mars' enigmatic mysteries. ■

REFERENCES AND NOTES

1. B. Knapmeyer-Endrun *et al.*, *Science* **373**, 438 (2021).
2. A. Khan *et al.*, *Science* **373**, 434 (2021).
3. S. C. Stähler *et al.*, *Science* **373**, 443 (2021).
4. E. von Rebeur-Paschwitz, *Nature* **40**, 294 (1889).
5. A. Mohorovičić, *Potres od 8. X. 1909* (Školska knjiga, 1910).
6. R. D. Oldham, *Q. J. Geol. Soc.* **62**, 456 (1906).
7. I. Lehmann, *Publications du Bureau Central Seismologique International* **A14**, 87 (1936).
8. D. Banfield *et al.*, *Nat. Geosci.* **13**, 190 (2020).
9. D. Giardini *et al.*, *Nat. Geosci.* **13**, 205 (2020).
10. C. F. Yoder, A. S. Konopliv, D. N. Yuan, E. M. Standish, W. M. Folkner, *Science* **300**, 299 (2003).
11. A. J. Stewart, M. W. Schmidt, W. van Westrenen, C. Liebske, *Science* **316**, 1323 (2007).
12. D. J. Stevenson, *Nature* **412**, 214 (2001).
13. C. L. Johnson *et al.*, *Nat. Geosci.* **13**, 199 (2020).
14. S. Zhong, M. T. Zuber, *Earth Planet. Sci. Lett.* **189**, 75 (2001).
15. H. Harder, U. R. Christensen, *Nature* **380**, 507 (1996).

10.1126/science.abj8914

QUANTUM COMPUTING

Coherent manipulation of a spin qubit

A new device opens new avenues for spin qubits

By Göran Wendin and Vitaly Shumeiko

Quantum computers (QCs) promise to exponentially speed up a number of problems in optimization, materials science, and chemistry. The caveat is that they may not even get the chance. One outstanding issue is the coherence time, which is how long a system can remain in a quantum state. We live in the era of noisy intermediate-scale quantum processors, and the time to execute the necessary number of gates in the quantum circuit may be longer than the coherence time of the quantum register. Extending coherence time can be accomplished by improving the properties for existing qubits, which is tedious and demands large resources but almost guarantees steady progress. Alternatively, looking for alternative qubits holds the promise of a breakthrough, if it does not end up becoming a wild goose chase. On p. 430 of this issue, Hays *et al.* (1) make progress on the latter approach for the Andreev spin qubit (ASQ).

A recent radical approach that raised great expectations was to go for Majorana qubits implemented in semiconductor nanowires (NWs) connected to superconductors. This type of Majorana qubit has turned out to be difficult to realize, but the design is ideally suited for implementing promising qubits on the basis of the Andreev physics of the Josephson effect (2), the Andreev level qubit (ALQ) (3–6), and the ASQ (7–9).

For QCs to be competitive, a quantum register, the quantum analog of a classical processor register that holds data being processed by the CPU, needs to have a few hundred perfect qubits. This increases to millions of qubits for quantum error correction. Scalable architectures are therefore all important, and a QC contender is a large array of spin qubits with control and readout circuits arranged in a three-dimensional (3D) architecture (10, 11). The background for the solid-state spin qubit was provided by creating few-electron quantum dots in 2D electron gas (2DEG) structures. Decoherence due to charge noise made it clear that long-lived qubits

might have to use spin degrees of freedom. The first experiments demonstrated spin-flip singlet-triplet transitions in 2DEG quantum dots. This is now one of the major types of qubits used in large-scale prototype quantum processors (10, 11). The other major type is the electron spin in impurities implanted in 2DEG structures (11).

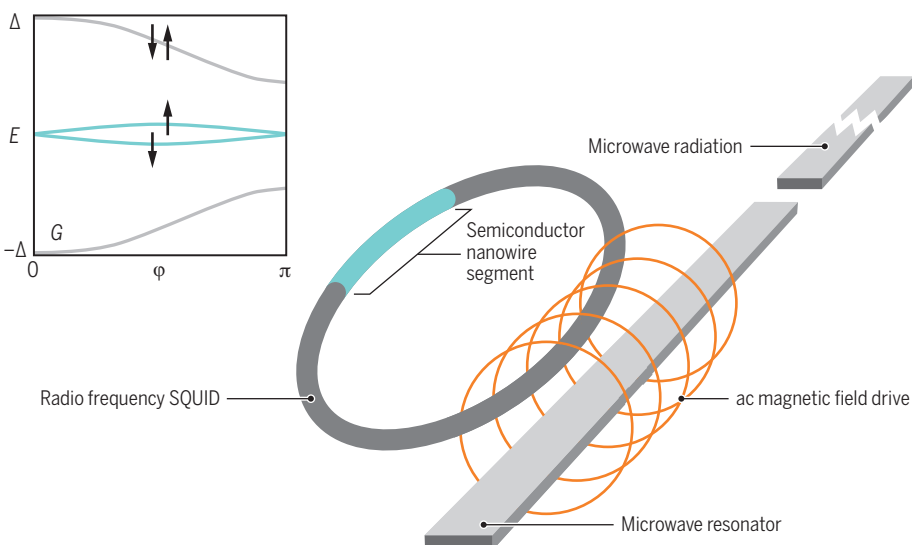
ALQ and ASQ use a very different kind of electron trap, a Josephson junction weak link in a superconducting (S) ring introduced by a short section of normal (N) semiconductor NW (3, 4). In such SNS structures, Andreev reflection from S electrodes forms a discrete set of bound states—Andreev levels (ALs)—within the superconducting energy gap. At SN and NS interfaces, the supercurrent is converted to current through Andreev bound states, allowing a Josephson supercurrent to flow through the junction. Both the energies of Andreev states and the supercurrent they carry depend on the superconducting phase difference across the junction. They can be tuned by the magnetic flux threading the ring, which allows for manipulation and measurement of the state populations.

The paradigmatic energy structure of the Andreev states in a conducting channel inside the energy gap consists of a ground state, an excited spin singlet state, and a spin doublet in the middle of the gap, which may be occupied by spurious single quasiparticles (see the figure). Spin-orbit interaction in the NW lifts spin degeneracy of the doublet and demonstrates the role of the electron spin and the possibility of implementing an ASQ.

The ALQ (2, 3) is operated like a radio-frequency superconducting quantum interference device through flux bias creating a phase difference that sets up a circulating supercurrent in the ring and adding a microwave field to create transitions between the ground state and the excited singlet Andreev state. Embedding ALQ in a microwave resonator allows for strong coupling to the qubit for fast operation and readout (4, 5). This is in line with the standard cavity or circuit quantum electrodynamics (cQED) way of operating and reading out transmon-type qubits that form the basis for most superconducting platforms.

Generating a spin qubit

Andreev level and spin qubits are created in a device where a short semiconductor nanowire forms a Josephson junction in a radio frequency superconducting quantum interference device (SQUID). The energies of the Andreev levels (graph) are tuned by a static magnetic flux through the SQUID ring, creating a phase difference across the junction. Spurious quasiparticles trapped in the semiconductor nanowire segment can have spin states with two distinct energies because of spin-orbit interaction within the nanowire. The Andreev spin qubit can be then manipulated and read out through coupling to the electromagnetic field of the resonator.



Chalmers University of Technology, 41296 Gothenburg, Sweden. Email: goran.wendin@chalmers.se

Recent work (9) precisely implements an extended version of this scheme, using a longer NW and a gate voltage to tune the number of conducting channels to two channels in the N-region. This establishes two quasiparticle doublets and allows excitation of spin-flip transitions between the different doublets. However, the intra-doublet spin-flip qubit transition was beyond reach.

The groundbreaking result is that Hays *et al.* (1) have been able to demonstrate that the device (9) can be operated as a spin qubit, ASQ, in a strong-coupling cQED setup, creating Rabi oscillation with 52 ns spin-echo coherence time and detecting it with single-shot readout. This establishes the ASQ as an interesting alternative but also heralds the need for advanced materials science.

For the ASQ to compete in a decisive manner, the coherence time must approach seconds. This criterion requires the NW N-region to be fabricated with semi-

“...looking for alternative qubits holds the promise of a breakthrough, if it does not end up becoming a wild goose chase.”

conductor material that does not create unwanted coupling to nuclear spins, while conserving the Andreev interface properties and spin-orbit coupling. Another challenge is to control and extend the quasiparticle trapping times into the subsecond domain. The bottom line is that there are no shortcuts to qubit paradigm shifts because basic physics and materials science will determine the progress of quantum computing. ■

REFERENCES AND NOTES

1. M. Hays *et al.*, *Science* **373**, 430 (2021).
2. A. Furusaki, M. Tsukada, *Phys. Rev. B Condens. Matter* **43**, 10164 (1991).
3. J. Lantz, V. S. Shumeiko, E. Bratus, G. Wendin, *Physica C* **368**, 315 (2002).
4. A. Zazunov, V. S. Shumeiko, E. N. Bratus, J. Lantz, G. Wendin, *Phys. Rev. Lett.* **90**, 087003 (2003).
5. C. Janvier *et al.*, *Science* **349**, 1199 (2015).
6. M. Hays *et al.*, *Phys. Rev. Lett.* **121**, 047001 (2018).
7. N. M. Chtchelkatchev, Y. V. Nazarov, *Phys. Rev. Lett.* **90**, 226806 (2003).
8. L. Tosi *et al.*, *Phys. Rev. X* **9**, 011010 (2019).
9. M. Hays *et al.*, *Nat. Phys.* **16**, 1103 (2020).
10. L. M. K. Vandersypen, M. A. Eriksson, *Phys. Today* **72**, 38 (2019).
11. A. Chatterjee *et al.*, *Nat. Rev. Phys.* **3**, 157 (2021).

10.1126/science.abk0929

BIOCHEMISTRY

The inner workings of an enzyme

A high-throughput mutation screen dissects the mechanistic basis of enzyme activity

By **Zachary T. Baumer** and **Timothy A. Whitehead**

Predictive understanding for how a particular amino acid sequence encodes enzymatic function is a grand challenge in molecular biology, with profound impacts in fields ranging from industrial biotechnology, computational protein design, and agriculture to predictive identification of disease mutations (1) and medicinal chemistry. Innovative methods for high-throughput and quantitative measurements of different aspects of enzymatic function are needed to achieve this goal. On page 411 of this issue, Markin *et al.* (2) describe a laboratory-on-a-chip platform called High-Throughput Microfluidic Enzyme Kinetics (HT-MEK) as a step in this direction. The technique allows high-fidelity in vitro biochemical and biophysical characterization of more than 1000 mutants of the model enzyme PafA (phosphate-irrepressible alkaline phosphatase of *Flavobacterium*). HT-MEK identifies partially overlapping yet distinct networks of amino acids that undergird individual reaction steps of PafA, illuminating the mechanistic basis of catalysis for this enzyme.

The ability to read and write DNA cheaply has enabled high-throughput, massively parallel experiments, broadly known as deep mutational scanning (3), which can be used to probe how a given enzyme sequence determines function. Typically, many different mutants of an enzyme are passed through some screen or genetic selection, changing the underlying frequency of mutants in that population. Next, the frequency change for each library variant is quantified and converted to a score. A key advantage of this method is scale: Tens of thousands of mutations can be evaluated in a single experiment. However, for enzymes there is usually, although not always (4), a narrow dynamic range of approximately fivefold in activity. This is sometimes advantageous in that small changes in activity can substantially enhance cellular fitness, which is useful for understanding steps in molecular evolution. Additionally, robust identification of small

improvements in enzymes are particularly valuable for engineering, in which many mutations conferring small improvements can be combined (5, 6). However, deciphering enzyme mechanisms requires activity measurements that span orders of magnitude. Another challenge in deep mutational scanning is that a given mutation may affect several catalytic or biophysical properties that are flattened to a single score. For example, a low score may be the result of protein folding, activity effects, or a combination of both. Thus, there are numerous challenges in interpreting these datasets that cloud a fuller understanding of mechanistic enzymology.

HT-MEK is a conceptually different approach than deep mutational scanning (see the figure). HT-MEK involves the construction of a reaction chamber around an immobilized DNA sequence that encodes an enzyme variant fused to a fluorescent protein. An in vitro reaction produces the enzyme, which is then assayed against different reactants. The fluorescent protein fusion allows quantification of the amount of enzyme in the reaction chamber, enabling determination of Michaelis-Menten kinetic parameters over a large dynamic range spanning several orders of magnitude. Segregation of enzymes into individual chambers affords measurement of enzyme activity as a function of environmental conditions, temperatures, inhibitors, and substrates. Integration of these experiments yields mechanistic dissection of the enzyme.

Alkaline phosphatases have been extensively studied for decades; what more was revealed by Markin *et al.*? The key data generated from HT-MEK—the mutational effects of individual PafA reaction steps for nearly 1000 mutants—allows precise delineation of networks of residues identified by a particular kinetic characteristic. In many cases, these networks extend large distances, from the active site to positions on the enzyme surface more than 20 Å away. Some of the findings have precedence in the literature, such as elucidating that the majority of residues with large catalytic effects are more than 10 Å beyond the active site (7). Yet the ability to separate the effects of protein folding from individual catalytic effects offers unparalleled clarity into these allosteric communication channels. HT-MEK thus offers a tantalizing possibility for the unbiased identification of allosteric networks in

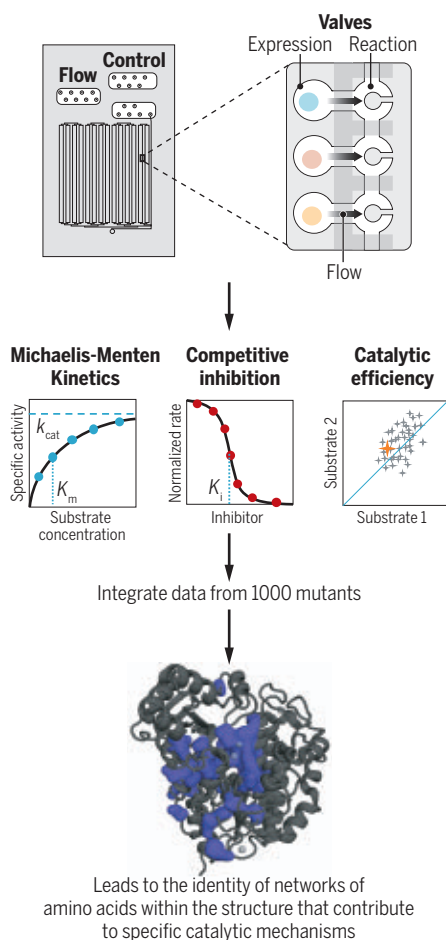
Department of Chemical and Biological Engineering,
University of Colorado, Boulder, CO 80305, USA.
Email: timothy.whitehead@colorado.edu

other enzymes, with applications in medicinal chemistry and synthetic biology. A new finding of Markin *et al.* is that mutations to PafA are more likely to result in tighter binding than weaker binding of the competitive inhibitor phosphate. This is unusual because mutations on average are expected to decrease binding. It suggests that strong evolutionary mechanisms in enzymes ensure rapid and efficient turnover (catalytic rate) without excessive end-product inhibition.

Although there are limitations in the demonstrations of HT-MEK, there are clear ways to improve generalizing the method and its throughput. HT-MEK requires product for-

Enzyme dissection at scale

The High-Throughput Microfluidic Enzyme Kinetics (HT-MEK) approach allows the simultaneous characterization of more than 1000 single-residue mutants of an enzyme. Each chamber expresses a single mutant that is flowed into an enzyme reaction compartment. Kinetic characterization [such as apparent unimolecular rate constant (k_{cat}), Michaelis constant (K_m), inhibition constant (K_i), and substrate specificity] of each mutant is integrated and mapped back onto the structure of the enzyme, revealing networks of residues involved in different catalytic and functional effects.



mation to be linked to fluorescence. Markin *et al.* used both a non-natural fluorescent substrate and a fluorescent biosensor (8) that detects the phosphate product. The biosensor can be used directly for many diverse enzyme classes that produce inorganic phosphate. Probing other enzyme classes will require innovative methods for high-affinity biosensor design in which the product is linked to fluorescence at a response rate faster than that of the internal enzyme kinetics. Also, PafA is a fast enzyme with a catalytic efficiency approximately 10-fold higher than an average enzyme (9). Evaluating less efficient enzymes, such as those involved in secondary metabolism, may result in a diminished dynamic range and sensitivity for assay measurements. In addition, Markin *et al.* used a clever experimental design to disentangle protein misfolding from protein activity by performing activity measurements at different temperatures and choosing mutations unlikely to result in global unfolding. PafA is more stable than typical enzymes, and whether this experimental design can be generalized is an open question. Also, the throughput tested by Markin *et al.* was a little more than 1000 mutants, or less than 10% of the throughput of deep mutational scanning. Several technologies can be envisioned to overcome this throughput limit, including parallelization of devices, advances in on-chip oligonucleotide synthesis of entire synthetic genes, and scale down of the reaction chambers.

Beyond exploring the limits of mechanistic enzymology, there are several near-term applications, such as the de novo design of enzymes (10). This approach has succeeded in generating active sites of enzymes with atomic resolution, but turnover rates are generally poor without substantial directed evolution; something is missing in the design concept. Teasing apart the mechanistic basis of activity and inactivity for hundreds or thousands of enzyme designs could help identify potentially missing factors. ■

REFERENCES AND NOTES

1. B. J. Livesey, J. A. Marsh, *Mol. Syst. Biol.* **16**, e9380 (2020).
2. C. J. Markin *et al.*, *Science* **373**, eabf8761 (2021).
3. C. L. Araya, D. M. Fowler, *Trends Biotechnol.* **29**, 435 (2011).
4. E. Firnberg *et al.*, *Mol. Biol. Evol.* **31**, 1581 (2014).
5. A. Goldenzweig, S. J. Fleishman, *Annu. Rev. Biochem.* **87**, 105 (2018).
6. E. E. Wrenbeck *et al.*, *ACS Synth. Biol.* **8**, 474 (2019).
7. E. E. Wrenbeck *et al.*, *Nat. Commun.* **8**, 15695 (2017).
8. M. Brune *et al.*, *Biochemistry* **33**, 8262 (1994).
9. A. Bar-Even *et al.*, *Biochemistry* **50**, 4402 (2011).
10. D. Hilvert, *Annu. Rev. Biochem.* **82**, 447 (2013).

ACKNOWLEDGMENTS

The authors are supported by the National Science Foundation (NSF) (Chemical, Bioengineering, Environmental and Transport Systems award 2030221 to T.A.W.; NSF Graduate Research Fellowship Program to Z.T.B.) and National Institutes of Health, National Institute of General Medical Sciences under award R21GM129559-01 to T.A.W.

10.1126/science.abj8346

SPECTROSCOPY

Proximity and single-molecule energetics

Scanning probes measure how nearby oxygen molecules affect triplet lifetimes of pentacene

By Linfei Li and Nan Jiang

Probing single molecules in their nanoenvironment can reveal site-specific phenomena that would be obscured by ensemble-averaging experiments on macroscopic populations of molecules. Particularly in the past decade, major technological breakthroughs in scanning probe microscopy (SPM) have led to unprecedented spatial resolution and versatility and enabled the interrogation of molecular conformation, bond order, molecular orbitals, charge states, spins, phonons, and intermolecular interactions. On page 452 of this issue, Peng *et al.* (1) use SPM to directly measure the triplet lifetime of an individual pentacene molecule and demonstrate its dependence on interactions with nearby oxygen molecules with atomic precision. In addition to allowing the local tuning and probing of spin-spin interactions between molecules, this study represents a notable advance in the single-molecule regime and provides insights into many macroscopic behaviors and related applications in catalysis, energy-conversion materials, or biological systems.

Single-molecule studies have benefited from the high resolution achieved with well-defined functionalized probes, especially with carbon monoxide-terminated atomic force microscopy (AFM) tips (2). The versatility and applicability of AFM have also been enhanced by biasing the tip with gate voltages and supporting molecules on insulating substrates. In this configuration, the conductive AFM tip serves as an atomically controlled charge injector with single-charge sensitivity. Such electrical addressing of electronic states of single molecules (3) allows for the study

Department of Chemistry, University of Illinois at Chicago, Chicago, IL 60607, USA. Email: njiang@uic.edu

of charge distribution and transport in single-molecule devices, organic electronics, and photovoltaics.

Beyond steady-state spectroscopy, excited-state dynamics of single molecules can be measured by using an ultrashort and high-intensity electric (voltage) or optical (laser) pulse (the “pump”) to excite the sample. After a nonequilibrium state is generated, a second weaker pulse (the “probe”) monitors the change of the excited state. By varying the time delay between the two pulses, the temporal evolution of the excited state can be mapped out.

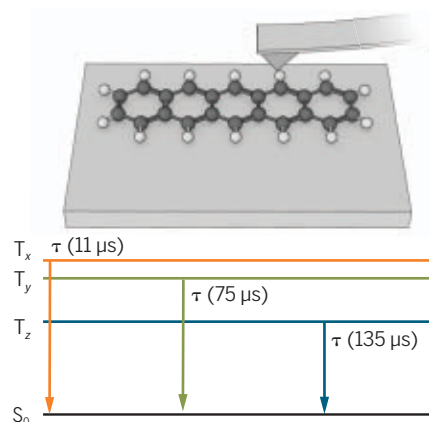
Peng *et al.* used the electronic pump-probe approach in AFM to measure the lifetime of the excited triplet state of an individual pentacene molecule with atomic precision (see the figure). They observed strong quenching of the triplet lifetime by co-adsorbed molecular oxygen (O_2). The electronic energy-transfer processes had an intriguing dependence on the arrangement of surrounding O_2 molecules, which they controlled by atomic manipulation with the tip. Spin-relaxation measurements of single molecules in space with atomic resolution provide insights into their local interactions with each other, as well as with their nanoenvironment. Such information could be useful for spin-based quantum-information storage or quantum computing (4).

Given the radiative relaxation of excited states, SPM-coupled optical spectroscopy provides a powerful tool to perform spatially and energy-resolved spectroscopic studies of single molecules. Specifically, site-resolved excitations of molecules can be induced by highly localized scanning tunnel microscopy (STM) current, and the resulting luminescence, which carries information that describes excited states, can be probed by integrated optical detection systems. This approach revealed redox state-dependent excitation of single molecules and intermolecular excitonic coupling interactions with atomic-scale spatial precision (5, 6). A study of electroluminescence demonstrated selective triplet formation by manipulating electron spin inside a molecule (7), which could provide a route to interrogate quantum spintronics and organic electronics at the single-molecule level.

Besides tunneling electrons, the interaction of photons with molecules can provide valuable structural information and chem-

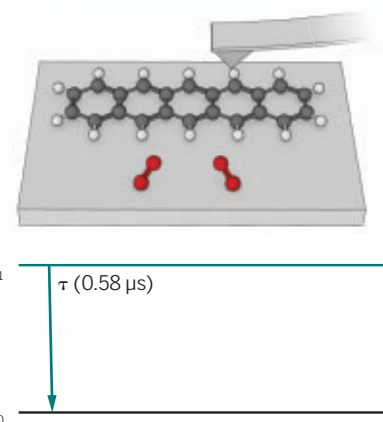
Atomically addressing excited single molecules

The effect of nearby oxygen molecules on the lifetimes (τ) of triplet states T_x , T_y , and T_z or T_1 decaying to the singlet state S_0 of individual pentacene molecules has been probed on an insulating salt surface.



Molecular interactions

Scanning probe microscopy was combined with electronic pump-probe techniques to follow pentacene excitation. The tip placed oxygen molecules next to pentacene.



Triplet lifetimes

Pentacene exhibits three triplet states, but nearby oxygen (red) quenches this excitation to one triplet state with a much shorter submicrosecond lifetime.

ical identification through measurements of absorption, emission, or scattering of light. In particular, by confining laser light at the atomic-scale SPM junction and taking advantage of plasmon-enhanced Raman scattering, tip-enhanced Raman spectroscopy can overcome the diffraction limit of conventional optical spectroscopy and thereby achieve sub-molecular chemical spatial resolution (8). Such capability provides in-depth insights into single-molecule chemistry and site-specific chemical effects at the spatial limit (9).

Most excited states induced by photon absorption are incredibly short-lived (on the order of picoseconds to femtoseconds), so time-resolved optical STM techniques have been developed with ultrafast lasers. For example, pump-probe terahertz laser pulses were used to induce state-selective ultrafast STM tunneling currents through a single molecule. This approach allowed the molecular orbital structure and vibrations to be measured directly on the femtosecond time scale (10). Optical STM further showed the capability to explore photon and field-driven tunneling with angstrom-scale spatial resolution and attosecond temporal resolution. This experimental platform can be used to study quasiparticle dynamics in superconductor and two-dimensional materials with exceptional resolutions (11).

Single-molecule studies could open avenues to access extremely transient states and chemical heterogeneity, such as the vibration of atoms within a molecule, the precession of a spin, ultrashort-lived complex reaction intermediates, and some key stochastic processes of reactions in chemistry and biology. For example, the study of Peng *et al.* relates to the reactivity of electronic excited states of organic molecules to O_2 (and thus air). These processes can affect various natural photochemical and photophysical processes undergoing excitation by sunlight that can lead to transformation, degradation, or aging (12). The insightful descriptions of molecular conformation, dynamics, and function provided by spatially resolved single-molecule studies could inform complex and emergent behaviors of populations of molecules or even cells. ■

REFERENCES AND NOTES

1. J. Peng *et al.*, *Science* **373**, 452 (2021).
2. L. Gross, F. Mohn, N. Moll, P. Liljeroth, G. Meyer, *Science* **325**, 1110 (2009).
3. S. Fatayer *et al.*, *Nat. Nanotechnol.* **13**, 376 (2018).
4. M. N. Leuenberger, D. Loss, *Nature* **410**, 789 (2001).
5. Y. Zhang *et al.*, *Nature* **531**, 623 (2016).
6. B. Doppagne *et al.*, *Science* **361**, 251 (2018).
7. K. Kimura *et al.*, *Nature* **570**, 210 (2019).
8. J. Lee, K. T. Crampton, N. Tallarida, V. A. Apkarian, *Nature* **568**, 78 (2019).
9. S. Mahapatra, L. Li, J. F. Schultz, N. Jiang, *J. Chem. Phys.* **153**, 010902 (2020).
10. T. L. Cocker, D. Peller, P. Yu, J. Repp, R. Huber, *Nature* **539**, 263 (2016).
11. M. Garg, K. Kern, *Science* **367**, 411 (2020).
12. P. R. Ogilby, *Chem. Soc. Rev.* **39**, 3181 (2010).

ACKNOWLEDGMENTS

We acknowledge support from the National Science Foundation (CHE-1944796).

“Spin-relaxation measurements of single molecules in space with atomic resolution provide insights into their local interactions with each other...”

SCIENCE AND LAW

Experimental jurisprudence

Psychologists probe lay understandings of legal constructs

By **Roseanna Sommers**

Historically, the role psychology has played in the legal system has been confined to discrete domains that lawyers and judges tend to recognize as psychological. For example, in trademark disputes, litigants seeking to establish “consumer confusion” often hire experts to collect survey data showing that consumers are apt to mistake one brand for another. This is a textbook example of “law and psychology”: bringing methodological rigor to traditional legal analyses. Recently, however, a growing number of “experimental jurisprudence” scholars have been studying the law from the outside—theorizing what its doctrines are doing, criticizing its doctrines for what they are not doing—rather than from the inside, helping to sharpen traditional legal analyses (1–2). These empiricists, moreover, have trained their sights on legal constructs that might not strike one as particularly psychological, such as causation, consent, reasonableness, ownership, punishment, contract, and even law itself (e.g., what makes the law the law as opposed to some other kind of social arrangement) (3–10). This new approach departs from traditional law and psychology in both its scope and ambition: Beyond providing narrow expertise on matters that lawyers readily recognize as psychological (e.g., confusion, memory, insanity), experimental jurisprudence aims to advance legal theory broadly.

Experimental jurisprudence examines how core legal concepts are understood by laypeople who know little about the law. Researchers then compare laypeople’s ordinary concepts against their legal counterparts (1–3). For example, using survey experiments, psychologists have discovered

that laypeople’s causal judgments are affected by counterfactual reasoning, which is in turn influenced by whether an agent behaved immorally (11). When an agent violates a moral norm, it increases the relevance of a counterfactual in which the agent behaved in a norm-abiding manner.

Scholars have taken this insight and deployed it to rethink tort law’s doctrine of proximate cause, which determines when

way to explain the twists and turns of the case law is to recognize that judges simply decide which party is ultimately morally blameworthy and assign proximate causation accordingly. Skeptics in this camp contend that proximate cause represents nothing more than judicial anarchy: Judges engage in outcome-driven reasoning, which they dress up, post hoc, in the language of causality (3).

But others have argued that given experimental research, legal causation is not as maddening or mysterious as it seems, with the doctrine of causal supersession best theorized not as judicial lawlessness but as roughly following the commonsense understanding of cause (3). This ordinary concept embeds judgments about the relevance of various counterfactual alternatives and therefore bakes in judgments about moral norms, which may be why it appears as if judges’ causal determinations are influenced by considerations of blameworthiness. By this view, the doctrine’s relationship to causation is not complete confabulation, contrary to the skeptics’ complaint. Instead, judges’ seemingly inconsistent decisions may reflect the ordinary folk notion of causation.

Recent experimental work uncovers that laypeople intuit that another key legal concept, consent, is compatible with certain forms of deception (5). For example, most American research participants believe

that a patient consents when he agrees to a medical procedure as the result of his doctor’s false statements. The same is true when a civilian allows police officers into her home because they lie about what they are searching for. Laypeople tend to report that the deceived targets have autonomously authorized invasions into their bodies and properties. This diverges from the standard autonomy-based understanding of consent, which holds that consent must be sufficiently knowing and informed to be valid. A court long ago concluded, for example, that a defendant could be held liable for kidnapping when he tricked a young woman into sailing to Panama on the promise that a job as a governess awaited, when his true intention was to employ her in a brothel. Her consent to board the ship was negated by her employer’s deceit just as if she had been held at gunpoint.

an intervening event will “break the causal chain” between an original negligent act and a subsequent injury, such that the original wrongdoer is relieved from liability (3). For instance, a classic causal supersession case involves a service station that negligently leaves car keys in an unlocked vehicle and a thief who steals the car and negligently injures a plaintiff in a collision. Should the service station be held liable for the plaintiff’s damages? Although this case may seem unusual, it is hardly exotic: Judges are often called upon to determine whether negligent actors are the “proximate” causes of—and ultimately liable for—injuries occurring far down the causal chain. When they do so, they rely on cases like these as precedents.

The hitch is that prior cases have taken inconsistent stances on the question of causal supersession. Several influential legal theorists have argued that the best



Yet the law has often been inconsistent, occasionally departing from the maxim that deception invalidates consent. Just 2 years after the would-be Panamanian governess was deemed to have been kidnapped, a notably similar case arising in the same jurisdiction reached the opposite result. A man was tricked into boarding a ship to Mexico on the understanding that he would be employed as a railroad worker at a rate of 35 dollars per month in US currency; in fact, the job would pay only 1 dollar per month in Mexican currency. The court concluded that he was not kidnapped, because the false promise of his wages was “a shabby trick, but not a crime” (5).

The classic doctrinal explanation for this discrepancy insists that there are two kinds of fraud: fraud in the factum, which pertains to the essence of an activity (e.g., lying about what job one is being recruited for), and fraud in the inducement, which pertains to a mere “inducement” (e.g., lying about how much a job will pay). In several instances, the common law treats only fraud in the factum as nullifying consent, even though both kinds of fraud may be upsetting to the individual who is deceived.

Experimental jurisprudential research can illuminate puzzling doctrinal quirks such as these. In one experiment using consent-by-deception scenarios that roughly corresponded to the two kinds of common-law fraud, participants read about a consumer who wanted to make a purchase so that he could earn credit card reward points that would enable him to book a free flight. In one version of the scenario, the store clerk lied to the consumer about what item he was ordering. In the other version, the clerk lied about whether the purchase qualified for reward points. Participants observed that the consumer personally cared more about the points than about what item he was buying, but they simultaneously regarded his consent as more undermined when the clerk lied to him about what product he was purchasing. They could discern what mattered most to the individual, but their judgments of consent tracked something else entirely: whether he was misled about something that went to the essence of the transaction. Their consent intuitions thus mirrored the peculiar factum versus inducement distinction found in the common law.

This line of research carries implications for the contemporary legal controversy over the so-called “riddle of rape-by-de-

ception”—why Anglo-American law refuses to treat fraudulently procured consent to sex as rape, except under extreme circumstances, such as when a doctor misrepresents sex as a medical procedure or when one person impersonates another, leading to sex with the wrong person. Prevailing explanations for this puzzle have rested on ideas about gender, suggesting that the law is punishing women who are seen as unchaste, by declaring that they have consented despite being deceived. This new research suggests that something more general may be going on. In a variety of domains, lay participants endorse the intuition that only essential fraud (“fraud in the factum”) defeats consent; their judgments follow this pattern not just for consent to sex but also for consent to medical procedures, tattoos, and contracts for sale (5).

This experimental approach to legal scholarship has raised some controversy (12). Some question why it makes any difference what laypeople think; legal theory, they object, cannot be crowdsourced. What is more, although subtle and surprising features of legal concepts are often shared by folk intuition (2), at other times experimental jurisprudence research has uncovered stark divergences between ordinary concepts and legal concepts (5, 12). For example, recent work on the

ordinary concept of contract shows that laypeople intuit that agreements must be in writing to be legally valid and that contractual terms will invariably be enforced as written, even when they are unenforceable or result from material, bad-faith fraud (13).

It would be a mistake to insist that where ordinary concepts and legal concepts diverge, the law has been refuted. Most legal constructs are not like obscenity, which is explicitly defined by reference to community standards. Although courts have shown some openness to considering public opinion polling when it comes to defining obscene speech, that does not mean that surveys to set the definitions of all legal concepts should be used. A more defensible approach would recognize that there may be good reason for the law in some areas not to look just the way people believe, expect, or prefer it to look. Thus, although it is a welcome development that psychologists are venturing beyond their historical confines and pushing into new territory not recognized as distinctly psychological, they must acknowledge the limits of their approach. Experiments cannot

settle deep, contested, normative questions about what the law should be.

Still, folk intuitions are worth studying. As a practical matter, in many jurisdictions ordinary people are empowered to determine what counts as causation or consent in legal cases. These are “questions of fact,” routinely decided by juries. Of course, research using hypothetical scenarios is limited in what it reveals about how people are likely to behave (14). Nonetheless, vignette studies are relatively well suited to investigating how people are likely to evaluate strangers whose situations they have no personal stake in—the task presented to jurors. Experimental jurisprudence studies often ask participants to evaluate scenarios that are based on real legal cases, rather than probing intuitions about fantastical thought experiments.

Beyond illuminating jury decision-making, experimental jurisprudence research can offer new insights into why the law is the way it is (1–3). And because experimental jurisprudence bears on broad legal concepts like causation, reasonableness, contract, and consent, it has the potential to take the field of law and psychology beyond its limited historical role and to establish it as a more central player in contemporary jurisprudential debates.

Looking to the future, experimental jurisprudence scholars are expected to use a growing set of methodologies: neuroimaging, computational and corpus linguistics, cross-cultural studies, and developmental work with participants from across the age spectrum (1, 7–8, 10, 15). As experimental jurisprudence gains traction, it may someday come to be an influential methodology that informs legal scholarship and practice, taking its place alongside economics, history, sociology, critical theory, and philosophy. ■

REFERENCES AND NOTES

1. K. P. Tobia, SSRN 10.2139/ssrn.3680107 (2021).
2. K. P. Tobia, in *Law and Mind: A Survey of Law and the Cognitive Sciences*, B. Brożek et al., Eds. (Cambridge Univ. Press, 2021), pp. 86–96.
3. J. Knobe, S. J. Shapiro, *Univ. Chic. Law Rev.* **88**, 165 (2021).
4. J. A. Macleod, *Indiana Law J.* **94**, 957 (2019).
5. R. Sommers, *Yale Law J.* **129**, 2232 (2020).
6. K. P. Tobia, *Ala. Law Rev.* **70**, 293 (2018).
7. S. E. Nancekivell, O. Friedman, S. A. Gelman, *Trends Cogn. Sci.* **23**, 102 (2019).
8. F. Margoni, J. Geipel, C. Hadjichristidis, L. Surian, *Exp. Psychol.* **65**, 105 (2018).
9. T. Wilkinson-Ryan, D. A. Hoffman, *Stanford Law Rev.* **67**, 1269 (2015).
10. I. R. Hannikainen et al., *PsyArXiv* 10.31234/osf.io/c2ytm (2021).
11. J. F. Kominsky, J. Phillips, T. Gerstenberg, D. Lagnado, J. Knobe, *Cognition* **137**, 196 (2015).
12. F. Jiménez, SSRN 10.2139/ssrn.3815405 (2021).
13. R. Sommers, SSRN 10.2139/ssrn.3767209 (2021).
14. R. Sommers, V. K. Bohns, *Yale Law J.* **128**, 1962 (2019).
15. K. P. Tobia, *Harv. Law Rev.* **134**, 726 (2020).

10.1126/science.abf0711

IMMUNOLOGY

Bespoke brain immunity

The brain and spinal cord maintain a distinct cache of immune cells

By Rita H. Nguyen^{1,2,3} and Paul Kubes^{2,3}

A primary defense strategy in response to infection or inflammation is to mobilize innate immune cells through the circulatory system to the affected organ. Patrolling monocytes and neutrophils constantly monitor healthy tissue and extravasate from vascular endothelium to infected or injured tissue through a stepwise mechanism (1). On page 409 of this issue, Cugurra *et al.* (2) demonstrate in a mouse model a pathway by which the central nervous system (CNS) bypasses this circulatory patrol system and supplies the meninges (the mem-

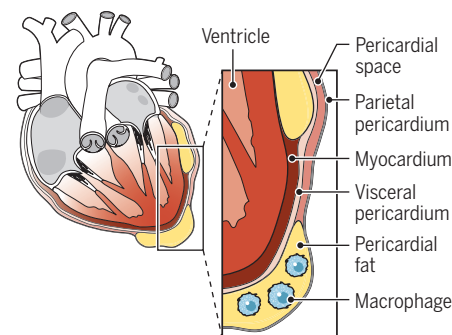
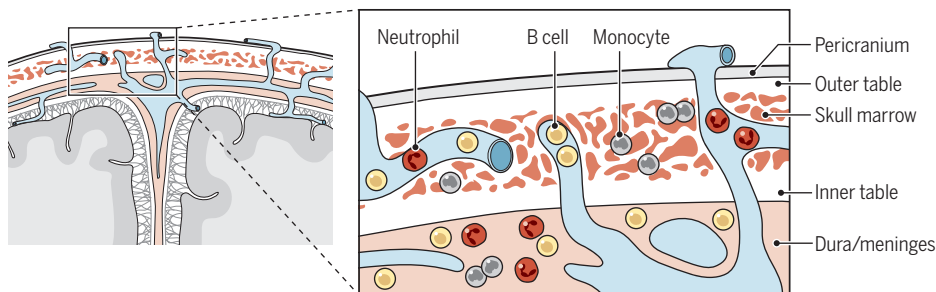
in a rich vascular network contained in the meninges and the skull. These vascular spaces have traditionally been thought to provide additional sites for cerebrospinal fluid resorption. However, recent animal studies have demonstrated that these skull-meninges connections have a much more dynamic role, and under pathological conditions such as stroke, they serve as highways for myeloid cells to quickly transmit from skull marrow to the brain parenchyma (4). Cugurra *et al.* expand these findings to demonstrate that even under homeostatic conditions, the skull marrow directly supplies myeloid cells to the meninges. These meningeal myeloid

monocytes in pathological conditions.

Brioschi *et al.* also show that the skull marrow supplies the meninges with B cells in mice. These meningeal B cells mature in the dura and learn to recognize and tolerate CNS antigens. However, in aging mice, the meninges become populated with antigen-experienced, aged B cells derived from the peripheral circulation that have the potential to disrupt the balance of the distinct CNS immune milieu. This finding invites intriguing hypotheses for the pathophysiology of CNS autoimmune diseases such as multiple sclerosis, for which B cell depletion has shown therapeutic benefit (5).

Organ-specific immune niches

The central nervous system is enclosed by the skull and vertebrae, the marrow of which provide a private source of immune cells that can invade the brain and spinal cord under pathological conditions through ossified vascular channels. Other organs harbor specialized immune cells in the tissue and spaces that surround them, such as specialized macrophages that can repair cardiac tissue after ischemia.



branes that enclose the brain and spinal cord) with functionally distinct myeloid cells through channels that traverse the skull bone marrow. Moreover, on page 408 of this issue, Brioschi *et al.* (3) demonstrate that the meninges are also populated with B cells directly derived from skull marrow hematopoiesis. These studies show that the brain is an immunologically distinct organ that is surrounded by its own cadre of immune cells.

The skull consists of two thin plates of outer and inner cortical bone tables separated by the diploic space of cancellous bone. The diploic space contains the fatty marrow and is highly vascular, filled with veins that span from the outer and inner tables. Emissary veins cross from the scalp, through the skull, to the meninges overlying the brain. Thus, the brain is enclosed

cells act as sentries of the brain, poised to respond to the first sign of perturbation.

Indeed, Cugurra *et al.* went on to study CNS–bone marrow–derived myeloid cells versus blood-derived myeloid cells in mice under three different pathological conditions: autoimmune encephalomyelitis (EAE), spinal cord injury (SCI), and optic nerve crush injury. The authors observed that monocytes that infiltrated the spinal cord or optic nerve in these models were primarily derived from CNS-marrow, suggesting that the meningeal monocytes preferentially protect the organ that it borders. Furthermore, gene expression analysis between CNS-marrow–derived versus blood-derived infiltrating monocytes in EAE reveal that blood-derived monocytes were more enriched for proinflammatory pathways, suggesting differential roles for these

Cugurra *et al.* and Brioschi *et al.* suggest that the brain is distinct in having a direct bone marrow pipeline of immune cells to the borders of the brain (see the figure). Although other organs such as heart, lungs, and visceral organs do not have this selective pipeline, it is often forgotten that all these organs do have an organized structure or reservoir of immune cells at their borders. In 1906, the omentum, which covers the visceral organs, was referred to as “the policeman of the abdomen” because it attenuates peritonitis and improves surgi-

¹Department of Clinical Neurosciences, Division of Neurosurgery, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. ²Department of Pharmacology and Physiology, University of Calgary, Calgary, Alberta, Canada. ³Snyder Institute for Chronic Diseases, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. Email: pkubes@ucalgary.ca

cal wound healing (6). Areas on the omentum called milky spots contain immune cells that can promote angiogenesis (7) and tissue repair (8).

Fat-associated lymphoid clusters in the peritoneal, pericardial, and pleural cavity of mice are storage sites for lymphoid cells (9) and have been implicated in myocardial fibrosis after myocardial infarct (heart attack) (10). In addition, a distinct population of GATA-binding protein 6-positive (GATA6⁺) macrophages reside in all three major cavities, depend on retinoic acid for their identity, and retain their cavity gene expression signature regardless of whether they are in the peritoneal, pleural, or pericardial space (11). These cells accumulate rapidly at sites of organ injury and affect healing from the outside (12). Comparably, Cugurra *et al.* and Brioschi *et al.* illustrate that the brain and spinal cord harbor an exclusive pool of immune cells in the meninges that influence CNS diseases. Clearly, many organs have an immune presence localized to their borders, but the degree to which these cells are solicited to the parenchyma is unclear.

The wealth of immune cells that surround various organs also raises the issue of immune cell recruitment. The canonical manner by which immune cells are recruited is through the vasculature and, in most cases, by extravasation from the post-capillary venules (13). However, Cugurra *et al.*, Brioschi *et al.*, and others (4, 12) have raised the possibility that immune cells could be recruited through the process of “invasion,” which involves migration into an organ from the perimeter, perhaps even by way of an avascular route. Having a pool of mature immune cells surrounding an organ provides a critical, immediately available reservoir of specific immune cells. For example, recruitment of monocytes from bone marrow to tissues where they become mature macrophages to initiate repair could take days, especially if new vasculature needs to be constructed. By contrast, a population of mature monocytes in the CNS, or mature GATA6⁺ macrophages in visceral cavities, are poised to instantly respond to brain, heart, or lung injury.

The findings of Cugurra *et al.* and Brioschi *et al.* suggest that the blood-brain barrier does not necessarily need to be disrupted for meningeal immune cells to infiltrate the brain parenchyma. The clinical implications are numerous. For example, gliomas are primary brain tumors that are notoriously difficult to treat. Infiltrating monocytes have been shown to promote tumorigenesis (14). It would be fascinating to exploit the skull-meninges connections to influence myeloid cell chemotaxis as

an immunotherapeutic option. Moreover, there is currently no medical treatment available for traumatic brain injury. Recent data show that myeloid cells promote vascular repair after traumatic brain injury (15). Perhaps the skull marrow myeloid cell reservoir can be harnessed as an immediate source of reparative cells.

It remains unknown whether there are specific CNS signaling molecules that preferentially recruit meningeal immune cells over blood-derived cells. Is this also the case for visceral organs, heart, and lungs? Furthermore, the temporal dynamics of infiltration of CNS-marrow-derived versus blood-derived cells versus cavity immune cells needs to be explored and evaluated against disease progression. In surgical interventions, inadvertent removal of the border pericardium (during heart surgery), fusion of the pleural space (to limit effusions), craniotomy (removal of part of the skull), or durotomy (perforation of the dura mater meningeal membrane) can oc-

“Having a pool of mature immune cells surrounding an organ provides a critical, immediately available reservoir of specific immune cells.”

cur. What are the implications of these procedures for these cell niches and the physiological responses of an organ? The studies of Cugurra *et al.* and Brioschi *et al.* remind us that there is a vast amount of immunity that surrounds each organ with a coterie of immune cells with distinct phenotypes. In the case of the brain, it provides yet another specialized layer that should be considered in the context of the CNS. ■

REFERENCES AND NOTES

1. C. Auffray *et al.*, *Science* **317**, 666 (2007).
2. A. Cugurra *et al.*, *Science* **373**, eabf7844 (2021).
3. S. Brioschi *et al.*, *Science* **373**, eabf9277 (2021).
4. F. Herisson *et al.*, *Nat. Neurosci.* **21**, 1209 (2018).
5. S. L. Hauser *et al.*, *N. Engl. J. Med.* **383**, 546 (2020).
6. R. Morison, *BMJ* **1**, 76 (1906).
7. I. García-Gómez *et al.*, *Neurol. Res.* **27**, 807 (2005).
8. S. Shah *et al.*, *PLOS ONE* **7**, e38368 (2012).
9. C. Bénézech *et al.*, *Nat. Immunol.* **16**, 819 (2015).
10. M. Horckmans *et al.*, *Circulation* **137**, 948 (2018).
11. J. F. Deniset *et al.*, *Immunity* **51**, 131 (2019).
12. J. Wang, P. Kubes, *Cell* **165**, 668 (2016).
13. E. Kolaczowska, P. Kubes, *Nat. Rev. Immunol.* **13**, 159 (2013).
14. D. H. Gutmann, H. Kettenmann, *Neuron* **104**, 442 (2019).
15. M. V. Russo, L. L. Latour, D. B. McGavern, *Nat. Immunol.* **19**, 442 (2018).

10.1126/science.abj8183

VIEWPOINT: COVID-19

Scent of a vaccine

Intranasal vaccination should block SARS-CoV-2 transmission at the source

By Frances E. Lund¹ and Troy D. Randall²

The highly contagious severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infects the respiratory tract and is transmitted, in part, by respiratory droplets and aerosols. Consequently, unvaccinated people are encouraged to wear masks in public, self-quarantine if symptomatic, and practice social distancing. Despite these precautions, millions are dying. As the pandemic takes its toll, vaccines are once again headline news, notably for the speed of their development and the success of messenger RNA (mRNA) vaccines. Given the respiratory tropism of the virus, however, it seems surprising that only seven of the nearly 100 SARS-CoV-2 vaccines currently in clinical trials are delivered intranasally. Advantages of intranasal vaccines include needle-free administration, delivery of antigen to the site of infection, and the elicitation of mucosal immunity in the respiratory tract.

The idea that intranasal vaccination preferentially protects the respiratory tract is not new: Development of the US Food and Drug Administration (FDA)-approved live attenuated influenza vaccine (LAIV) began in the 1960s. Immunologists have long known that nasal infection or vaccination elicits an immunoglobulin A (IgA) response in both serum and respiratory fluids, whereas intramuscular vaccines primarily elicit serum IgG. IgA is particularly important in the upper airways and nasal passages, where it is actively transported across the epithelium and released into the airway lumen as a dimer bound to secretory component, a stabilizing configuration that allows it to more effectively neutralize viruses like SARS-CoV-2 (1). By contrast, IgG enters and protects the lower lung through passive transudation across the thin alveolar epithelium (2). IgG is also found in the upper respiratory tract and nasal passages,

¹Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL, USA. ²Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA. Email: randallt@uab.edu

perhaps carried from the lower lung by the mucociliary escalator. However, protection of the nasal passages by IgG is only achieved at high serum concentrations (2). Consequently, intramuscular vaccines that elicit high titers of serum IgG can reduce viral titers in the lungs and nasal passages.

CD8⁺ T cells are another important component of antiviral immunity and directly kill virus-infected cells, thereby reducing viral replication and accelerating viral clearance and recovery. Some activated CD8⁺ T cells develop into memory cells, which by themselves do not prevent infection, but are poised for rapid reactivation and effector function. Notably, B and T cells primed by mucosal vaccination or infection express receptors that promote homing to mucosal sites as long-lived antibody-secreting cells or as tissue-resident memory cells. Resident memory B and T cells in the lung and nasal passages act as nonredundant, first responders to challenge infection and are essential for rapid virus clearance (3, 4). The placement of tissue-resident memory cells in the respiratory tract requires that they encounter antigen in the respiratory tract (3, 5), meaning that vaccines designed to recruit resident memory cells to the respiratory tract should be administered intranasally.

Compared to intramuscular vaccines, intranasal vaccines provide two additional layers of protection: Vaccine-elicited IgA and resident memory B and T cells in the respiratory mucosa provide an effective barrier to infection at those sites; and, even if infection does occur, perhaps by a viral variant, cross-reactive, resident memory B and T cells, which encounter antigen earlier and respond more quickly than systemic memory cells, impede viral replication and reduce viral shedding and transmission (see the figure).

Of the seven SARS-CoV-2 vaccines being tested for intranasal delivery, six are live-attenuated viruses or virus-vectored vaccines and one is a protein subunit vaccine (see the table). Attenuated viruses and viral vectors that encode vaccine antigens are particularly useful for intranasal immunization because the infection process effectively breaches the epithelium and is intrinsically immunogenic. Because vaccine antigens are expressed by infected cells, antigen presentation occurs via the class I pathway and efficiently triggers CD8⁺ T cell responses—an advantage over protein subunit vaccines that poorly engage CD8⁺ T cells.

Preclinical studies of adenovirus-vectored vaccines expressing the SARS-CoV-2 spike host receptor protein or its receptor binding domain (RBD) demonstrate that

intranasal delivery triggers long-lasting, virus-neutralizing serum IgG responses as well as antigen-specific IgA and CD8⁺ T cells in the respiratory tract (6–8). Moreover, both intranasal and intramuscular vaccination with adenovirus-vectored vaccines protect against pneumonia and weight loss after a challenge infection. However, animals vaccinated intramuscularly still shed virus from the nasal passages, whereas animals vaccinated intranasally have reduced viral replication and shedding in both the lungs and the nasal passages (8).

“...it seems surprising that only seven of the nearly 100 SARS-CoV-2 vaccines currently in clinical trials are delivered intranasally.”

Adenoviruses are natural human pathogens, and many adults have been exposed to one or more strains, meaning that they may have antivector antibodies that impair vaccine efficacy (negative interference). However, Ad5-vectored intranasal influenza vaccine (NasoVAX), administered at high doses, works similarly in Ad5 seropositive and seronegative individuals (9), perhaps because the inoculating volume dilutes local antibody concentrations. Nevertheless, in an attempt to avoid any potential negative interference, some developers are using rare strains of human adenoviruses or chimp adenoviruses, to which most humans have not been exposed.

The influenza-vectored SARS-CoV-2 vaccine being developed by the University of Hong Kong may face related hurdles. The deletion of the influenza virus gene encod-

ing nonstructural protein 1 (NS1) strongly attenuates the vector and allows developers to replace NS1 with the SARS-CoV-2 spike-RBD. Like adenovirus-vectored vaccines, this one should also elicit mucosal IgA against RBD and place resident memory cells in the respiratory tract. However, negative interference from preexisting antibodies against the influenza vector may impair its effectiveness. Similarly, Meissa Vaccines developed a live attenuated respiratory syncytial virus (RSV) vector in which it replaced the RSV F and G host receptor proteins with SARS-CoV-2 spike. Delivered intranasally, the chimeric virus should elicit mucosal immunity. Notably, the change in surface proteins will likely alter the cellular tropism of the virus and perhaps its immunogenicity. Preexisting antibodies against RSV should not interfere with vaccination, but preexisting antibodies against spike may neutralize it.

Live attenuated SARS-CoV-2 intranasal vaccines should also effectively elicit mucosal IgA responses and resident-memory cells in the respiratory tract. Unlike vectored vaccines that express only spike or RBD, live attenuated SARS-CoV-2 has the advantage of expressing (and potentially eliciting immune responses against) all viral proteins, thereby conferring broad-spectrum immunity that should cross-react with and provide some level of immunity against variant strains of SARS-CoV-2. Although modern molecular techniques minimize the risk of reversion, live attenuated viruses retain replicative capacity and are contraindicated for infants <2 years, people aged >49 years, or immune-compromised persons. Live attenuated SARS-CoV-2 and spike-expressing RSV may also face scrutiny over their potential to cause neuronal symptoms (10).

Past experience with LAIV will be relevant to these live attenuated vaccines. In

Intranasal SARS-CoV-2 vaccines in clinical trials

NAME	DEVELOPER	TYPE (ANTIGEN)	CLINICAL TRIAL
ChAdOx1-S	University of Oxford	Chimp adenovirus vector (spike)	NCT04816019 (phase 1)
AdCOVID	Altimmune	Adenovirus 5 vector (RBD)	NCT04679909 (phase 1)
BBV154	Bharat Biotech	Simian adenovirus vector (spike)	NCT04751682 (phase 1)
DeINS1-nCoV-RBD LAIV	University of Hong Kong	Live attenuated influenza virus (RBD)	NCT04809389 (phase 1)
MV-014-212	Meissa Vaccines	Live attenuated RSV (spike)	NCT04798001 (phase 1)
COVI-VAC	Codagenix	Live attenuated SARS-CoV-2	NCT04619628 (phase 1)
CIBG-669	Center for Genetic Engineering and Biotechnology, Cuba	Protein subunit AgnHB (RBD)	RPCEC00000345 (phase 1/2)

HB, hepatitis B virus; RBD, receptor binding domain; RSV, respiratory syncytial virus; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

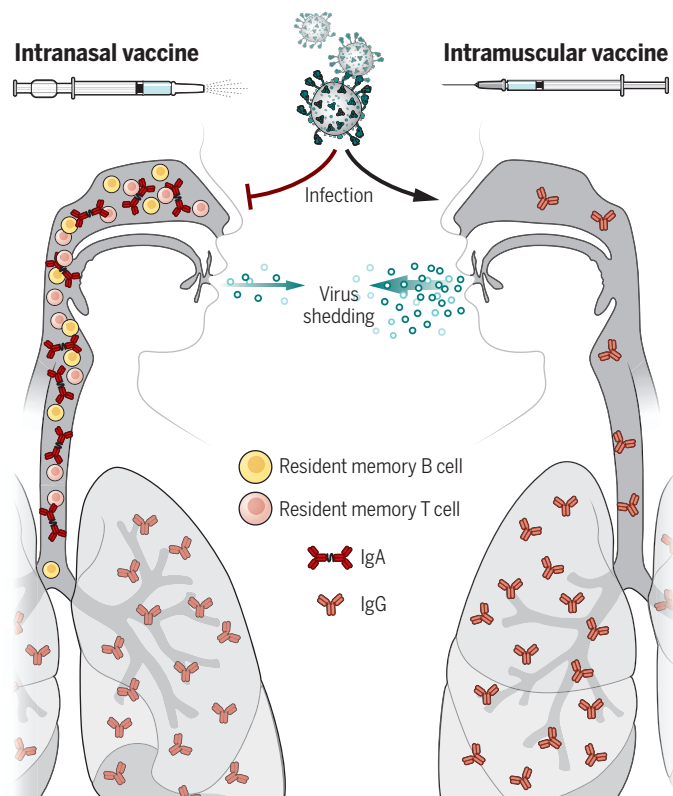
children, intranasal LAIV is generally superior to intramuscular vaccination (11). This success likely reflects the immunological naïveté of children (most have not been exposed to influenza virus). As a result, there is no immune barrier to LAIV infection in the nasal passages and vaccine “take” is efficient, leading to robust mucosal IgA responses and the placement of tissue-resident memory cells in the airways. LAIV is also effective in adults, but not necessarily better than intramuscular vaccination (11), in part because prior influenza virus infection has established a baseline of immunity that impairs the infectivity of LAIV. Consequently, live attenuated SARS-CoV-2 vaccines may elicit robust protection in naïve individuals, but preexposed individuals may have sufficient immunity to neutralize the vaccine, rendering it ineffective even as a booster.

Only one of the intranasal vaccines in clinical trials is inert—Cuba’s CIBG-669, which consists of RBD linked to the hepatitis B virus core antigen, a potent stimulator of T cells. Because inert vaccines do not rely on infection or gene expression, they cannot be neutralized by preexisting antibodies. However, soluble proteins delivered to the nasal passages do not efficiently breach the epithelium. Instead, they must be transported across the epithelial barrier by specialized microfold (M) cells (12), which deliver antigens to immune cells underneath the epithelium.

Notably absent from the list of intranasal vaccines are those formulated as lipid-encapsulated mRNA. Delivered intramuscularly, mRNA vaccines elicit high titers of serum IgG against encoded antigens. Rodent studies suggest that mRNA vaccines are also efficacious when delivered intranasally (13). However, it is important to distinguish intranasal delivery and nasal vaccination. Rodents are often anesthetized for intranasal vaccination and infection, causing them to take slow, deep breaths that deliver the inoculum all the way into the lung. As a result, much of the literature (including some cited here) on intranasal vaccination in rodents actually refers to intrapulmonary vaccination, which may provide more complete protection than strictly nasal vaccination.

Routes of vaccination

Immunoglobulin A (IgA) and resident memory B and T cells in the nasal passages and upper airways are elicited by intranasal vaccination and prevent infection and reduce virus shedding. Serum IgG elicited by intramuscular vaccination transudates into the lungs and prevents pulmonary infection but allows infection in the nasal passages and virus shedding.



Nevertheless, resident memory cells in the nasal passages can prevent virus dissemination to the lung (4). Given that vaccine delivery to the lower respiratory tract may directly cause inflammation or may exacerbate conditions such as asthma or chronic obstructive pulmonary disease (COPD), intranasal vaccines are typically administered to humans in a way that prevents antigen delivery to lungs.

Lipid formulation is critical for mRNA vaccine stability, for cell targeting, and for releasing mRNA to the cytosol. Thus, the future success of intranasal mRNA vaccines will likely hinge on developing lipid nanoparticles that target the appropriate cell types in the nasal passages. Unlike viruses and viral vectors, lipid nanoparticles lack proteins on their surface and should not be neutralized by antibodies, making the same formulation viable for repeated vaccination. However, adverse events such as fatigue and malaise are frequently linked to mRNA vaccination. Therefore, intranasal mRNA vaccines should be developed cautiously to avoid side effects and reactogenicity.

Ultimately, the goal of vaccination is to elicit long-lived protective immunity.

However, the duration of serum antibody responses varies considerably, depending on poorly understood attributes of the initiating antigen (14). Mucosal antibody responses are often considered short-lived, but their actual duration may depend on how antigen is encountered. Similarly, recirculating central-memory T cells are self-renewing and persist for long periods, whereas lung-resident memory T cells wane relatively rapidly—more so for CD8⁺ T cells than for CD4⁺ T cells. Thus, intranasal vaccines may have to balance the goal of local immunity in the respiratory tract with the longevity of systemic immunity. However, effective vaccination strategies need not be restricted to a single route. Indeed, memory cells primed by intramuscular vaccination can be “pulled” into mucosal sites by subsequent mucosal vaccination (15). Thus, the ideal vaccination strategy may use an intramuscular vaccine to elicit a long-lived systemic IgG response and a broad repertoire of central memory B and T cells, followed by an intranasal booster that recruits mem-

ory B and T cells to the nasal passages and further guides their differentiation toward mucosal protection, including IgA secretion and tissue-resident memory cells in the respiratory tract. ■

REFERENCES AND NOTES

1. Z. Wang et al., *Sci. Transl. Med.* **13**, eabf1555 (2021).
2. K. B. Renegar et al., *J. Immunol.* **173**, 1978 (2004).
3. S. R. Allie et al., *Nat. Immunol.* **20**, 97 (2019).
4. A. Pizzolla et al., *Sci. Immunol.* **2**, eaam6970 (2017).
5. S. R. McMaster et al., *Mucosal Immunol.* **11**, 1071 (2018).
6. A. O. Hassan et al., *bioRxiv* 10.1101/2021.05.08.443267 (2021).
7. R. G. King et al., *bioRxiv* 10.1101/2020.10.10.331348 (2020).
8. N. van Doremalen et al., *bioRxiv* 10.1101/2021.01.09.426058 (2021).
9. S. Tasker et al., *Vaccines (Basel)* **9**, 224 (2021).
10. T. Solomon, *Nat. Rev. Neurol.* **17**, 65 (2021).
11. K. G. Mohn et al., *Hum. Vaccin. Immunother.* **14**, 571 (2018).
12. H. Kiyono, S. Fukuyama, *Nat. Rev. Immunol.* **4**, 699 (2004).
13. J. C. Lorenzi et al., *BMC Biotechnol.* **10**, 77 (2010).
14. I. J. Amanna et al., *N. Engl. J. Med.* **357**, 1903 (2007).
15. H. Shin, A. Iwasaki, *Nature* **491**, 463 (2012).

ACKNOWLEDGMENTS

F.E.L. and T.D.R. are consultants for, and receive research funding from, Altimmune Inc.

RETROSPECTIVE

Ei-ichi Negishi (1935–2021)

Pioneering chemist who used transition metals to make and break bonds

By Shengming Ma^{1,2,3}

Ei-ichi Negishi, a groundbreaking chemist in the field of organometallic chemistry directed toward organic syntheses, died on 6 June at the age of 85. Negishi pioneered the use of transition metals to break and create covalent bonds. His technique using palladium to catalyze a general reaction, now known as the Negishi coupling, earned him the Nobel Prize in Chemistry in 2010 (shared with Richard F. Heck and Akira Suzuki). This reaction of organozinc reagents with halocarbons revolutionized the syntheses of natural products, bioactive compounds, drug molecules, and molecules for materials sciences. He also pioneered the use of zirconium to form a carbon-carbon bond, break a carbon-carbon double or triple bond, and form a carbon-aluminum bond.

Negishi, a Japanese citizen, was born in northern China on 14 July 1935. He received a bachelor's degree in chemistry from the University of Tokyo in 1958 and subsequently joined Teijin, Ltd., where he conducted research on polymer chemistry. He then moved to the United States on a Fulbright-Smith-Mundt scholarship. In 1963, he received his PhD in organic chemistry from the University of Pennsylvania. During his graduate studies, he was inspired by a lecture by Herbert C. Brown, one of two 1979 Nobel laureates in chemistry. After returning to Teijin, Ltd., for 3 more years, Negishi joined Brown's lab at Purdue University in West Lafayette, Indiana, as a postdoctoral fellow. In 1972, he became an assistant professor at Syracuse University in New York and began his lifelong study of transition metal-catalyzed reactions. He returned to Purdue University in 1979 and remained there until his retirement in 2019.

Between 1976 and 1980, Negishi's research group laid the foundation for the palladium-catalyzed cross-coupling reactions involving aluminum, zirconium, boron, and zinc reagents that would lead to his Nobel Prize. In addition, he made seminal contributions to organozirconium chemistry. His group created the highly reactive butene-ZrCp₂, now

known as the Negishi reagent. His zirconium-catalyzed carboalumination of alkynes and alkenes has reshaped the methods for the syntheses of various types of organozirconium and organoaluminum reagents.

I joined Negishi's group at Purdue University in 1993 as a postdoctoral fellow, although I had been familiar with his work since reading his publications on carbopalladation several years earlier. He was serious about instilling fastidious habits in his lab members, who would quickly learn to stay organized and keep detailed documentation. Negishi required us to analyze crude reaction mixtures before any means of separation to ensure that we did not lose useful scientific information, which, he reminded us, could



lead to serendipitous discoveries. Before every group meeting, we were required to fill out information sheets on the results of our recent experiments, making it easy for Negishi or colleagues to check our progress. He also instructed us to collect all the hard copies of crude nuclear magnetic resonance spectra and other characterization data in the order of the corresponding laboratory notebook page numbers and keep them in file folders, making them easy to find. His documentation was so well organized that I could find collections of my own spectra in his laboratory when I returned to Purdue University in 2016, almost two decades after recording those data. In addition to meticulous record-keeping, Negishi taught me

how to write effective scientific papers by showing unexpected observations to demonstrate originality.

Always open to unexpected observations from the lab and new ideas from students, Negishi was an excellent mentor. He taught me that a scientist should focus on just one or two topics and joked that if one has too many children and does not pay enough attention to them as they are growing up, they might decide to call someone else their father or mother. At the time I joined his group, many chemists were working on alkenes and alkynes, so he and I decided to pursue the unpopular topic of allenes. When I left West Lafayette, we both felt that the chemistry of allenes had great potential, and he encouraged me to continue studying them. His advice molded my career; I have now been studying allenes for more than 20 years.

Always energetic, Negishi often took group members on skiing excursions and out to eat at local restaurants—he enjoyed Chinese food, but tasting the authentic cuisine during a visit to Shanghai in the early 1990s put him off the Chinese food available in the United States. He loved playing the piano and was always ready to perform a piece or two. His favorites were classics by Beethoven, Bach, Pachelbel, and Mozart. He even conducted an orchestra during the closing ceremony of the Pacificchem 2015 conference, the world's largest meeting of the chemistry community.

The American Chemical Society honored Negishi in 1998 with the Award in Organometallic Chemistry and in 2010 with the Award for Creative Work in Synthetic Organic Chemistry. The Royal Society of Chemistry awarded him the Sir Edward Frankland Prize in 2000. He received Japan's Order of Culture in 2010. These are just a few of the many awards and honors that recognized his work, which so drastically changed the way chemists make and break the carbon bonds in modern medicines, agrochemicals, and functional materials.

Negishi was curious, frank, fair, and always ready to debate. Organized and open-minded, he was an admirable role model. He will surely be remembered by his family, students, and colleagues for his characteristic charm and loving nature, and his work will continue to inspire scientists to develop new chemical compounds to better the lives of human beings. ■

¹State Key Laboratory of Organometallic Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai, China. ²Department of Chemistry, Zhejiang University, Hangzhou, China. ³Department of Chemistry, Fudan University, Shanghai, China. Email: masm@sioc.ac.cn



Training is crucial in reducing noise in medical diagnostic contexts, argue the authors.

normative frameworks or known outcomes. One does not need to know the true value of a proposition—and, in fact, there may be no “true” value—to observe noisy results.

Consider the pronouncements of wine tasters or movie critics, for example. We might reasonably be suspicious of there being a “correct answer” in matters of taste and might observe that individuals tend to diverge in their evaluations. In these situations, divergences are inconsequential and expected, but in others, they can be both surprising and devastating.

The authors of *Noise* are especially adept at gathering examples of the latter type of noise, and they produce evidence of it in the most disparate of fields, ranging from hiring practices and performance evaluations to cancer diagnoses, financial forecasts, and even forensic science. They find that experts not only regularly disagree with each other in situations laypeople consider matters of fact but also sometimes disagree with their own earlier pronouncements.

The authors dedicate a great deal of attention to characterizing and accounting for sources and variations of noise, but their work aims to be constructive as well as critical. They prescribe what they call “decision hygiene,” urging decision-makers to regularly assume external perspectives, making use of relevant statistics instead of mere impressions; establish common anchors for shared decision-making frameworks to minimize scale subjectivity; and segment and strategically sequence the presentation of different dimensions of complex decisions to avoid cascades of influence, in which an early impression colors all subsequent assessments.

Noise is, by the authors’ own admission, more anecdotal than their previous works. Whereas both Kahneman’s *Thinking, Fast and Slow* and Sunstein’s *Nudge* (with Richard Thaler) were the culmination of well-vetted research programs, *Noise* is quite tentative at times, relying on ad hoc analyses and unpublished research. Furthermore, although it is impressive in its scope and ambition, statisticians are apt to be disappointed with the occasionally clumsy handling of concepts critical to their field. With this compromise, however, comes the assurance that the text will likely appeal to a broader audience than it might otherwise have and that such readers will discover a great deal about the sources, types, and means of minimizing decision-making noise. ■

REFERENCES AND NOTES

1. M. E. Frankel, *Current Contents* 25, 14 (1986).

10.1126/science.abj8353

BOOKS *et al.*

PSYCHOLOGY

The elusive quest to make consistent calls

Variability in high-stakes decision-making is more prevalent than we might like to believe

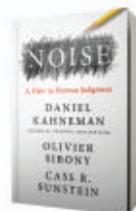
By **Darren Frey**

It is not uncommon to read about criminals receiving drastically different sentences for seemingly similar crimes. We may rationalize these differences, supposing we do not know all the relevant facts or the nuances of the laws applied.

We generally trust the expertise of the judges tasked with sentencing the accused. Yet, shortly after Judge Marvin Frankel first donned his robe and was required to render a verdict in a federal case, it occurred to him that he “had no experience or special knowledge whatever about the subjects of penology, criminology, sentencing philosophies, or any other pertinent learning” (1).

Unsettled by this reflection, Frankel began to observe sentencing patterns among his colleagues in the 1970s and found enormous discrepancies in judgments of very similar crimes. In one early study he oversaw, 50 judges from different districts were asked to evaluate a set of identical hypothetical cases, and the variability of their opinions was alarming. For example, the very same extortion case was

judged to merit anywhere from 3 years of imprisonment, in the most lenient case, to 20 years, in the harshest, with certain judges additionally imposing considerable fines. Subsequent legal analysis found variation of this sort to be the rule rather than the exception. How is it that such serious judgments could essentially amount to a high-stakes lottery?



Noise: A Flaw in Human Judgment
Daniel Kahneman,
Olivier Sibony,
Cass R. Sunstein
Little, Brown Spark,
2021. 464 pp.

Daniel Kahneman, Olivier Sibony, and Cass Sunstein’s new book, *Noise*, offers a sustained series of responses to this question while cataloging similar phenomena across a number of other domains. Their primary contention is that variability in decision-making contexts where one expects consistency is both more prevalent and more consequential than one would imagine. In their words, “wherever there is judgment, there is noise, and more of it than you think.”

To statisticians, the concept of noise is familiar, often the product of measurement limitations. If you are observing a series of measurements and you find that repeated readings (absent any environmental changes) are inconsistent, then the measurements are noisy.

Unlike bias, which can be thought of as a systematic deviation from a given outcome, noise occurs in decision contexts without

The reviewer is a lecturer at Sciences Po, Paris, France, and director at Ensemble Insight. Email: darrenf@post.harvard.edu

HISTORY OF SCIENCE

An atomic warning

Early efforts to expand nuclear energy were rife with racism and peril, reminds a historian

By **M. V. Ramana**

“Those who cannot remember the past are condemned to repeat it,” wrote the philosopher George Santayana in 1905. Heeding this aphorism, in *The Wretched Atom*, historian Jacob Darwin Hamblin seeks to remind readers of the misguided 20th-century effort launched by the United States, its allies, and international agencies to expand nuclear energy around the world. The compelling narrative should lead readers to realize the importance of preventing a repeat of the follies that marked the early decades of the atomic age.

Hamblin covers a vast amount of material in this well-written book, which en-

The events described in this book took place over many continents, with actors ranging from the right-wing media mogul (and nuclear energy supporter) Matsutaro Shoriki of Japan and Kwame Nkrumah, head of the state of the newly decolonized Ghana, to the plant breeder M. S. Swaminathan in India and Hans Blix, the Swedish director general of the International Atomic Energy Agency. The range of players reflects the many facets of our world tied to nuclear energy that Hamblin explores.

Nuclear weapons are, of course, the best-known technology that is closely tied to nuclear energy, and quite a few of Hamblin's chapters highlight how various geopolitical and economic goals outweighed concerns about nuclear weapons prolifera-



The United States helped apartheid South Africa establish the Pelindaba nuclear research center in 1965.

compasses the period from soon after the close of the Second World War through the end of the 20th century. He divides this history into three eras: “Atomic Promises,” a period defined by numerous ideas for civilian applications of atomic energy; “Atomic Propaganda,” when these ideas resulted in the spread of atomic technology, albeit with mixed success; and “Atomic Prohibition,” when non-Western countries, China and India in particular, acquired atomic bombs, and Western countries adopted new norms to govern nuclear commerce.

The reviewer is at the Liu Institute for Global Issues, School of Public Policy and Global Affairs, University of British Columbia, Vancouver, BC, Canada. Email: m.v.ramana@ubc.ca

tion in key decisions. But the importance of Hamblin's book lies in its exploration of the intersection of nuclear energy with seemingly unrelated topics, such as racism, colonialism and neocolonialism, propaganda, and surveillance and control.

Hamblin describes, for example, major differences in how the United States treated Ghana and South Africa in the 1960s. Both African countries were seriously interested in nuclear energy. Even as it passed civil rights legislation domestically, the Johnson administration supported apartheid South Africa's nuclear program, and US companies helped build its first nuclear reactor in 1965. In contrast, the United States not only refused to help Ghana, but, according

**The Wretched Atom:
America's Global
Gamble with Peaceful
Nuclear Technology**
Jacob Darwin Hamblin
Oxford University Press,
2021. 328 pp.



to Hamblin, the country also helped “crush” the “prospect of an ambitious peaceful nuclear program by an independent African nation not ruled by whites.”

Such instances are not mere historical oddities. Today, for example, some companies are working to develop so-called nuclear batteries, sealed reactors that contain enough fuel for their predicted operational lifetimes. Although seemingly a technological choice, the neocolonial politics of this endeavor lie in the envisioned geographical distribution: The reactors are to be assembled and fueled in predominantly white nations, while developing countries with predominantly nonwhite populations are to be at the receiving end. The implicit message is that the latter are not to be trusted with operating a reactor or replacing its fuel.

Historical amnesia is also at play in the US nuclear industry's recent attempts to sell nuclear power plants to Saudi Arabia despite Saudi leaders openly citing Iran's potential development of nuclear weapons as part of the reason for their interest in establishing a nuclear program of their own (1). Hamblin describes how the promise of nuclear electricity in India and Pakistan and desalinated water in Israel were used as justifications for nuclear programs in these countries, but these seemingly beneficial initiatives ultimately contributed to the advancement of each nation's nuclear weapons programs. Remembering this past could uncover the potential long-term repercussions of exporting nuclear reactors to Saudi Arabia.

All books involve editorial decisions about which topics to emphasize. I found myself wishing that *The Wretched Atom* had explored more thoroughly, for example, the corporate propaganda used to promote nuclear energy and increase profits, as well as the efforts by US governmental institutions to portray nuclear technology as risk free. But such is the hallmark of a good book—readers are left wanting more, not less. ■

REFERENCES AND NOTES

1. A. Murphy, M. V. Ramana, “The Trump administration is eager to sell nuclear reactors to Saudi Arabia. But why?” *Bulletin of the Atomic Scientists*, 16 April 2019; <https://thebulletin.org/2019/04/the-trump-administration-is-eager-to-sell-nuclear-reactors-to-saudi-arabia-but-why>.

10.1126/science.abj6811

Saccharomyces cerevisiae yeast, one example of microfungi, is crucial to the production of bread, beer, and wine.

Edited by Jennifer Sills

Include all fungi in biodiversity goals

We applaud the call for mycologists and decision-makers to seize the opportunity to include macrofungi in the post-2020 global biodiversity targets (“Include macrofungi in biodiversity targets,” Y. Cao *et al.*, Letters, 11 June, p. 1160). It is shocking that only a meager 425 of the millions of fungal species on the planet have been evaluated for The International Union for Conservation of Nature (IUCN) Red List of Threatened Species (1, 2). However, the fact that most of the fungi assessed on the IUCN Red List are macrofungi—those forming easily observed spore-bearing structures above- or belowground (3)—reflects a bias toward well-known species that hinders efforts to characterize global extinction risk for fungi (4). Microfungi deserve equal consideration.

Although people associate fungi with mushrooms, most fungi do not produce reproductive structures visible to the human eye. For example, the vitally important arbuscular mycorrhizal fungi colonize the roots of 80% of all plants, a symbiosis that helped plants establish on land (5). Molds, such as those from which penicillin was isolated, are also microfungi (6). *Saccharomyces* yeasts, which give us bread, beer, and wine, are single-celled (7, 8).

Lack of knowledge about which fungi are most at risk of extinction hampers our ability to inform conservation actions to support those species and ultimately provide fungi-based solutions to tackle pressing global challenges (9). Therefore, we are

calling for the parties of the Convention on Biological Diversity meeting at the UN Biodiversity Conference (COP15) later this year to explicitly include all fungi in the designated targets. Most working documents discuss the conservation of flora and fauna (10); incorporating the funga (11) will write Kingdom Fungi into conservation frameworks while unlocking funding for mycological research, surveys, and educational programs (12). Fungi underpin all life on Earth. We cannot afford to neglect them in our efforts to halt biodiversity loss.

Susana C. Gonçalves^{1*}, Danny Haelewaters^{2,3*}, Giuliana Furci⁴, Gregory M. Mueller⁵

¹Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, 3000-456 Coimbra, Portugal. ²Department of Biology, Ghent University, 9000 Ghent, Belgium. ³Faculty of Science, University of South Bohemia, 370 05 České Budějovice, Czech Republic. ⁴Fungi Foundation, Brooklyn, NY 11216, USA. ⁵Negaunee Institute for Plant Conservation Science and Action, Chicago Botanic Garden, Glencoe, IL 60022, USA.

*Corresponding author. Email: scgoncal@uc.pt (S.C.G.); danny.haelewaters@gmail.com (D.H.)

REFERENCES AND NOTES

- D.L. Hawksworth, R. Lücking, *Microbiol. Spectr.* **5**, FUNK-0052–2016 (2017).
- “The IUCN Red List of Threatened Species. Version 2021-1” (International Union for Conservation of Nature, 2021).
- G.M. Mueller *et al.*, *Biodivers. Conserv.* **16**, 37 (2007).
- E. Nic Lughadha *et al.*, *Plants People Planet* **2**, 389 (2020).
- W.R. Rimington *et al.*, *Proc. R. Soc. B* **285**, 20181600 (2018).
- J. Houbakken, J.C. Frisvad, R.A. Samson, *IMA Fungus* **2**, 87 (2011).
- B. Gallone *et al.*, *Cell* **166**, 1397 (2016).
- P.J. Boynton, D. Greig, *Yeast* **31**, 449 (2014).
- A. Antonelli *et al.*, “State of the world’s plants and fungi 2020” (Royal Botanic Gardens, 2020).
- Convention on Biological Diversity, “Update of the zero draft of the post-2020 global biodiversity framework” (2020); www.cbd.int/doc/c/3064/749a/0f65ac7f9def86707f4eaeaf/post2020-prep-02-01-en.pdf.
- F. Kuhar, G. Furci, E.R. Drechsler-Santos, D.H. Pfister, *IMA Fungus* **9**, 71 (2018).
- G. Furci, M. Sheldrake, C. Rodríguez-Garavito, “Fungi are critical to human, ecosystem, and planetary well-being: It’s time to include them within conservation frameworks,” *FaunaFloraFunga* (2021); www.faunaflorafunga.org/.

10.1126/science.abk1312

Climate lawsuits could protect Brazilian Amazon

In the Brazilian Amazon, illegal deforestation increased by 37.5% between 2016 and 2020 (1). Since 2019, 50% of annual forest destruction has occurred on public lands, perpetrated by land grabbers (2). Protected areas and undesignated public forests have been invaded and illegally deforested by land grabbers who sell the land for profit and use its natural resources (3). The government’s actions have been ineffective in curbing deforestation (4). From January to May, deforestation in the region increased by 25% compared to the same period a year earlier (1). Land grabbing and deforestation have been promoted by loosening protection regulations (5–8). Fortunately, in the absence of effective strategies by the executive branch of government, the Brazilian judiciary has been creative.

An unprecedented lawsuit (9) filed by the Brazilian Federal Prosecution Office (BFPO) against a land grabber requires compensation for climate damages. From 2011 to 2018, about 2400 ha of pristine forest were illegally deforested in a protected area (the Antimary Extractive Reserve) in the South of Amazonas State (9). Using the free access platform Carbon Calculator (10)

LIFE IN SCIENCE

The hidden Olympic spectator

From my perch deep in the rocky terrain of the Yanshan Mountains, 90 km northwest of Beijing, I watch as preparations are made for the 2022 Winter Olympics. Through a gap in the tree line, I can see the construction of the official Olympic alpine ski course across the valley. With just a few kilometers between my secluded field site and the slopes, I could return as a distant Olympic spectator. Yet, I would not be alone.

I found this scenic spot while working to document wildlife in greater Beijing. My group is especially interested in the leopard cat—an elusive felid slightly larger than a house cat. When leopards went extinct in the Beijing area in the 1990s, the leopard cat became the top predator. To examine the state of the local population, we set up cameras along 15 km of this mountainous trail.

One crystal clear winter morning, we captured an image of a leopard cat walking across fresh snow. In the background, the Olympic ski slope looms. The species persists, even in this human-dominated landscape so close to the megacity of Beijing.

In February, I will serve on the Olympic alpine skiing referee team. To be surrounded by hundreds of athletes and spectators will be a thrill—a marked contrast to my solitary fieldwork. Humans are capable of extraordinary feats, beautifully showcased by Olympic events. Such achievements bring me hope that we can leverage our strengths to ensure continued coexistence between wilderness and human settlements. As we watch the skiers race down the slopes, I know that the striding leopard cat will be watching us.

Shu-Jin Luo

Peking University, Beijing, China. Email: luo.shujin@pku.edu.cn

10.1126/science.abl3669



A leopard cat passes the distant site of future Olympic ski slopes near Beijing, China.

provided by the Amazon Environmental Research Institute (IPAM Amazon), the BFPO estimated the amount of carbon emission (1.5 million tons CO₂) due to that illegal deforestation. The BFPO projected the due compensation for climate damages at USD 9 million, 181.3% more than the compensatory value would be for local environmental harms alone. This amount was established by multiplying the CO₂ emitted by the illegal deforestation by USD 5 per ton of CO₂—a price referenced in the Amazon Fund, a REDD+ mechanism that invests in deforestation prevention and forest sustainability (11).

The Court has already issued a preliminary injunction in favor of the BFPO (12), recognizing the existence of illegal deforestation and embargoing the sale of cattle raised in the area. This innovative lawsuit sets a valuable precedent for other cases against land grabbers in the Brazilian Amazon. Strong monetary disincentives imposed by the courts could help win the battle against illegal deforestation and its adverse impacts on regional and global climate balance.

Ana Carolina Haliuc Bragança¹, Paulo Moutinho^{2*}, Rafael da Silva Rocha¹, Ane Alencar², Livia Laureto², Isabel Castro², Claudia Azevedo-Ramos³

¹Brazilian Public Prosecutor's Office (Ministério Público Federal), 69060-000 Manaus, Brazil.

²Amazon Environmental Research Institute (IPAM), 70863-520 Brasília, Brazil. ³Center for Higher Amazonian Studies (NAEA)/Federal University of Para, 66075-110, Belém, Brazil.

*Corresponding author.

Email: moutinho@ipam.org.br

REFERENCES AND NOTES

- PRODES—Amazônia (2020); www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes [in Portuguese].
- M. Fellows, A. Alencar, M. Bandeira, I. Castro, C. Guyot, "Amazon on fire: Deforestation and fire in indigenous lands" (2021); <https://ipam.org.br/wp-content/uploads/2021/04/Amazon-on-Fire-ILs.pdf>.
- C. Azevedo-Ramos et al., *Land Use Pol.* **99**, 104863 (2020).
- J. Alix-Garcia, L. L. Rausch, J. L'Roë, H. K. Gibbs, J. Munger, *Conserv. Lett.* **11**, e12414 (2018).
- B. Brito, P. Barreto, A. Brandao, S. Baima, P. H. Gomes, *Environ. Res. Lett.* **14**, 10.1088/1748-9326/ab1e24 (2019).
- W. D. Carvalho et al., *Perspect. Ecol. Conserv.* **17**, 122 (2019).
- L. Ferrante, P. M. Fearnside, *Environ. Conserv.* **46**, 261 (2019).
- B. W. Nelson et al., *For. Ecol. Manage.* **117**, 149 (1999).
- Public Civil Action, Civil Inquiry no. 1005885-78.2021.4.01.3200 (2021); <http://apps.mpf.mp.br/aplusmpf/index2#/detalhe/920000000000016005461?modulo=0&sistema=portal> [in Portuguese].
- "Carbon Calculator—Amazon," CCAL (2021); www.carboncal.org.br.
- Amazon Fund (2020), www.amazonfund.gov.br/en/home/.
- Ministério Público Federal v. de Rezende (2021); <http://climatecasechart.com/climate-change-litigation/non-us-case/ministerio-publico-federal-v-de-rezende/>.

10.1126/science.abk1981

RESEARCH

IN SCIENCE JOURNALS

Edited by Michael Funk



ANIMAL CULTURE

Birds in the bin

It is by now well accepted that humans are not the only animal to have complex culture, and we have also found that ecological novelty can lead to cultural innovation. Klump *et al.* documented the emergence of an evolving set of behaviors in response to human-generated resources, specifically garbage bins, in sulphur-crested cockatoos. This finding both documents the existence and spread of complex foraging culture among parrots, a lineage known for high-level cognitive function, and illuminates how the spread of a cultural innovation can lead to regional distinct variations. —SNV *Science*, abe7808, this issue p. 456

Cockatoos have devised methods of opening trash bins, and individuals can learn these methods from each other.

STRUCTURAL BIOLOGY

How to catch a dynamic state

AMP-activated protein kinase (AMPK) is a key sensor of energy status in eukaryotes. Its dynamic structure is regulated by allosteric factors including phosphorylation and binding of nucleotides and metabolites. Yan *et al.* developed conformation-specific antibodies that trap AMPK in a fully inactive state that has experienced a large, domain-level rotation. Biophysical experiments in cells and in vitro are consistent

with the structural work and support a model in which the activation loop is fully exposed in the completely inactive, dephosphorylated state. These structures inform our understanding of the complex allosteric behavior in this crucial metabolic regulator. —MAF

Science, abe7565, this issue p. 413

PHOTOPHYSICS

Straining for high photoluminescence

The photoluminescence quantum yield in monolayer

transition metal dichalcogenides generally drops at high emission intensities because the excitons undergo nonradiative annihilation. Kim *et al.* show that this process is resonantly amplified in these materials by van Hove singularities in their joint density of states. However, application of small mechanical strains (~0.5%) shifted the van Hove singularities and suppressed the nonradiative processes. Near-unity photoluminescence quantum yield at high exciton densities was seen in exfoliated monolayers of molybdenum sulfide, tungsten sulfide, and

tungsten selenide, as well as centimeter-scale tungsten sulfide monolayers grown by chemical vapor deposition.

—PDS

Science, abi9193, this issue p. 448

PLANT SCIENCE

Calcium signaling for host cell death

In response to microbial pathogens, some plants kill off their own cells to limit further spread of infection. The Toll/Interleukin-1 receptor/Resistance class of nucleotide-binding leucine-rich repeat receptors (known as NLRs) function in plants as immune receptors. These NLRs work together with a dedicated set of helper proteins. Jacob *et al.* reveal the structure of one of these helpers known as NRG1 (N REQUIREMENT GENE 1). The structure resembles a known animal cation channel. The authors demonstrate that helper NLRs directly control calcium ion influx to initiate host cell death, providing a mechanism for NLR outputs. —PJH

Science, abg7917, this issue p. 420

SPECTROSCOPY

Triplet-state lifetime quenched by oxygen

Little is known about the atomistic mechanism that nature uses to mitigate the destructive interaction of triplet-excited pigment chromophores with omnipresent oxygen. Peng *et al.* tackled this challenge by developing a technique based on conducting atomic force microscopy to populate and track triplets in a single pentacene molecule, a model π -conjugated system, placed on a sodium chloride surface (see the Perspective by Li and Jiang). The authors show how the triplet-state lifetime can be quenched in controllable manner by atomic-scale manipulations with oxygen co-adsorbed in close vicinity. The presented single-molecule spectroscopy paves the way for further atomically resolved studies of triplet excited states that

play an important role in many other fields, such as organic electronics, photocatalysis, and photodynamic therapy. —YS

Science, abh1155, this issue p. 452;
see also abj5860, p. 392

BIOCHEMISTRY

Stimulating and suppressing HIFs

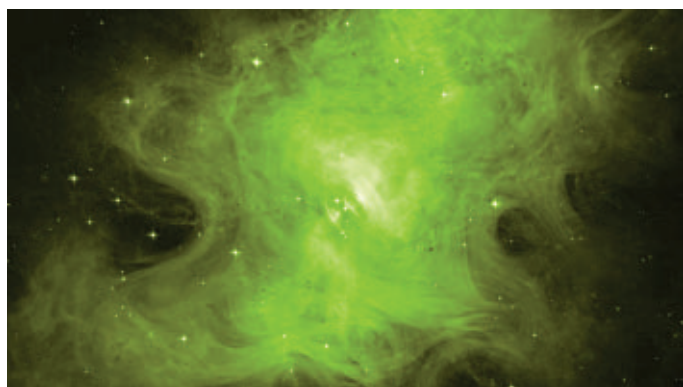
Cells respond and adapt to hypoxia (low oxygen) by activating the α subunits of the hypoxia-inducible factor (HIF) family of transcription factors. Daly *et al.* explored the protein modifications and binding partners for HIF-1 α and HIF-2 α . Their findings, which included a phosphorylated cysteine residue and interactions with mitochondrial proteins, indicate that the oxygen-dependent regulation of HIF activity is more complex than was previously appreciated. This dataset will help to delineate the selective regulation of these closely related isoforms in fine-tuning the cellular response to hypoxia. —LKF

Sci. Signal. **14**, eabf6685 (2021).

ASTROPARTICLE PHYSICS

High-energy photons from the Crab Nebula

The Crab Nebula contains a pulsar that excites the surrounding gas to emit high-energy radiation. The combination of the pulsar's youth and nearby location makes the nebula the brightest gamma-ray source in the sky. Cao *et al.* report



The Crab Nebula—seen here in a colored, visible light image—produces extremely high-energy gamma rays.

observations of this source at energies of tera- to peta-electron volts, extending the spectrum of this prototypical object. They combine these data with observations at lower energies to model the physics of the emission process. The multiwavelength data can be explained by a combination of synchrotron radiation and inverse Compton scattering. —KTS

Science, abg5137, this issue p. 425

MALARIA

An antimalarial advance

There is great need for antimalarial drugs in the face of growing resistance to existing therapies. Muriithi *et al.* characterized MMV688533, an acylguanidine identified from compounds inhibiting known human drug targets that were screened for activity against *Plasmodium falciparum*. MMV688533 showed rapid in vitro killing of multiple *P. falciparum* strains as well as *P. vivax*. A single dose rapidly reduced parasitemia in a *P. falciparum* severe combined immunodeficient mouse model of infection, and this agent displayed favorable pharmacokinetic and toxicity profiles. MMV688533 selected for only low-grade resistance, with resistant parasites remaining sensitive to existing antimalarials. These findings suggest that MMV688533 is a promising antimalarial candidate with a low resistance risk and the promise of single-dose cure, which merits further study. —CNF

Sci. Transl. Med. **13**, eabg6013 (2021).

IN OTHER JOURNALS

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



Trees rehydrate and grow mostly at night, when soil and air are more humid.

GENOMICS

Australian disappearances

One of the highest levels of extinction occurred in Australia, where 34 mammals have disappeared since European colonization. One suggestion for this vulnerability is the small population sizes and thus lack of genetic diversity, an explanation that could be predictive of future losses. To investigate this possible mechanism, Roycroft *et al.* obtained genomic data from eight extinct rodent species from museum samples in Australia collected ~150 years ago and compared them with 42 closely related extant species. Unexpectedly, the authors found relatively high genetic diversity among the specimens, indicating that the populations were in fact quite large. From these data, the authors estimate that ~10 million years of unique evolutionary history has been lost in Australia since

European colonization. Large population sizes and diversity of species seem to have offered little protection against the depredations of humans. —LMZ

Proc. Natl. Acad. Sci. U.S.A. **118**, e2021390118 (2021).

CANCER

Cancer-killing viruses to the brain

Oncolytic viruses can selectively infect cancer cells to stimulate cell death and antiviral and antitumor immune responses. Fares *et al.* report a phase 1 trial in patients with high-grade glioma treated with an engineered oncolytic adenovirus delivered by neural stem cells (NSCs). Delivery of oncolytic viruses in NSCs has been previously shown in mouse models to improve the delivery and distribution of virus particles within gliomas. The 12 patients underwent surgical resection, and the NSCs containing oncolytic virus



PLANT SCIENCE

Growing by night

Tree rings provide an annual report on how trees grow. Using continuous dendrometer measurements, Zweifel *et al.* were able to measure radial growth in tree trunks moment by moment. Deciduous or evergreen trees were surveyed in a selection of Swiss forests that had been mostly managed sustainably since 1876. The trees mainly grew at night, most rapidly just before dawn. Midday to afternoon was evidently the down time for radial growth. Growth in girth was aided by soil moisture and air humidity; indeed, trees needed several hours of night to slake the thirst they acquired during the day. The authors suggest that radial stem growth depends on nocturnal atmospheric conditions, whereas carbon assimilation depends more on daytime conditions. —PJH

New Phytol. 10.1111/nph.17552 (2021).

deposition is usually transferred to other substrates for device applications. Dry transfer can be achieved with rubbery polymer overlays, but this usually requires etching of the metal, and the relatively stiff polymer can mechanically introduce defects into monolayer graphene. Seo *et al.* spin-coated an n-doped adhesive gel formed from an amine-rich poly-ethylenimine cross-linked with glutaraldehyde onto graphene monolayers grown on copper. After capping with a silica-on-silicon substrate, a gel formed after thermal curing, and peeling of substrate removed the graphene layer from the metal. The charge transfer into the graphene helped to prevent the mechanical shock that could have wrinkled or torn the graphene and also imparted high conductivity and maintained optical transparency. —PDS
ACS Nano 10.1021/acsnano.0c10798 (2021).

were injected into the brain. The patients were then treated with chemotherapy plus radiotherapy. The trial showed that the NSC-oncolytic virus treatment was safe, and there was some evidence of immune stimulation and potential survival benefit, although this was not a placebo-controlled study. —GKA

Lancet Oncol. 10.1016/S1470-2045(21)00245-X (2021).

MATERIALS SCIENCE

Polymeric fire-extinguishing coatings

Although lightweight polymeric foams work well for thermal insulation, they tend to be flammable and do not bond well with most fire-retardant coatings. Drawing inspiration from the adhesion capabilities of snails and tree frogs, Ma *et al.* designed a polymeric coating based on a copolymer of hydroxyethyl acrylate

(HEA) and sodium vinylsulfonate (VS). This waterborne, phase-separated polymer can be sprayed onto foams, where the hydroxyl groups in the HEA lead to strong H-bonding and the sodium sulfonate groups in the VS provide fire retardancy. The coating does not affect the thermal insulation properties of the foam but can self-extinguish a flame with low heat and smoke release because of its tendency to char. —MSL
ACS Nano 10.1021/acsnano.1c02254 (2021).

SCIENTIFIC WORKFORCE

Research experiences and social justice

Past diversity training initiatives have not yet led to equity in the STEM field. Camacho *et al.* propose a new tool to help approach that goal. BUILD PODER is an undergraduate biomedical research training program rooted in Critical Race

Theory (CRT). The program unites students and faculty around biomedical research questions while contextualizing these questions around issues of social justice. The integration of research and social justice allows for a curriculum directed at transparency, respect, and correcting the historical abuses in science. Analysis of outcomes among a sample of undergraduate Latinx seniors revealed that BUILD PODER participants reported higher levels of science personal identity and science social identity upon graduation, demonstrating a positive effect of CRT-informed research experiences on Latinx students' science identity. —MMc

CBE Life Sci. Ed. 20, ar23 (2021).

GRAPHENE TRANSFER

Coat, dope, peel

Graphene grown on metal surfaces by chemical vapor

PHYSIOLOGY

Albumin's many talents

Albumin is a common soluble protein with many physiological roles, including hormone transport, antioxidant activity, and regulation of cellular pressure and intracellular pH. Zhao *et al.* found that albumin binding is needed for the activation of the human voltage-gated protein channel (hHv1). Binding and mutagenesis analyses and modeling by molecular dynamics simulation showed that albumin binds across the two voltage sensor domains of hHv1 to enhance the open-state probability of the channel and increase the proton current. In the uterus, high albumin levels promote binding to the hHv1 channel to initiate sperm capacitation. Albumin likely has similar stimulatory activities on the hHv1 channel in other tissues, for example, in sustaining pathological levels of reactive oxygen species and proteases from neutrophils in some disease states. —BAP

Nat. Commun. 12, 3855 (2021).

ALSO IN SCIENCE JOURNALS

Edited by Michael Funk

NEUROIMMUNOLOGY

Getting around the blood–brain barrier

The meninges comprise three membranes that surround and protect the central nervous system (CNS). Recent studies have noted the existence of myeloid cells resident there, but little is known about their ontogeny and function, and whether other meningeal immune cell populations have important roles remains unclear (see the Perspective by Nguyen and Kubes). Cugurra *et al.* found in mice that a large proportion of continuously replenished myeloid cells in the dura mater are not blood derived, but rather transit from cranial bone marrow through specialized channels. In models of CNS injury and neuroinflammation, the authors demonstrated that these myeloid cells have an immunoregulatory phenotype compared with their more inflammatory blood-derived counterparts. Similarly, Brioschi *et al.* show that the meninges host B cells that are also derived from skull bone marrow, mature locally, and likely acquire a tolerogenic phenotype. They further found that the brains of aging mice are infiltrated by a second population of age-associated B cells, which come from the periphery and may differentiate into autoantibody-secreting plasma cells after encountering CNS antigens. Together, these two studies may inform future treatment of neurological diseases. —STS

Science, abf7844, abf9277, this issue p. 409, p. 408; see also abj8183, p. 396

CORONAVIRUS

Intranasal vaccines

Among the many vaccines against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), intranasal forms fill an important gap. Most vaccines are administered

intramuscularly, where they elicit systemic immune responses and central immune memory. But respiratory viruses predominantly enter the nasal passage first, suggesting that a strong mucosal immune response in the nasal cavity, as is induced by intranasal vaccines, may be beneficial. In a Perspective, Lund and Randall discuss the SARS-CoV-2 intranasal vaccines that are in clinical trials, the rationale for their use, and whether they may be viable for immunization or as a booster to improve protection from SARS-CoV-2. —GKA

Science, abg9857, this issue p. 397

PHYSIOLOGY

Intestinal HDL is hepatoprotective

High-density lipoprotein (HDL) is important for cholesterol metabolism and may have anti-inflammatory and antimicrobial properties. Although HDL is mainly produced by the liver, the intestine is also a source. Han *et al.* show in mice that intestinal HDL is not routed to the systemic circulation. Rather, in the form of HDL3, it is directly transported to the liver through the hepatic portal vein. There, it sequesters bacterial lipopolysaccharide from the gut that can trigger inflammation and liver damage. In various models of liver injury, loss of enteric HDL exacerbated pathology. By contrast, drugs elevating intestinal HDL improved disease outcomes. HDL3 is enriched in human portal venous blood, suggesting that enteric HDL may be targetable for the treatment of liver disease. —STS

Science, abe6729, this issue p. 410

ENZMOLOGY

Go big or you'll get lost

Rational mutagenesis is a common approach to investigating or engineering enzyme function in vitro, but the ease with which one can manipulate protein

sequences belies many pitfalls in connecting sparse activity data to an enzyme's true functional landscape. Using a high-throughput platform, Markin *et al.* expressed, purified, and performed an array of kinetic measurements on a target esterase, collecting data from >1000 mutations spanning the entire protein (see the Perspective by Baumer and Whitehead). Protein misfolding into an inactive state, rather than decreased equilibrium stability, was a crucial factor in negatively affected variants spread throughout the protein. When combined with prior mechanistic understanding and structures, four “functional components” help to rationalize the otherwise complex spatial pattern of effects of mutations on different aspects of enzyme function, all of which would be invisible from mutagenesis of just a few residues. —MAF

Science, abf8761, this issue p. 411; see also abj8346, p. 391

NEURODEVELOPMENT

Developing neurons practice for real life

As a mouse runs forward across the forest floor, the scenery that it passes flows backwards. Ge *et al.* show that the developing mouse retina practices in advance for what the eyes must later process as the mouse moves. Spontaneous waves of retinal activity flow in the same pattern as would be produced days later by actual movement through the environment. This patterned, spontaneous activity refines the responsiveness of cells in the brain's superior colliculus, which receives neural signals from the retina to process directional information. —PJH

Science, abd0830, this issue p. 412

QUANTUM DEVICES

Superconducting spin qubit

To date, the most promising solid-state approaches for developing quantum information-processing systems have been based on the circulating supercurrents of superconducting circuits and manipulating the spin properties of electrons in semiconductor quantum dots. Hays *et al.* combined the desirable aspects of both approaches, the scalability of the superconducting circuits and the compact footprint of the quantum dots, to design and fabricate a superconducting spin qubit (see the Perspective by Wendin and Shumeiko). This so-called Andreev spin qubit provides the opportunity to develop a new quantum information processing platform. —ISO

Science, abf0345, this issue p. 430; see also abk0929, p. 390

PLANETARY SCIENCE

Single seismometer structure

Because of the lack of direct seismic observations, the interior structure of Mars has been a mystery. Khan *et al.*, Knapmeyer-Endrum *et al.*, and Stähler *et al.* used recently detected marsquakes from the seismometer deployed during the InSight mission to map the interior of Mars (see the Perspective by Cottaar and Koelemeijer). Mars likely has a 24- to 72-kilometer-thick crust with a very deep lithosphere close to 500 kilometers. Similar to the Earth, a low-velocity layer probably exists beneath the lithosphere. The crust of Mars is likely highly enriched in radioactive elements that help to heat this layer at the expense of the interior. The core of Mars is liquid and large, ~1830 kilometers, which means that the mantle has only one rocky layer rather than two like the Earth has. These results provide

a preliminary structure of Mars that helps to constrain the different theories explaining the chemistry and internal dynamics of the planet. —BG

Science, abf2966, abf8966, abi7730, this issue p. 434, p. 438, p. 443
see also abj8914, p. 388

MATERNAL IMMUNOLOGY

Aire-expressing cells defend the fetus

The normal maternal immune system adapts to pregnancy by learning to tolerate both allogeneic paternal and pregnancy-associated antigens expressed by the fetus and placenta. Gillis-Buck *et al.* investigated the contribution of cells expressing the autoimmune regulator (*Aire*) gene to maintaining fetal health using a mouse model in which *Aire*-expressing cells were deleted during early pregnancy. Ablation of *Aire*-expressing cells did not impair fetal implantation, but resulted in intrauterine fetal growth restriction during both allogeneic and syngeneic pregnancies and was associated with increased activation of effector T cells. Selective deletion of extra-thymic *Aire*-expressing cells (eTACs) while preserving *Aire*-expressing thymic epithelial cells revealed that eTACs are required for maintaining tolerance and fetal health. These findings demonstrate an unanticipated role for eTACs as guardians of fetal health during pregnancy. —IRW

Sci. Immunol. **6**, eabf1968 (2021).

ultraviolet laser with molecules of deuterated water (HOD). The theoretical expectation was that dissociation to H+OH would be more common, but they found instead that dissociation to D+OH was more common and that the ratio of H+OH to D+OH reactions was highly wavelength dependent. This phenomenon may have played a key role in determining the chemical structure of the material from which Earth was formed. —ES

Sci. Adv. 10.1126/sciadv.abg7775
(2021).

ASTROPHYSICS

Understanding the origin of Earth's water

The deuterium to hydrogen (D/H) ratio in Earth's oceans differs from the bulk composition of the protosolar nebula, which suggests that either some of Earth's water was deposited by comets and asteroids or that large spatial fluctuations in D/H existed in the initial nebula. Luo *et al.* performed detailed laboratory measurements of the interaction of a tunable

RESEARCH ARTICLE SUMMARY

NEUROIMMUNOLOGY

Heterogeneity of meningeal B cells reveals a lymphopoietic niche at the CNS borders

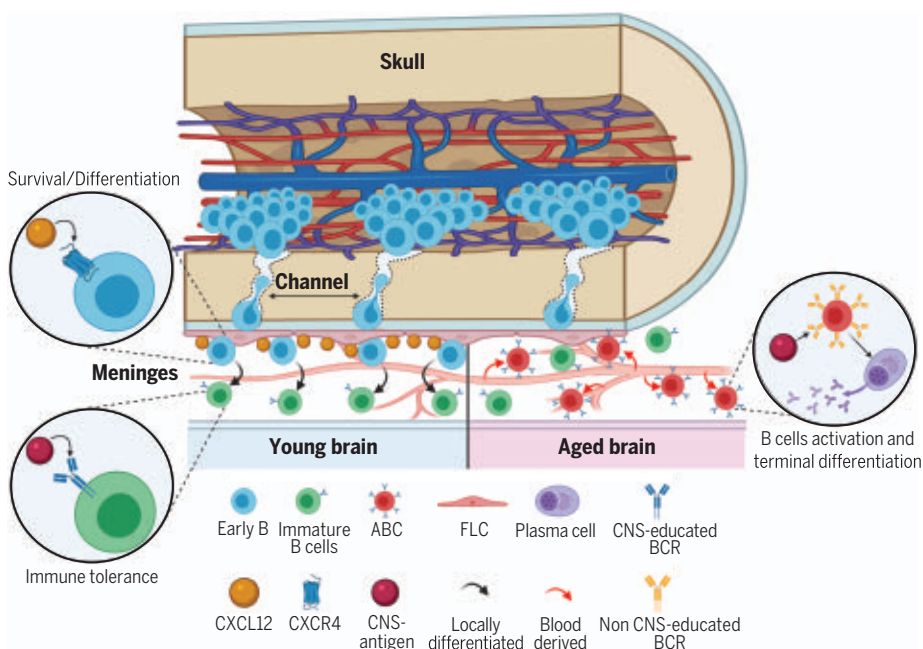
Simone Brioschi[†], Wei-Le Wang[†], Vincent Peng[†], Meng Wang, Irina Shchukina, Zev J. Greenberg, Jennifer K. Bando, Natalia Jaeger, Rafael S. Czepielewski, Amanda Swain, Denis A. Mogilenko, Wandy L. Beatty, Peter Bayguinov, James A. J. Fitzpatrick, Laura G. Schuettelpelz, Catrina C. Fronick, Igor Smirnov, Jonathan Kipnis, Virginia S. Shapiro, Gregory F. Wu, Susan Gilfillan, Marina Cella, Maxim N. Artyomov, Steven H. Kleinstein, Marco Colonna*

INTRODUCTION: The meninges are a three-layer membrane that covers the central nervous system (CNS): The outermost layer, the dura mater, is attached onto the skull periosteum; the two inner layers, the arachnoid and the pia mater, cover the brain cortex. The meningeal compartment hosts both innate and adaptive immune cells, providing constant immunosurveillance of the CNS border regions. Given its strategic location, meningeal immunity is a key component of neuroimmune disorders. Recent studies have investigated the origin and dynamics of meningeal myeloid cells. However, little is known about meningeal B cells.

RATIONALE: To investigate the phenotype of meningeal B cells, we used different high-

throughput techniques, such as single-cell RNA sequencing (scRNA-seq), cytometry by time of flight (CyTOF), and single-cell B cell receptor sequencing (scBCR-seq). We further corroborated our finding by flow cytometry and confocal imaging. We determined the origin of meningeal B cells by bone marrow transplantation (BMT) and parabiosis experiments. Lastly, we investigated how meningeal B cells change during aging.

RESULTS: We characterized the mouse meninges using scRNA-seq, which revealed that meningeal B cells encompass multiple stages of their development, spanning pro-B to mature B cells. Identical subsets were found in the bone marrow (BM), but not in the blood. CyTOF and flow



Composition of meningeal B cells in young and aged mice. Meningeal B cells originate from the calvaria and migrate to the meninges through skull vascular channels. These B cells complete their development locally, where fibroblast-like cells (FLCs) provide critical factors for survival and differentiation of early B cells (i.e., CXCL12). At this stage, CNS-antigen experience induces immune tolerance. In aged mice, age-associated B cells (ABCs) infiltrate the meninges from the periphery and differentiate into Ig-secreting plasma cells upon encounter with CNS antigens.

cytometry further confirmed this result, demonstrating that early B cell subsets that are normally present in the BM are also found in the meninges under homeostasis. BMT with selective reconstitution of the skull BM showed that meningeal B cells are derived from the calvaria; namely, the hematopoietic region located within the cranial flat bones. We then performed parabiosis between wild-type and CD19-Tomato mice, which express the tdTomato fluorescent protein specifically in B cells. This experiment demonstrated that circulating B cells minimally infiltrated the mouse meninges under homeostasis. Using confocal imaging, we showed that B cells migrate from the calvaria to the meninges through specialized vascular channels traversing the inner skull bone. Interactome analysis of our scRNA-seq data highlighted a network of molecular communications between meningeal B cells and dura fibroblast-like cells (FLCs). Notably, we showed that FLCs express high levels of *Cxcl12*, whereas dura early B cells express its receptor *Cxcr4*. Expression of these molecules in the respective populations was validated by imaging and flow cytometry. The CXCL12–CXCR4 axis is required for the survival and differentiation of early B cells in the BM, and the same mechanism may be present in the dura. We also showed that aged mice (~2 years of age) accumulate age-associated B cells (ABCs) and plasma cells in the dura. Compared to naïve B cells, ABCs featured important transcriptional changes, as well as reduced diversity of the V-region repertoire and accumulation of somatic mutations, thus indicating antigen experience. Analysis of B cells clones by scBCR-seq showed that dura ABCs infiltrated from the periphery. On the basis of BCR clonality, we also suggest that dura ABCs may locally undergo terminal differentiation into immunoglobulin (Ig)-secreting plasma cells.

CONCLUSION: This study reveals that mouse meninges harbor a lymphopoietic niche specific for the CNS borders. B cell development in the meninges may induce immune tolerance against CNS antigens, thus preserving immune privilege within the CNS. However, blood-derived ABCs accumulate in the meninges over time. Peripheral ABCs are not educated by the CNS antigens and may locally differentiate into CNS-reactive plasma cells. This condition may endanger the immune-privileged CNS environment during aging. ■

The list of author affiliations is available in the full article online.

*Corresponding author. Email: mcolonna@wustl.edu

[†]These authors contributed equally to this work.

Cite this article as S. Brioschi et al., *Science* 373, eabf9277 (2021). DOI: 10.1126/science.abf9277

READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abf9277>

RESEARCH ARTICLE

NEUROIMMUNOLOGY

Heterogeneity of meningeal B cells reveals a lymphopoietic niche at the CNS borders

Simone Brioschi^{1†}, Wei-Le Wang^{1†}, Vincent Peng^{1†}, Meng Wang², Irina Shchukina¹, Zev J. Greenberg³, Jennifer K. Bando⁴, Natalia Jaeger¹, Rafael S. Czepielewski¹, Amanda Swain¹, Denis A. Mogilenko¹, Wandy L. Beatty⁵, Peter Bayguinov⁶, James A. J. Fitzpatrick^{6,7,8}, Laura G. Schuettelpelz³, Catrina C. Fronick⁹, Igor Smirnov¹, Jonathan Kipnis¹, Virginia S. Shapiro¹⁰, Gregory F. Wu¹¹, Susan Gilfillan¹, Marina Cella¹, Maxim N. Artyomov¹, Steven H. Kleinstein^{2,12}, Marco Colonna^{1*}

The meninges contain adaptive immune cells that provide immunosurveillance of the central nervous system (CNS). These cells are thought to derive from the systemic circulation. Through single-cell analyses, confocal imaging, bone marrow chimeras, and parabiosis experiments, we show that meningeal B cells derive locally from the calvaria, which harbors a bone marrow niche for hematopoiesis. B cells reach the meninges from the calvaria through specialized vascular connections. This calvarial-meningeal path of B cell development may provide the CNS with a constant supply of B cells educated by CNS antigens. Conversely, we show that a subset of antigen-experienced B cells that populate the meninges in aging mice are blood-borne. These results identify a private source for meningeal B cells, which may help maintain immune privilege within the CNS.

The central nervous system (CNS) is enveloped by the meninges, which harbor different immune cell types that provide constant surveillance at the CNS border (1–3). Although meningeal lymphocytes are thought to derive exclusively from the systemic circulation, recent findings question this premise. Vascular connections between the calvaria (the flat bones forming the top dome of the skull) and meninges have recently been described (4–6), and two multi-dimensional studies identified a cluster of developmentally immature B cells in the mouse CNS (7, 8). We hypothesized that these B cells may derive from calvarial hematopoiesis. Here, we show that meningeal B cells encompass

multiple stages of B cell development, spanning pro-B to mature B cells. Using parabiotic mice and bone marrow (BM) chimeras with selective reconstitution of the skull BM, we demonstrate that most meningeal B cells originate from the calvaria. We envision that calvaria-derived B cells are locally educated by CNS-derived antigens to prevent the generation of immunoglobulins (Igs) with high affinity for CNS epitopes. Conversely, a population of age-associated B cells (ABCs) infiltrates the mouse meninges from the circulation during aging. This study sheds light on the origin and phenotypes of meningeal B cells in homeostasis and during aging, challenging the widely accepted idea that meningeal adaptive immunity originates exclusively from systemic circulation.

Results

Meningeal B cells are extravascular and can exit the CNS compartment through the dura lymphatics

Meninges are formed by three membranes. The dura mater is the outermost layer attached onto the skull periosteum, whereas the two inner layers, the arachnoid and pia mater (leptomeninges), cover the brain cortex (Fig. 1A). This region is enriched in different immune cell types, and B cells represent about 30% of the total CD45⁺ cells in mouse meninges (Fig. 1B and fig. S1A). Because the dura contains a relative abundance of immune cells and can be more easily dissected than the leptomeninges, we focused most of our subsequent studies on dural B cells. By flow cytometry, B cells were found in samples of both the dura mater and the brain and leptomeninges (fig. S1B). The majority of meningeal B cells in young-adult

mice were B2 type, whereas innate B cells (B1a and B1b) represented a minor population (fig. S1C). We surveyed B cells by confocal imaging in the brain and spinal cord of *Cd19*^{Cre}; *Rosa26*^{tdTomato} mice (hereafter CD19-Tomato), which express the tdTomato fluorescent protein specifically in CD19⁺ cells (9). Although no B cells were found in the brain parenchyma (fig. S1D), B cells were present in the leptomeninges along the brain surface (fig. S1E), indicating that brain B cells are extraparenchymal. Two-photon in vivo imaging in the subdural space of the CD19-Tomato mice showed that most meningeal B cells were localized in the extravascular compartment (Fig. 1, C and D) and appeared relatively immobile as compared to intravascular B cells (Fig. 1, E and movie S1). The dura mater contains blood and lymphatic vessels (LVs) along the sagittal and transverse sinuses (fig. S2A). These areas were particularly enriched in B cells, some of which were located within dura LVs (fig. S2B). Dura lymphatics drain immune cells and molecules to the cervical lymph nodes (cLNs) (10, 11), suggesting that meningeal B cells may undergo a similar fate. To test this hypothesis, we introduced CD19-Tomato splenocytes into the cerebrospinal fluid (CSF) of wild-type mice by intracisterna magna (ICM) injection (Fig. 1F), and 24 hours later CD19-Tomato B cells were found accumulating in the cLNs (Fig. 1, G and H). Donor-derived B cells appeared to be located in both the dura LVs and cLN B cell zone (Fig. 1, I and J). Thus, LVs may serve as a migratory route for B cells exiting the CNS compartment.

Meningeal B cells are phenotypically similar to bone marrow B cells

Although B cells represent a main immune population in the meninges, little is known about their composition under homeostasis. A previous single-cell RNA sequencing (scRNA-seq) study on mouse brain immune cells identified two distinct B cell subsets, labeled as “mature” and “immature” (7). We reclustered these transcriptomic data (GSE98969) and identified B cells based on the enrichment of Ig transcripts (Fig. 2A). This population uniformly expressed pan-B cell signature genes (*Cd79b* and *Cd19*). However, mature B cell genes (*H2-Aa* and *Ms4a1*) and early B cell genes (*Rag1* and *Cd93*) appeared unevenly distributed in this cluster (Fig. 2B), suggesting that B cells in the CNS encompass different stages of their development. Next, we compared the phenotype of meningeal B cells to that of B cells in the BM, blood, and spleen (Fig. 2C). Based on the Hardy fraction (12, 13), early B cells were identified as CD19⁺B220^{lo}CD43^{hi} (bona fide fraction B-C), late B cells as CD19⁺B220^{lo}CD43^{lo} (bona fide fraction D-E), and mature B cells as CD19⁺B220^{hi}CD43⁺ (fraction F). BM and meningeal B cells appeared

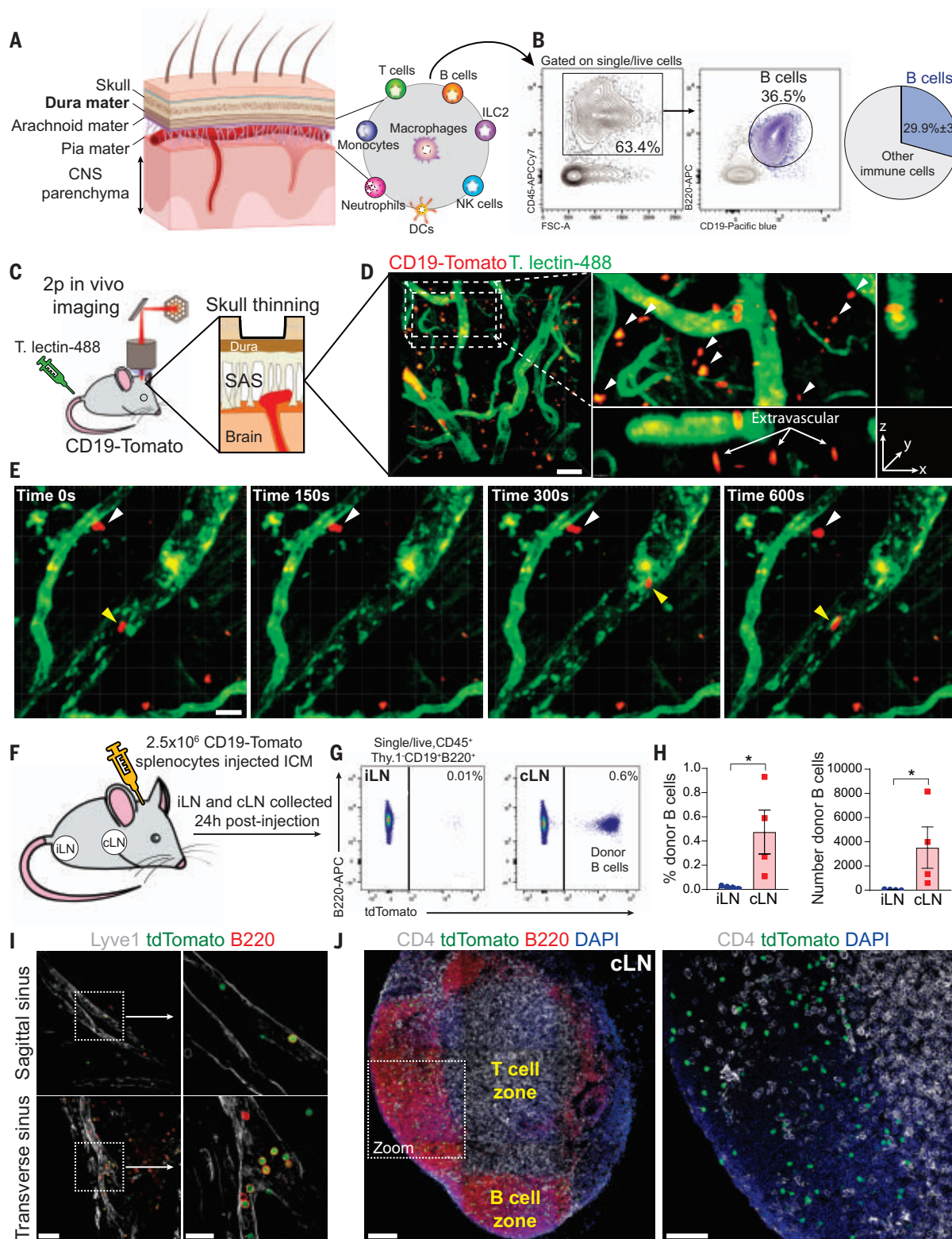
¹Department of Pathology and Immunology, Washington University School of Medicine, Saint Louis, MO 63110, USA. ²Interdepartmental Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT 06511, USA. ³Department of Pediatrics, Washington University School of Medicine, Saint Louis, MO 63110, USA. ⁴Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA. ⁵Department of Molecular Microbiology, Center for Infectious Disease Research, Washington University School of Medicine, Saint Louis, MO 63110, USA. ⁶Washington University Center for Cellular Imaging, Washington University School of Medicine, Saint Louis, MO 63110, USA. ⁷Departments of Cell Biology and Physiology and Neuroscience, Washington University School of Medicine, Saint Louis, MO 63110, USA. ⁸Department of Biomedical Engineering, Washington University in Saint Louis, Saint Louis, MO 63130, USA. ⁹McDonnell Genome Institute, Washington University School of Medicine, Saint Louis, MO 63110, USA. ¹⁰Department of Immunology, Mayo Clinic, Rochester, MN 55905, USA. ¹¹Department of Neurology, Washington University in Saint Louis, Saint Louis, MO 63110, USA. ¹²Department of Pathology, Yale School of Medicine, New Haven, CT 06520, USA.

*Corresponding author. Email: mcolonna@wustl.edu

†These authors contributed equally to this work.

Fig. 1. B cells represent a major immune cell type in mouse meninges and are capable of trafficking through meningeal lymphatics.

(A) Cartoon representing the structural organization of the meninges. (B) Representative flow cytometry plot showing the proportion of B cells within the overall CD45⁺ population in mouse dura (average of $n = 4$ mice, data generated from a single experiment). (C) Schematic depiction of the experimental approach to perform in vivo two-photon imaging in the subdural space of CD19-Tomato mice. (D) Representative two-photon image of extravascular B cells in CD19-Tomato mouse meninges (scale bar: 50 μm). Images on the right are enlarged from box. (E) Two-photon time-lapse imaging in the meninges of a CD19-Tomato mouse. Intra-vascular B cells: yellow arrowheads; extravascular B cells: white arrowheads (scale bar: 20 μm). (F) Schematic depiction of the experimental approach (ICM: intracisterna magna). (G) Flow cytometry analysis of donor (CD19-Tomato)-derived B cells in inguinal lymph nodes (iLNs) and cervical lymph nodes (cLNs) 24 hours after injection. (H) Frequency and absolute number of donor-derived B cells in iLNs and cLNs (mean \pm SEM; $n = 4$ mice; Mann-Whitney U test, $*P < 0.05$; data generated from a single experiment). (I) Confocal images of donor (CD19-



Tomato) B cells trafficking through the dura lymphatics 24 hours after injection (representative of $n = 4$ mice, data generated from a single experiment) [scale bars: 50 μm (left) and 20 μm (right)]. (J). Confocal image of donor (CD19-Tomato)-derived B cells in cLNs 24 hours after injection (representative of $n = 4$ mice, data generated from a single experiment) [scale bars: 100 μm (left) and 50 μm (right)].

Fig. 2. Mouse meninges harbor a heterogeneous B cell population that encompasses multiple developmental stages.

(A) Uniform manifold approximation and projection (UMAP) of CD45⁺ brain cells from a deposited scRNA-seq dataset (GSE98969). Normalized expression for *Ighm*, *Igkc*, and *Iglc2* was color coded for transcript counts. The cell cluster highly enriched for the gene set was inferred as B cells. (B) Enrichment for transcripts linked to different B cell maturation stages. Pan-B cell: *Cd79b* and *Cd19*; mature B cell: *H2-Aa* and *Ms4a1*; and immature B cell: *Rag1* and *Cd93*.

(C) Flow cytometry analysis of B cells from the BM, brain, dura, blood, and spleens of C57BL/6 mice. B cells were gated as CD45⁺CD19⁺ and lineage (CD3, CD11b, F4/80, and Gr-1)-negative (left column) and were further divided into early (B220^{lo}CD43^{hi}), late (B220^{lo}CD43^{lo}), and mature (B220^{hi}CD43⁺) subset (middle column). CD93

and IgM staining is shown for the three subsets (right column). (D) Flow cytometry analysis of *Rag1*^{-/-} mice. B cells derived from BM, brain, and dura were gated as CD19⁺ and lineage (CD3, CD11b, F4/80, and Gr-1)-negative (left column) and were further divided into early (B220^{lo}CD43^{hi}), late (B220^{lo}CD43^{lo}), and mature (B220^{hi}CD43⁺) subsets (right column). (E) Quantification of B cell subsets in BM, brain and dura of C57BL/6 and *Rag1*^{-/-} mice (*n* = 8 and 7 mice, respectively; data generated from two independent experiments). (F) Representative confocal images of the dura mater from *Rag1*^{GFP} mice (scale bar: 100 μm). (G) Representative confocal image of IgM⁺ (yellow arrowhead) and IgM⁻ B cells (red arrowhead) in CD19-Tomato mice [scale bars: 1 mm (left) and 50 μm (right)].

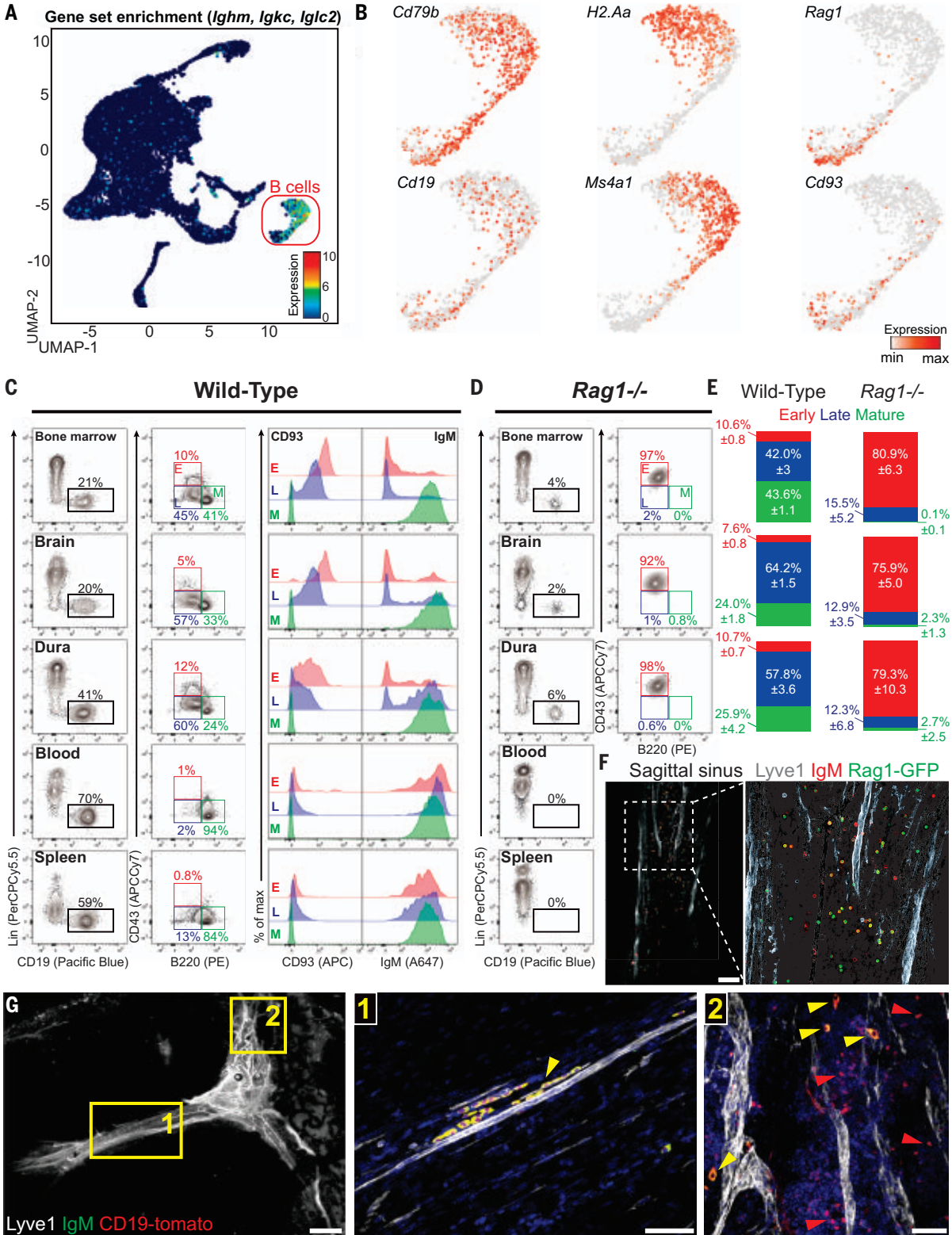
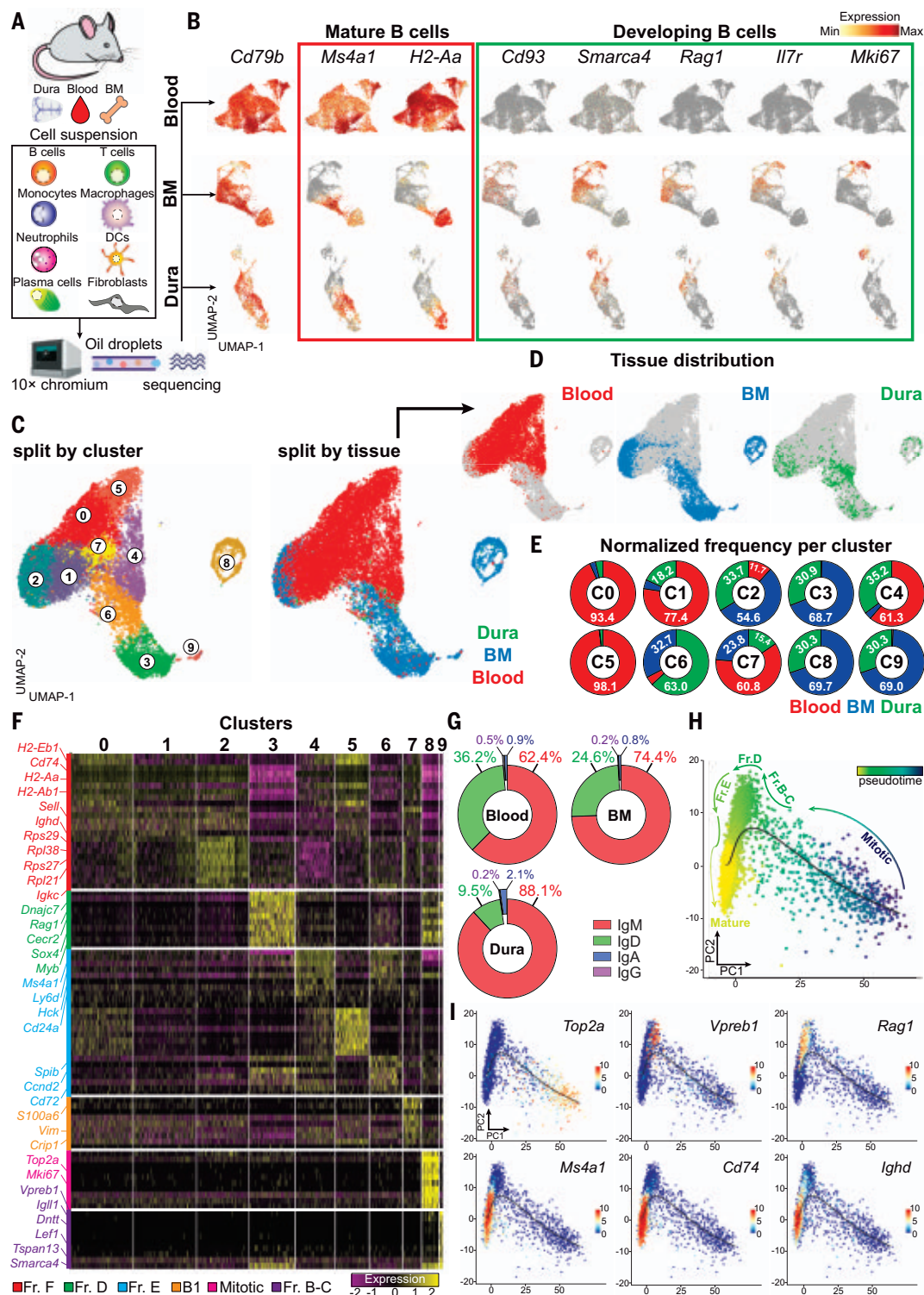


Fig. 3. scRNA-seq analysis reveals a similar transcriptomic pattern in dura and BM B cells.

(A) Schematic depiction of the experimental design related to scRNAseq. (B) Expression of featured genes denoting mature and developing B cells in blood, BM, and dura. (C) UMAP of 13,281 B cells aggregated from blood, BM, and dura collected from three C57BL6 mice and colored by cluster (left) or tissue origin (right) (data generated from a single experiment). (D) UMAP plots split by tissue showing the distribution of blood, BM, and dura B cells. (E) Proportional contribution of the three tissues (blood, BM, and dura) to these 10 B cell clusters. (F) Gene expression heatmap of the top 10 signature genes per cluster. (G) Frequency of C-region usage per tissue determined by scBCRseq. (H) Developmental trajectory displayed on PCA (principal component analysis) map and colored by slingshot pseudotime. (I) Enrichment of key transcripts differentially expressed throughout B cell maturation.



similarly distributed into these three subsets. In both compartments, early B cells were IgM⁺CD93⁺, mature B cells were IgM⁺CD93⁺, and late B cells had an intermediate phenotype. By contrast, most blood and spleen B cells exhibited a mature phenotype. We then repeated this analysis in *Rag1*^{-/-} mice, which lack mature B cells (Fig. 2D), and found a

population of B220^{lo}CD43^{hi} B cells in the BM, brain, and dura, whereas the blood and spleen completely lacked B cells. Thus, the composition of meningeal B cells closely resembles that of the BM, both in wild-type and *Rag1*^{-/-} mice (Fig. 2E). As a validation, we imaged the dura mater from *Rag1*^{GFP} knock-in mice (14) and found *Rag1*-expressing B cells,

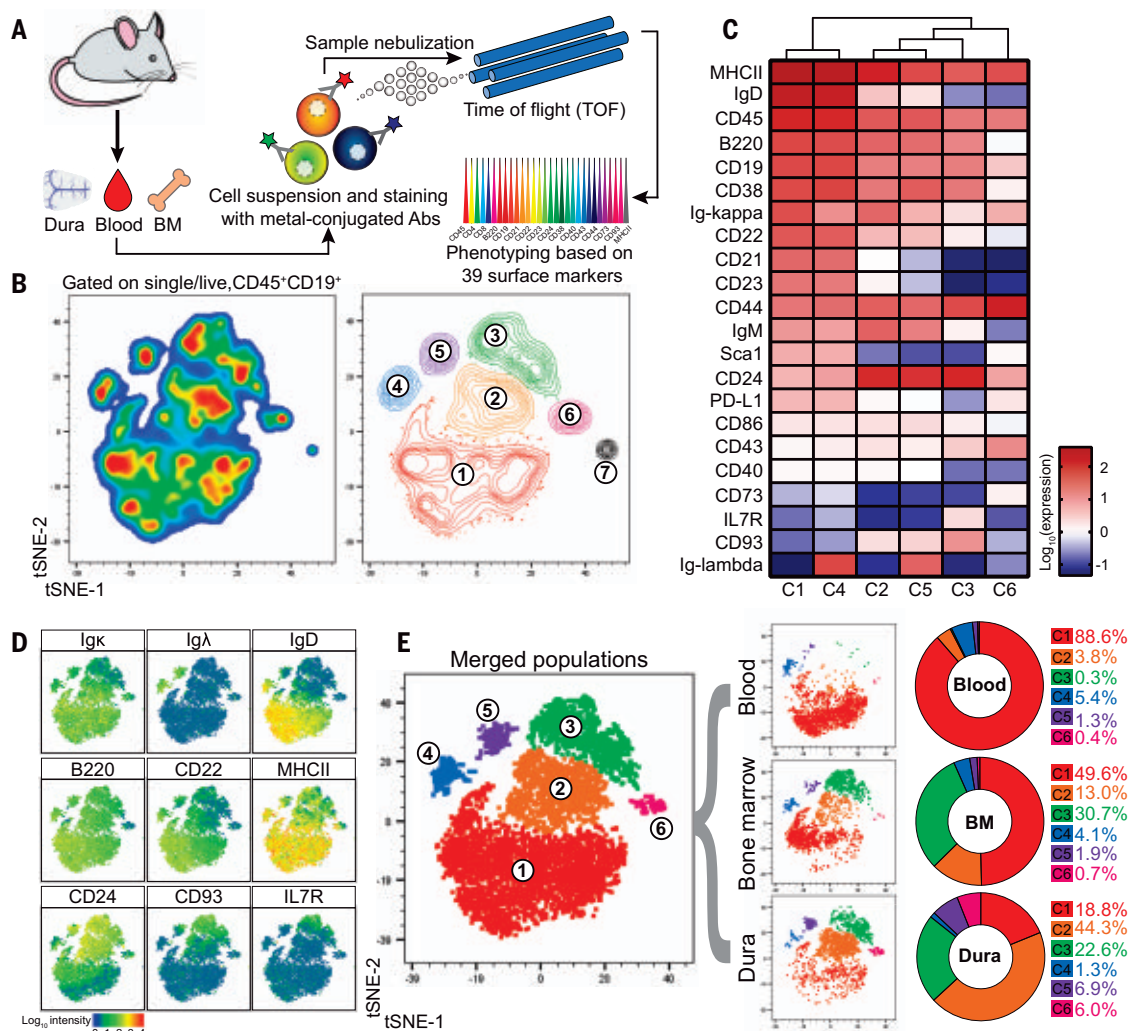
preferentially located along the sagittal sinus (Fig. 2F and fig. S3A). We also performed confocal imaging of the dura mater from CD19-Tomato mice stained for immunoglobulin M (IgM) and found distinct subsets of IgM⁺ and IgM⁻ B cells (Fig. 2G). Immunofluorescent imaging of CD93, Ki67, GL7 (markers of early B cells), and CD20 (marker of mature B cells) in

Fig. 4. Mass cytometry confirms developmental heterogeneity of dura B cells. (A) Schematic depiction of the experimental design related to cytometry by time of flight (CyTOF) analysis.

(B) Unsupervised clustering by means of t-distributed stochastic neighbor embedding (t-SNE) of 6000 live singlet CD45⁺CD19⁺ B cells obtained from concatenation of three tissues (blood, BM, and dura) collected from three C57BL/6 mice (data generated from a single experiment). Cells are displayed by pseudo-color (left) and contour plots (right).

(C) Heatmap showing the surface protein expression of selected B cell markers among the six B cell clusters (cluster 7 was arbitrarily excluded because of low cell number). (D) Staining enrichment of representative surface markers for mature and developing B cells.

(E) Merged t-SNE plot colored by cluster (left) and t-SNE plot split by tissue (middle). Frequency of each cluster in blood, BM, and dura (right).



the dura mater confirmed that early B cells are present in mouse meninges under homeostasis (fig. S3B). Lastly, we analyzed the expression of various markers differentially expressed across B cell development and confirmed that BM and meninges harbor overlapping B cell phenotypes (fig. S4).

Single-cell analyses resolve multiple developmental stages of dura B cells

To gain in-depth insight into the diversity of meningeal B cells, we performed scRNA-seq of cells isolated from dura mater, blood, and BM. We analyzed unsorted cell suspensions to minimize experimental artifacts due to cell stress and to include both immune and non-immune cells (Fig. 3A). Cells that passed the quality control underwent unsupervised clustering and were displayed on the Uniform Manifold Approximation and Projection (UMAP) space (figs. S5 to S7). In the dura mater, we resolved clusters of neutrophils (43.7%), B cells (26.3%), T cells (3.2% double negative, 2.8% CD8⁺, and 1.7% CD4⁺), monocytes (2.5% Ly6C⁺, 1.2% Ly6C⁻), natural killer cells (1.6%), macro-

phages (1.3%), mast cells (1.1%), plasmacytoid dendritic cells (pDCs) (1%), classical dendritic cells (cDCs) (0.9%), type 2 innate lymphoid cells (ILC2) (0.4%), and plasma cells (0.3%) (fig. S5). Dura fibroblasts (5.7%) were identified from the expression of *Col1a* and *Mgp*, as recently reported (15). Focusing on the B cell compartment, we found multiple B cell clusters in all analyzed tissues. Indeed, all B cells in the blood exhibited a rather homogeneous expression of the mature markers *Ms4a1* and *H2-Aa*, whereas the early B cell markers *Cd93*, *Smarca4*, *Rag1*, *IL7r*, and *Mki67* were not detected (Fig. 3B). The diversity in blood B cells was mostly driven by the expression of κ or λ light chains (fig. S7). By contrast, B cells in the BM and dura exhibited a nonoverlapping enrichment for both mature and early markers (Fig. 3B), confirming the presence of multiple stages of B cell development in these niches.

Next, we reclustered all B cells from the three tissues on a single UMAP space (Fig. 3C) and found that B cells from the BM and dura occupied overlapping territories, whereas most of the blood B cells clustered separately (Fig.

3D). This analysis yielded 10 different clusters (C0 to C9). C0 and C5 were almost entirely (>90%) formed by blood B cells; C3, C6, C8, and C9 contained B cells from both BM and dura but were largely depleted of blood B cells (<5%); C1 and C4 contained B cells from both blood and dura but were depleted of BM B cells (<5%); and C2 and C7 contained B cells from all compartments (Fig. 3E). To determine the identity of each cluster, we performed a differential gene expression analysis (cluster versus total; cutoff = $\log_2FC > 0.5$ where FC is fold change) (Fig. 3F and table S1), followed by enrichment analysis of Gene Ontology (GO) biological process (fig. S8, A and B). From these analyses, we determined that blood contained primarily mature naïve B cells (C0, C1, C2, C5; fraction F), a small subset of immature B cells (C4; fraction E), and B1b cells (C7). By contrast, BM and dura contained mature naïve B cells (C2; fraction F), immature B cells (C6; fraction E), pre-B cells (C3; fraction D), pro-B cells (C9; fraction B-C), and mitotic B cells (C8). B cell receptor sequencing (BCR-seq) analysis showed that 88.1% of dura B cells with

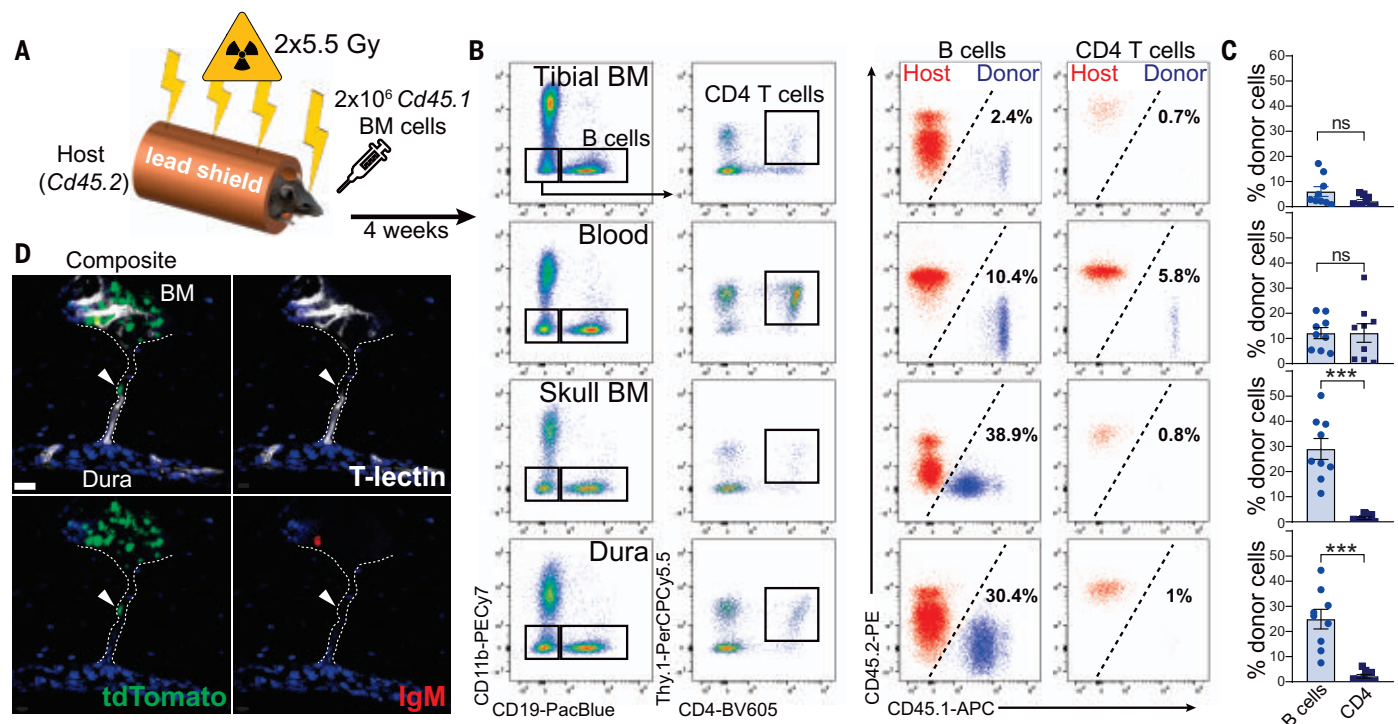


Fig. 5. Skull BM chimeras demonstrate that dura B cells originate from the calvaria. (A) Schematic depiction of the experimental design of calvaria BM transplantation. (B) Representative flow cytometry plots of B cells (gated on CD19⁺CD11b⁺ cells) and CD4 T cells (gated on CD19⁺CD11b⁺Thy1.1⁺CD4⁺ cells) from multiple compartments in recipient mice (left). The percentages of donor-derived (CD45.1⁺) B and CD4

T cells are shown (right). (C) Frequency of donor-derived B cells and CD4 T cells per compartment (mean ± SEM; n=9 mice; Mann-Whitney U test ***P<0.001; data generated from two independent experiments). (D) Representative confocal image of an IgM⁺ B cell trafficking from the calvarial BM toward meninges through a skull vascular channel (white arrowhead) (scale bar: 20 μm).

productive V(D)J rearrangement were IgM type, again confirming their early stage of development (Fig. 3G). Finally, we performed pseudotime analysis (Fig. 3H) and identified a developmental trajectory that originated from mitotic B cells, then traversed pro-B and pre-B intermediate stages, before eventually giving rise to mature B cells (Fig. 3I). To empirically test this model, we performed bromo-deoxyuridine (BrdU) pulse-chase experiments focusing on BM and dura B cells (fig. S8C). At 24 hours post-BrdU pulse, early B cells exhibited the maximal incorporation of BrdU. At 3 days, the percentage of BrdU⁺ early B cells was sharply reduced, but it increased in both late and mature B cell populations. At 5 days, almost no BrdU⁺ B cells could be detected, presumably because BrdU-labeled cells underwent apoptosis, further differentiated, or were progressively diluted by newly generated B cells. Thus, the molecular signature of dura B cells highlights a stepwise developmental continuum, comprehensive of multiple intermediate phenotypes that are normally found in the BM but not in the periphery.

To validate the scRNA-seq data, we performed a mass cytometry study of dura, blood, and BM B cells (Fig. 4A and fig. S9). Unsupervised clustering of the CD45⁺CD19⁺ popula-

tion using t-distributed stochastic neighbor embedding (t-SNE) identified seven distinct clusters (C1 to C7) (Fig. 4B and table S2). We analyzed clusters C1 to C6 in detail (Fig. 4, C and D), whereas C7 was excluded from further analysis because it contained only 17 cells. C1 and C4 exhibited a typical mature phenotype (MHC-II⁺IgM⁺IgD⁺) but differentially expressed κ or λ light chains. C2 and C5 had an immature B cell phenotype (IgD⁺CD22⁺CD21⁺CD23⁺CD24⁺), whereas κ or λ light chains were differentially expressed. C3 had an early B cell phenotype (IgD⁺CD21⁺CD23⁺IgM⁺CD24⁺CD43⁺IL7R⁺CD93⁺). C6 was negative or low for both naïve B cell markers and pre-B markers but expressed CD43, CD44, CD73, and PD-L1, indicating an activated phenotype (16). Next, we determined the distribution of B cells from each compartment among the six clusters (Fig. 4E). Blood B cells were maximally enriched in C1 (88.6%), BM B cells were distributed among all clusters (C1= 49.6%, C2 = 13%, C3 = 30.7%), and dura B cells were primarily concentrated in C2 (44.3%), with substantial fractions in C1 (18.8%) and C3 (22.6%). Blood and BM were equally represented in C4 (5.4 and 4.1%, respectively), whereas the dura was more abundant than either blood or BM in C5 (6.9% versus 1.3 and 1.9%). Thus, mass cytometry analysis confirms

that the dura mater contains a large proportion of developing B cells.

Meningeal B cells mostly originate from the calvaria

Because only a few immature B cells were present in the blood, it was unlikely that developing B cells in the dura originated from the systemic circulation. We hypothesized that dura B cells originated in the calvarial BM. Indeed, the caudal region of the skull harbored a large hematopoietic niche (fig. S10, A and B), containing a variety of immune cells, including IL7R⁺ early B cells and LSK stem cells (fig. S10C). Electron microscopy (EM) and confocal imaging highlighted a complex cellular system in this area, containing both myeloid and lymphoid cells (fig. S10, D and E). To test our hypothesis, we performed an atypical BM transplantation experiment. During irradiation, the mouse body was protected with a lead shield, leaving only the head fully exposed (Fig. 5A). Head-irradiated Cd45.2 mice received Cd45.1 BM cells and 4 weeks later, the percentage of B cell chimerism was assessed in different tissues (Fig. 5, B and C). In both tibial BM and peripheral blood, only a negligible percentage of donor-derived B cells was found (6 and

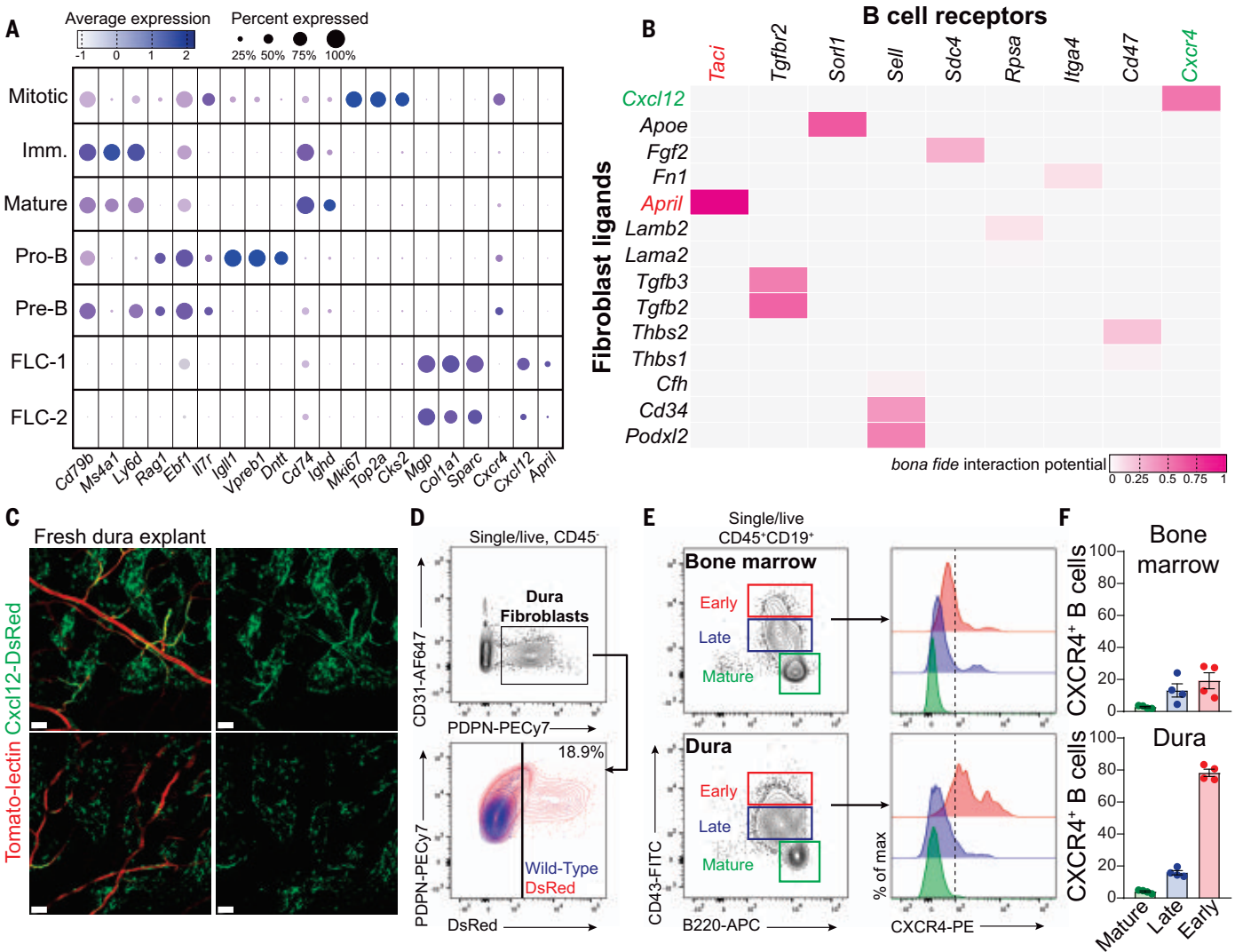


Fig. 6. Ligand-receptor interactions between CXCL12⁺ fibroblasts and CXCR4⁺ early B cells in the dura. (A) Dot plot presenting the expression level of featured genes in fibroblasts and B cells at different maturation stages. (B) NicheNet analysis displaying the putative interaction score between fibroblasts-derived ligands and B cell receptors. (C) Representative confocal images of freshly prepared dura explant from a *Cxcl12*^{DsRed} reporter mouse

(tomato lectin 488 injected i.v. before perfusion) (scale bar: 50 μ m). (D) Flow cytometry analysis of dura fibroblasts from *Cxcl12*^{DsRed} reporter mice (representative of three mice; data generated from a single experiment). (E) Flow cytometry analysis of CXCR4 expression in early, late, and mature B cells from dura and BM. (F) Frequency of CXCR4⁺ B cells in the three B cells subsets from dura and BM (mean \pm SEM; $n = 4$ mice; data generated from a single experiment).

12.1% respectively), whereas a large population of donor-derived B cells was present in skull BM and dura (29 and 24.9%, respectively). By contrast, almost no donor-derived CD4⁺ T cells were found in these compartments. We performed the reverse experiment, using a *Cd45.1* host, transplanted with *Cd45.2* BM cells and obtained comparable results (fig. S11A). To determine whether circulating B cells may potentially contribute to the pool of meningeal B cells under homeostasis, we performed a parabiosis experiment using wild-type and CD19-Tomato mice. After 4 weeks of shared circulation, the percentage of tdTomato-positive B cells was assessed in wild-type mice. A large population of parabiont-derived B cells

was present in the blood and spleen (39.2 and 36.4%, respectively), whereas a minor infiltration could be found in brain and dura (11.3 and 8.0%, respectively) (fig. S11, B and C). Comparable results were obtained by adoptive transfer of CD19-Tomato splenocytes into wild-type recipient mice (fig. S11, D and E). Thus, most meningeal B cells appear to originate in the calvarial BM and not from the circulation. It has been recently shown that vascular channels form a direct communication between the calvaria and the meningeal space, thus allowing migration of immune cells independently of the systemic circulation (4, 6). We performed confocal imaging of skull cryosections after immunofluorescent staining for

endothelial markers (tomato lectin, CD34, and CD31). B cells were visualized by means of the CD19-Tomato reporter. In agreement with these studies, we found vascular channels through the inner skull bone (fig. S12A). As a complementary technique, we performed x-ray tomography on intact skull specimens impregnated with radiopaque metals. The resulting three-dimensional (3D) image displayed in z series reveals the presence of radiolucent areas (seemingly vessels) running through the calvaria and opening at the base of the sagittal sinus (fig. S12B and movie S2). Finally, we provide evidence of IgM⁺ B cells trafficking through these channels, confirming that calvaria-derived B cells reach the meningeal compartment at an early stage of

their development (Fig. 3D and fig. S12, C and D).

Dural fibroblasts produce crucial factors for B cell development

Notably, one of the top signature genes of fibroblast-like cells (FLCs) in the dura was *Cxcl12* (17), whereas dural early B cells expressed *Cxcr4* (Fig. 6A). NicheNet analysis highlighted a network of molecular communications between FLC and dura B cells. Among the several ligands detected in dural fibroblasts, *Cxcl12*, *April*, *Tgfb2*, and *ApoE* were assigned a high probability score (Fig. 6B). *Cxcl12* expression in dural FLCs was confirmed by imaging and flow cytometry using *Cxcl12*^{DsRed} reporter mice (Fig. 6, C and D). Surface expression of CXCR4 in early B cells was also confirmed by flow cytometry (Fig. 6E). The majority of early B cells in the dura had CXCR4 expressed on the cell surface (Fig. 6F), suggesting that calvaria-derived B cells may require this receptor for their homing to the meningeal compartment. Finally, confocal imaging demonstrated B cells in close contact to *Cxcl12*-expressing FLCs in the sagittal sinus (fig. S13). As the CXCL12–CXCR4 axis is crucial for B cell development, we propose that dural FLCs may support the survival and differentiation of early B cells derived from the calvarial BM.

ABCs and plasma cells accumulate in mouse meninges during aging

Lymphocytes exhibit clonal and phenotypical imbalance during aging (18, 19). Indeed, we found a marked increase in the total number of B cells in the dura of aged mice (fig. S14, A and B). We then investigated the impact of aging on meningeal B cells by comparing young (8- to 12-week-old) and aged (20- to 25-month-old) mice by scRNA-seq and scBCR-seq (Fig. 7A). An accumulation of clonal B cells was found in the aged dura (fig. S14C), as well as an increased clonal overlap with blood B cells (fig. S14D), suggesting infiltration from the periphery. All B cell clusters in the dura were equally represented in young and aged mice, except for one cluster that was almost entirely (96.1%) derived from aged mice (Fig. 7, B and C) and was therefore annotated as “age-associated B cells” (ABCs). Differential gene expression analysis between ABCs and mature B cells ($\log_2FC > 0.5$; adjusted $P < 0.01$) revealed 105 differentially expressed genes (table S3). The top up-regulated transcripts included *ApoE*, *Ly6a*, *Ighm*, *Igkc*, *Cd2*, *Lgals1*, *Zbtb20*, and *Syk*, whereas the top down-regulated transcripts contained *Fcer2a*, *Cr2*, *Cd55*, *Sell*, and *Ehfl* (Fig. 7, D and E). Using flow cytometry, we confirmed a significant expansion of B220^{hi}CD23[−]CD25⁺Sca1⁺ ABCs in the dura of aged mice (Fig. 7, F and G). Consistent with the scRNA-seq data, this cell population also

expressed increased amounts of Syk and ApoE protein (Fig. 7H and fig. S14E). Furthermore, dura ABCs had increased numbers of *Ighm* heavy-chain transcripts compared to mature B cells (Fig. 7I). scBCR-seq further confirmed increased *Ighm*, and reduced *Ighd*, C-region usage in ABCs (Fig. 7J). Additionally, ABCs exhibited greater similarity in their V gene usage profiles (fig. S14F), as well as accumulation of somatic mutations in the BCRs of cells expressing the *Ighm* heavy chain (fig. S14G). Thus, dural ABCs are antigen-experienced B cells. Using these BCR data, we determined the percentage of B cells in clones that were shared between dura and blood (Fig. 8A). We found that dural B cells had no clonal overlap with circulating B cells in young mice, whereas a minor overlap could be observed in aged mice. By contrast, many dural ABCs were members of clones shared with the blood (30.8%), suggesting that these cells had trafficked from the periphery.

Next, we focused on dural plasma cells (fig. S15, A and B), which also appeared expanded in the aged dura (Fig. 7C). Differential gene expression analysis revealed that the expression of Ig isotypes was markedly shifted between the two age groups. In young mice, dural plasma cells were predominantly IgA⁺ but became mostly IgM⁺ in aged mice (fig. S15C). By confocal imaging, we surveyed plasma cells located along the sagittal sinus (fig. S15D) and confirmed the massive increase in the number of these cells in aged mice (fig. S15E), especially IgG and IgM types, whereas the number of IgA plasma cells was unchanged (fig. S14F). Lastly, we assessed the redundancy of plasma cell clones detected by scBCR-seq to gain insight into their origin (Fig. 8B). Although only a few plasma cell clones could be detected in the dura mater of young mice, we found an important clonal overlap with blood, indicating that plasma cells (mostly IgA) from young mice infiltrate the dura from the periphery, as recently demonstrated (20). By contrast, plasma cell clones in the aged dura (mostly IgM) exhibited a negligible overlap with the blood repertoire, indicating that most of these cells were not derived from the blood. Notably, we found a clonal overlap (15.4%) between dural plasma cells and dural ABCs, suggesting that, in aged mice, some dural ABCs may locally undergo terminal differentiation into IgM-secreting plasma cells. The top 10 most expanded clones detected in the aged dura specifically mapped onto both ABC and plasma cell clusters (Fig. 8C), supporting the possibility of a clonal relatedness between these two populations.

Discussion

Recent studies have shed light on the origin and phenotypic diversity of the myeloid cell landscape at the CNS borders (8, 21–24). Yet,

little is known about B cells resident at CNS interfaces. Using complementary techniques, we demonstrated that meningeal B cells encompass multiple stages of the development. These B cells originate in the calvaria and infiltrate the meninges through a network of channels uncoupled from the systemic circulation. Early B cells may complete their maturation within the meningeal compartment, wherein dura fibroblasts can provide critical molecules for B cell development, such as CXCL12 (25, 26). Using *Nur77*^{GFP} mice, in which green fluorescent protein (GFP) expression reflects BCR engagement by self-antigens (27), we showed that transitional and mature meningeal B cells are equipped with a functional BCR, evinced by GFP expression in these populations (fig. S16, A to C). Thus, the calvaria may provide an early and rapid source of B cells that develop in the dura, resulting in negative selection of B cells with high affinity for local self-epitopes. In support of this hypothesis, transgenic mice carrying the IgH chain of the mAb 8.18C5 specific for myelin oligodendrocyte glycoprotein (MOG) (28) showed a significant reduction of MOG-specific B cells in dura compared to tibial BM, suggesting that negative selection may occur locally (fig. S16, D to F). The calvaria was recently described as a source of meningeal neutrophils (4), and a companion paper shows that calvarial BM supplies the CNS with myeloid cells, both under homeostasis and inflammation (29). We further integrate these findings showing that meningeal B cells share a similar origin.

In aged mice, we identified a distinct population of meningeal B cells corresponding to ABCs (30), which have been mostly found in the spleen (19, 31–33), but never in the CNS. scBCR-seq indicates that dural ABCs are antigen-experienced B cells infiltrated from the systemic circulation. Additionally, the aged dura mater features a robust expansion of IgM⁺ and, to a lesser extent, IgG⁺ plasma cells. By contrast, meningeal IgA⁺ plasma cells, which derive from the gut (20, 34), appeared unaffected by aging. The infiltration of blood-borne ABCs, as well as the accumulation of plasma cells within the meninges, may endanger the immune-privileged CNS environment during aging. Because B cells are key players in neuroinflammatory and autoimmune disorders (35, 36), these findings may help better explain the origin of self-reactive B cells in these pathologic conditions.

Materials and methods

Animals

All mice used in this study were C57BL/6J housed under specific pathogen-free conditions at Washington University School of Medicine animal facility. Homozygous *Cd19*^{Cre} mice were

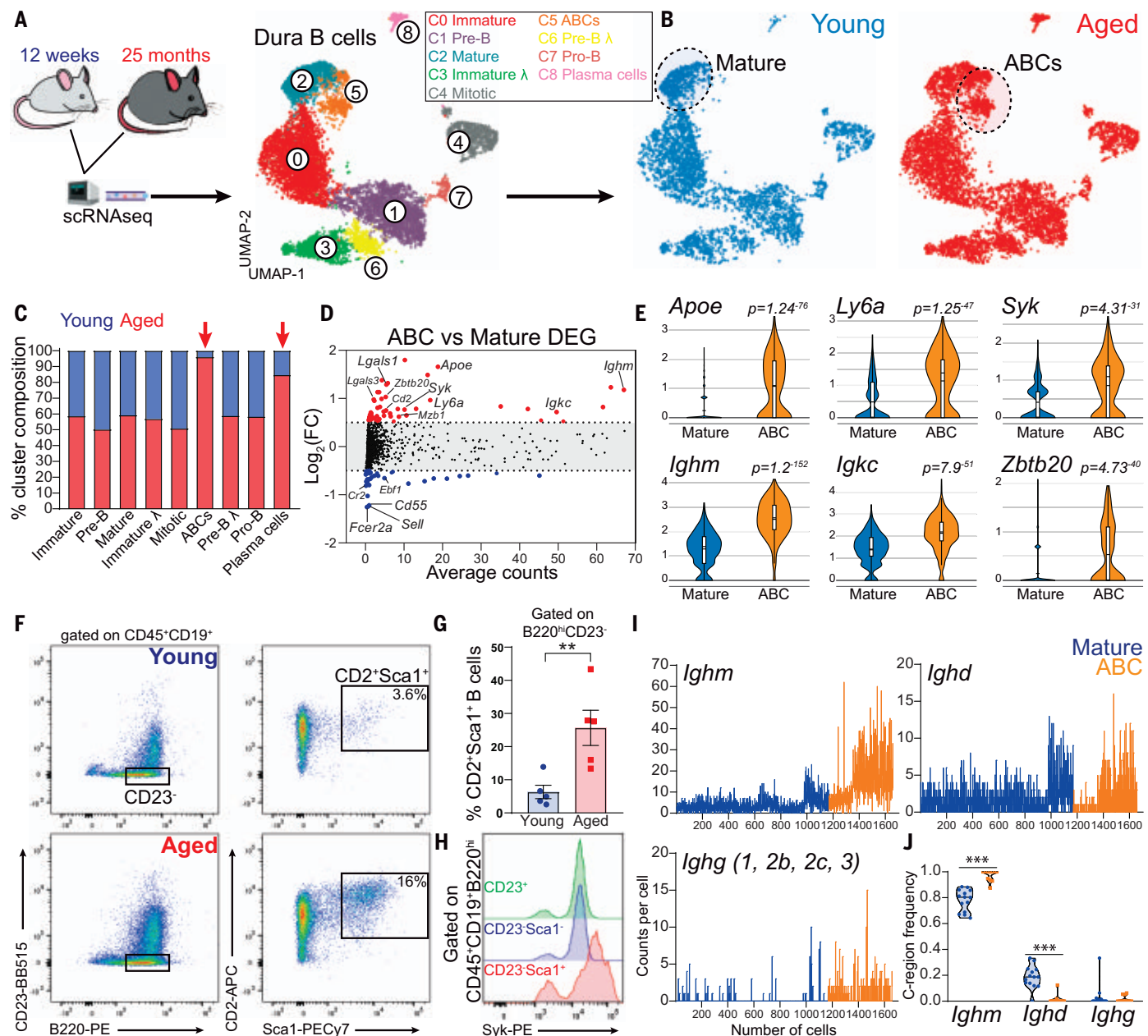


Fig. 7. Age-associated B cells disseminate throughout the dura of aged mice. (A) Schematic depiction of the experimental design of scRNAseq comparing young and aged mice. UMAP of 9352 B cells aggregated from seven 12-week-old and seven 25-month-old C57BL/6 female mice (data generated from two independent experiments). (B) Distribution of B cells from young and aged mice. (C) Contribution of young versus aged mice to each cluster. (D) Differential gene expression analysis of ABCs versus mature B cells. (E) Violin plots showing top up-regulated genes in ABCs compared to mature B cells. (F) Dura ABCs gated as the B220^{hi}CD23⁺CD2⁺Sca1⁺ cells. (G) ABC

population is significantly increased in the dura of aged mice compared to young mice (mean ± SEM; $n = 5$ mice; unpaired Student's t test, $^{**}P < 0.01$; data generated from two independent experiments). (H) Flow cytometry histogram showing increased amounts of Syk protein in dura ABCs (representative of three 18-month-old female mice; data generated from a single experiment). (I) Ig heavy-chain transcript counts per cell in ABCs and mature B cells. (J) Frequency of heavy-chain usage determined by BCRseq in ABCs and mature B cells (violin plot; $n = 11$ to 14 mice per group; two-way ANOVA and Bonferroni post-hoc test $^{***}P < 0.001$).

purchased from the Jackson Laboratory (JAX stock #006785) and mated with homozygous *Rosa26-STOP^{tdTomato}* (Ai14) mice bred in house. Heterozygous *Cxcl12^{DSRed}* mice were purchased from the Jackson Laboratory (JAX stock #022458) and mated with C57BL/6J mice bred in house. *Nur77^{GFP}* mice were provided

by Dr. P. Allen (Department of Pathology and Immunology, Washington University School of Medicine). *IGH^{MOG}* mice were provided by Dr. G.F. Wu (Department of Neurology, Washington University in Saint Louis). All experiments involving laboratory animals were performed under the approval of the

Institutional Animal Care and Use Committee at Washington University in St. Louis (protocol #19-0981). For flow cytometry and imaging experiments, both sexes were used and equally distributed among the experimental groups. For CyTOF, scRNAseq, and parabiosis experiments, females only were

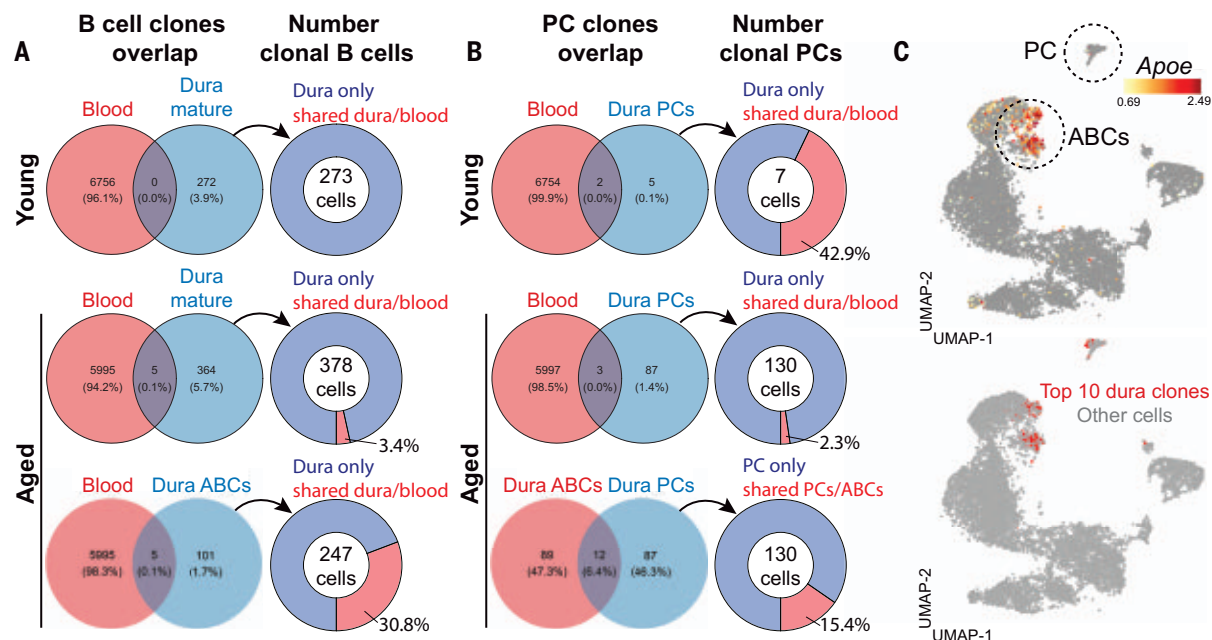


Fig. 8. Age-associated B cells in the aged dura originate from the periphery. (A) Venn diagrams representing the proportion of detected B cell clones shared between dura mature B cells and blood in young and aged mice, or between dura ABCs and blood. The connected pie chart represents the percentage of dura B cells (mature or ABCs) belonging to shared clonotypes. (B) Venn diagrams representing the proportion of detected plasma cell (PC) clones shared between dura and blood in young and aged mice, or between dura plasma cells and dura ABCs. The connected pie chart represents the percentage of dura plasma cells belonging to shared clonotypes. (C) UMAP showing the distribution of the 10 most frequent clones in the dura.

used to minimize biological variability. For studies on young-adult mice, 8- to 12-week-old mice were used. For studies on aged mice, 20-25-month-old mice were used. On the day of sacrifice, mice received a lethal dose of ketamine-xylazine injected intraperitoneally (i.p.) (respectively, 300 mg and 30 mg per kilogram of body weight). After complete loss of the paw-pinch reflex, blood samples were collected by heart puncture. All other tissues analyzed in this study (dura, brain, spleen, and bone marrow) were collected upon perfusion with 30 ml of ice-cold phosphate-buffered saline (PBS). Parabiotic mice were prepared by the surgery core in the Hope Center for Neurological Disorders (Washington University, St. Louis, MO) as previously described (37).

Intravenous injection of CD19-Tomato splenocytes

Five-month-old CD19-Tomato mice were sacrificed with CO₂ and spleens were immediately collected into ice-cold PBS. Splenocytes were mechanically extracted by mashing the spleen on a 70- μ m strainer and collected in a 50-ml conical centrifuge tube. Samples were pelleted by centrifugation at 300g for 15 min at 4°C, followed by red blood cells lysis on ice for 2 min. Splenocytes were then washed in sterile PBS, pelleted by centrifugation at 300g for 5 min at 4°C, and resuspended in 1 ml of sterile PBS. Before counting, cells were filtered

through a 35- μ m sterile strainer. Three-month-old gender-matched C57BL/6 mice were used as recipients. Mice were anesthetized with a dose of 80 mg per kilogram ketamine and 10 mg per kilogram xylazine administered by i.p. injection. After complete loss of the paw-pinch reflex, mice received a retro-orbital injection of 2×10^7 freshly prepared splenocytes from CD19-Tomato mice. Recipient mice were sacrificed one week later and blood, spleen, brain, and dura were analyzed by flow cytometry.

Intra cisterna magna injection of CD19-Tomato splenocytes

A single CD19-Tomato mouse was perfused under sterile conditions and spleen was immediately collected into ice-cold PBS. Splenocytes were mechanically extracted by mashing the spleen on a 70- μ m strainer and collected in a 50-ml conical centrifuge tube. Samples were pelleted by centrifugation at 300g for 15 min at 4°C, followed by red blood cell lysis on ice for 2 min. Splenocytes were then washed, and resuspended in sterile PBS at a final concentration of 5×10^5 cells/ μ l. Injection of 2.5×10^6 CD19-Tomato splenocytes (5 μ l) in the cisterna magna of 3-month-old gender-matched C57BL/6 mice was performed as previously described (10). After injection, mice were sutured and monitored until they completely recovered. The dura mater, cervical lymph nodes, and inguinal lymph nodes were collected 24 hours after injection.

Sample preparation for flow cytometry or scRNAseq

Blood samples underwent red blood cell lysis at room temperature for 5 min. Hereafter, sample preparation was entirely carried out either on ice or at 4°C. No enzymatic digestion was used in this study. The dura mater was stripped off the inner skull surface using precision tweezers. Single-cell suspensions of brain, dura, and calvarial bone marrow were generated via mechanical dissociation using a dounce homogenizer. Cells were then filtered through a 70- μ m strainer and collected in a 50-ml conical centrifuge tube. Splenocytes were mechanically extracted by mashing the spleen on a 70- μ m strainer and collected in a 50-ml conical centrifuge tube. Tibial bone marrow was extracted from open tibias into a 1.5-ml microcentrifuge tube by centrifugation at top-speed for 20 s using a bench centrifuge. Spleen and bone marrow samples underwent red blood cells lysis on ice for 2 min. Samples were pelleted by centrifugation at 300g for 15 min at 4°C. Dura pellets were resuspended in 5 ml of 75% isotonic percoll, overlaid with 3 ml of PBS. Stromal cells and debris were depleted by centrifugation at 1000g for 30 min at 4°C (acceleration 0, break 1). The interface percoll-PBS was collected for analysis. Brain pellets were resuspended in 5 ml of 30% isotonic percoll, overlaid with 2 ml of PBS. Myelin was depleted by centrifugation at 1000g for 30 min

at 4°C (acceleration 0, break 1). The pellet was collected for analysis.

Flow cytometry analysis

Single-cell suspensions were washed in PBS followed by live/dead staining (Zombie Aqua or Zombie UV, Biolegend) at 1:1000 dilution for 15 min on ice. Fc-receptor blockade was performed using CD16/32 blocking antibody (clone 93, Biolegend; or clone 197, made in house) incubated 10 min on ice. Surface staining was always performed between 30 min and 1 hour on ice (complete antibody list in table S4). For staining of intracellular markers, fixation and permeabilization were performed using either the BD Cytofix/Cytoperm kit (BD Bioscience), or the Fcγ3/Transcription Factor Staining kit (eBioscience), according to the product instructions. For ApoE intracellular staining, 1 mg/ml of purified anti-ApoE (clone HJ6.3, kindly provided by David Holtzman) was conjugated with AlexaFluor647 and fixed/permeabilized cells were stained overnight (1:100 dilution) at 4°C. For staining of IGH^{MOG} BCR, B cells were incubated with 2 mg/ml of biotin-conjugated recombinant MOG (1:100 dilution) for 30 min on ice prior to surface markers staining. Flow cytometry analysis was performed on BD X20, BD LSR Fortessa, or BD Canto-II (BD Bioscience). Raw data were analyzed with FlowJo v10.

Skull bone marrow transplantation

Twelve-week-old C57BL/6J (*Cd45.2*) mice were anesthetized with a dose of 80 mg per kilogram ketamine and 10 mg per kilogram xylazine administered by i.p. injection. Mice were inserted into a 1-inch-thick lead shield (1 inch-thick Lead Vial Shield, 50 ml, Pinestar Technology) leaving only the head exposed. Each mouse received 11 Gy of gamma irradiation, split into two doses 4 hours apart. Immediately after the second dose, mice were reconstituted with 2×10^6 *Cd45.1* bone marrow cells injected intravenously (i.v.) Mice were left undisturbed in their home cage for the following 4 weeks. On the day of experiment, samples were processed as described above. The percentage of chimerism within different immune populations was determined by staining for CD45.2 and CD45.1.

Immunofluorescence staining

Brains, spinal cords, and skull caps were fixed in 4% paraformaldehyde (PFA) at 4°C overnight. Fixed specimens were then dehydrated in 30% sucrose solution for at least 48 hours and then cut into 60- to 100-μm-thick sections at the cryostat (Leica). Staining on free-floating sections was performed. Cryosections were blocked for 4 hours in PBS with 5% bovine serum albumin (BSA) and 0.5% Triton X-100. Primary antibody staining was performed in PBS with 1% BSA and 0.5% Triton

X-100 for 48 hours at 4°C. Secondary staining with fluorochrome-conjugated antibodies was performed at room temperature for 2 hours (complete antibody list in table S4). Immunostained sections were mounted on Superfrost glass slides (Fisher Scientific) and embedded in Prolong Glass anti-fade mounting media (ThermoFisher). Lymph nodes were fixed and dehydrated as above and sliced into 20 μm-thick sections at the cryostat. Sections were directly mounted on the Superfrost glass slides and air dried. Blocking was performed for 1 hour (5% BSA, 0.5% Triton X-100), and antibody staining for 24 hours at 4°C (1% BSA, 0.5% Triton X-100). For whole-mount preparation of dura mater, skull caps were fixed as above, and then decalcified in 0.5 M pH 8.0 EDTA solution (Corning) for 48 hours at 4°C. Blocking and staining were performed as above. The whole skull was mounted on Superfrost glass slides and embedded in Fluoromount-G anti-fade mounting media (Southern Biotech). Stained samples (both cryosections and whole mount preparations) were covered with 1.5H high-precision cover glass (Marienfeld Superior) and left to dry overnight at room temperature before imaging.

Confocal imaging

Confocal imaging of dura mater whole mounts preparations and brain cryosections was performed using a Zeiss LSM880 airyscan inverted confocal microscope equipped with a 34-channel GaAsP (gallium arsenide phosphide) detector. Cryosections were imaged with a 40×/1.4 oil-immersion objective. Dura whole mounts were imaged with a 40×/1.2 water-immersion objective. Images were acquired at 2048 × 2048-pixel resolution, 20- to 50-μm-thick *z*-stack *z*-step = 1 μm, line averaging = 2, using ZEN Black (ZEISS Efficient Navigation) software (Zeiss). When needed, tile-scan mode was used to generate large filed-of-view images. Confocal imaging of leptomeninges was performed on fresh unfixed brains from young adult *CD19^{Cre}:Rosa26^{tdTomato}* mice. Briefly, mice were anesthetized with a lethal dose of ketamine and xylazine administered by i.p. injection. After complete loss of the paw-pinch reflex, blood vessels were fluorescently labeled by retro-orbital injection of 100 μl of DyLight 488-labeled tomato lectin (Vector Laboratories, 1:1 dilution). Two minutes later, the mouse was decapitated, and the brain was immediately rinsed in ice-cold PBS and imaged using a Leica SP8 inverted confocal microscope. Images were acquired with a 25×/0.95 water-immersion objective, at 2048 × 2048- or 4096 × 4096-pixel resolution, 70 μm-thick *z*-stack, *z*-step = 5 μm, line and frame average=3, using Leica Application Suite, LAS X (Leica Microsystems). Maximal projections were rendered in Imaris V8.3 (Bitplane, Zurich, Switzerland).

Two-photon imaging

Young adult *CD19^{Cre}:Rosa26^{tdTomato}* mice were anesthetized with a dose of 80 mg per kilogram ketamine and 10 mg per kilogram xylazine administered by i.p. injection. After complete loss of the paw-pinch reflex, fur on the head and neck was shaved and the skin on the parietal skull bone was surgically removed. Skull bone was thinned using an electrical micro drill and the head was fixed on a metal holder to minimize movements during the live imaging. Immediately before imaging, blood vessels were fluorescently labeled by retro-orbital injection of 100 μl of DyLight 488-labeled tomato lectin (Vector Laboratories, 1:1 dilution). Time-lapse in vivo imaging was performed with a Leica SP8 2-photon imaging system. DyLight-488 and tdTomato were excited with Mai Tai Deepsee and Insight Deepsee lasers (Spectra Physics, Santa Clara Ca, USA) optimally tuned to 920 nm and 1050 nm respectively. Emission spectra were collected on ultrasensitive hybrid detectors as follows: >560 nm (tdTomato), 495-560 nm (DyLight 488), and <458 nm (second harmonic). The subdural space was identified using the second harmonic generation and the blood vessels as territory landmarks. Time-lapse imaging was performed with a 25×/0.95 water-immersion objective, at 1024 × 1024-pixel resolution, *z*-stacks = 40 μm, *z*-step = 2.5 μm, acquisition speed = 4 frames per min, imaging time = 30 min, using Leica Application Suite, LAS X (Leica Microsystems). Maximal projection and video editing were performed with Imaris V8.3 (Bitplane, Zurich, Switzerland).

Mass cytometry (CyTOF) analysis

Single-cell suspension from dura mater, tibial bone marrow and blood were obtained as described above. Cells were resuspended in Cy-FACS buffer (PBS, Rockland; 0.1% BSA, Sigma A3059; 0.02% NaN₃; 2mM EDTA). Fc-receptor blockade was performed using CD16/32 blocking antibody (clone 93, Biolegend, 1:100). Cells were stained for 1 hour on ice with the surface staining cocktail (complete antibody list in table S4). Viability staining was performed by resuspending cells in 2.5 μM Cisplatin for 1 min. Cells were then fixed with 4% PFA for 30 min on ice followed by overnight DNA staining with iridium-labeled intercalator (1:3000). Cells were acquired with the CyTOF2 mass cytometry system (Fluidigm) and analyzed using Cytobank. Immune cells were hierarchically gated as follows: nucleated cells, exclusion EQ calibration beads, live cells, single cells, and CD45⁺. Dimensionality reduction analysis was performed on CD45⁺ cells using the tSNE toolX in Cytobank to apply the Barnes-Hut implementation of the t-SNE algorithm. Unsupervised clustering of the total CD45⁺ cells was performed using the following parameters: iterations = 2000;

perplexity = 100; Theta = 0.5. This setup has been chosen as it demonstrated to produce the best separation of biologically meaningful populations in the three compartments. Next, the CD45⁺CD19⁺ population from the blood, BM, and dura were further analyzed with FlowJo v10. B cells populations from independent replicates were aggregated and down-sampled to 2000 single/live B cells per compartment. Finally, downsampled files were concatenated and reclustered by t-SNE algorithm using default settings (iterations = 1000; perplexity = 20; Theta = 0.5).

BrdU staining for cell proliferation

Mice received two i.p. injections 6 hours apart of 5 mg BrdU (Sigma) dissolved in PBS. Mice were randomized and sacrificed at multiple time points after injection (24 hours, 3 days, 5 days, and 7 days). Samples were prepared as described above, followed by BrdU staining using the BrdU flow Kit (BD Pharmingen). Briefly, after surface staining, cells were fixed with fix/perm BD buffer for 20 min on ice, followed by 10 min nuclear permeabilization with BD perm-plus buffer, and a second fixation for 5 min at room temperature. Fixed/permeabilized cells were incubated with deoxyribonuclease I (DNase I) (300 µg/ml) for 1 hour at 37°C, followed by intranuclear BrdU staining (Biolegend clone 3D4, 5 µl BrdU-PE per sample) for 30 min at room temperature. Flow cytometry analysis was performed on BD LSR Fortessa.

Transmission electron microscopy

After perfusion with ice-cold PBS, skull caps were removed and fixed with 4% paraformaldehyde (PFA), 0.1% glutaraldehyde (Polysciences Inc., Warrington, PA) in phosphate-buffer 0.1 M, pH 7.2 overnight at 4°C. Samples were decalcified in 0.5 M pH 8.0 EDTA solution (Corning) for 48 hours at 4°C and fixed again in 2% PFA, 2.5% glutaraldehyde in 100 mM sodium cacodylate buffer, pH 7.2 for 2 hours at room temperature. Samples were washed in sodium cacodylate buffer at room temperature and postfixed in 2% osmium tetroxide (Polysciences Inc.) for 2 hours at room temperature. Samples were then rinsed extensively in dH₂O prior to en bloc staining with 1% aqueous uranyl acetate (Ted Pella Inc., Redding, CA) for 1 hour at room temperature. After several rinses in dH₂O, samples were dehydrated in a graded series of ethanol and embedded in Eponate 12 resin (Ted Pella Inc.). Thick sections of 400 nm were cut with a Leica Ultracut UCT ultramicrotome (Leica Microsystems Inc., Bannockburn, IL) and stained with 1% toluidine blue, 2% sodium borate. Brightfield images were acquired on a Zeiss AxioObserver D1 inverted microscope (Carl Zeiss Inc. Thornwood, NY) equipped with an AxioCam 503 color camera. Areas of interest were identified, and images were

acquired with 20×/0.5 and 40×/1.2 objectives. Ultrathin sections of 95 nm from the areas of interest were then collected on Cu grids, stained with uranyl acetate and lead citrate, and viewed on a JEOL 1200 EX transmission electron microscope (JEOL USA Inc., Peabody, MA) equipped with an AMT 8-megapixel digital camera and AMT Image Capture Engine V602 software (Advanced Microscopy Techniques, Woburn, MA).

X-ray tomography

Upon perfusion, mouse cranium was fixed in 4% PFA overnight at 4°C and stained with 10% Lugols iodine solution (38) at 21°C for 5 days prior to imaging. Following incubation, samples were embedded in 2% agarose, and imaged using a Zeiss Versa 520 (Carl Zeiss Microscopy, White Plains, NY) using either a 0.4× flat panel detector or a 4× objective at 80 kV. Final tomograms were visualized, and resultant images were generated in Zeiss XM3DViewer v1.2.9. Z-projections were generated in ImageJ.

Single-cell RNA sequencing and single-cell BCR sequencing library preparation

For scRNAseq seven 10-week-old and seven 25-month-old female mice were used. Single-cell suspension from dura mater, tibial bone marrow and blood were obtained as described above. No sorting or enrichment was used as this experiment aimed to achieve a comprehensive transcriptomic representation of both immune and stromal cells. Samples were resuspended into low binding microcentrifuge tubes with PBS + 0.04% BSA at a final concentration of ~1000 cells/µl. Single cells gene-expression and V(D)J libraries were prepared by the McDonnell Genome Institute (MGI) at Washington University using the 5' Single Cell with V(D)J Enrichment Reagents Kit from 10x Genomics. Using the 10x Chromium VDJ + 5' Gene expression v1 platform, up to 17,500 cells were partitioned into nanoliter Gel-bead-in-Emulsions (GEMs) droplets. Each GEM underwent retro-transcriptase reaction to generate single-cell cDNA and received a unique 12-nt cell barcode and unique molecular identifier (UMI). GEM cDNA was amplified for 11 cycles before being purified using SPRIselect beads. For scRNAseq, 10 µl of purified cDNA was used to generate the library for RNA sequencing. Libraries were sequenced on a NovaSeq S4 (Illumina, 300 cycle kit) flow cell, targeting 50K read pairs per cell. For scBCRseq, 2 µl of purified cDNA was used to generate the sequencing of the V(D)J region. Sequencing was performed using a NovaSeq S4 (Illumina, 300 cycle kit) flow cell, targeting 5K read pairs per cell.

Single-cell RNAseq analysis

Cell Ranger Software Suite (v3.1.0) from 10X Genomics was used for sample demultiplex-

ing, barcode processing, and single-cell counting. Cell Ranger count was used to align samples to the reference genome GRCm38 (mm10), quantify reads, and filter reads and barcodes. Contamination by ambient RNA in droplets was normalized using SoupX (v1.4.5) (<https://github.com/constantAmateur/SoupX>). Furthermore, contaminating genes from erythrocytes and platelets (*Clu*, *Gng11*, *Gp1bb*, *Gp9*, *Hba-a1*, *Hba-a2*, *Hbb-b1*, *Hbb-b2*, *Nrgn*, *Pf4*, *Ppbbp*, and *Tubb1*) were filtered out from the dataset. The Seurat (v3.2.2) package in R was used for downstream analysis. For quality control, cells with mitochondrial content >20% were removed. Cells with low UMI and gene number per cell were filtered out. Cutoffs for UMI and gene number were determined on the basis of histograms showing cell density as a function of UMI per gene counts. For dura and bone marrow samples, cutoffs of >500 UMI and >250 genes were applied. For blood samples, cutoffs of >500 UMI and >1000 genes were applied. Genes expressed in fewer than 10 cells were removed from the dataset.

Samples from different tissues were clustered separately. For each tissue, data were normalized using the SCTransform method regressed on mitochondrial gene percentage and integrated using FindIntegrationAnchors function and Canonical Correlation Analysis (CCA) (39, 40). Principal component analysis was performed, and the top 40 principal components were selected for dimensionality reduction using the Uniform Manifold Approximation and Projection (UMAP) algorithm. For identification of marker genes and differential expression analysis, we used the MAST algorithm implemented through the Seurat R package (41). Marker genes were identified by comparing each cluster against all other clusters using the FindConservedMarkers function with default settings (log-fold change threshold of 0.25 and >10% cells expressing the gene) and age as the grouping variable. Cell clusters from each tissue were annotated based on marker gene expression and clusters consisting of doublets were manually removed. For further analyses of B cell heterogeneity, B cell clusters were extracted and re-clustered. For reclustering analyses, data were normalized using a scaling factor of 10,000 and mitochondrial read percentage was regressed with a negative binomial model. Principal component analysis was performed using the top 2000 most variable genes and UMAP analysis was performed using the top 10 principle components. Clustering was performed using a resolution of 0.4. For data visualization and figure preparation of scRNA-seq data BBrowser version 2.7.5 was used (42).

Single-cell pseudotime trajectories were inferred using diffusion map algorithms implemented through the R package destiny (43). Normalized expression values were used as

input for the generation of diffusion maps. Cells were ordered based on the first diffusion component. To further visualize B lineage differentiation, pseudotime was inferred using the slingshot R package (44). Gene Ontology analysis for biological processes enriched in different B cell populations was performed on the top 50 transcripts in each cluster using Metascape (<https://metascape.org>) (45). Receptor-ligand interactions were mined from the scRNA-seq data using the NicheNet algorithm (46). In brief, NicheNet analysis was performed using the marker genes for all clusters of dura B cell and dura fibroblast using default settings. After calculation of interaction scores between possible receptor-ligand combinations, pairs were filtered for those bona fide interactions that were documented in the literature and publicly available databases.

Single-cell BCR repertoire analysis

Sample demultiplexing and barcode processing was performed using the Cell Ranger Single-Cell Software Suite (10x Genomics). CellRanger-5.0.12 was used to align reads to the reference genome (vdj_GRCm38_alts_ensembl-5.0.0) and assemble BCRs. BCR sequencing data was processed using the Immcantation framework (imccantation.org) (47, 48). Germline V(D)J gene annotation was performed for all 10X Genomics BCR sequences using IgBLAST v1.16.0, with a curated set of Ig reference alleles specific for the C57BL/6 strain of *Mus musculus* (49). IgBLAST output was parsed using Change-O v0.4.6 (47). Additional quality control required sequences to be productively rearranged and have valid V and J gene annotations, consistent chain annotation (excluding sequences annotated with H chain V gene and L chain J gene), and a junction length that is a multiple of 3. Cells with exactly the same barcode and BCR sequences between different samples were excluded. Furthermore, only cells with exactly one H chain sequence paired with at least one L chain sequence were kept. After processing, there were 52,509 cells with paired H and L chains of which 19,447 had paired gene expression data.

B cell clonal lineages were inferred using hierarchical clustering with single linkage (50, 51). Cells were first partitioned based on common H and L chain V and J gene annotations and junction region lengths. Within each partition, cells whose H chain junction distance from each other were clustered as clones. This distance threshold was determined by manual inspection to identify the local minimum between the two modes of the bimodal distance-to-nearest distribution. The clones were further split based on the common L chain V and J gene.

Mutation frequency was calculated for cells by counting the number of nucleotide mis-

matches from the germline sequence in the H chain variable segment leading up to the CDR3. The calculation was performed using the *calcObservedMutations* function from *SHazaM* v1.0.2.

Statistical analysis

For flow cytometry and imaging data, graphs and statistics were produced using the GraphPad Prism 8 software package. Statistical difference between two groups was determined by either two-tailed unpaired Student's *t* test (when both groups passed the Kolmogorov-Smirnov normality test) or two-tailed Mann-Whitney *U* test (when at least one group did not pass the Kolmogorov-Smirnov normality test). When the effects of two independent variables were considered, the two-way analysis of variance (ANOVA) with Bonferroni post-hoc test was used. When more than two groups were compared, one-way analysis of variance (ANOVA) with Bonferroni post-hoc test was used. All statistical analyses display individual values for each biological replicate, mean value, and standard error of the mean (SEM). Statistical significance was set at $P < 0.05$.

REFERENCES AND NOTES

1. K. Alves de Lima, J. Rustenhoven, J. Kipnis, Meningeal Immunity and Its Function in Maintenance of the Central Nervous System in Health and Disease. *Annu. Rev. Immunol.* **38**, 597–620 (2020). doi: 10.1146/annurev-immunol-102319-103410; pmid: 32340575
2. B. Engelhardt, P. Vajkoczy, R. O. Weller, The movers and shapers in immune privilege of the CNS. *Nat. Immunol.* **18**, 123–131 (2017). doi: 10.1038/ni.3666; pmid: 28092374
3. R. Rua, D. B. McGavern, Advances in Meningeal Immunity. *Trends Mol. Med.* **24**, 542–559 (2018). doi: 10.1016/j.molmed.2018.04.003; pmid: 29731353
4. F. Herisson et al., Direct vascular channels connect skull bone marrow and the brain surface enabling myeloid cell migration. *Nat. Neurosci.* **21**, 1209–1217 (2018). doi: 10.1038/s41593-018-0213-2; pmid: 30150661
5. R. Cai et al., Panoptic imaging of transparent mice reveals whole-body neuronal projections and skull-meninges connections. *Nat. Neurosci.* **22**, 317–327 (2019). doi: 10.1038/s41593-018-0301-3; pmid: 30598527
6. H. Yao et al., Leukaemia hijacks a neural mechanism to invade the central nervous system. *Nature* **560**, 55–60 (2018). doi: 10.1038/s41586-018-0342-5; pmid: 30022166
7. H. Keren-Shaul et al., A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* **169**, 1276–1290.e17 (2017). doi: 10.1016/j.cell.2017.05.018; pmid: 28602351
8. S. Mundt et al., Conventional DCs sample and present myelin antigens in the healthy CNS and allow parenchymal T cell entry to initiate neuroinflammation. *Sci. Immunol.* **4**, eaau8380 (2019). doi: 10.1126/sciimmunol.aau8380; pmid: 30679199
9. R. C. Rickert, J. Roes, K. Rajewsky, B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic Acids Res.* **25**, 1317–1318 (1997). doi: 10.1093/nar/25.6.1317; pmid: 9092650
10. A. Louveau et al., CNS lymphatic drainage and neuroinflammation are regulated by meningeal lymphatic vasculature. *Nat. Neurosci.* **21**, 1380–1391 (2018). doi: 10.1038/s41593-018-0227-9; pmid: 30224810
11. J. H. Ahn et al., Meningeal lymphatic vessels at the skull base drain cerebrospinal fluid. *Nature* **572**, 62–66 (2019). doi: 10.1038/s41586-019-1419-5; pmid: 31341278
12. J. W. Tung, M. D. Mrzsek, Y. Yang, L. A. Herzenberg, L. A. Herzenberg, Phenotypically distinct B cell development pathways map to the three B cell lineages in the mouse. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6293–6298 (2006). doi: 10.1073/pnas.0511305103; pmid: 16606838
13. R. R. Hardy, C. E. Carmack, S. A. Shinton, J. D. Kemp, K. Hayakawa, Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J. Exp. Med.* **173**, 1213–1225 (1991). doi: 10.1084/jem.173.5.1213; pmid: 1827140
14. N. Kuwata, H. Igarashi, T. Ohmura, S. Aizawa, N. Sakaguchi, Cutting edge: Absence of expression of RAG1 in peritoneal B-1 cells detected by knocking into RAG1 locus with green fluorescent protein gene. *J. Immunol.* **163**, 6355–6359 (1999). pmid: 10586023
15. J. DeSisto et al., Single-Cell Transcriptomic Analyses of the Developing Meninges Reveal Meningeal Fibroblast Diversity and Function. *Dev. Cell* **54**, 43–59.e4 (2020). doi: 10.1016/j.devcel.2020.06.009; pmid: 32634398
16. F. Schena et al., Dependence of immunoglobulin class switch recombination in B cells on vesicular release of ATP and CD73 ectonucleotidase activity. *Cell Rep.* **3**, 1824–1831 (2013). doi: 10.1016/j.celrep.2013.05.022; pmid: 23770243
17. J. Rustenhoven et al., Functional characterization of the dural sinuses as a neuroimmune interface. *Cell* **184**, 1000–1016.e27 (2021). doi: 10.1016/j.cell.2020.12.040; pmid: 33508229
18. N. Almanzar et al., Tabula Muris Consortium, A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. *Nature* **583**, 590–595 (2020). doi: 10.1038/s41586-020-2496-1; pmid: 32669714
19. D. A. Mogilenko et al., Comprehensive Profiling of an Aging Immune System Reveals Clonal GZMK⁺ CD8⁺ T Cells as Conserved Hallmark of Inflammation. *Immunity* **54**, 99–115.e12 (2021). doi: 10.1016/j.immuni.2020.11.005; pmid: 33271118
20. Z. Fitzpatrick et al., Gut-educated IgA plasma cells defend the meningeal venous sinuses. *Nature* **587**, 472–476 (2020). doi: 10.1038/s41586-020-2886-4; pmid: 33149302
21. T. Goldmann et al., Origin, fate and dynamics of macrophages at central nervous system interfaces. *Nat. Immunol.* **17**, 797–805 (2016). doi: 10.1038/ni.3423; pmid: 27135602
22. M. J. C. Jordão et al., Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. *Science* **363**, eaat7554 (2019). doi: 10.1126/science.aat7554; pmid: 30679343
23. H. Van Hove et al., A single-cell atlas of mouse brain macrophages reveals unique transcriptional identities shaped by ontogeny and tissue environment. *Nat. Neurosci.* **22**, 1021–1035 (2019). doi: 10.1038/s41593-019-0393-4; pmid: 31061494
24. D. Mrdjen et al., High-Dimensional Single-Cell Mapping of Central Nervous System Immune Cells Reveals Distinct Myeloid Subsets in Health, Aging, and Disease. *Immunity* **48**, 380–395.e6 (2018). doi: 10.1016/j.immuni.2018.01.011; pmid: 29426702
25. C. Fistonich et al., Cell circuits between B cell progenitors and IL-7⁺ mesenchymal progenitor cells control B cell development. *J. Exp. Med.* **215**, 2586–2599 (2018). doi: 10.1084/jem.20180778; pmid: 30158115
26. Y. Nie et al., The role of CXCR4 in maintaining peripheral B cell compartments and humoral immunity. *J. Exp. Med.* **200**, 1145–1156 (2004). doi: 10.1084/jem.20041185; pmid: 15520246
27. J. Zikherman, R. Parameswaran, A. Weiss, Endogenous antigen tunes the responsiveness of naive B cells but not T cells. *Nature* **489**, 160–164 (2012). doi: 10.1038/nature11311; pmid: 22902503
28. T. Litzenburger et al., B lymphocytes producing demyelinating autoantibodies: Development and function in gene-targeted transgenic mice. *J. Exp. Med.* **188**, 169–180 (1998). doi: 10.1084/jem.188.1.169; pmid: 9653093
29. J. K. A. Cugurra et al., Skull and vertebral bone marrow are myeloid reservoirs for the meninges and CNS parenchyma. *Science* **373**, XXX (–YYY) (2021).
30. M. P. Cancro, Age-Associated B Cells. *Annu. Rev. Immunol.* **38**, 315–340 (2020). doi: 10.1146/annurev-immunol-092419-031130; pmid: 31986068
31. A. V. Rubtsov et al., Toll-like receptor 7 (TLR7)-driven accumulation of a novel CD11c⁺ B-cell population is important for the development of autoimmunity. *Blood* **118**, 1305–1315 (2011). doi: 10.1182/blood-2011-01-331462; pmid: 21543762
32. Y. Hao, P. O'Neill, M. S. Naradikian, J. L. Scholz, M. P. Cancro, A B cell subset uniquely responsive to innate stimuli accumulates in aged mice. *Blood* **118**, 1294–1304 (2011). doi: 10.1182/blood-2011-01-330530; pmid: 21562046
33. M. Manni et al., Regulation of age-associated B cells by IRF5 in systemic autoimmunity. *Nat. Immunol.* **19**, 407–419 (2018). doi: 10.1038/s41590-018-0056-8; pmid: 29483597
34. O. L. Rojas et al., Recirculating Intestinal IgA-Producing Cells Regulate Neuroinflammation via IL-10. *Cell* **176**, 610–624.e18 (2019). doi: 10.1016/j.cell.2018.11.035; pmid: 30612739

35. S. L. Hauser *et al.*, B-cell depletion with rituximab in relapsing-remitting multiple sclerosis. *N. Engl. J. Med.* **358**, 676–688 (2008). doi: [10.1056/NEJMoa0706383](https://doi.org/10.1056/NEJMoa0706383); pmid: [18272891](https://pubmed.ncbi.nlm.nih.gov/18272891/)
36. S. L. Hauser *et al.*, Ofatumumab versus Teriflunomide in Multiple Sclerosis. *N. Engl. J. Med.* **383**, 546–557 (2020). doi: [10.1056/NEJMoa1917246](https://doi.org/10.1056/NEJMoa1917246); pmid: [32757523](https://pubmed.ncbi.nlm.nih.gov/32757523/)
37. Y. Wang *et al.*, TREM2-mediated early microglial response limits diffusion and toxicity of amyloid plaques. *J. Exp. Med.* **213**, 667–675 (2016). doi: [10.1084/jem.20151948](https://doi.org/10.1084/jem.20151948); pmid: [27091843](https://pubmed.ncbi.nlm.nih.gov/27091843/)
38. B. D. Metscher, MicroCT for developmental biology: A versatile tool for high-contrast 3D imaging at histological resolutions. *Dev. Dyn.* **238**, 632–640 (2009). doi: [10.1002/dvdy.21857](https://doi.org/10.1002/dvdy.21857); pmid: [19235724](https://pubmed.ncbi.nlm.nih.gov/19235724/)
39. T. Stuart *et al.*, Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019). doi: [10.1016/j.cell.2019.05.031](https://doi.org/10.1016/j.cell.2019.05.031); pmid: [31178118](https://pubmed.ncbi.nlm.nih.gov/31178118/)
40. C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019). doi: [10.1186/s13059-019-1874-1](https://doi.org/10.1186/s13059-019-1874-1); pmid: [31870423](https://pubmed.ncbi.nlm.nih.gov/31870423/)
41. G. Finak *et al.*, MAST: A flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278 (2015). doi: [10.1186/s13059-015-0844-5](https://doi.org/10.1186/s13059-015-0844-5); pmid: [26653891](https://pubmed.ncbi.nlm.nih.gov/26653891/)
42. T. Le, et al., BBrowser: Making single-cell data easily accessible (2020). bioRxiv 2020.12.11.414136 [Preprint]. 11 December 2020. doi: [10.1101/2020.12.11.414136](https://doi.org/10.1101/2020.12.11.414136)
43. P. Angerer *et al.*, destiny: Diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241–1243 (2016). doi: [10.1093/bioinformatics/btv715](https://doi.org/10.1093/bioinformatics/btv715); pmid: [26668002](https://pubmed.ncbi.nlm.nih.gov/26668002/)
44. K. Street *et al.*, Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018). doi: [10.1186/s12864-018-4772-0](https://doi.org/10.1186/s12864-018-4772-0); pmid: [29914354](https://pubmed.ncbi.nlm.nih.gov/29914354/)
45. Y. Zhou *et al.*, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019). doi: [10.1038/s41467-019-09234-6](https://doi.org/10.1038/s41467-019-09234-6); pmid: [30944313](https://pubmed.ncbi.nlm.nih.gov/30944313/)
46. R. Browaeys, W. Saelens, Y. Saeys, NicheNet: Modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020). doi: [10.1038/s41592-019-0667-5](https://doi.org/10.1038/s41592-019-0667-5); pmid: [31819264](https://pubmed.ncbi.nlm.nih.gov/31819264/)
47. N. T. Gupta *et al.*, Change-O: A toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356–3358 (2015). doi: [10.1093/bioinformatics/btv359](https://doi.org/10.1093/bioinformatics/btv359); pmid: [26069265](https://pubmed.ncbi.nlm.nih.gov/26069265/)
48. J. A. Vander Heiden *et al.*, pRESTO: A toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics* **30**, 1930–1932 (2014). doi: [10.1093/bioinformatics/btu138](https://doi.org/10.1093/bioinformatics/btu138); pmid: [24618469](https://pubmed.ncbi.nlm.nih.gov/24618469/)
49. W. B. Alsoussi *et al.*, A Potently Neutralizing Antibody Protects Mice against SARS-CoV-2 Infection. *J. Immunol.* **205**, 915–922 (2020). doi: [10.4049/jimmunol.2000583](https://doi.org/10.4049/jimmunol.2000583); pmid: [32591393](https://pubmed.ncbi.nlm.nih.gov/32591393/)
50. N. T. Gupta *et al.*, Hierarchical Clustering Can Identify B Cell Clones with High Confidence in Ig Repertoire Sequencing Data. *J. Immunol.* **198**, 2489–2499 (2017). doi: [10.4049/jimmunol.1601850](https://doi.org/10.4049/jimmunol.1601850); pmid: [28179494](https://pubmed.ncbi.nlm.nih.gov/28179494/)
51. J. Q. Zhou, S. H. Kleinstein, Cutting Edge: Ig H Chains Are Sufficient to Determine Most B Cell Clonal Relationships. *J. Immunol.* **203**, 1687–1692 (2019). doi: [10.4049/jimmunol.1900666](https://doi.org/10.4049/jimmunol.1900666); pmid: [31484734](https://pubmed.ncbi.nlm.nih.gov/31484734/)

ACKNOWLEDGMENTS

We thank E. Lantelme for her excellent management of the flow cytometry core facility at the department of Pathology and Immunology (Washington University in St. Louis). We thank D. Bender and R. Lin from the Bursky Center for Human Immunology and Immunotherapy Programs (CHiIPs) for their invaluable support of sample preparation for CyTOF analysis and acquisition of CyTOF data. We thank B. Saunders, B. Zinselmeyer, and G. Randolph for their assistance with two-photon in vivo imaging. We thank Y. Chen and D. Holtzman for providing the anti-ApoE antibody. We also thank L. Adamo for providing CD19-Tomato mice for preliminary experiments, P. Allen for providing *Nur77^{GFP}* mice, A. Archambault for providing *IGH^{MOG}* mice, N. Sakaguchi for providing *Rag1^{GFP}* mice, and K. Ayasoufi for

helping with the perfusion and shipment of *Rag1^{GFP}* samples. Lastly, we thank R. Pelanda, J. Andersson, F. Melchers, A. Pernis, P. Kendall, and A. Ellebedy for their critical comments and helpful suggestions during the preparation of this manuscript.

Funding: This research was supported in part by NIH RF1 AG051485, NIH RF1 AG059082 (M. Colonna), NIH R01AI104739 (S.H.K.), NIH F30 DK127540-01, NIH T32 DK 77653-28 (V.P.), the Gruber Foundation (M.W.), NIH K99 DK118110 (J.K.B.), and FA-2020-01-IBD-1 fellowship from the Lawrence C. Pakula, MD IBD Education and Innovation Fund (R.S.C.). **Author contributions:** S.B, W.W, V.P: conceptualization, data curation, formal analysis, methodology, investigation, software, validation, visualization, and writing the original draft. M.W., I.S.: formal analysis, methodology, and software. Z.J.G, J.K.B., N.J., R.S.C., A.S., D.A.M.: formal analysis and methodology. W.L.B., P.B., J.A.J.F., L.G.S., C.C.F., I.S., J.K., V.S., G.W., S.G.: methodology and resources. M.C., M.N.A., S.H.K.: supervision. M.Colonna: funding acquisition, supervision, project administration, manuscript review, and editing. **Competing interests:** S.H.K. receives consulting fees from Northrop Grumman. All other authors declare no conflict of interest. **Data and materials availability:** scRNA-seq and scBCR-seq data available from NCBI GEO with accession number GSE174836. All other data and materials are available in the paper and supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/eabf9277/suppl/DC1
Materials and Methods
Figs. S1 to S16
Tables S1 to S4
MDAR Reproducibility Checklist
Movies S1 and S2

29 November 2020; resubmitted 1 April 2021
Accepted 24 May 2021
Published online 3 June 2021
[10.1126/science.abf9277](https://doi.org/10.1126/science.abf9277)

RESEARCH ARTICLE SUMMARY

NEUROIMMUNOLOGY

Skull and vertebral bone marrow are myeloid cell reservoirs for the meninges and CNS parenchyma

Andrea Cugurra[†], Tornike Mamuladze[†], Justin Rustenhoven[†], Taitea Dykstra, Giorgi Beroshvili, Zev J. Greenberg, Wendy Baker, Zach Papadopoulos, Antoine Drieu, Susan Blackburn, Mitsuhiro Kanamori, Simone Bríoschi, Jasmin Herz, Laura G. Schuettelpelz, Marco Colonna, Igor Smirnov, Jonathan Kipnis*

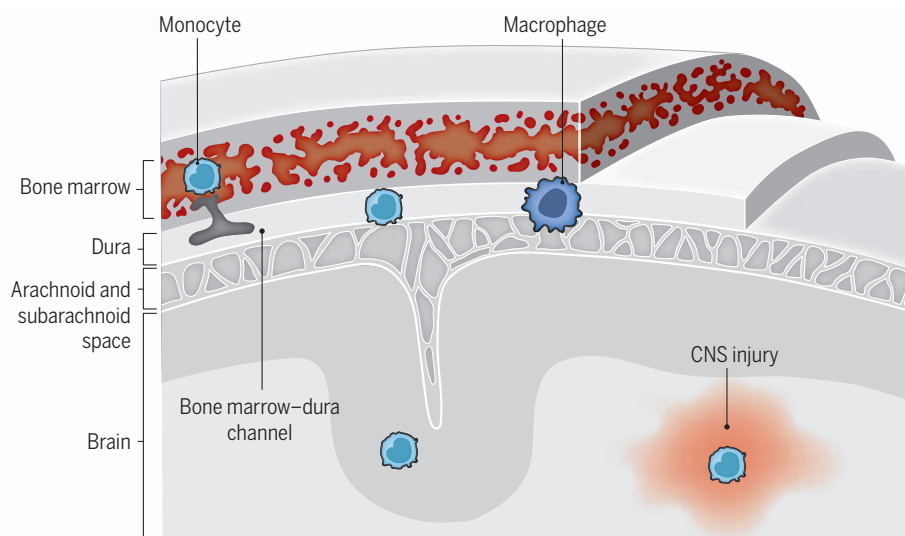
INTRODUCTION: Myeloid cells are highly heterogeneous, with distinct ontogenies and immune functions. The diverse role of myeloid populations, including monocytes, macrophages, and neutrophils, in homeostasis of central nervous system (CNS) function is increasingly recognized. Brain borders, including the dura mater of the meninges, host a substantial pool of myeloid cells that mediate immune surveillance of the CNS. During injury and neuroinflammation, myeloid cells infiltrate the CNS and display heterogeneous functions from microglia, the CNS resident macrophages. Understanding myeloid ontogeny in CNS borders and in the CNS parenchyma under homeostasis and during perturbations is critical for our understanding of the basic mechanisms underlying immune responses in the diseased CNS and for the design of therapeutic approaches for neurological conditions associated with inflammation, such as injury, chronic neurodegenerative diseases, CNS infections, and brain tumors.

RATIONALE: Myeloid cells are critical contributors to CNS function, yet the origins of these cells at CNS borders and the routes of their migration into the CNS parenchyma under inflammatory or neurodegenerative conditions remain unclear. Under certain circumstances—for example, monocyte infiltration in experimental autoimmune encephalomyelitis (EAE) or after CNS injury—myeloid populations are playing protective and pathological roles. We rationalized that these heterogeneous functions may represent diverse cellular origins and sought to investigate the sources of myeloid cell infiltrates and the routes by which they access the injured or inflamed parenchyma. Channels connecting the skull bone marrow to the dura to allow neutrophil migration during stroke were previously described. Thus, we explored whether these channels and the skull bone marrow niche also allow for maintenance of the brain dural myeloid pool under homeostasis and in mouse models of CNS diseases. We further extended our ob-

servations to the vertebrae bone marrow niche as a potential source of myeloid cells for the spinal dura.

RESULTS: Using the parabiotic pairing of WT-UBC-GFP or CD45.1-CD45.2 mice that share blood circulation, we observed a substantial pool of monocytes and neutrophils in the brain and spinal cord dura that were not blood derived and had not arisen from tissue intrinsic progenitors or local proliferation. Using several approaches, including calvaria bone-flap transplantation and selective irradiation regimens with head or body shielding and bone marrow transfers, we found that these non-blood-derived myeloid cells had originated from skull and vertebrae bone marrow. We reconfirmed the existence of skull bone marrow-dura channels and described the presence of similar channels connecting spinal cord vertebrae bone marrow to spinal dura, allowing an anatomical route for constant myeloid migration from local bone marrow pools. After various CNS injury models, including spinal cord injury, EAE, and optic nerve crush, we showed that CNS-associated bone marrow contributes to parenchymal CNS infiltration. Using single-cell RNA sequencing, we demonstrated potentially nonredundant roles for blood and CNS bone marrow-derived monocytes infiltrating the spinal cord during spinal cord injury and EAE, with blood-derived cells displaying a more inflammatory and possibly pathogenic phenotype.

CONCLUSION: Our results show that bone marrow niches adjacent to the brain and spinal cord supply monocytes to the meninges under homeostasis and after CNS injury or neuroinflammatory disease. Moreover, we demonstrate that CNS border-derived myeloid cells can migrate through meningeal barriers into parenchyma under injurious and inflammatory conditions. CNS border-derived myeloid cells exhibit a less inflammatory and more regulatory phenotype when compared with their blood-derived counterparts in models of CNS injury and neuroinflammation. Thus, it is possible that these cells are playing critical roles in the regulation—and possibly attenuation—of immune responses in other contexts, such as CNS infections and brain tumors. Understanding the function of brain border-derived myeloid cells in CNS immune surveillance in physiology and pathology may lead to development of therapeutic avenues to treat neurological diseases. ■



Myeloid cells from skull and vertebrae bone marrow supply brain borders and the CNS during health and disease. Under homeostasis, the dural meninges host a pool of monocytes derived from local bone marrow sources that can differentiate into tissue-resident macrophages. During CNS injury, bone marrow-derived monocytes can infiltrate CNS parenchyma from brain borders.

The list of author affiliations is available in the full article online.

*Corresponding author. Email: kipnis@wustl.edu

[†]These authors contributed equally to this work.

Cite this article as A. Cugurra *et al.*, *Science* 373, eabf7844 (2021). DOI: 10.1126/science.abf7844

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abf7844>

RESEARCH ARTICLE

NEUROIMMUNOLOGY

Skull and vertebral bone marrow are myeloid cell reservoirs for the meninges and CNS parenchyma

Andrea Cugurra^{1,2,†}, Tornike Mamuladze^{3,†}, Justin Rustenhoven^{3,†}, Taite Dykstra³, Giorgi Beroshvili³, Zev J. Greenberg⁴, Wendy Baker², Zach Papadopoulos^{3,5}, Antoine Drieu³, Susan Blackburn³, Mitsuhiro Kanamori³, Simone Brioschi³, Jasmin Herz³, Laura G. Schuettelpelz⁴, Marco Colonna³, Igor Smirnov³, Jonathan Kipnis^{1,2,3,5,*}

The meninges are a membranous structure enveloping the central nervous system (CNS) that host a rich repertoire of immune cells mediating CNS immune surveillance. Here, we report that the mouse meninges contain a pool of monocytes and neutrophils supplied not from the blood but by adjacent skull and vertebral bone marrow. Under pathological conditions, including spinal cord injury and neuroinflammation, CNS-infiltrating myeloid cells can originate from brain borders and display transcriptional signatures distinct from their blood-derived counterparts. Thus, CNS borders are populated by myeloid cells from adjacent bone marrow niches, strategically placed to supply innate immune cells under homeostatic and pathological conditions. These findings call for a reinterpretation of immune-cell infiltration into the CNS during injury and autoimmunity and may inform future therapeutic approaches that harness meningeal immune cells.

Myeloid cells, including monocytes, neutrophils, and macrophages, display extraordinary heterogeneity and diverse functions depending on their ontogeny and local niche (1–3). Macrophages represent an essential component of the myeloid niche and take on distinct transcriptional signatures to meet the demands of the tissues in which they reside (4). Research on microglia, the macrophages of the central nervous system (CNS), led to the discovery that these tissue-resident macrophages are derived from the yolk sac during embryonic development (5–7). Unlike several other tissue macrophages that can be replaced by blood-borne monocytes, according to specific anatomical and developmental constraints (8), microglia are self-maintained and are not replenished by blood-derived monocytes other than under defined experimental conditions (5, 9–12). In such cases, myeloid cells from the periphery may also invade the CNS tissue, although their origin, as well as their stability within the CNS, is poorly understood. The perivascular spaces and the meningeal membranes that cover the borders of the

CNS are also populated by a variety of myeloid cells. Based on elegant fate-mapping strategies and multidimensional tools, recent work has begun to reveal the richness of this myeloid landscape, both at steady state and in different diseases (13–20). Less well studied, however, is the source and role of these brain border-resident monocytes, monocyte-derived macrophages, and neutrophils during homeostasis and CNS dysfunction.

Recent work described the existence of direct ossified vascular channels connecting skull bone marrow to meninges that are capable of dispersing neutrophils during inflammation (21–23). We hypothesized that these connections may also allow homeostatic myeloid trafficking between adjacent meningeal and bone marrow niches. Here, we demonstrate that CNS-associated bone marrow niches in the skull and vertebrae are myeloid reservoirs for the meninges and CNS parenchyma. Under homeostasis, these bone marrow pools supply the brain and spinal dural meninges with monocytes and neutrophils via direct dural-bone marrow connections. After injury and neuroinflammation, these cells can mobilize to infiltrate the CNS parenchyma and display distinct phenotypes from their blood-derived counterparts.

Results

CNS borders host a substantial pool of monocytes and neutrophils that are not blood derived

In an attempt to understand the origin of border myeloid cells under homeostasis, we parabiotically joined the circulations of green fluorescent protein under human ubiquitin C

promoter (UBC-GFP) and wild-type (WT) mice, allowing us to visualize host and fluorescent donor-derived cells, and performed flow cytometric analyses of the blood, brain-associated tissues (cranial dura mater and spinal dura mater), hematopoietic organs (skull bone marrow, vertebrae bone marrow, and femur bone marrow), and peripheral control tissues (spleen and liver) after 60 days of pairing (Fig. 1, A to C). To ensure that our analysis only examined true parenchymal populations and not blood contaminants, mice received intravenous anti-CD45 antibodies 2 min before euthanasia, which labeled blood and vascular-associated leukocytes that were subsequently excluded in our gating strategy (Fig. 1B). Analysis of WT parabionts demonstrated that although blood chimerism for total CD45⁺ cells reached a 50:50 ratio (fig. S1A), neither neutrophils nor monocytes ever reached perfect chimerism in the blood. After 2 months, a GFP:WT cell ratio of ~30:70 was achieved (Fig. 1, D and E). The reduced chimerism of these myeloid cells was likely due to the short half-life of these cells as well as active myelopoiesis occurring in the bone marrow.

Despite imperfect blood chimerism of Ly6C⁺ monocytes and neutrophils, GFP⁺ cell proportions detected in the cranial dura and spinal dura were significantly lower than that of blood (Fig. 1, D and E). By comparison, the myeloid chimerism displayed in other peripheral tissues were more akin to the proportions present in the blood (Fig. 1, D and E). These data suggested that there was a substantial pool of monocytes and neutrophils located along the brain borders that had not originated from the UBC-GFP parabiont. Cranial and spinal dural CD4⁺ T cells from the UBC-GFP parabiont were not significantly different from the blood chimerism, suggesting a blood origin as described previously (24) (Fig. 1F). Furthermore, GFP⁺Ly6c⁺ monocytes and B cells displayed different levels of chimerism in distinct tissues. GFP⁺ B cells specifically appeared to be less represented at the brain borders, also suggesting a nonblood origin (fig. S1, B and C). This phenomenon is described in more detail in an accompanying manuscript (25). Additional components of the dural myeloid compartment—including macrophages, conventional dendritic cells (cDC)1s, and cDC2s—also displayed a low proportion of GFP⁺ cells compared with the non-border-associated spleen pool (fig. S1, D to F). Analysis of several bone marrow niches confirmed that most of the detected LSK (lineage[−]SCA-1⁺c-Kit⁺) hematopoietic progenitors were GFP⁺ cells, as expected from an active WT proliferative niche (fig. S1G). We corroborated these flow cytometry data by immunohistochemistry (IHC) and demonstrated only a minor proportion of GFP⁺CCR2⁺ and GFP⁺Ly-6B.2⁺ cells in WT host tissues (Fig. 1G). Using CD45.1 and CD45.2 pairs as parabionts, we observed a similar imbalance

¹Gutenberg Research Fellowship Group of Neuroimmunology, Focus Program Translational Neuroscience (FTN) and Immunotherapy (FZI), Rhine Main Neuroscience Network (rmn²), University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany. ²Department of Neuroscience, School of Medicine, University of Virginia, Charlottesville, VA, USA. ³Department of Pathology and Immunology, Washington University in St. Louis, St. Louis, MO, USA. ⁴Department of Pediatrics, Washington University School of Medicine, Saint Louis, MO, USA. ⁵Neuroscience Graduate Program, School of Medicine, Washington University in St. Louis, St. Louis, MO, USA.

*Corresponding author. Email: kipnis@wustl.edu

†These authors contributed equally to this work.

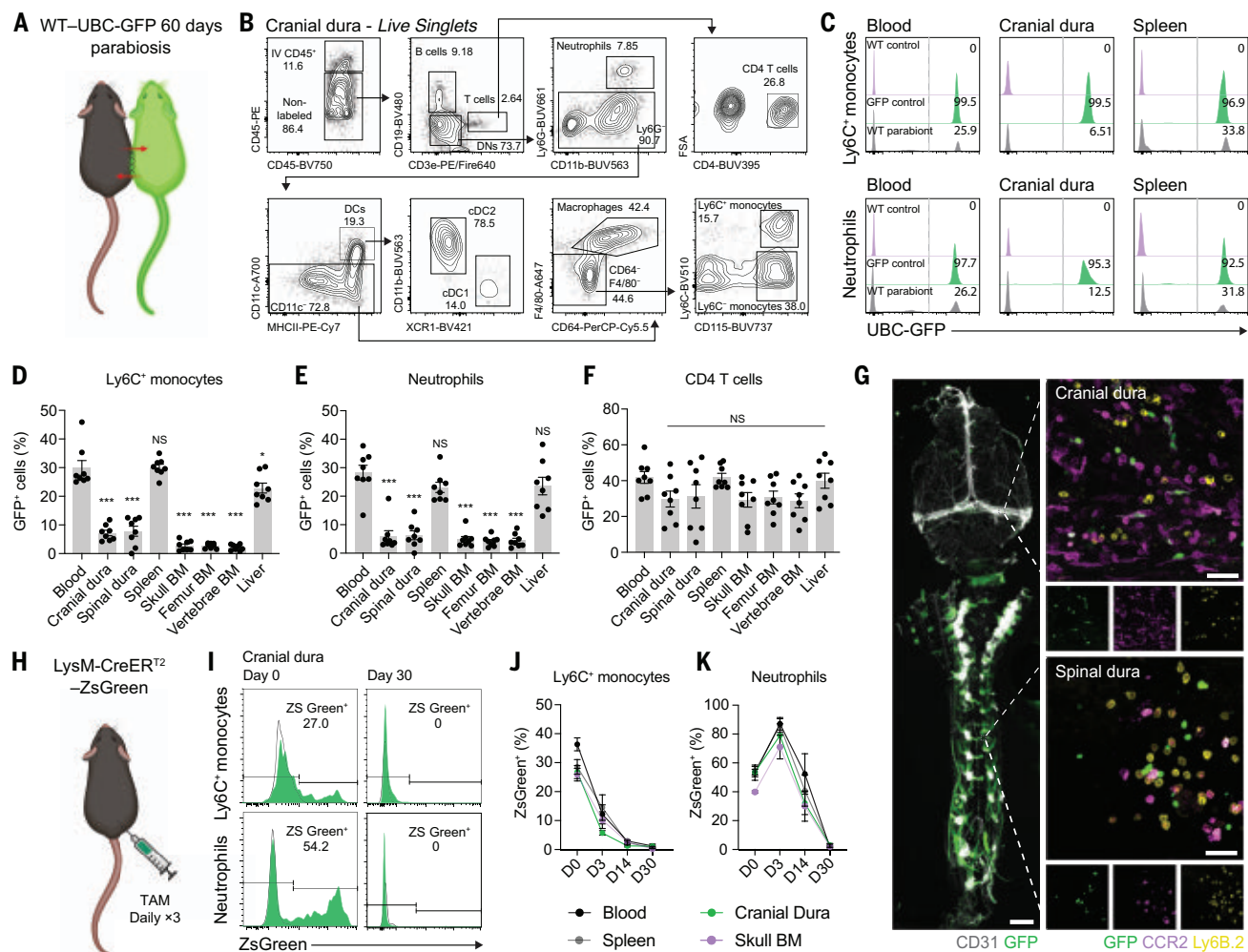


Fig. 1. CNS borders host a substantial pool of monocytes and neutrophils that is not blood derived. (A) Experimental design for parabiosis experiments. WT mice were parabiotically joined with UBC-GFP mice and analyzed 60 days later. (B) Representative gating strategy for flow cytometry of immune cells from the cranial dura. (C) Representative histograms of GFP⁻ and GFP⁺ Ly6C⁺ monocytes and neutrophils across tissues. WT and UBC-GFP control mice without intravenous anti-CD45 antibody injection were used to assign gates. (D to F) Quantification of flow cytometry analysis for GFP⁺ Ly6C⁺ monocytes, GFP⁺ neutrophils, and GFP⁺ CD4 T cells across tissues. NS indicates not significant; ****P* < 0.001 versus blood [one-way analysis of variance (ANOVA) with Dunnett's post hoc test]; *n* = 8 mice per tissue, representative of three independent

experiments. Data are means ± SEM. (G) Representative images of cranial and spinal dura whole mounts from three WT parabionts 60 days after parabiotic pairing. Scale bars: 2000 μm (left) and 50 μm (right). (H to K) Myeloid fate mapping strategy using LysM-CreERT2:ZsGreen reporter mice to assess the half-life of meningeal Ly6C⁺ monocytes and GFP⁺ neutrophils. Mice received daily tamoxifen injections for 3 days and were then analyzed at four different time points. Cre⁻ animals were used as controls to assign gates. Black outlined histograms represent Cre⁻ controls and green histograms represent LysM-CreERT2:ZsGreen reporter mice. *n* = 3 or 4 mice per group, representative of two independent experiments. Data are means ± SEM. D, day.

toward host-derived monocytes and neutrophils within the dura compared with the blood chimerism, suggesting that the observed findings were consistent across different parabiotic pairs (fig. S1, H and I). Thus, a substantial pool of non-blood-derived monocytes and neutrophils appears to populate the cranial and spinal dura.

Next, we investigated whether the dura harbored monocyte progenitor cells that could maintain a local monocyte pool. Using flow cytometry, we could efficiently identify monocyte-dendritic cell progenitors (MDPs) and

monocyte-committed progenitors (cMoPs) in the skull bone marrow. These populations were entirely absent from the cranial dura, however (fig. S1, J to S). To exclude the possibility that the low levels of meningeal chimerism observed for Ly6C⁺ monocytes and neutrophils could be attributed to the proliferation and self-renewal of those cells in the tissue, we explored the turnover of these myeloid cells using a tamoxifen-inducible fate-mapping Cre recombinase system driven by the pan-myeloid promoter LysM to express the fluorescent protein ZsGreen (Fig. 1H). After tamoxifen injections

over three consecutive days, blood monocytes and neutrophils reached a recombination efficiency of 36 and 87%, respectively. After 30 days, almost no circulating ZsGreen⁺ cells could be detected (Fig. 1, I to K). Notably, the fate-mapping efficiency of the blood was mirrored by that in the dura across all investigated time points, indicating that the half-lives of the myelomonocytic cells in the blood and in the meningeal spaces were comparable (Fig. 1, H to K; and fig. S2, A to D). To rule out the possibility that the unlabeled fraction of monocytes and neutrophils provides a local

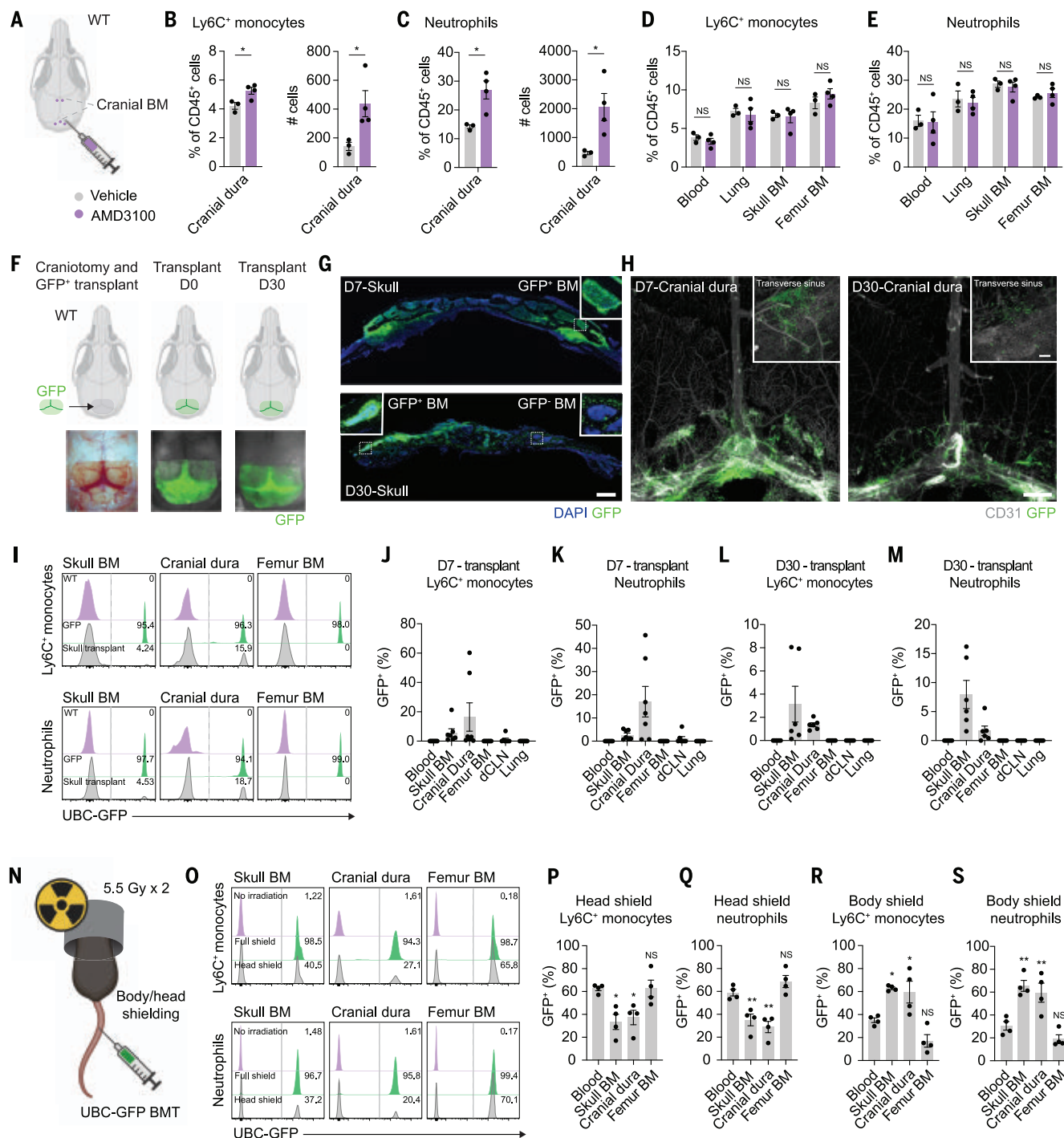


Fig. 2. Skull bone marrow supplies brain borders with myeloid cells.

(A) Experimental design for bone marrow egress experiments. The outer periosteal layer of the skull bone was thinned near the bone marrow (BM) sites, and 1 μ l of CXCR4 antagonist AMD3100 (1 μ g/ml) or a vehicle was applied to five sites for 5 min. (B and C) Quantification for flow cytometry analysis of IV-CD45⁺ Ly6C⁺ monocytes and neutrophils in cranial dura 1 day after AMD3100 administration. **P* < 0.05 (Student's *t* test); *n* = 3 or 4 mice per group, representative of two independent experiments. Data are means \pm SEM. (D and E) Quantification for flow cytometry analysis of IV-CD45⁺ Ly6C⁺ monocytes and neutrophils in control tissues 1 day after AMD3100 administration. NS indicates not significant (two-way ANOVA with Sidak's post hoc test); *n* = 3 or 4 mice per group, representative of two independent experiments.

Data are means \pm SEM. (F) Schematic representation of the calvaria-flap transplantation experiments. Images were taken on anesthetized mice by stereomicroscopy immediately after craniotomy (left), after the subsequent calvaria-flap transplantation (middle), and 30 days after the transplantation (right). (G) Representative images from three mice of GFP⁺ calvaria flaps 7 and 30 days after the transplantation. Insets show GFP⁺ bone marrow 7 and 30 days after the transplantation. Dead bone marrow was replaced by non-GFP⁺ cells. Scale bar: 500 μ m. DAPI, 4',6-diamidino-2-phenylindole. (H) Representative images from four mice of cranial dura 7 and 30 days after the transplantation showing GFP⁺ cells below the transfer site. Scale bar: 1000 μ m. (I) Representative histograms of GFP⁺ and GFP⁻ Ly6C⁺ monocytes and neutrophils across tissues. WT and UBC-GFP control mice without intravenous

anti-CD45 antibody injection were used to assign gates. (**J to M**) Quantification of flow cytometry analysis of IV-CD45⁺GFP⁺Ly6C⁺ monocytes and GFP⁺ neutrophils from calvaria flap-transplanted mice; $n = 6$ or 7 mice per group, representative of two independent experiments for day 7 and three independent experiments for day 30. Data are means \pm SEM. (**N**) Experimental design of irradiation and bone marrow transplantation experiments with different shielding strategies. Four million GFP⁺ bone marrow cells were intravenously transplanted into head-shielded or body-shielded WT mice after split-dose 11 gray (Gy)

irradiation. (**O**) Representative histograms of GFP⁺ and GFP⁺ Ly6C⁺ monocytes and neutrophils across tissues. Fully irradiated and nonirradiated mice transplanted with UBC-GFP cells without intravenous anti-CD45 antibody injection were used as control animals and to assign gates. (**P to S**) Quantification of flow cytometry analysis for IV-CD45⁺GFP⁺Ly6C⁺ monocytes and GFP⁺ neutrophils in BMT experiments. NS indicates not significant; * $P < 0.05$ and ** $P < 0.01$ versus blood (one-way ANOVA with Dunnett's post hoc test); $n = 4$ mice per tissue, representative of two independent experiments. Data are means \pm SEM.

pool of self-sustaining tissue-resident cells, we performed *in vivo* 5-ethynyl-2'-deoxyuridine (EdU) labeling and Ki67 staining. Analysis of the blood, cranial dura, and skull bone marrow confirmed that 24 hours after two EdU pulses, most of the Ly6C⁺ monocytes and neutrophils in the dura incorporated EdU but displayed low levels of Ki67 (fig. S2, E to J). Thus, these cells appeared to originate from a proliferating source but likely experienced rapid turnover in these tissues and low proliferation capabilities outside the bone marrow niche. Moreover, dural neutrophils displayed the same level of EdU incorporation as the skull bone marrow pool, whereas blood neutrophils were almost entirely devoid of EdU⁺ cells, suggesting that dural neutrophils may arise from the nearby bone marrow sources (fig. S2, E to J). Finally, we investigated if the dural Ly6C⁺ monocyte and neutrophil pool used meningeal lymphatics to egress from the tissues. For this, we used KikGR transgenic mice with KikGreen-to-KikRed photoconvertible fluorescent protein (26). Although neutrophils and Ly6C⁺ monocytes were present in the draining deep cervical lymph nodes 24 hours after the final photoconversion, no significant differences in KikRed cells could be detected in the deep cervical lymph nodes, compared with unconverted controls (fig. S2, K to N). Thus, replenishment of dural meningeal Ly6C⁺ monocytes and neutrophils appears to come from a local source and is not due to homeostatic maintenance within the tissue itself or from the blood.

Skull bone marrow supplies brain borders with myeloid cells

Direct connections through ossified channels harboring blood vessels extend from the skull bone marrow to the dura and allow neutrophils and tumor cell migration during inflammatory or pathological states (21–23). A prior strategy for neutrophil tracking used injections into skull bone marrow to label brain-infiltrating cells (21). However, on reproducing this approach in an attempt to assess a bone marrow origin for homeostatic dural myeloid cells, we found it unsuitable for our purpose because it resulted in immediate leakage of the tracers into the dura mater and their subsequent drainage by meningeal lymphatic vessels (fig. S3, A to N). Despite the caveats that come with these direct bone marrow injections,

we found that a brief topical application of AMD3100, a CXCR4 antagonist, to cranial bone marrow through a thinned skull promoted myeloid cell egress. This resulted in significant increases in Ly6C⁺ monocytes and neutrophils in underlying dura mater, without altering proportions in the blood, lung, or other bone marrow niches, suggesting direct migration from local skull bone marrow (Fig. 2, A to E). We did not observe a change in the number of myeloid cells in the skull bone marrow 24 hours after AMD3100 administration. This is likely due to the number of monocytes and neutrophils infiltrating the dura representing only a very small proportion of the CNS-associated bone marrow niche.

These data prompted us to look closely into the possibility that the skull bone marrow is a source for meningeal myeloid cells. Therefore, we developed a calvaria bone-flap transplantation, in which a piece of the calvarium (~6 mm by 4 mm) containing a substantial bone marrow reservoir was transplanted from UBC-GFP mice to cover a skull window of similar size created in WT mice whose dura mater was carefully left intact (Fig. 2F). IHC analysis of transplanted calvaria flaps 7 and 30 days after transplantation revealed the continued presence of GFP⁺ bone marrow, though some sites were partially repopulated by non-GFP-expressing cells (Fig. 2G and fig. S4, A and B). Consistent with sterile injury-induced dural angiogenesis (27), vascular remodeling was observed at 7 days but returned to the coverage of naïve mice after 30 days (Fig. 2H and fig. S4, C and D). The transplanted flap contained cranial sutures necessary for skull repair (28). Indeed, after 1 month, transplanted bone was incorporated into the WT skull (fig. S4E and movie S1). We detected GFP⁺ cells—including CCR2⁺ monocytes, IBA1⁺ macrophages, and CD31⁺ vasculature—in the underlying dura of the transplanted calvaria flap (fig. S4, F and G). Flow cytometric analysis of blood, cranial dura, and skull bone marrow confirmed that although the donor pool of GFP⁺ cells in the transplanted skull bone marrow niche was low, it was able to give rise to dural Ly6C⁺ monocytes and neutrophils even 30 days after transplant (Fig. 2, I to M; and fig. S4, H and I). Demonstrating a direct route for dural myeloid infiltration from the transplanted skull, we observed previously described vascular channels connecting the transplanted skull

to the dura mater (fig. S4, J and K; and movies S2 and S3). Analysis of the underlying cortex revealed no microglial activation by Sholl analysis after 30 days (fig. S4, L and M), and after an initial weight loss after transplantation, mice resumed a normal weight gain, suggesting minimal chronic perturbations using this model (fig. S4N).

We next used an irradiation regimen to further exclude the possibility that the observed phenotype after parabiosis or calvaria bone-flap transplantation strategies is inflammation driven (Fig. 2N). Mice underwent lethal irradiation coupled with different shielding strategies and were reconstituted with UBC-GFP-derived bone marrow. As expected, cranial dural Ly6C⁺ monocytes and neutrophils in head-shielded mice, whose bodies were irradiated, showed similar levels of chimerism as skull bone marrow. Chimerism at both sites was significantly lower than that in the blood and femur bone marrow (Fig. 2, O to Q). The body-shielding strategy, in which the head was irradiated, displayed the opposite effect, with a significantly higher proportion of GFP⁺ cells in the skull bone marrow and cranial dura compared with blood or femur bone marrow (Fig. 2, R and S).

To explore dural-derived factors that could recruit bone marrow-derived myeloid cells, we used RNA-magnet algorithms from whole-dura single-cell RNA sequencing (scRNA-seq) analysis (24) to unbiasedly identify ligand expression for monocyte and neutrophil chemokine receptors (fig. S5, A and B). We then performed semiquantitative chemokine protein expression analysis from whole-dura homogenates to confirm the true presence of predicted ligands (fig. S5C). We identified high expression of CCL2, CCL12, and CCL8, which can recruit monocytes through CCR2 signaling, as well as CCL6, which can recruit neutrophils via CCR1. This suggested that the homeostatic dura contains appropriate signals for local myeloid recruitment from adjacent bone marrow (fig. S5, C to E). Thus, under homeostatic conditions, the myeloid niche distributed along the brain borders receives a substantial input from the skull bone marrow, which appears to act as a critical dispenser of myeloid cells. This provides an example of a healthy tissue hosting myeloid cells that are continuously replenished by a source that does not use blood as a major route.

The inflamed CNS is infiltrated by blood and CNS-adjacent bone marrow–derived myeloid cells

Intrigued by the large proportion of non-blood-derived monocytes and neutrophils residing in the dura, we speculated that these cells may be positioned for a prompt response in the event of parenchymal CNS damage. We proceeded to investigate this scenario using parabiotic pairing in three CNS-injury models: experimental autoimmune encephalomyelitis (EAE), spinal cord injury, and optic nerve crush injury. For EAE experiments, WT and UBC-GFP mice were joined parabiotically for 60 days, and both groups were immunized with myelin oligodendrocyte glycoprotein peptide (35–55) (MOG_{35–55}) in complete Freund's adjuvant to induce EAE. Lumbar spinal cord, lumbar vertebral bone marrow, spinal dura, and blood were collected from the WT parabionts 15 days after EAE induction and stained for flow cytometry and IHC (Fig. 3A and fig. S6, A and B). In agreement with their homeostatic presence, a substantial pool of dural and CNS-infiltrating Ly6C⁺ monocytes did not originate from the blood, indicated by their significantly lower GFP⁺ proportion compared with the blood chimerism (Fig. 3, B and C). Interestingly, the same was not observed for neutrophils, CD4 T cells, and Ly6C[−] monocytes, suggesting blood as the major source of these infiltrates at this time point (Fig. 3D and fig. S6, C to E). Dural and spinal cord GFP⁺ macrophages were also observed, suggesting that infiltrating monocytes can differentiate under these inflammatory conditions, though not at a sufficient amount to explain the chimerism discrepancy (fig. S6F). In agreement with flow cytometry analysis, IHC for cranial dura, spinal dura, and spinal cord revealed that infiltrating GR1⁺ cells within WT tissues were rarely GFP⁺ (Fig. 3E). In the cranial dura, GR1⁺ cells were mostly associated with the dural sinuses, and we observed evidence of intraluminal lymphatic localization, suggesting the possibility of lymphatic trafficking of myeloid cells during neuroinflammatory diseases as previously described (29) (fig. S6, G to I).

After spinal cord injury in CD45.1-CD45.2 parabionts, CNS-infiltrating Ly6C⁺ monocytes (but not neutrophils, CD4 T cells, or Ly6C[−] monocytes) also displayed a significantly lower proportion of host-derived cells in the spinal cord and spinal dura compared with blood, suggestive of a local bone marrow origin (fig. S7, A to G). GFP⁺ monocyte-derived macrophages were present, but their proportion in the injured spinal cord was low (fig. S7H). Similar results were obtained for another CNS injury, whereby optic nerve-infiltrating Ly6C⁺ monocytes and neutrophils after optic nerve crush in WT-UBC-GFP parabionts displayed a significantly lower proportion of donor-derived

cells compared with the blood chimerism (fig. S7, I to L). Importantly, the same was not observed in a periphery injury model using ear skin puncture that lacks a local bone marrow reservoir and shows similar chimerism in the site of injury to that observed in the blood (fig. S7, M to P).

To identify a direct route that allows non-blood-derived monocytes to invade spinal cord tissue, we explored whether bone marrow channels identified for the skull were also present in spinal cord vertebrae. Using hematoxylin and eosin (H&E) staining and IHC in naïve and EAE mice 15 days after induction, we observed channels directly connecting vertebrae bone marrow to spinal dura, with cells present within the channels (Fig. 3F and fig. S8, A to C). Further, we observed that spinal cord infiltrates of GR1⁺ cells were frequently associated with the presence of these bone marrow–spinal dura channels (Fig. 3F). The exact anatomical connections of these channels remain controversial, but we found that channels terminated within the dura mater and did not directly contact the CNS tissue (Fig. 3F and fig. S8, A to E). Exactly how myeloid cells traverse the underlying arachnoid and pia maters to penetrate the CNS parenchyma remains to be established, but like the blood-brain barrier, it is plausible that these barriers, consisting of tight junctions, are similarly disrupted under neuroinflammatory conditions (30).

We next sought to determine whether CNS-infiltrating monocytes of different origins (i.e., blood versus bone marrow or meningeal), may have different phenotypes based on the route by which they gain access to the CNS. To explore this, we induced EAE in WT-UBC-GFP parabionts, isolated GFP⁺ and GFP[−] spinal cord–infiltrating IV-CD45[−]CD45^{hi} cells from WT mice 15 days after induction, and performed scRNA-seq (fig. S9, A to C). Most of the sorted cells were monocytes, but we also obtained populations of neutrophils, dendritic cells (DCs), B cells, T cells, microglia and macrophages, and an undefined proliferating population cluster of cells from both GFP⁺ and GFP[−] populations (Fig. 3, G and H). In agreement with previous EAE experiments, monocytes made up most of the GFP[−] pool, consistent with a nonblood origin (Fig. 3I and fig. S9C). Gene ontology pathway analysis for biological processes of differentially expressed genes in GFP⁺ versus GFP[−] monocytes revealed down-regulated pathways involved in anion- and lipid-transport metabolism pathways and wound healing (Fig. 3J and table S1). Up-regulated pathways were largely involved in leukocyte migration, adhesion, and T cell activation (Fig. 3K and table S1). Analysis of differentially expressed genes revealed a significant induction of diverse myeloid and lymphocyte chemokines (*Ccl2*, *Ccl5*, *Cxcl9*, and

Cxcl16) and proinflammatory cytokines (*Il6*, *Il1a*, *Il1b*, *Tnf*, and *Ifng*) (Fig. 3L and table S2). These data suggest a potential skewed pro-inflammatory role for blood compared with bone marrow–derived monocytes in EAE, in line with the recently described pathogenic subset of CXCL10⁺/CXCL9⁺ monocytes in neuroinflammatory disease (31). These cells may be involved in enhanced leukocyte trafficking, for example by CCL5-, CXCL9-, and CXCL16-mediated T cell recruitment and proinflammatory cytokine production, which are critical factors in EAE pathogenesis (32).

Notably, scRNA-seq analysis of infiltrating IV-CD45[−]CD45^{hi} GFP⁺ versus GFP[−] monocytes from the spinal cord of WT parabionts 3 days after spinal cord injury revealed a similar up-regulation of pathways involved in leukocyte migration and activation, with several conserved up-regulated chemokines and cytokines in GFP⁺ blood-derived monocytes (fig. S9, D to L; and tables S3 and S4). Thus, blood and bone marrow–derived monocytes may have differential roles in diverse CNS injury and inflammatory conditions, perhaps similar to those observed for choroid plexus compared with blood trafficking during spinal cord injury (33).

Discussion

In this study, we demonstrate that the steady-state myeloid niche located at the CNS borders in mice consists of a substantial pool of monocytes and neutrophils that are continuously replenished from bone marrow niches in the adjacent skull or vertebrae. Our model of calvaria bone-flap transplantation—used here as a proof-of-concept prototype—and various irradiation regimens demonstrate that skull bone marrow can supply the adjacent meningeal tissue with myelomonocytic cells without using the blood route. Moreover, our results indicate that these bone marrow–derived monocytes can differentiate into meningeal macrophages. This revealed a myeloid niche that is open to homeostatic peripheral inputs that, in contradiction to the classic Van Furth dogma, does not use the blood as a major route.

It is tempting to speculate that channels aligned between the bone marrow and the dura might, under homeostatic conditions, represent a principal route for the migration of monocytes and neutrophils. In our skull transplantation model, we detected GFP⁺ vascular elements sprouting from the bone marrow niche and extending toward the dural parenchyma. It is therefore reasonable to conceive of a scenario wherein myelomonocytic cells crawl abluminally along these vessels using the vasculature as a scaffold, as is the case of neuroblasts during CNS development (34), glioma tumor spread in the brain parenchyma (35), or T cells finding their zone in the spleen (36).

Beyond homeostatic trafficking, we demonstrated that in the context of either injury or

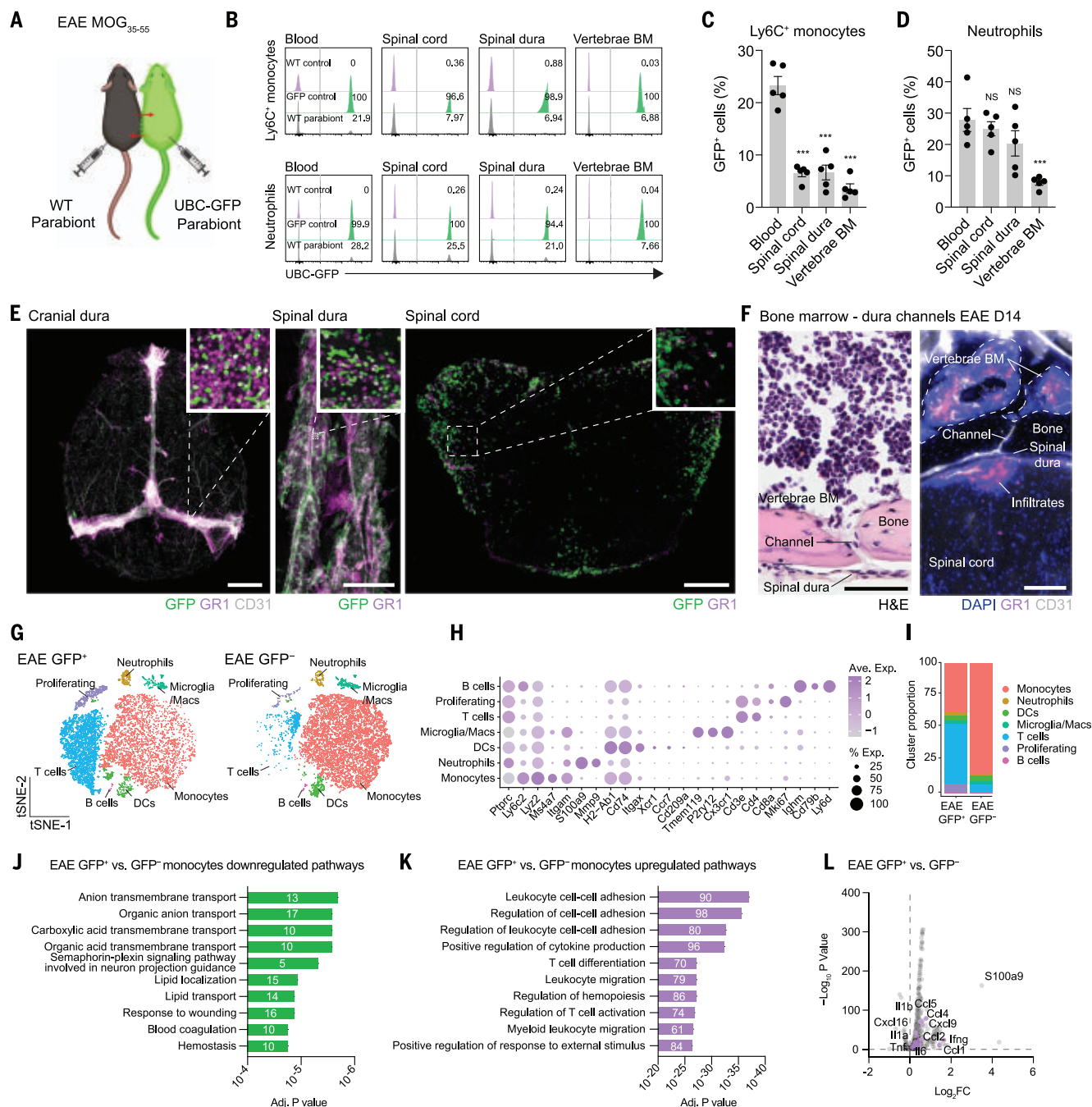


Fig. 3. The inflamed CNS is infiltrated by blood and CNS-adjacent bone marrow-derived myeloid cells. (A) Experimental design for EAE induction in parabiotic mice. Sixty days after WT and UBC-GFP mice were paired, EAE was induced in both mice by MOG₃₅₋₅₅ immunization, and the WT mice were analyzed 15 days after induction. (B) Representative histograms of GFP⁺ and GFP⁺ Ly6C⁺ monocytes and neutrophils across tissues. WT and UBC-GFP control mice without intravenous anti-CD45 antibody injections were used to assign gates. (C and D) Quantification of flow cytometry analysis of IV-CD45⁺GFP⁺Ly6C⁺ monocytes and GFP⁺ neutrophils in EAE-induced WT parabionts. NS indicates not significant; ****P* < 0.001 versus blood (one-way ANOVA with Dunnett's post hoc test); *n* = 5 mice per tissue, representative of two independent experiments. Data are means ± SEM. (E) Representative images from three mice of cranial dura, spinal dura, and spinal cord sections of WT parabionts 15 days after the EAE induction. Only a minor portion of GFP⁺ cells are costained with GR1. Scale bars: 2000 μm (left and middle) and 200 μm (right). (F) Representative H&E and IHC images from five mice of vascular channels found in vertebrae connecting bone marrow to underlying spinal

dura. On the left, cells can be seen within channels connecting bone marrow and spinal dura. On the right, GR1⁺ infiltrates are closely associated with the CD31⁺ vascular channel. Scale bars: 100 μ m. **(G)** t-distributed stochastic neighbor embedding (tSNE) visualizations of color-coded scRNA-seq analysis based on cell types for CD45^{hi}IV-CD45⁻GFP⁻ or GFP⁺ populations from the spinal cord of WT parabionts 15 days after EAE induction. $n = 3$ pooled mice per sample. Macs, macrophages. **(H)** Dot plots demonstrating scaled gene expression and percentage of cells expressing these genes for cluster phenotyping markers. **(I)** Cluster distributions detailing proportions of cell types in IV-CD45^{hi}CD45⁻GFP⁻ and GFP⁺ samples from scRNA-seq analysis. **(J and K)** Top 10 down-regulated and up-regulated gene ontology biological process terms describing differentially expressed genes between GFP⁺ and GFP⁻ monocytes from scRNA-seq analysis. Numbers in bar graphs represent the number of differentially expressed genes belonging to that pathway. **(L)** Volcano plot showing differentially expressed genes between GFP⁺ and GFP⁻ monocytes from scRNA-seq analysis. Inflammatory chemokines and cytokines are highlighted. FC, fold change.

autoimmune neuroinflammation, monocytes can invade the brain and spinal cord parenchyma without the need for blood, and thus breakdown of the blood-brain barrier. In this sense, the CNS seems to behave as a distinctive organ, in that it is able to hold a privileged dialogue with its surrounding bone marrow niches. Indeed, we envisage these monocytes and neutrophils as a myeloid reservoir situated at the borders of the CNS under homeostatic conditions and ready to leap to its defense when needed, thereby complementing blood-borne inflammatory cells in supporting the CNS. Whether dura-derived chemokines differentially favor migration from local bone marrow over blood or, indeed, whether the two routes use different recruitment factors entirely remains to be clarified. Intriguingly, this could allow for the manipulation of specific subsets of myeloid cells, with potentially distinctive phenotypes. From our scRNA-seq analysis, it appears that infiltrating Ly6c⁺ cells from the blood and the bone marrow may have distinct roles in EAE pathology, potentially through differences in immune cell recruitment or exacerbation of inflammatory events, though their exact roles remain to be seen. It will also be intriguing to further explore the kinetics of blood- and CNS-associated bone marrow-derived myeloid infiltration during EAE development, to explore whether one population represents the “first responders” to CNS disruption.

The skull and vertebral bone marrow-derived meningeal myeloid reservoir described here reshapes our interpretation of the neuroimmunological events occurring at the CNS borders under both physiological and pathological conditions. It appears likely that this reservoir of immune cells may also be of relevance in other scenarios of brain pathology, such as neurodegeneration, brain tumors, CNS or meningeal infection, and aging, or in the context of microglial replacement under specific ablative regimens.

Materials and methods

Mice

Mice were housed in the animal facility with four to five animals per cage in a temperature- and humidity-controlled environment with a 12-hour light-dark cycle. Mice were provided with rodent chow and water *ad libitum*. All mice were housed for habituation for at least 1 week before the start of the experiment. Unless specified otherwise, mice were 2 to 3 months of age at the beginning of the experiments. The following strains were used: C57BL/6 (WT; JAX 000664), C57BL/6-Tg(UBC-GFP)30Scha/J (UBC-GFP; JAX 004353), B6.SJL-Ptprca Pepcb/BoyJ (CD45.1; JAX 002014), Lyz2 tm1(CRE^{ER-T2})Grtn/J (LysM-CRE^{ER-T2}; JAX 031674), B6.Cg-Gt(ROSA)26Sor tm6(CAG-ZsGreen1)Hze/J (ZsGreen, JAX 007906), Prox1-CRE^{ER-T2} (JAX 022075), B6.Cg-Gt(ROSA)26Sor tm14(CAG-tdTomato)

Hze/J (Ai14; JAX 007914), and Tg(CAG-KikGR)33Hadj/J (KikGR, JAX 013753). Mice were bred in-house or purchased from the Jackson Laboratory. All experiments were approved by the Institutional Animal Care and Use Committee of the University of Virginia and/or Washington University in St. Louis. Experiments were only performed in institutions for which experimental approval was granted.

Parabiosis

Mice were anesthetized with an intraperitoneal injection of a ketamine [100 mg per kilogram of body weight (mg/kg)]-xylazine (10 mg/kg) cocktail. Fur was shaved on the corresponding lateral aspect of each mouse and cleaned with alternating betadine solution and alcohol swabs. After muscles were fully relaxed, matching skin incision from the olecranon to the knee joint was made on each side of the incision, the subcutaneous fascia was bluntly dissected to create 0.5 cm of free skin. To reduce direct pulling on the skin sutures, the olecranon and knee were sutured together by an absorbable 5-0 Vicryl suture. After pressing together, the dermis of the parabiotic partners, skin incisions were closed with 5-0 Nylon suture. After surgery, systemic analgesics (2.5 mg of ketoprofen per kilogram of body weight) and antibiotics [2.5 mg of Baytril (Bayer) per kilogram of body weight] were administered. One single parabiotic pair was housed per cage, wet food was provided, and mice were monitored daily for pain and distress. Unless specified otherwise, all parabiotic pairs were sacrificed after 60 days, disjointed, and separately perfused.

Tissue collection and processing for IHC

Mice were sacrificed using a lethal dose of anesthetics (Euthasol, 10% v/v) by intraperitoneal injection and transcardially perfused with 0.025% heparin in phosphate-buffered saline (PBS) solution. Cranial meninges and brain were harvested by immediate decapitation posterior to the occipital bone; skin, mandible, and muscles were removed; and skull was fixed in 4% paraformaldehyde (PFA) at 4°C for 24 hours. After 24 hours, the meninges were peeled from the skull cup and placed in PBS with 0.25% sodium azide at 4°C until used. Brains were fixed for additional 24 hours in PFA at 4°C, transferred in PBS with 30% sucrose until fully immersed, and then embedded in optimal cutting temperature (OCT) compound (Fisher), frozen rapidly over dry ice, and stored at -20°C until use. Forty-micrometer-thick sections (coronal) were sliced on a cryostat (Leica). Calvarium and femur were fixed in 4% PFA at 4°C for 24 hours, transferred in decalcification solution (PBS with 0.1 M of EDTA) at agitation for 4 to 6 days at 37°C with daily change of the solution, and then transferred in PBS with 30% sucrose over-

night. Calvarium and femur were then embedded in OCT, frozen rapidly over dry ice, and stored at -20°C until use. Thirty-micrometer-thick sections (coronal or sagittal) were sliced onto gelatin-coated slides on a cryostat (Leica). Vertebral laminae were cut on both sides, vertebral foramen was exposed, and spinal cord was dissected using fine forceps and fixed in 4% PFA at 4°C for 48 hours. Meninges were cut along the ventral side, and whole-mount meninges were peeled from spinal cord and placed in PBS with 0.25% sodium azide at 4°C until used. Spinal cord was transferred in PBS with 30% sucrose overnight, embedded in OCT, frozen rapidly over dry ice, and stored at -20°C until use. Twenty-five-micrometer-thick sections were sliced onto gelatin-coated slides on a cryostat (Leica). Liver, lung, and lymph nodes were harvested using forceps, placed in 4% PFA overnight, transferred into PBS containing 30% sucrose overnight, and then embedded in OCT and frozen rapidly over dry ice and stored at -20°C until use. Twenty-five-micrometer-thick sections were sliced onto gelatin-coated slides on a cryostat (Leica). For immunofluorescence staining, tissues were rinsed in PBS and washed with PBS with 0.2% Triton X-100 for 10 min, followed by incubation in blocking and permeabilization buffer containing PBS with 0.2% Triton X-100 and 2% normal chicken serum for 1 hour at room temperature with gentle agitation. If applicable, endogenous biotin was blocked using a streptavidin-biotin blocking kit (Vector Laboratories SP-2002) following the manufacturer's instructions. To quench autofluorescence of femur and skull bone marrow, Trueblack (Biotium, 23007) was used following the manufacturer's instructions. Sections were then incubated with primary antibodies in the blocking and permeabilization buffer for 24 hours at 4°C with gentle agitation, washed three times with PBS with 0.2% Triton X-100, and incubated with secondary antibodies in blocking and permeabilization buffer for 2 hours at room temperature. Slices were washed three times with PBS containing 0.2% Triton X-100, and nuclei were then counterstained with 4',6-diamidino-2-phenylindole (DAPI). Free-floating sections were mounted on Superfrost Plus slides and coverslipped with Aqua-Mount (Lerner) or ProLong Gold Antifade Mountant (Molecular Probes) and glass coverslips. Gelatin-coated slides were mounted with Aqua-Mount (Lerner) and glass coverslips.

For H&E staining, formalin fixed spines were embedded in paraffin, and 10-μm-thick sections were cut on a microtome (Leica), mounted onto slides, deparaffinized, and stained with Hematoxylin and Eosin Stain Kit (Vector Laboratories H-3502) according to the manufacturer's instructions.

For cleared skull-dura whole-mount imaging, the vDISCO tissue clearing and processing protocol was used as described (23). Briefly,

after perfusion, skull with attached meninges was dissected and placed in PBS 4% PFA overnight, followed by decolorization solution for 24 hours. After washing with PBS, samples were placed in decalcification solution [10% w/v EDTA (Carl Roth, 1702922685) in 0.1 M PBS] for 2 days at room temperature. Samples were then permeabilized, stained, and cleared using the vDISCO protocol with passive tissue immersion as described (23). Cleared samples were immersed in ethyl cinnamate (Sigma, W243019) and placed in a chambered coverglass (Thermo, 155360) for imaging.

Confocal microscopy and image analysis

Tissues slides were stored at 4°C for no more than 1 week until images were acquired using confocal microscopy (Leica, TCS SP8 or Leica, Stellaris) with 10× (0.4 NA), 20× (0.75 NA), and 40× (1.3 NA) objectives (Leica) and widefield microscopy (Olympus Slideview VS200) with 10× (0.4 NA), 20× (0.8 NA), and 40× (0.95 NA) objectives (Olympus) and a Hamamatsu ORCA Flash4.0 V3 digital camera. Quantitative analysis of acquired images was performed using Fiji package for ImageJ. Five to 15 representative images were acquired per sample, and the results were averaged to generate the value used for a single mouse.

Single-cell isolations

Mice were injected with a lethal dose of Euthasol (10% v/v) and transcardially perfused with ice-cold PBS with 0.025% heparin. Blood was collected before perfusion from the retro-orbital sinus. Blood was centrifuged at 420g for 4 min. Red blood cell lysis was then performed by resuspension in 1 ml of ammonium-chloride-potassium (ACK) lysis buffer (Quality Biological) for 10 min at room temperature, and 2 ml of PBS was added to inactivate ACK buffer. After centrifugation at 420g for 4 min, samples were resuspended in fluorescence-activated cell sorting (FACS) buffer [PBS with 2% bovine serum albumin (BSA)] and kept at 4°C until use. Tissues were harvested as described above. Briefly, dura and lymph nodes were digested for 15 min and lung, liver, and skin were digested for 30 min at 37°C in prewarmed digestion buffer [Dulbecco's minimum essential medium (DMEM) with 1 mg/ml Collagenase VIII (Sigma-Aldrich) supplemented with 0.5 mg/ml of deoxyribonuclease I (DNase I) (Sigma-Aldrich), 2% fetal bovine serum (FBS)]. After digestion, tissues underwent several steps of mechanical homogenization followed by filtration through a 70-µm cell strainer. Enzymes were inactivated with DMEM with 10% FBS, and cells were then centrifuged at 420g for 4 min and washed in FACS buffer. Single-cell suspensions were kept in FACS buffer on ice until use. After dissecting the meninges, spinal cord and brains were harvested, followed by a digestion step for 30 min in pre-

warmed digestion buffer, triturated with a 10-ml serological pipette, digested for an additional 20 min, and again triturated with a 5-ml serological pipette. Cells were passed through a 70-µm cell strainer, enzymes were inactivated by adding an equal volume of DMEM with 10% FBS, and samples were centrifuged at 420g for 4 min. Myelin was removed by adding a 1:1 ratio of 22% BSA in PBS to DMEM, and samples were centrifuged at 1000g for 10 min, with no brake. The upper myelin-containing layer and underlying BSA were aspirated, and the single-cell suspension was kept in FACS buffer until use. For the isolation of skull bone marrow, meninges were peeled from the skull cup as described above. The calvarium was then cut into small pieces using sterile scissors and mechanically dissociated in FACS buffer with a pestle, followed by a filtration step through a 70-µm cell strainer. After centrifugation at 420g for 4 min, red blood cells were removed by adding 1 ml of ACK lysis buffer for 2 min at room temperature and centrifugation at 420g for 4 min, and the cell pellet was resuspended in FACS buffer until use. Femur bone marrow was harvested by dissecting the left femur. The proximal and distal epiphysis were cut using sterile scissors, and bone marrow was then flushed out from the femur using a 23-gauge needle filled with PBS with 2% BSA onto a 70-µm cell strainer and gently pressed. After centrifugation, red blood cells were lysed with 1 ml ACK lysis buffer for 2 min, centrifuged at 420g for 4 min, and resuspended in ice-cold FACS buffer until use.

Flow cytometry and FACS

Single-cell suspensions were washed and resuspended in FACS buffer. Fc receptors were blocked for 5 min on ice using an anti-CD16/CD32 antibody cocktail (Fc block; 2 µg/ml, Biolegend), followed by a 10-min incubation at room temperature with antibody cocktails for surface staining. The list of flow antibodies is provided in table S5. To determine cell viability, a Zombie Aqua or Zombie NIR fixable Viability kit was used following the manufacturer's instructions. Samples were acquired with a Gallios flow cytometer (Beckman Coulter) or an Aurora spectral flow cytometer (Cytek). Data were analyzed using FlowJo version 10. For FACS, a BD FACSaria II (BD Biosciences) was used.

Intravenous immune labeling

To label intravascular leukocytes, mice were retro-orbitally injected with 2 µg of phycoerythrin (PE)-conjugated anti-CD45 antibody (Biolegend, clone 30-F11) under isoflurane anesthesia 2 min before euthanasia. In all instances, proper intravenous administration was confirmed by complete labeling of blood leukocytes.

Tamoxifen-induced Cre recombination

To induce Cre recombination in LysMCreER^{T2} mice, a 20 mg/ml stock solution of tamoxifen (Sigma Aldrich) was prepared by dissolving in corn oil at 37°C for 24 hours followed by sterile filtration. Transgenic mice were injected intraperitoneally daily with 100 µl of tamoxifen at 8 weeks of age for three consecutive days. Mice were then sacrificed at 0, 3, 14, or 30 days after the final tamoxifen injection.

Intracalvaria and intratibial bone marrow injection

All procedures were performed on mice anesthetized with a ketamine (100 mg/kg)–xylazine (10 mg/kg) mixture. Body temperature was maintained throughout the surgery by placing mice on a heating pad at 37°C while they were restrained in a stereotaxic apparatus. The head of the mouse was shaved, a skin midline incision was made, and the skull was exposed. Injections were performed similarly as described previously (21). Briefly, an electric drill was used to gently erode the compact bone covering bone marrow niches on occipital and frontal bones until a thin layer of periosteum was left. A Hamilton syringe complemented with a 34-gauge blunt needle RN 0.37 PT3 was used to manually poke the final hole for the injection. All injections were performed manually. Every injection consisted of a maximum volume of 2 µl. A single injection took around 45 s. After the injection, the needle was maintained in place for 15 s to prevent excessive leakage of the injected solution. Several injections were performed per mouse for a total final volume injected that ranged between 4 and 10 µl. Several types of injections were performed: (i) UBC-GFP progenitor cells freshly flushed from femur (roughly 10,000 cells injected per mouse), (ii) Evans Blue, (iii) CD45-PE antibody (Biolegend, clone 30-F11), and (iv) CellTracker Deep red (C34565, Molecular Probes). The Hamilton was filled directly with the concentrated antibody solution from the Biolegend/Molecular Probes vials. For intratibial bone marrow injection, a skin incision was made below the knee, and using a 30-gauge needle, bone was perforated between tuberosity and crest of tibia. Three microliters of CellTracker CMFDA (green) (C7025 Molecular Probes) dye was manually injected over 30 s. After the procedure, the skin was sutured, and mice were gently placed on a heating pad and monitored until awake. Once awake, mice were injected with ketoprofen and antibiotics (enrofloxacin, 2.5 mg/kg). After the bone marrow injections, mice were sacrificed 2 or 24 hours later.

Skull bone marrow delivery of AMD3100 was performed as previously described (37). Briefly, mice were anesthetized with a ketamine (100 mg/kg)–xylazine (10 mg/kg) mixture. The head of the mouse was shaved, a skin midline

incision was made, and skull was exposed. Using an electrical drill, the outer periosteal layer was thinned on top of skull bone marrow near the bregma and lambda on five spots, without damaging the bone marrow. One microliter of 1 mg/ml AMD3100 (ab120718, Abcam) or vehicle was applied on each spot for 5 min. The skin was sutured, and mice were sacrificed 24 hours later.

Calvarium bone-flap transplantation

Mice were anesthetized with a ketamine (100 mg/kg)-xylazine (10 mg/kg) mixture. Body temperature was maintained throughout the surgery by placing mice on a heating pad at 37°C. The heads of WT mice were shaved, skin midline incisions were made, and the skull was exposed. Using an electrical drill, a 4-mm-by-6-mm cranial window was made on parietal and interparietal bones. During bone drilling, periodic pauses were made to cool the bone with 0.9% sodium chloride solution. After the cranial window was detached from surrounding bones, craniotomy was performed with the help of fine forceps, and extra caution was taken to not damage the underlying dura mater. Craniotomy was performed with abundant 0.9% sodium chloride solution bathing the brain. If dura mater, brain, or sinuses were damaged during the craniotomy, mice were sacrificed and were not used for our analysis. In parallel, a sex-matched UBC-GFP mouse, whose skull bone marrow was used as a donor of calvarium, was euthanized, decapitated, and a piece of calvarium comparable to the window drilled in the host mouse was obtained. Connective tissue and pericranium layers were left attached on the bone flap to increase survival of the transplanted bone. Before the transplantation, meninges were peeled from the donor skull. The donor piece of skull containing the bone marrow was carefully placed within the surgical window created in the host mouse. Then, the donor skull was fixed to the host skull using cyanoacrylate glue and a trimmed 10-0 nylon suture needle. The skin was sutured, and mice were maintained on a heating pad until they awoke. Mice were observed daily and injected with ketoprofen (2.5 mg/kg) and antibiotics [2.5 mg of Baytril (Bayer) per kilogram of body weight] once per day for the first 5 days after the transplantation day.

EAE

For active induction of EAE in parabiotic pairs, both mice joined in parabiosis were immunized by subcutaneous injection of 200 µg of MOG₃₅₋₅₅ (CSBio) in complete Freund's adjuvant (Sigma Aldrich) with 4 mg/ml of *Mycobacterium tuberculosis* (H37Ra) (BD Biosciences). Mice received 200 ng of pertussis toxin (List Biological Laboratories) intraperitoneally on day 0 (immunization day) and day 1 (24 hours after immunization).

Spinal cord injury

Mice were anesthetized with a ketamine (100 mg/kg)-xylazine (10 mg/kg) mixture. Skin over the upper thoracic area was shaved and cleaned with betadine and alcohol. A 15-mm midline skin incision was made, and connective tissue and muscles were bluntly dissected. After a Th9 laminectomy the spinal cord was crushed with fine forceps. After the injury, muscles and skin were sutured separately and allowed to recover from anesthesia on heating pads. Once awake mice were injected daily intraperitoneally with buprenorphine (0.05 mg/kg) and Baytril (2.5 mg/kg) (Bayer) for three consecutive days. Seventy-two hours after surgery, mice were sacrificed for further analysis.

Skin injury

Mice in parabiosis were briefly anesthetized by placing them in an isoflurane induction chamber for 3 min, then a fast ear punch was performed on the left ear of each parabiont (WT and UBC-GFP mouse). The ear puncher was cleaned with a solution of 70% ethanol between each couple of parabionts. After the punch, mice were injected intraperitoneally with ketoprofen analgesia (2.5 mg/kg) and placed back in their cages to wake. Mice were sacrificed 24 hours after skin injury.

Optic nerve injury

Mice in parabiosis were anesthetized by placing them in an isoflurane induction chamber for 3 min. The soft tissue around the eye was dissected to expose the optic nerve. The optic nerve was crushed with an N5 self-closing forceps. Mice were injected intraperitoneally with ketoprofen analgesia (2.5 mg/kg), placed back in their cages, and allowed to recover on a heating pad. Mice were sacrificed 24 hours after optic nerve injury.

Irradiation and bone marrow transplantation

All mice received split dose irradiation with a first dose of 550 rads [5.5 gray (Gy)] followed by another dose of 550 rads (5.5 Gy) 6 to 7 hours later. For full-body irradiation, mice were irradiated without anesthetics. For head shielding, mice were anesthetized with a ketamine (100 mg/kg)-xylazine (10 mg/kg) mixture. Mice heads were shielded using a 1-inch-thick lead shield (1-inch Lead Vial Shield, 50 ml, Pinestar Technology), exposing the body below the neck. Alternatively, for body shielding, mice were inserted into a lead shield exposing only the head. All mice were injected intravenously with 4×10^6 (100 µl in PBS) freshly isolated bone marrow cells from UBC-GFP donors 6 hours after the last irradiation. Mice were analyzed 4 weeks after the bone marrow transplantation.

EdU and Ki67 staining for proliferation analysis

Mice received two intraperitoneal injections of 10 mg of EdU per kilogram of body weight

24 hours apart and were sacrificed 24 hours after the final injection. After generation of single-cell suspensions for flow cytometry and surface staining as described earlier, fixation, permeabilization, and EdU staining were performed following the manufacturer's instruction (Click-iT Plus EdU Alexa Fluor 488 Flow Cytometry Assay Kit, C10632, ThermoFisher Scientific). Intracellular staining for Ki67 and additional staining for PE conjugates was performed for 10 min at room temperature after EdU staining.

Photoconversion of KikGR

Mice were anesthetized with a ketamine (100 mg/kg)-xylazine (10 mg/kg) mixture. The heads were shaved, skin midline incisions were made, and the skull was exposed. Photoconversion was performed using an Optogenetics-LED-Violet module (Prizmatix) with the light source placed 1 cm above the skull for 2 min at an intensity of 5. Mice were then sutured and allowed to recover on a heating pad until responsive.

Proteome profiler chemokine array

Chemokines were semiquantitatively determined from 2-month-old WT mouse dural homogenates using the Proteome Profiler Mouse Chemokine Array Kit (R&D Systems) as per manufacturer's instructions. Briefly, peeled cranial dura was peeled from the skull and placed into PBS with 1X Complete Mini ETA Protease inhibitor cocktail (Sigma Aldrich) and homogenized using a Mini Bead-beater (BioSpec Products) and 2.3-mm zirconia-silica beads (BioSpec Products). Triton X-100 was added to a final concentration of 1%, were frozen at -80°C, thawed, and centrifuged at 1000g for 10 min to pellet tissue. Three pooled dural homogenates were used for each sample, and samples were run in duplicate. Semiquantitative analysis was performed by densitometric measurements using FIJI.

EAE scRNA-seq

Three pairs of female WT and UBC-GFP C57BL/6J mice were parabiotically joined at 2 months of age. Mice were allowed to recover for 60 days, whereafter EAE was induced in both pairs by MOG₃₅₋₅₅ immunization. After 15 days, mice were sacrificed, and spinal cord tissue was collected from the WT parabionts. Tissues were dissociated, stained, and sorted using a BD FACSaria II (BD Biosciences) based on the following phenotypes: DAPI⁺CD45^{hi}IV-CD45⁻GFP⁻ singlets or DAPI⁺CD45^{hi}IV-CD45⁻GFP⁺ singlets. Each sample represents GFP⁻ or GFP⁺ subsets from three pooled female mice at 4 to 5 months of age. Cells were sorted into 1.5-ml tubes with DMEM, pelleted, viability determined using trypan blue exclusion, and resuspended in 0.04% nonacetylated BSA. Sample

loading and library construction were performed using the 10X Genomics Chromium platform and Chromium Single Cell 3' Library and Gel Bead Kit v3 as previously described (24). Data preprocessing, dimensionality reduction and clustering, differential expression analysis, and pathway enrichment were performed as previously described (24).

Spinal cord injury scRNA-seq

Four pairs of female WT and UBC-GFP C57BL/6J mice were parabiotically joined at 2 months of age. Mice were allowed to recover for 60 days, whereafter spinal cord injury was performed in both pairs by a crush at Th9 level with fine forceps. After 3 days, mice were sacrificed, and spinal cord tissue was collected from the WT parabionts. Tissues were dissociated, stained, and sorted using a BD FACSaria II (BD Biosciences) based on the following phenotypes: DAPI⁺CD45^{hi}IV-CD45⁻GFP⁻ singlets or DAPI⁺CD45^{hi}IV-CD45⁻GFP⁺ singlets. Each sample represents pooled GFP⁻ or GFP⁺ subsets from four pooled female mice at 4 to 5 months of age. Cells were sorted into 1.5-ml tubes with DMEM, pelleted, viability determined using trypan blue exclusion, and resuspended in 0.04% nonacetylated BSA. Sample loading and library construction were performed using the 10X Genomics Chromium platform and Chromium Single Cell 3' Library and Gel Bead Kit v3 as previously described (24). Data preprocessing, dimensionality reduction and clustering, differential expression analysis, and pathway enrichment were performed as previously described (24).

Whole-dura scRNA-seq analysis

Whole-dura scRNA-seq analysis was performed on samples as previously described (24). To determine potential cell-cell ligand-receptor interactions promoting monocyte and neutrophil infiltration, the RNAMagnet package (38) with RNAMagnetSignaling was used, and the top signaling pair molecules were examined for both monocytes and neutrophils with every other cell type present in the dataset as previously described (24).

Experimental study design

All experiments were blinded, where possible, for at least one of the independent experiments. For irradiation and bone marrow transfer experiments and skull transplantation studies, blinding was not possible because of obvious differences in the physical appearance of mice. No data were excluded, with the exception of skull transplantation studies, where mice were excluded if they displayed no GFP signal in the skull bone marrow after 7 or 30 days, demonstrating that the surgery was unsuccessful. Animals from different cages were randomly assigned to different experimental groups. Statistical methods were not

used to recalculate or predetermine study sizes but were based on similar experiments previously published (24, 39, 40, 41). Statistical tests for each experiment are provided in the figure legends. All data represent biological replicates. Statistical analysis was performed using Prism 8.0 (GraphPad Software, Inc.).

REFERENCES AND NOTES

- K. Bassler, J. Schulte-Schrepping, S. Wamstatter, A. C. Aschenbrenner, J. L. Schultze, The myeloid cell compartment—Cell by cell. *Annu. Rev. Immunol.* **37**, 269–293 (2019). doi: [10.1146/annurev-immunol-042718-041728](https://doi.org/10.1146/annurev-immunol-042718-041728); pmid: 30649988
- M. Guillemins, A. Mildner, S. Yona, Developmental and functional heterogeneity of monocytes. *Immunity* **49**, 595–613 (2018). doi: [10.1016/j.immuni.2018.10.005](https://doi.org/10.1016/j.immuni.2018.10.005); pmid: 30332628
- L. G. Ng, R. Ostuni, A. Hidalgo, Heterogeneity of neutrophils. *Nat. Rev. Immunol.* **19**, 255–265 (2019). doi: [10.1038/s41577-019-0141-8](https://doi.org/10.1038/s41577-019-0141-8); pmid: 30816340
- S. Gordon, A. Plüddemann, Tissue macrophages: Heterogeneity and functions. *BMC Biol.* **15**, 53 (2017). doi: [10.1186/s12915-017-0392-4](https://doi.org/10.1186/s12915-017-0392-4); pmid: 28662662
- F. Ginhoux et al., Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science* **330**, 841–845 (2010). doi: [10.1126/science.1194637](https://doi.org/10.1126/science.1194637); pmid: 20966214
- K. Kierdorf et al., Microglia emerge from erythromyeloid precursors via Pu.1- and Irf8-dependent pathways. *Nat. Neurosci.* **16**, 273–280 (2013). doi: [10.1038/nn.3318](https://doi.org/10.1038/nn.3318); pmid: 23334579
- C. Schulz et al., A lineage of myeloid cells independent of Myb and hematopoietic stem cells. *Science* **336**, 86–90 (2012). doi: [10.1126/science.1219179](https://doi.org/10.1126/science.1219179); pmid: 22442384
- M. Guillemins, C. L. Scott, Does niche competition determine the origin of tissue-resident macrophages? *Nat. Rev. Immunol.* **17**, 451–460 (2017). doi: [10.1038/nri.2017.42](https://doi.org/10.1038/nri.2017.42); pmid: 28461703
- B. Ajami, J. L. Bennett, C. Krieger, W. Tetzlaff, F. M. V. Rossi, Local self-renewal can sustain CNS microglia maintenance and function throughout adult life. *Nat. Neurosci.* **10**, 1538–1543 (2007). doi: [10.1038/nn2014](https://doi.org/10.1038/nn2014); pmid: 18026097
- M. R. P. Elmore et al., Colony-stimulating factor 1 receptor signaling is necessary for microglia viability, unmasking a microglia progenitor cell in the adult brain. *Neuron* **82**, 380–397 (2014). doi: [10.1016/j.neuron.2014.02.040](https://doi.org/10.1016/j.neuron.2014.02.040); pmid: 24742461
- J. C. Cronk et al., Peripherally derived macrophages can engraft the brain independent of irradiation and maintain an identity distinct from microglia. *J. Exp. Med.* **215**, 1627–1647 (2018). doi: [10.1084/jem.20180247](https://doi.org/10.1084/jem.20180247); pmid: 29643186
- F. C. Bennett et al., A combination of ontogeny and CNS environment establishes microglial identity. *Neuron* **98**, 1170–1183.e8 (2018). doi: [10.1016/j.neuron.2018.05.014](https://doi.org/10.1016/j.neuron.2018.05.014); pmid: 29861285
- J. Herz, A. J. Filiano, A. Smith, N. Yogev, J. Kipnis, Myeloid cells in the central nervous system. *Immunity* **46**, 943–956 (2017). doi: [10.1016/j.immuni.2017.06.007](https://doi.org/10.1016/j.immuni.2017.06.007); pmid: 28636961
- M. J. C. Jordão et al., Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. *Science* **363**, eaat7554 (2019). doi: [10.1126/science.aat7554](https://doi.org/10.1126/science.aat7554); pmid: 30679343
- S. Mundt et al., Conventional DCs sample and present myelin antigens in the healthy CNS and allow parenchymal T cell entry to initiate neuroinflammation. *Sci. Immunol.* **4**, eaau8380 (2019). doi: [10.1126/sciimmunol.aau8380](https://doi.org/10.1126/sciimmunol.aau8380); pmid: 30679199
- D. Mridjen et al., High-dimensional single-cell mapping of central nervous system immune cells reveals distinct myeloid subsets in health, aging, and disease. *Immunity* **48**, 380–395.e6 (2018). doi: [10.1016/j.immuni.2018.01.011](https://doi.org/10.1016/j.immuni.2018.01.011); pmid: 29426702
- S. G. Utz et al., Early fate defines microglia and non-parenchymal brain macrophage development. *Cell* **181**, 557–573.e18 (2020). doi: [10.1016/j.cell.2020.03.021](https://doi.org/10.1016/j.cell.2020.03.021); pmid: 32259484
- H. Van Hove et al., A single-cell atlas of mouse brain macrophages reveals unique transcriptional identities shaped by ontogeny and tissue environment. *Nat. Neurosci.* **22**, 1021–1035 (2019). doi: [10.1038/s41593-019-0393-4](https://doi.org/10.1038/s41593-019-0393-4); pmid: 31061494
- B. Korin et al., High-dimensional, single-cell characterization of the brain's immune compartment. *Nat. Neurosci.* **20**, 1300–1309 (2017). doi: [10.1038/nn.4610](https://doi.org/10.1038/nn.4610); pmid: 28758994
- B. Ajami et al., Single-cell mass cytometry reveals distinct populations of brain myeloid cells in mouse neuroinflammation and neurodegeneration models. *Nat. Neurosci.* **21**, 541–551 (2018). doi: [10.1038/s41593-018-0100-x](https://doi.org/10.1038/s41593-018-0100-x); pmid: 29507414
- F. Herisson et al., Direct vascular channels connect skull bone marrow and the brain surface enabling myeloid cell migration. *Nat. Neurosci.* **21**, 1209–1217 (2018). doi: [10.1038/s41593-018-0213-2](https://doi.org/10.1038/s41593-018-0213-2); pmid: 30150661
- H. Yao et al., Leukaemia hijacks a neural mechanism to invade the central nervous system. *Nature* **560**, 55–60 (2018). doi: [10.1038/s41586-018-0342-5](https://doi.org/10.1038/s41586-018-0342-5); pmid: 30022166
- R. Cai et al., Panoptic imaging of transparent mice reveals whole-body neuronal projections and skull-meninges connections. *Nat. Neurosci.* **22**, 317–327 (2019). doi: [10.1038/s41593-018-0301-3](https://doi.org/10.1038/s41593-018-0301-3); pmid: 30598527
- J. Rustenhoven et al., Functional characterization of the dural sinuses as a neuroimmune interface. *Cell* **184**, 1000–1016.e27 (2021). doi: [10.1016/j.cell.2020.12.040](https://doi.org/10.1016/j.cell.2020.12.040); pmid: 33508229
- S. Brioschi et al., Heterogeneity of meningeal B cells reveals a lymphopoietic niche at the CNS borders. *Science* **373**, eaab9277 (2021). doi: [10.1126/science.aba9277](https://doi.org/10.1126/science.aba9277)
- S. Nowotzsch, A.-K. Hadjantonakis, Use of KikGR a photoconvertible green-to-red fluorescent protein for cell labeling and lineage analysis in ES cells and mouse embryos. *BMC Dev. Biol.* **9**, 49 (2009). doi: [10.1186/1471-213X-9-49](https://doi.org/10.1186/1471-213X-9-49); pmid: 19740427
- M. V. Russo, L. L. Latour, D. B. McGavern, Distinct myeloid cell subsets promote meningeal remodeling and vascular repair after mild traumatic brain injury. *Nat. Immunol.* **19**, 442–452 (2018). doi: [10.1038/s41590-018-0086-2](https://doi.org/10.1038/s41590-018-0086-2); pmid: 29662169
- M. Yu et al., Cranial suture regeneration mitigates skull and neurocognitive defects in craniosynostosis. *Cell* **184**, 243–256.e18 (2021). doi: [10.1016/j.cell.2020.11.037](https://doi.org/10.1016/j.cell.2020.11.037); pmid: 33417861
- G. J. Randolph, Emigration of monocyte-derived cells to lymph nodes during resolution of inflammation and its failure in atherosclerosis. *Curr. Opin. Lipidol.* **19**, 462–468 (2008). doi: [10.1097/MOL.0b013e32830d5f09](https://doi.org/10.1097/MOL.0b013e32830d5f09); pmid: 18769227
- P. Mastorakos, D. McGavern, The anatomy and immunology of vasculature in the central nervous system. *Sci. Immunol.* **4**, eaav0492 (2019). doi: [10.1126/sciimmunol.aav0492](https://doi.org/10.1126/sciimmunol.aav0492); pmid: 31300479
- A. Giladi et al., Cxcl10⁺ monocytes define a pathogenic subset in the central nervous system during autoimmune neuroinflammation. *Nat. Immunol.* **21**, 525–534 (2020). doi: [10.1038/s41590-020-0661-1](https://doi.org/10.1038/s41590-020-0661-1); pmid: 32313246
- C. A. Wagner, P. J. Roqué, J. M. Goverman, Pathogenic T cell cytokines in multiple sclerosis. *J. Exp. Med.* **217**, e20190460 (2020). doi: [10.1084/jem.20190460](https://doi.org/10.1084/jem.20190460); pmid: 31611252
- R. Shechter et al., Recruitment of beneficial M2 macrophages to injured spinal cord is orchestrated by remote brain choroid plexus. *Immunity* **38**, 555–569 (2013). doi: [10.1016/j.immuni.2013.02.012](https://doi.org/10.1016/j.immuni.2013.02.012); pmid: 23477737
- S. Bovetti et al., Blood vessels form a scaffold for neuroblast migration in the adult olfactory bulb. *J. Neurosci.* **27**, 5976–5980 (2007). doi: [10.1523/JNEUROSCI.0678-07.2007](https://doi.org/10.1523/JNEUROSCI.0678-07.2007); pmid: 17537968
- V. A. Cuddapah, S. Robel, S. Watkins, H. Sontheimer, A neurocentric perspective on glioma invasion. *Nat. Rev. Neurosci.* **15**, 455–465 (2014). doi: [10.1038/nrn3765](https://doi.org/10.1038/nrn3765); pmid: 24946761
- A. Chauveau et al., Visualization of T cell migration in the spleen reveals a network of perivascular pathways that guide entry into T zones. *Immunity* **52**, 794–807.e7 (2020). doi: [10.1016/j.immuni.2020.03.010](https://doi.org/10.1016/j.immuni.2020.03.010); pmid: 32298648
- T. L. Roth et al., Transcranial amelioration of inflammation and cell death after brain injury. *Nature* **505**, 223–228 (2014). doi: [10.1038/nature12808](https://doi.org/10.1038/nature12808); pmid: 24317693
- C. Baccin et al., Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nat. Cell Biol.* **22**, 38–48 (2020). doi: [10.1038/s41556-019-0439-6](https://doi.org/10.1038/s41556-019-0439-6); pmid: 31871321
- A. Louveau et al., CNS lymphatic drainage and neuroinflammation are regulated by meningeal lymphatic vasculature. *Nat. Neurosci.* **21**, 1380–1391 (2018). doi: [10.1038/s41593-018-0227-9](https://doi.org/10.1038/s41593-018-0227-9); pmid: 30224810
- S. Da Mesquita et al., Functional aspects of meningeal lymphatics in ageing and Alzheimer's disease. *Nature* **560**, 185–191 (2018). doi: [10.1038/s41586-018-0368-8](https://doi.org/10.1038/s41586-018-0368-8); pmid: 30046111
- S. Da Mesquita et al., Meningeal lymphatics affect microglia responses and anti-Ab immunotherapy. *Nature* **593**, 255–260 (2021). doi: [10.1038/s41586-021-03489-0](https://doi.org/10.1038/s41586-021-03489-0); pmid: 33911285

ACKNOWLEDGMENTS

We thank S. Smith for editing the manuscript and A. Impagliazzo for the graphical summary. We thank all the members of the Kipnis laboratory for their valuable comments during numerous discussions of this work. We would also like to thank McDonnell Genome Institute for processing single-cell RNA sequencing and the Flow Cytometry Core of the Department of Pathology and Immunology, School of Medicine, Washington University in St. Louis, for assistance with cell sorting. A.C. would like to dedicate this work to the memory of his father, Giorgio Cugurra. **Funding:** This work was conducted with funding from National Institutes of Health grants AT010416, AG034113, and AG057496 (J.K.). **Author contributions:** Initial conceptualization: A.C.;

Intellectual contributions: A.C., T.M., J.R., G.B., Z.P., A.D., M.K., J.H., J.K.; Methodology: A.C., T.M., J.R., G.B., Z.G., Z.P., A.D., M.K., S.B., L.S., J.H., M.C., I.S., J.K.; Formal analysis: A.C., T.M., J.R., G.B., T.D., A.D., Z.P.; Investigation: A.C., T.M., J.R., G.B., Z.G., W.B., Z.P., A.D., M.K., S.B., J.H., I.S.; Resources: S.B., L.S., M.C., J.K.; Data curation: T.D.; Writing: A.C., T.M., J.R., J.K.; Reviewing and editing: A.C., T.M., J.R., J.K.; Visualization: A.C., T.M., J.R., T.D.; Supervision: J.K.; Funding acquisition: J.K. **Competing interests:** J.K. is a member of a scientific advisory group for PureTech Health. **Data and materials availability:** scRNA-seq data are available from Gene Expression Omnibus under accession number GSE161034. All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/eabf7844/suppl/DC1
Figs. S1 to S9
Tables S1 to S5
MDAR Reproducibility Checklist
Movies S1 to S3

24 November 2020; resubmitted 16 April 2021
Accepted 24 May 2021
Published online 3 June 2021
10.1126/science.abf7844


ScienceRobotics.org

DOESN'T YOUR RESEARCH DESERVE THE BEST READERS?

Submit your research: [cts.ScienceMag.org](https://cts.sciencemag.org)

ScienceRobotics



 Twitter: @SciRobotics

 Facebook: @ScienceRobotics

RESEARCH ARTICLE SUMMARY

PHYSIOLOGY

Enterically derived high-density lipoprotein restrains liver injury through the portal vein

Yong-Hyun Han*, Emily J. Onufer, Li-Hao Huang, Robert W. Sprung, W. Sean Davidson, Rafael S. Czepielewski, Mary Wohltmann, Mary G. Sorci-Thomas, Brad W. Warner, Gwendalyn J. Randolph*

INTRODUCTION: High-density lipoprotein (HDL) participates in cholesterol homeostasis and may also have anti-inflammatory or antimicrobial roles through its interaction with numerous plasma proteins. The liver synthesizes most HDL in the body, but the intestine also produces HDL. However, a role for intestinal HDL distinct from that produced by the liver has not been identified. While remodeling its cargo, HDL particles circulate through tissue spaces, but so far, HDL trafficking within tissues has been scarcely studied.

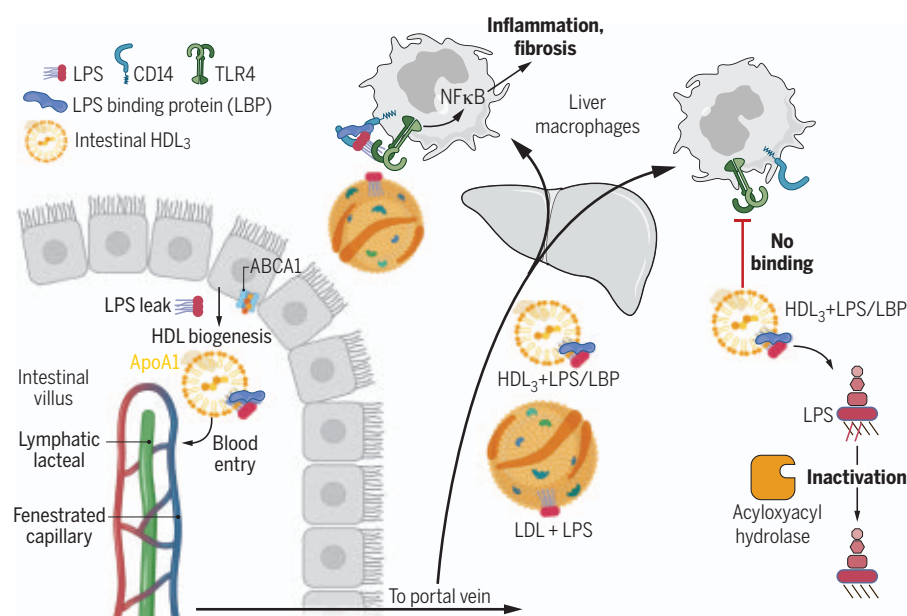
RATIONALE: We reasoned that understanding HDL-trafficking patterns might bring insight into its roles in health and disease, including whether HDL made by the intestine is functionally redundant with that produced by the liver. Using a knock-in mouse that we previously generated to phototag HDL in any tissue location, we aimed to trace the fate of HDL synthesized by the intestine.

RESULTS: Phototagged HDL derived from small bowel enterocytes was generated most abundantly by the ileum and did not travel into draining lymphatic vessels as enterocyte-derived chylomicrons do. Instead, intestinal HDL rapidly entered the portal vein, the major blood supply to the liver. This finding raised the issue of whether the liver might benefit from intestinal HDL and pointed us to an older concept that HDL might neutralize a key microbial signal that can escape a permeable gut: lipopolysaccharide (LPS) from Gram-negative bacteria. Past studies using multiple models have shown that LPS engagement of its receptor, Toll-like receptor 4 (TLR4), in the liver drives significant liver pathology, including inflammation that progresses to fibrosis. Using biochemical, proteomic, and functional approaches, we observed that the intestine produces a particular subspecies of HDL called HDL₃. Unlike another HDL subspecies (HDL₂), HDL₃ sequestered LPS so efficiently that it

could not bind to TLR4⁺ liver macrophages. In this way, HDL₃ produced by the intestine protected the liver from the inflammation and fibrosis observed in a variety of mouse models of liver injury that parallel clinically relevant conditions in humans, including surgical resection of the small bowel, alcohol consumption, or high-fat diets. Administration of an oral drug targeting the transcription factor liver X receptor, the master regulator of genes associated with HDL biogenesis, raised enteric HDL levels and protected the mice from liver pathology. This protection was lost if mice did not express enterically derived HDL, indicating that intestinal HDL was a key target of the drug. Six samples of human portal venous blood with matched systemic venous blood confirmed the enrichment of HDL₃.

Mechanistically, LPS-binding protein (LBP) was enriched in HDL₃ particles and was required for HDL₃ to mask LPS from detection by TLR4. This finding was unexpected because LBP otherwise promotes TLR4 signaling by shuttling LPS to CD14, which then shuttles it to TLR4. Thus, HDL₃ interacts with a known component of the TLR4-signaling platform, LBP, to hide LPS from detection. Without binding to TLR4, the HDL₃-LBP-LPS complex was not retained in liver. Instead, it exited the liver while the LPS associated with it was inactivated. The enzyme acyloxyacyl hydrolase, which is produced in part by liver macrophages and which deacylates critical fatty acid residues in LPS for TLR4 activation, could still access and act upon HDL₃-associated LPS to detoxify it. Low-density lipoprotein bound LPS, but not LBP, and was thus unable to prevent LPS activation of liver macrophages. LBP is in the same family of lipid-binding proteins as phospholipid transfer protein and cholesterol ester transfer protein, which have well-established roles in remodeling the lipid configuration of HDL. Another microbial lipid, lipoteichoic acid from Gram-positive bacteria, is known to bind LBP. We found that it too complexed with HDL₃ and suppressed the activation of liver macrophages.

CONCLUSION: The production of HDL by small bowel enterocytes in a form that potently masks LPS comprises a disease tolerance strategy to protect the liver from injury of enteric origin. Enteric HDL may thus be a suitable pharmacologic target for protecting the liver against gut-derived LPS leakage in alcoholic and non-alcoholic settings. ■



Trafficking and functional properties of enteric HDL. Enterocytes express ABCA1 to promote HDL biogenesis. The nascent HDL enters portal venous blood bearing LBP, which allows it to hide LPS from recognition by TLR4⁺ macrophages. Failed recognition prevents macrophage activation. Although its ability to trigger macrophages is suppressed by HDL₃, LPS in the HDL₃ complex can still be inactivated by acyloxyacyl hydrolase. Figure was drawn with BioRender.

The list of author affiliations is available in the full article online.

*Corresponding author. Email: girandolph@wustl.edu

(G.J.R.); yhan1015@kangwon.ac.kr (Y.-H.H.)

Cite this article as Y.-H. Han et al., *Science* 373, eabe6729 (2021). DOI: 10.1126/science.abe6729

READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abe6729>

RESEARCH ARTICLE

PHYSIOLOGY

Enterically derived high-density lipoprotein restrains liver injury through the portal vein

Yong-Hyun Han^{1,2*}, Emily J. Onufer³, Li-Hao Huang¹, Robert W. Sprung⁴, W. Sean Davidson⁵, Rafael S. Czepielewski¹, Mary Wohltmann¹, Mary G. Sorci-Thomas⁶, Brad W. Warner³, Gwendalyn J. Randolph^{1*}

The biogenesis of high-density lipoprotein (HDL) requires apoA1 and the cholesterol transporter ABCA1. Although the liver generates most of the HDL in the blood, HDL synthesis also occurs in the small intestine. Here, we show that intestine-derived HDL traverses the portal vein in the HDL₃ subspecies form, in complex with lipopolysaccharide (LPS)-binding protein (LBP). HDL₃, but not HDL₂ or low-density lipoprotein, prevented LPS binding to and inflammatory activation of liver macrophages and instead supported extracellular inactivation of LPS. In mouse models involving surgical, dietary, or alcoholic intestinal insult, loss of intestine-derived HDL worsened liver injury, whereas outcomes were improved by therapeutics that elevated and depended upon raising intestinal HDL. Thus, protection of the liver from injury in response to gut-derived LPS is a major function of intestinally synthesized HDL.

The portal vein collects venous drainage from the intestine, carrying nutrients and metabolites of host and microbiome origin to the liver (1). Through this route, components of the microbiome may drive liver steatohepatitis and fibrosis (2, 3). Enterically derived lipopolysaccharide (LPS) from Gram-negative bacteria triggers Toll-like receptor 4 (TLR4)-dependent injury in the liver after insult to the intestine (3–7).

Mechanisms to limit LPS-mediated liver injury through the gut-portal axis remain incompletely defined. We hypothesized that high-density lipoprotein (HDL) may have an overlooked role in protecting the liver through its potential to neutralize LPS (8–10). Indeed, it is unclear why HDL is synthesized by the intestine rather than solely by the liver. HDL-cholesterol (HDL-C) is the smallest lipoprotein particle in the blood and is best known for its role in cholesterol transport. Only two tissues produce the core protein component of HDL-C, apolipoprotein A1 (apoA1): the liver and the small intestine (11). When intestinal epithelial cells selectively delete the gene encoding the cholesterol transporter ABCA1, which is essential for HDL biogenesis, an ~25% reduction in

plasma HDL-C ensues. Conversely, an ~75% reduction in HDL-C occurs after liver-specific loss of ABCA1 (12), leaving investigators to regard the intestine as simply a second source of HDL-C.

An obstacle to considering a role for intestinal HDL in the gut-liver axis is the paucity of knowledge concerning how enterically derived HDL is delivered to the liver. HDL typically mobilizes from tissues through lymphatic vessels (13–15), which do not route to the liver from the intestine (16). However, an earlier study failed to demonstrate that enterically produced HDL-C entered lymphatics (17). Here, we show that enterically derived HDL-C alternatively traverses portal blood, and that nearly all HDL-C found in the portal vein arises from the intestine. Intestinal epithelial cells produced small HDL particles (HDL₃) (18) with potent LPS-neutralizing properties.

Results

Enterically derived HDL is the main source of HDL in portal blood

Although albumin levels were constant between portal and systemic blood (from the inferior vena cava) (fig. S1A), apoA1 was ~40% lower in portal versus systemic plasma in humans and mice (Fig. 1A). To determine whether this reduction resulted from diminished recirculation of HDL into portal blood, we traced HDL using photoactivatable green fluorescent protein (GFP) apoA1 knock-in mice (*Pgat^{KI/+}*) (15). When phototagged in the skin, HDL appeared in systemic blood but was very low in portal blood (Fig. 1B). Its appearance in mesenteric lymph to a concentration approaching that in the systemic circulation (Fig. 1B) suggested that it left the bloodstream to access lymph before entering the portal vein (Fig. 1B).

Within 5 min after phototagging HDL in the small intestinal lumen, fluorescence was strong in portal blood but not lymph (Fig. 1C). By 30 min, these compartments equilibrated (Fig. 1C). HDL phototagged along the intestinal exterior appeared in lymph but not portal blood (Fig. 1D). These patterns were unaffected by dietary composition or fasting (fig. S1, B to D). Thus, HDL tagged at the intestinal epithelium first enters the portal vein and is not observed in lymph until cargo in the portal vein passes through the liver and enters the systemic circulation. Upon reapproaching the gut, it appears to traverse into the interstitium and then into lymph.

Separate phototagging of the duodenum, jejunum, and ileum revealed the ileum as the major site of enteric HDL biogenesis (Fig. 1E). Like apoA1, HDL-C in portal plasma of humans or mice was present at lower concentrations than in systemic blood (Fig. 1F). HDL-C was decreased by >75% in portal blood of intestine-specific ABCA1-knockout mice (*Vil^{Cre}-Abca1^{fl/fl}*, *Abca1^{ΔVil}*) (Fig. 1G). However, in systemic blood of *Abca1^{ΔVil}* mice, HDL-C dropped by only 25% (Fig. 1H) (12). By contrast, a marked reduction of HDL-C in systemic but not portal blood was observed in liver-specific ABCA1-knockout mice (*Alb^{Cre}-Abca1^{fl/fl}*, *Abca1^{ΔAlb}*) (Fig. 1, G and H). Thus, two distinct blood compartments for HDL exist: one entering the portal drainage governed by intestinal production of HDL and the other in systemic vessels governed by liver production of HDL.

Portal blood HDL is mainly HDL₃ and strongly suppresses Kupffer cell inflammatory responses

Portal venous HDL was relatively small in size (~8 nm) in humans (Fig. 2A) and mice (Fig. 2B), suggestive of a subspecies of HDL called HDL₃. Small-sized HDL₃ and large-sized HDL₂ particles carry distinct accessory proteins, with paraoxonase 1 (PON1) enriched in the former and apoB in the latter (18). Thus, we compared HDL₂ or HDL₃ species separated by ultracentrifugation from pooled human systemic blood with that of the larger or smaller HDL species from portal or systemic blood isolated by size-exclusion fast protein liquid chromatography (FPLC) and affinity purification (Fig. 2C). This approach yielded four samples from the same individual, putative HDL₂ and HDL₃, each from both systemic and portal blood, allowing evaluation of how venous location affects HDL composition. Proteomic analysis revealed >250 proteins in each sample. A heatmap depicting relative abundance of proteins revealed that small-portal venous HDL shared a protein profile with HDL₃ from human systemic blood (Fig. 2D and table S1). Similarities included known enrichments of PON1, PON3, α_1 -antitrypsin (SerpinA1), and PLTP in all HDL₃ fractions regardless of portal or systemic blood origin

¹Department of Pathology, Washington University School of Medicine, St. Louis, MO 63110, USA. ²Laboratory of Pathology and Physiology, College of Pharmacy, Kangwon National University, Chuncheon 24341, South Korea.

³Department of Surgery, Washington University School of Medicine, St. Louis, MO 63110, USA. ⁴Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA. ⁵Department of Pathology and Laboratory Medicine, University of Cincinnati, Cincinnati, OH 45237, USA. ⁶Department of Medicine, Division of Endocrinology, Pharmacology and Toxicology, and Blood Research Institute, Medical College of Wisconsin, Milwaukee, WI 53226, USA.

*Corresponding author. E-mail: girandolph@wustl.edu (G.J.R.); yhan1015@kangwon.ac.kr (Y.-H.H.)

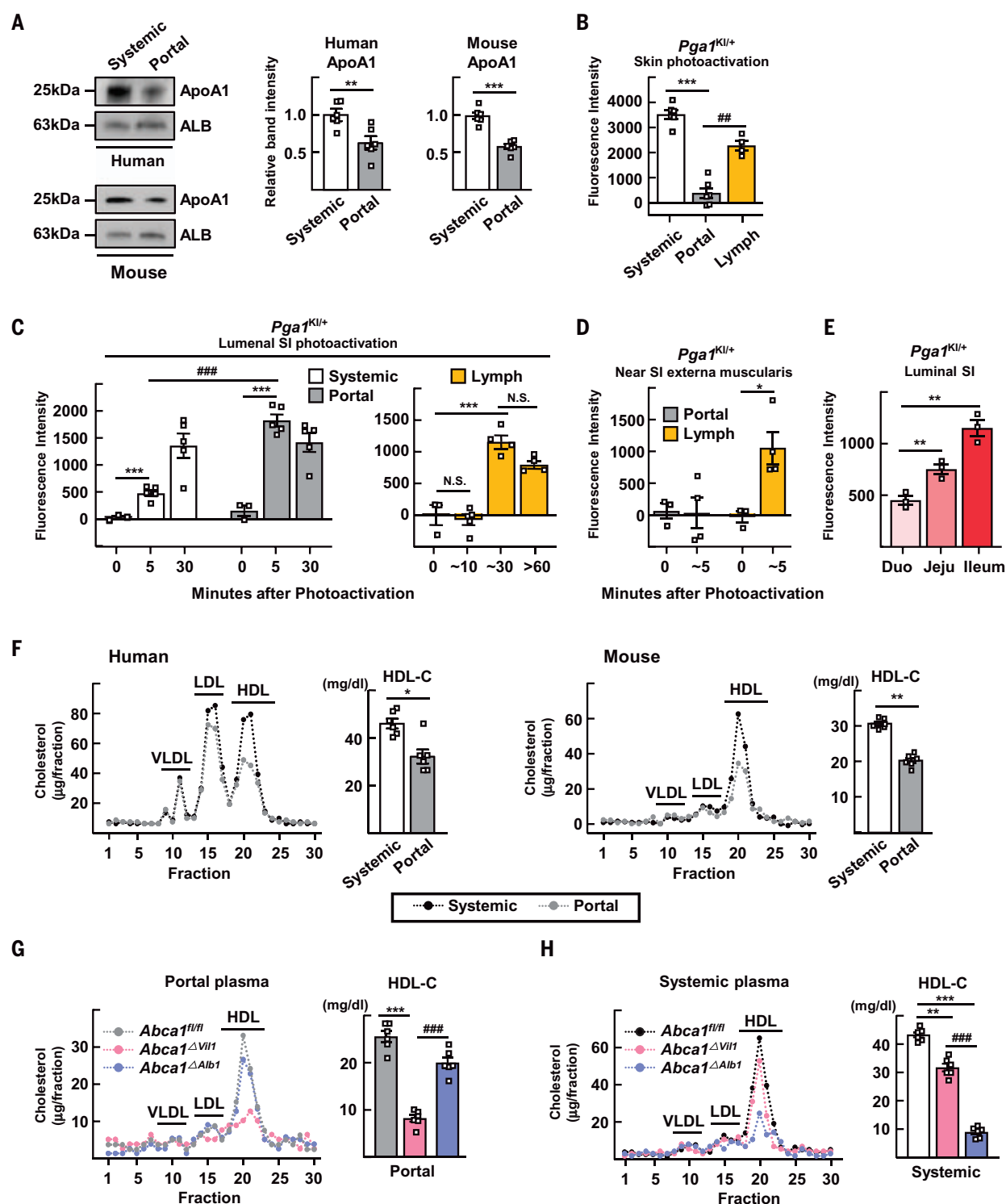


Fig. 1. Intestinal HDL transits through the portal vein and accounts for most of the HDL in the portal blood. (A) Immunoblot for apoA1 in systemic and portal serum from humans and mice. ALB, albumin. (B) Plasma and lymph fluorescence measured 2 hours after phototagging HDL in skin. (C and D) Portal plasma and mesenteric lymph fluorescence after phototagging the lumen (C) or externa muscularis (D) of the small intestine (SI) of *Pga1^{KI/+}* mice. (E) HDL phototagged in different small bowel regions separately. (F) Lipoprotein profiles from humans fasted overnight or mice fasted for 4 hours. (G and H) Lipoprotein profiles and HDL-C quantification of portal (G) or systemic (H) plasma after 4 hours of fasting. Plots show mean \pm SEM. (A) and (F) show paired data from

six human subjects or six WT mice (same subjects in both panels). (B) to (E) show data for 65 *Pga1^{KI/+}* mice. Each symbol designates different mice, except that systemic versus portal blood in (B) and systemic versus portal blood from the same time points in (C) were from the same mice (paired). Most panels show one experiment, except for (C), which combines two experiments. (G) and (H) show five individual mice from each of three genotypes from one experiment, with paired portal (G) and systemic (H) plasma within the same genotype. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, ### $P < 0.01$, #### $P < 0.001$. Data in (A) and (F) were analyzed for statistical significance using a paired t test; data in all other panels were analyzed using one-way ANOVA.

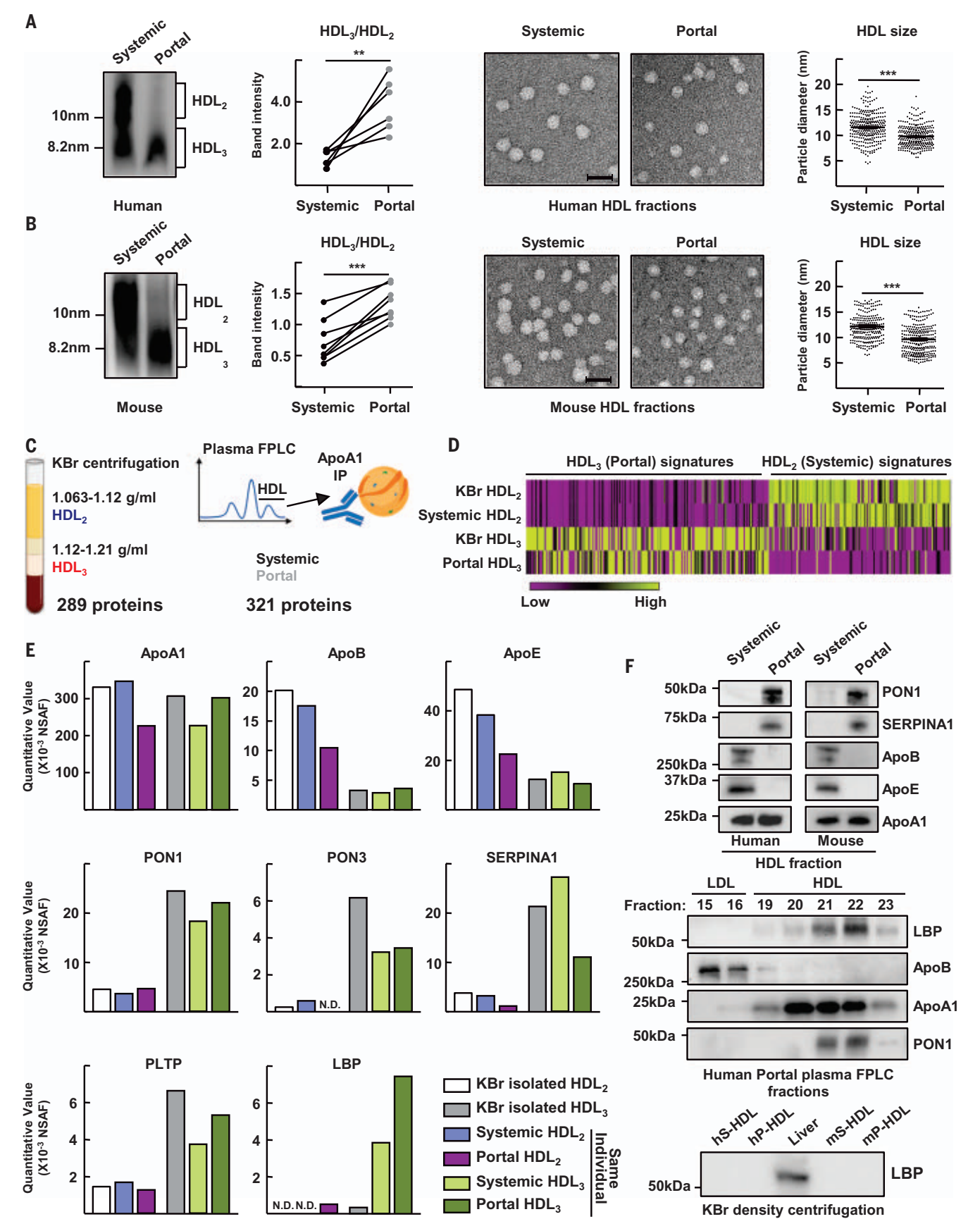


Fig. 2. Portal vein HDL is enriched in small HDL₃ particles. Samples (from Fig. 1A) of systemic or portal plasma from humans (A) or mice (B) immunoblotted for apoA1 after electrophoresis under nondenaturing conditions. Representative gels and plots show HDL₃/HDL₂ ratios in six paired samples. Also shown are representative electron microscopy images of negative-stained HDL fractions. Scale bars, 20 nm. HDL diameter measurements, plotted as individual symbols, combine assessments from four of six humans or mice (right). For proteomics (C to F), density ultracentrifugation-purified HDL₂ or HDL₃ yielded 289 associated proteins (C, left). Size-exclusion FPLC with immunopurified HDL identified 321 proteins (C, right). This experiment was performed once using four paired samples

derived from the same individual (HDL₂ or HDL₃ from two vascular beds). Additional samples were HDL₂ or HDL₃ from commercially available pooled human plasma isolated by density ultracentrifugation. (D) Heatmap of protein abundance. (E) Normalized spectral abundance factor plotted for selected proteins. (F) Select proteins immunoblotted from portal or systemic plasma HDL (top), using indicated fraction numbers collected after FPLC separation (middle). Immunoblot of LBP from human or mice systemic (hS- and mS-) and portal (hP- and mP-) plasma or whole liver lysate (bottom). LBP quantification is shown in fig. S2A. N.D., not detected. ***P* < 0.01, ****P* < 0.001. Paired *t* test was used for statistical evaluation in (A); unpaired *t* test was used in (B).

or method of isolation (Fig. 2E). Some proteins, such as LPS-binding protein (LBP), were especially enriched in immunopurified portal blood HDL₃ but notably absent after ultracentrifugation (Fig. 2E). The distribution of LBP was confirmed by immunoblotting of the FPLC fractions used to purify HDL₂ (fraction 20) versus HDL₃ (fraction 22) from human portal vein (Fig. 2F), systemic blood (fig. S2), or murine portal blood (fig. S2). When these serum samples were subjected to HDL isolation using ultracentrifugation, LBP was absent (Fig. 2F), suggesting its dissociation during ultracentrifugation (19).

Because LBP delivers LPS to CD14 to facilitate TLR4 signaling (20), we investigated whether the association of portal venous HDL₃ with LBP affects LPS signaling. We isolated primary liver Kupffer cells (KCs) (21, 22) (fig. S3A) from wild-type (WT) or *Tlr4*^{-/-} mice, and observed that portal vein-derived HDL from humans or mice neutralized LPS-induced pro-inflammatory responses in these cultures more effectively than HDL from systemic blood or no HDL at all (Fig. 3, A and B). Whole portal plasma obtained from *Abca1*^{ΔVil1} mice less effectively protected against induction of inflammation (Fig. 3C), implicating gut-derived HDL in this activity. Similarly, comparison of human HDL₂ or HDL₃ from human peripheral blood showed that HDL₃ more strongly suppressed LPS-induced inflammatory genes (Fig. 3D) than HDL₂ in a concentration-dependent, TLR4-dependent manner (Fig. 3D). Activation of KCs by lipoteichoic acid was also robust and strongly blocked by HDL₃. CpG oligodeoxynucleotides stimulated KCs less strongly, but HDL₃ did have some inhibitory effect (fig. S3B). Modest inhibition was observed when the cytokines interleukin 1β (IL-1β) or tumor necrosis factor (TNF) were used to induce inflammatory activation of KCs (fig. S3C). In these assays (Fig. 3, A to D), LBP was included as an exogenous additive to restore the LBP removed during centrifugation. Excluding LBP abrogated the anti-inflammatory effects of HDL₃ (Fig. 3E). Similarly, HDL₃ suppressed LPS bioactivity but only in the presence of LBP (Fig. 3F). Thus, portal blood HDL₃ inhibits LPS signaling in TLR4⁺ macrophages in an LBP-dependent manner.

Portal HDL₃ efficiently binds LBP and LPS to prevent LPS binding to KC TLR4

We next evaluated binding between HDL and LPS using LPS conjugated with biotin at its inner core (23–25). HDL₃ more robustly bound LPS compared with HDL₂, but the presence of LBP was required (Fig. 4A). HDL₂-bound LPS readily transferred to low-density lipoprotein (LDL) or very-low-density lipoprotein (VLDL), whereas most HDL₃-bound LPS remained associated with HDL₃ (Fig. 4B). In an immunoblotting assay, HDL₃ was indeed shown to be more effective than HDL₂ and LDL in binding LBP (Fig. 4C), and LDL did not compete (Fig. 4, C and D). LDL bound LPS efficiently (Fig. 4D) but did not bind LBP (Fig. 4C). LDL did not neutralize LPS bioactivity (Fig. 4E), nor did it dampen inflammatory gene expression (fig. S4A) with or without LBP (Fig. 4E). Reconstituted HDL also did not neutralize LPS (Fig. 4E). In the presence of LBP, interactions between biotin-LPS and HDL₃ generally reduced the detection of biotin-LPS by streptavidin capture (Fig. 4F). This reduction was caused by efficient masking of the inner core biotin label of LPS, because disruption of HDL₃ with detergent reexposed the biotin (Fig. 4F). Thus, binding of LBP to HDL₃ promotes the sequestration of LPS.

To determine whether HDL₃ sequesters LPS from KCs, we cocultured KCs with biotin-LPS and monitored surface binding (Fig. 4G). This binding largely depended upon TLR4 (Fig. 4H) and was unaffected by the absence or blockade of lipoprotein or scavenger receptors (fig. S4B). HDL₃ robustly prevented LPS interaction with KCs but only in the presence of LBP. Reconstituted HDL (Fig. 4G) or LDL in the presence or absence of LBP (fig. S4C) did not diminish binding, whereas HDL₂ partially blocked it (Fig. 4G). Peak interaction occurred within 3 hours of incubation with KCs (Fig. 4I). Substantial internalization of biotin-LPS by macrophages in the presence or absence of HDL₃ was not detected during this time (fig. S4D).

We next wondered whether, after failing to bind to cells, LPS associated with HDL₃ might later be inactivated. Using a low dose of HDL₃ (20 μg/ml) that only weakly suppressed LPS activity in a cell-free system (Fig. 3F), we observed that LPS activity was further lowered in

the presence of macrophages or macrophage supernatant. This was true as long as the KCs were not held at 4°C (Fig. 4J), suggesting that a soluble product from KCs inactivated LPS. The enzyme acylglycerol hydrolase (AOAH), which is produced in part by KCs, deacylates and thereby inactivates LPS (26). Depletion of AOAH from KC-conditioned medium increased the recovery of LPS bioactivity (Fig. 4K). Because AOAH would not remove the biotin from the inner core of LPS upon deacylation, we calculated LPS bioactivity normalized to recovered biotin after disrupting HDL₃ with detergent, finding that AOAH indeed drove inactivation of HDL₃-bound LPS (Fig. 4L). We performed a similar experiment in vivo, injecting a constant dose of biotin-LPS into the portal vein in complex with HDL₃, HDL₂, or LDL. After 30 min, when enterically derived HDL had passed through the liver to access the systemic circulation (Fig. 1C), we drew systemic venous blood to recapture and assess LPS. More of the recovered LPS was inactive when complexed with HDL₃ than with HDL₂ or LDL (Fig. 4M and fig. S4E), indicating the inactivation of HDL₃-associated LPS. The administration of LPS-loaded lipoproteins into the portal vein caused acute elevation in aspartate aminotransferase (AST), a measure that was lowest when HDL₃ was the vehicle carrying LPS and highest when LDL carried LPS (fig. S4F). Thus, HDL₃ masks LPS to limit inflammation by blocking LPS binding to KCs, but HDL₃-associated LPS remains susceptible to inactivation by AOAH, allowing its permanent inactivation.

The LPS receptor TLR4 on KCs drives liver injury and fibrosis

To determine whether HDL neutralization of LPS modulates liver inflammation in vivo, we studied a model of small bowel resection that promotes marked liver fibrosis (27, 28). Small bowel resection removed 50 or 75% of the small intestine (Fig. 5A), sparing the proximal part of the small intestine and the terminal ileum except when indicated otherwise. Within 3 months, these resections induced morphological changes in the liver (Fig. 5B), elevated the plasma AST (Fig. 5C), and enhanced the infiltration of myeloid cells (F4/80⁺

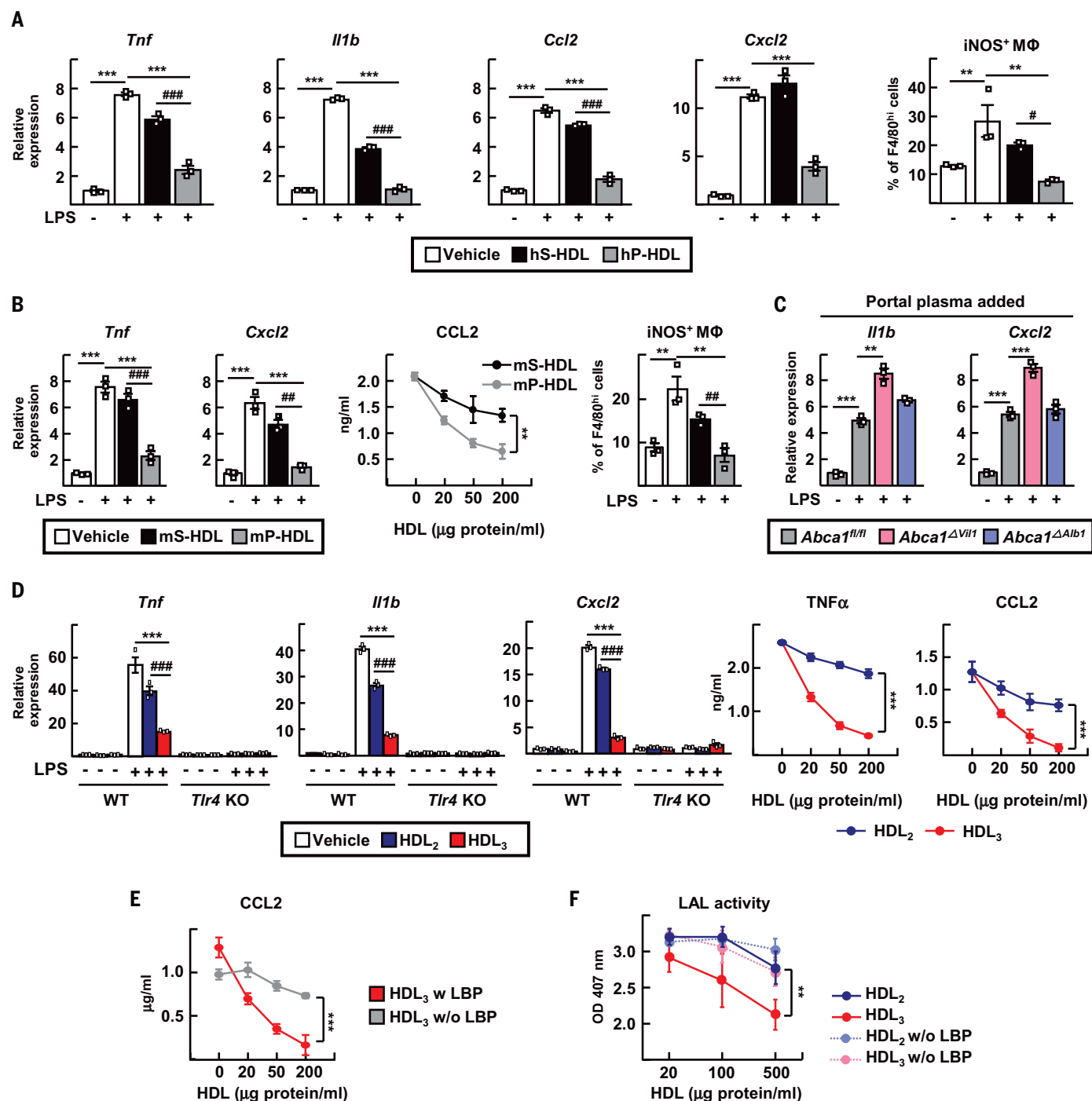


Fig. 3. Portal blood HDL₃ strongly inhibits LPS activation of KCs in an LBP-dependent manner. (A and B) LPS-treated KCs were incubated with 100 μg/ml of HDL from human (hS-HDL) or mouse (mS-HDL) systemic plasma or with human (hP-HDL) or mouse (mP-HDL) portal plasma. Analysis included RT-PCR for inflammatory mediators, flow cytometry phenotyping (iNOS⁺ F4/80^{hi} macrophages), and CCL2 ELISA. (C) RT-PCR from LPS-treated KCs incubated with 5% portal vein-derived plasma from *Abca1*^{fl/fl}, *Abca1*^{ΔVil1}, and *Abca1*^{ΔAlb1} mice. (D) WT and *Tlr4*^{-/-} KCs were incubated in LBP-containing medium with or without 20 ng/ml of LPS and 100 μg/ml of HDL₂, HDL₃, or vehicle control before RT-PCR or ELISA to detect the depicted mediators. (E) LPS-treated KCs were incubated with 100 μg/ml of HDL₃ with or without 1 μg/ml of LBP before CCL2 ELISA. (F) Endotoxin

LAL activity after 0.5 Ehrlich units (EU)/ml of *E. coli* LPS were preincubated with HDL with or without 1 μg/ml of LBP. HDL fractions in the experiments in this figure were isolated by density ultracentrifugation. Plots show mean ± SEM. Each symbol represents independent preparations of KCs from different mice. For (A) and (B), *n* = 3 different KC preparations (three experiments). For (C), *n* = 9 different KC preparations, with *n* = 3 per genotype (three experiments). In (D) and (E), *n* = 6 KC preparations (*n* = 3 each WT or *Tlr4* knockout). In (F), there were three independent technical replicates (three experiments) per condition per time point. ***P* < 0.01, ****P* < 0.001, #*P* < 0.05, ###*P* < 0.01, ####*P* < 0.001. Statistical analysis on data presented as bar graphs used one-way ANOVA; data presented as concentration curves were analyzed using two-way ANOVA.

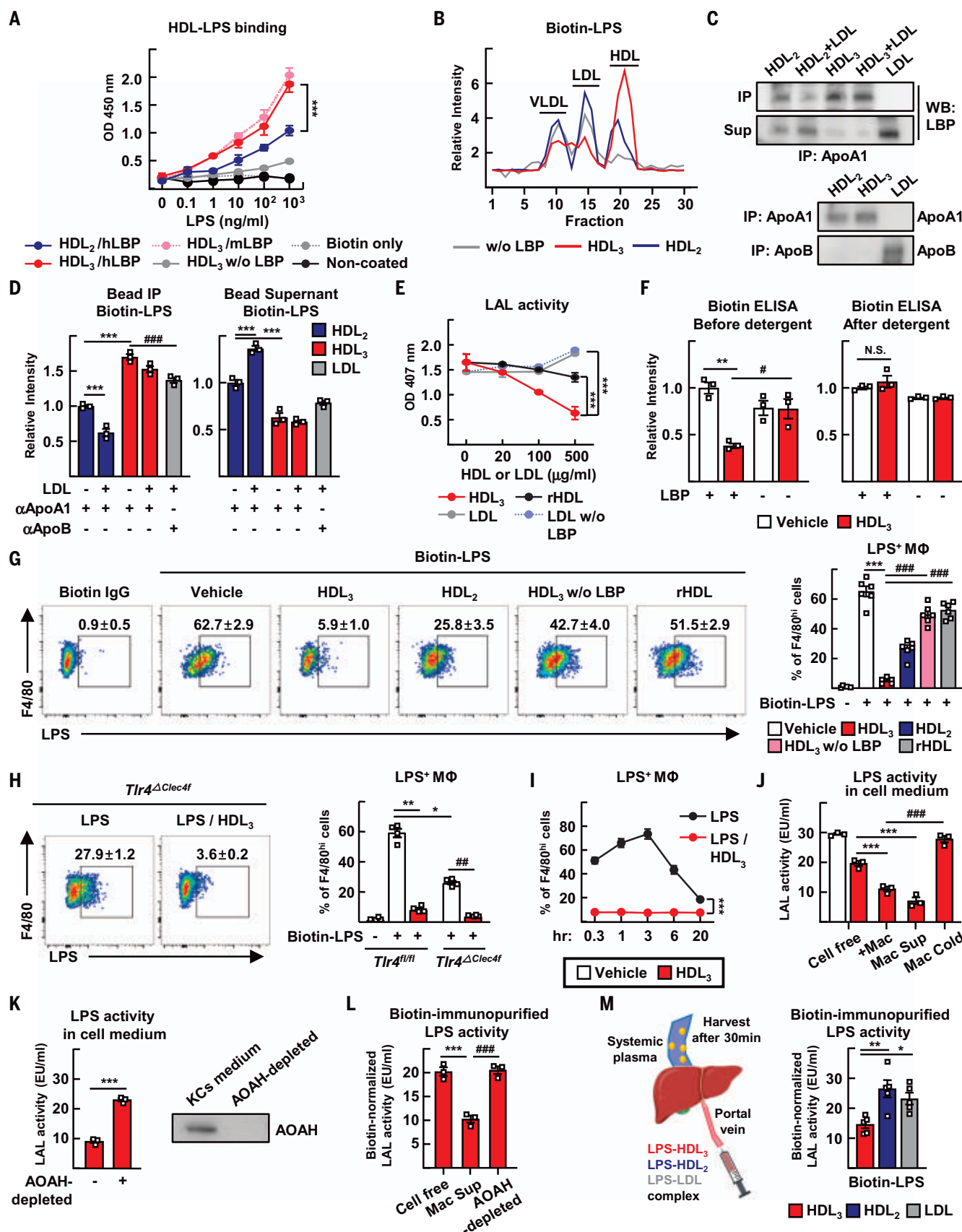


Fig. 4. HDL₃ binds LPS and masks its capacity to bind to TLR4⁺ KCs while allowing enzymatic inactivation by AOA. (A) ELISA to assess HDL and

biotin-LPS binding with or without LBP added. (B) HDL₃ or HDL₂ preincubated with biotin-LPS and incubated for 2 hours with LDL or VLDL. Biotin-LPS was then

detected in the various lipoproteins separated by FPLC. **(C and D)** Lipoproteins incubated with LBP and biotin-LPS for 2 hours and then retrieved using apoA1 or apoB IP. The amounts of LBP protein (C) and biotin-LPS (D) in resulting pellets (IP) versus supernatants (Sup) were measured. **(E)** Endotoxin LAL activity assessed after 0.5 EU/ml of *E. coli* LPS preincubated with different concentrations of HDL₃, rHDL, or LDL with or without 1 µg/ml LBP. **(F)** Biotin-LPS incubated with HDL₃ with or without LBP. The complex was then disrupted with SDS, and biotin-LPS was measured by streptavidin ELISA. **(G)** KCs cultured with HDL₃, HDL₂, or rHDL with or without LBP in the presence of biotinylated LPS for 3 hours, followed by streptavidin detection. Flow cytometry plots (left) and quantification (right) of binding to KCs are shown. **(H)** Binding of LPS to *Tlr4*^{fl/fl} and *Tlr4*^{ΔClec4e} KCs cultured with biotin-LPS (white) or biotin-LPS/HDL₃ (red). **(I)** Time course of KC binding to biotin-LPS. **(J)** LPS and HDL₃ (20 µg/ml) coincubated for 3 hours in cell-free medium, with KCs (+Mac), in medium conditioned by KCs (Mac Sup), or with KCs held at 4°C (Mac Cold). LAL activity was then assessed. **(K)** Endotoxin LAL activity assessed after LPS and HDL₃ were incubated in KC-conditioned medium from which

AOAH was depleted or not. **(L)** Biotin-LPS and HDL₃ incubated in the indicated medium. LAL activity was assessed, and SDS denaturation was conducted to allow for total biotin measurements. Relative LAL activity shown is normalized to a constant amount of biotin. **(M)** HDL₃, HDL₂, or LDL complexed with biotin-LPS and injected into the portal vein. Systemic plasma was harvested after 30 min. Endotoxin LAL activity and the amount of biotin were measured and normalized as in (L). All HDL fractions studied in this figure were isolated by density ultracentrifugation and, unless indicated [as in (B) to (G)], coincubated with 1 µg/ml of LBP during the assays (H to M). (A) to (F) depict three independent technical replicates (three experiments) per condition or time point. (G) to (L) used two to six different primary KC cultures (each prepared from different mice), with each symbol depicting data arising from one of the KC cultures. (M) is from one experiment using 15 WT mice (*n* = 5 mice per condition). **P* < 0.05, ***P* < 0.01, ****P* < 0.001, #*P* < 0.05, ##*P* < 0.01, ###*P* < 0.001. Statistical analysis on data presented as bar graphs used one-way ANOVA except for (K), for which a *t* test was used; data presented as concentration curves were analyzed using two-way ANOVA.

and S100A9⁺) into the liver (Fig. 5D). LPS bioactivity in the portal vein was elevated by small bowel resection (Fig. 5E) and was associated with increased intestinal permeability (Fig. 5F), reductions in the epithelial junction proteins ZO-1 and occludin (*Ocln*) (Fig. 5G), and elevated plasmalemma vesicle-associated protein 1 (PV1) in villus capillaries (Fig. 5H) draining into the portal vein (29).

Reduced liver fibrosis and inflammatory changes characterized KC-specific *Clec4f*-Cre \times *Tlr4*^{fl/fl} male mice compared with littermate controls (30) (Fig. 5, I to M). Female mice (fig. S5) developed disease similarly to males. Bone marrow transplants in which mice received *Tlr4*^{-/-} bone marrow confirmed the role of TLR4 (fig. S6). Thus, TLR4 expressed by KCs participates critically in liver fibrosis after small bowel resection.

Disruption of enterically derived HDL exacerbates liver injury

Portal venous HDL-C decreased after small bowel resection (Fig. 6A), possibly a consequence of the loss of bowel mass that might normally contribute to HDL biogenesis. Expression of ABCA1 sharply increased from the proximal to distal small bowel, whereas apoA1 modestly rose, overall fitting with the ileum as the main site for HDL production in the small bowel (Fig. 6B). When we modified the region of the bowel resected to remove the proximal 50% or distal 50% portion of the small intestine, HDL-C in portal blood decreased more substantially after distal resection (Fig. 6C). Accordingly, liver injury and inflammatory markers were greater in response to distal resection (Fig. 6D and fig. S7, A and B).

We compared liver injury outcomes in *Abca1*^{ΔVil1} mice versus control *Abca1*^{fl/fl} mice or *Abca1*^{ΔAlb1} mice after small bowel resection. HDL-C in portal blood was further reduced in *Abca1*^{ΔVil1} mice (Fig. 6E), and these mice indeed exhibited greater liver injury, fibrosis, and inflammation (Fig. 6, F to I). Sham

surgery did not provoke liver injury (fig. S7, C to F). After small bowel resection, LPS bioactivity was elevated in *Abca1*^{ΔVil1} mice (Fig. 6J, upper bar graph). However, the absolute amount of LPS in portal plasma matched that of other groups (Fig. 6J, lower bar graph), suggesting that the lower HDL-C in *Abca1*^{ΔVil1} mice led to increased LPS activity for a given quantity of LPS because of reduced neutralization.

Liver inflammation was elevated in *Abca1*^{ΔVil1} mice over control mice after perturbations including 12 weeks of high-fat diet (HFD) feeding or 4 weeks of the Lieber–DiCarli alcohol diet (ALD) (Fig. 6, K to M). These additional liver injury models are associated with elevated LPS translocation across the intestinal barrier (31, 32), elevations that we verified and that were in keeping with reduced HDL-C in the portal vein and apparent increased fat storage in the liver of the HFD model (fig. S7, G to I). Thus, enterically derived HDL protects against injury in multiple mouse models of liver damage.

Activation of LXR in the intestine increases HDL output and protects against liver injury

Liver X receptors (LXRs) are transcription factors that govern the expression of HDL-related genes such as *Abca1*. Low-dose LXR agonists such as GW3965, when administered orally, bypass activation of LXRs in the liver while targeting the intestine (33, 34). We thus administered GW3965 orally at a low dose in mice subjected to 75% small bowel resection and followed gene expression in the ileum and liver (Fig. 7, A to E, and fig. S8). GW3965 treatment prompted increases in *Abca1* and *Apoa1* mRNA in the ileum (Fig. 7, A and C). The impact of low-dose oral GW3965 on these genes and on LXR target genes associated with de novo lipogenesis was minimal in the liver (Fig. 7, B and D). However, inflammatory and collagen-remodeling genes in the liver were markedly down-regulated in response to GW3965 in *Abca1*^{fl/fl} mice that retained expression of in-

testinal HDL (Fig. 7E and fig. S9). These genes remained elevated in *Abca1*^{ΔVil1} mice (Fig. 7E), confirming that the effect of GW3965 depended on intestinal HDL. Oral, low-dose GW3965 did not affect *Abca1* or related genes in peritoneal macrophages adjacent to the portal venous drainage in the mesentery (fig. S8A). Some intestinal macrophages were positive for ABCA1 (fig. S8B), but deletion of ABCA1 in macrophages neither affected gene expression in the ileum nor altered portal HDL-C (fig. S8C). No changes to the liver were apparent in *Abca1*^{ΔVil1} mice compared with littermate *Abca1*^{fl/fl} controls receiving sham operations (fig. S10).

Functionally, portal venous HDL-C, remaining predominantly in the form of HDL₃, was increased by GW3965 in *Abca1*^{fl/fl} mice but not in *Abca1*^{ΔVil1} mice (Fig. 7, F and G). Indeed, oral GW3965 prominently reduced fibrosis and inflammation in the liver after small bowel resection (Fig. 7, H to K) but was unable to do so in *Abca1*^{ΔVil1} mice (Fig. 7, H to K). Thus, orally delivered low-dose GW3965 protects the liver from inflammation and fibrosis in a manner that depends upon its capacity to increase enteric HDL.

Discussion

Intestinal epithelial cells produce HDL particles (11), but neither the fate nor the function of intestinal HDL has been clear. We show here that the intestine produces the small form of HDL called HDL₃ and that it is enriched in LBP. This HDL is shuttled to the liver through the portal vein. By the time it reaches the portal blood, it is already complexed with LBP, and the LPS that it carries is masked so that it does not bind to liver KCs, circumventing the induction of proinflammatory and profibrotic genes. The intestinal epithelial location for HDL production allows the local capture of LPS before it gains access to downstream tissue.

That HDL may prevent LPS from binding to cells in the liver has been noted earlier by

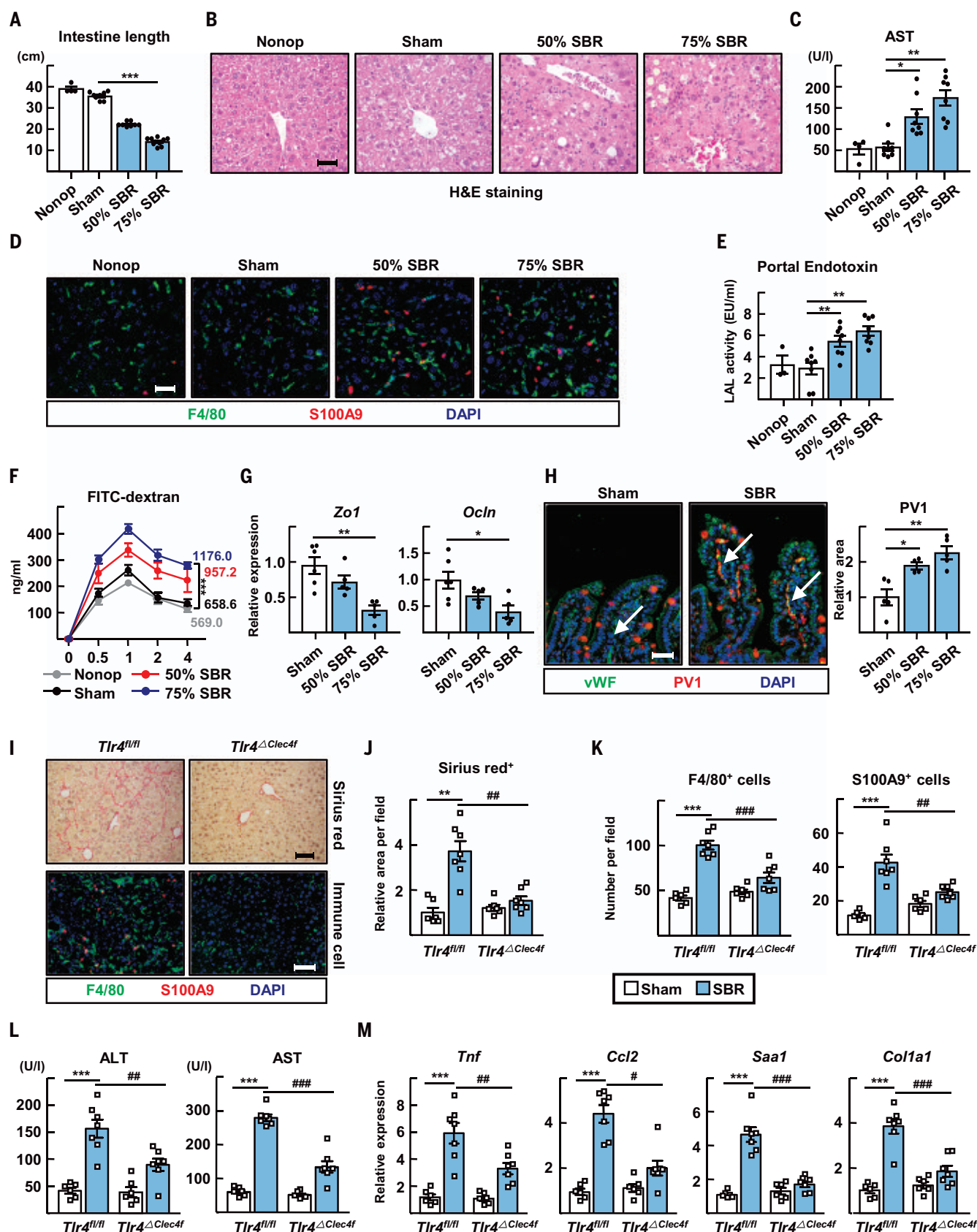


Fig. 5. Small bowel resection triggers TLR4-mediated liver inflammation.

(A to H) Small bowel resection (SBR) operations were conducted on WT mice. Nonoperated (Nonop) ($n = 4$), sham ($n = 8$), 50% SBR ($n = 8$), and

75% SBR ($n = 8$) mice were euthanized 12 weeks later. (A) Total length of the remaining small intestine. (B) Representative hematoxylin and eosin (H&E)-stained liver sections. Scale bar, 50 μ m. (C) Plasma AST levels.

(D) F4/80⁺ macrophages and S100A9⁺ neutrophils in liver sections. Scale bar, 50 μ m. (E) LAL endotoxin activity in portal plasma. (F) Kinetics of FITC-dextran translocation from intestine to peripheral blood. Area under the curve was measured. (G) qRT-PCR for intestinal mRNA transcripts encoding tight junction proteins. (H) Staining for von Willebrand Factor (vWF) (blood vessel) and PV1 in intestinal sections (left). Note that goblet cell mucin stains with the PV1 antibody, possibly nonspecifically. Relative increase in PV1 staining of vWF⁺ vessels (highlighted by white arrows) after SBR (right). Scale bar, 50 μ m. (I to M) Sham or 75% SBR operations were performed on *Tlr4*^{fl/fl} and *Tlr4* ^{Δ Clec4f} male mice, which were euthanized after 10 weeks. (I) Sirius red and immunostaining of liver sections. Scale bars, 100 μ m. (J) Relative sirius

red-positive area per field. (K) Numbers of F4/80⁺ macrophages and S100A9⁺ neutrophils per field. (L) Plasma ALT and AST levels. (M) Hepatic mRNA transcripts of inflammatory genes analyzed by qRT-PCR. (A) to (G) are the results of analysis of 30 WT mice combined from two experiments, with $n = 4$ to 10 mice per condition (nonoperated, sham, 50% SBR, or 75% SBR). (I) to (M) depict one experiment arising from analysis of 26 mice ($n = 6$ to 7 per genotype with or without SBR). Each symbol represents data from an individual mouse. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, # $P < 0.05$, ## $P < 0.01$, ### $P < 0.001$. Statistical analysis on data presented as bar graphs used one-way ANOVA; data presented as concentration curves were analyzed using two-way ANOVA.

Munford and colleagues (35), who pioneered our understanding of the role of AOA in inactivating LPS (26). Indeed, we show that whereas LBP-containing HDL₃ suppresses recognition of LPS by KCs, it remains accessible to AOA inactivation. It is unknown how HDL₃ prevents TLR4 recognition of LPS but not that of AOA. The difference may relate to the role of LBP. Because TLR4 depends upon LBP shuttling to CD14 to interact with LPS, the HDL₃-LBP complex may most effectively mask this interaction by masking the critical epitope that would support LBP-mediated handoff of LPS to CD14. However, AOA activity does not require LBP or CD14 and thus may recognize LPS in the HDL₃ particle through a region of LPS not masked by LBP.

HDL suppressed the LPS-mediated activation of KC, as well as the proinflammatory action of mediators such as lipoteichoic acid, which also interacts with HDL through LBP (36). Enterically derived HDL may bind and neutralize other microbial cargo not yet identified. Furthermore, the absence of HDL-mediated neutralization of microbial lipids such as LPS may affect the course of inflammation in locations other than the liver. For example, in Crohn's disease, a major inflammatory bowel disease, apoA1 is the most substantially down-regulated gene in the affected ileal tissue (37). In cardiovascular disease and sepsis, HDL₃ rather than HDL₂ levels correlate with better health outcomes (38, 39). The connection between disease pathogenesis in these various conditions and enteric HDL is ripe for future investigations.

Unanswered questions arise from this study. First, how do LPS and LBP interact with HDL₃ in a manner that masks the bioactivity of LPS? Although LBP is a critical promoter of TLR4 signaling, it conversely mediates suppressed signaling in the presence of HDL₃. The structural basis of this unexpected result deserves future attention. Moreover, studies in whole-body knockout mice likely obscure the anti-inflammatory contribution of LBP when bound to HDL₃ because of its other well-characterized proinflammatory role. Finally, are other components of HDL₃ needed to support the LPS-masking action of HDL₃ that we identify here?

Yet another mystery is why most portal venous HDL-C derives from the intestine. We had expected that portal venous blood HDL would arise from both portal and systemic sources. However, although systemic HDL clearly arrived to the gut or mesentery, as evidenced by efficient entry into intestine-draining mesenteric lymphatics, it was not strongly detected in the portal vein. These data suggest the existence of unknown trafficking steps, including the possible extravasation of systemic HDL near or within the intestine. After extravasation, HDL₂ that has entered or formed in the intestinal interstitium from the periphery may be too large to enter the fenestrated blood vessels that drain to the portal blood such that only enterically derived HDL₃ gains efficient access. Consistent with this possibility, mesenteric lymph is relatively deficient in the smaller HDL₃ particles but relatively enriched in HDL₂ (40). In contrast to our findings and those of others (12), studies in rats have found that the intestine routes HDL to lymph (41, 42). Although this discrepancy may be a species difference, studies in humans are more consistent with our present findings (43, 44) than with those in the rat. Moreover, it has been proposed that the use of lecithin-cholesterol acyltransferase inhibitors in these rat studies aberrantly affected the results (12).

We believe that this research has strong translational potential. In humans, like mice, portal blood was enriched in HDL₃ and potentially suppressed KC activation in response to LPS. We used three murine models of liver injury involving nutritional, alcoholic, or surgical insult to the intestine. All showed that intestinally derived HDL reduced liver injury. From a therapeutic perspective, oral delivery of LXR agonists proved effective in protecting the liver by upregulating HDL within the intestine, consistent with another recent study that engineered mice so that LXR activity was genetically augmented selectively in the intestine (45). Furthermore, our profiling and functional analysis revealed that the intestinal epithelium must express ABCA1, which is critical for HDL biogenesis, in order for a low-dose, oral LXR agonist to protect the liver. LXR agonists have failed to find utility in the

clinical setting to date, but orally restricted LXR agonists remain promising (46, 47). Our findings highlight the possibility that enteric HDL-raising LXR agonists have appeal for the treatment of various forms of liver injury. However, if suitable LXR agonists cannot be developed for application in humans, then other approaches to elevating intestinal HDL should be explored.

MATERIALS AND METHODS

Mice

C57BL/6 WT, *Tlr4*^{-/-} (B6.B10ScN-*Tlr4*^{lps-del/JthJ}; JAX #007227), *Vil1*-Cre (B6.Cg-Tg(*Vil1-cre*)997Gum/J; JAX #004586), *Alb1*-Cre (B6.FVB(129)-Tg(*Alb1-cre*)1Dlr/J; JAX #016832), *Abca1*^{fl/fl} (B6.129S6-*Abca1*^{tm1Jp}/J; JAX #028266), *Tlr*^{fl/fl} (B6(Cg)-*Tlr4*^{tm1.1Karp}/J; JAX #024872), *Clec4f*-Cre (C57BL/6J-*Clec4f*^{tm1(crc)Glass}/J; JAX #033296), *Abca1*/*gt*^{fl/fl} (B6.Cg-*Abca1*^{tm1Jp}*Abcg1*^{tm1Tail}/J; JAX #021067), and *Ly2z*-Cre (B6.129P2-*Ly2z*^{tm1(crc)lfo}/J; JAX #004781) mice (7–10 weeks of age) were originally purchased from Jackson Laboratories and housed in a specific-pathogen-free room at 22 to 24°C and 50 to 60% humidity with a 12-hour light-dark cycle. We previously generated and described PGA^{KJ/+} mice (15). *Ly2z*-Cre \times *Abca1*/*gt*^{fl/fl} breeders were provided by Dr. R. Apte and *Cd36*-knockout mice (both strains also on C57BL/6 background) by Dr. N. Abumrad at Washington University. All experiments were performed in a blinded and randomized fashion. Mice were housed on a 12-hour light-dark cycle in a temperature-controlled, specific-pathogen-free unit with food and water provided ad libitum. The studies were approved by the Washington University Animal Studies Committee (protocols 20170154, 20170252, and 20-0032) or the Medical College of Wisconsin Animal Studies Committee (protocol AUA00004173) in accordance with the National Institutes of Health laboratory animal care and use guidelines.

Small bowel surgery and associated treatments or transplants

For small bowel resection experiments, mice underwent a 50% proximal (jejunal) bowel resection, 75% proximal bowel resection, or

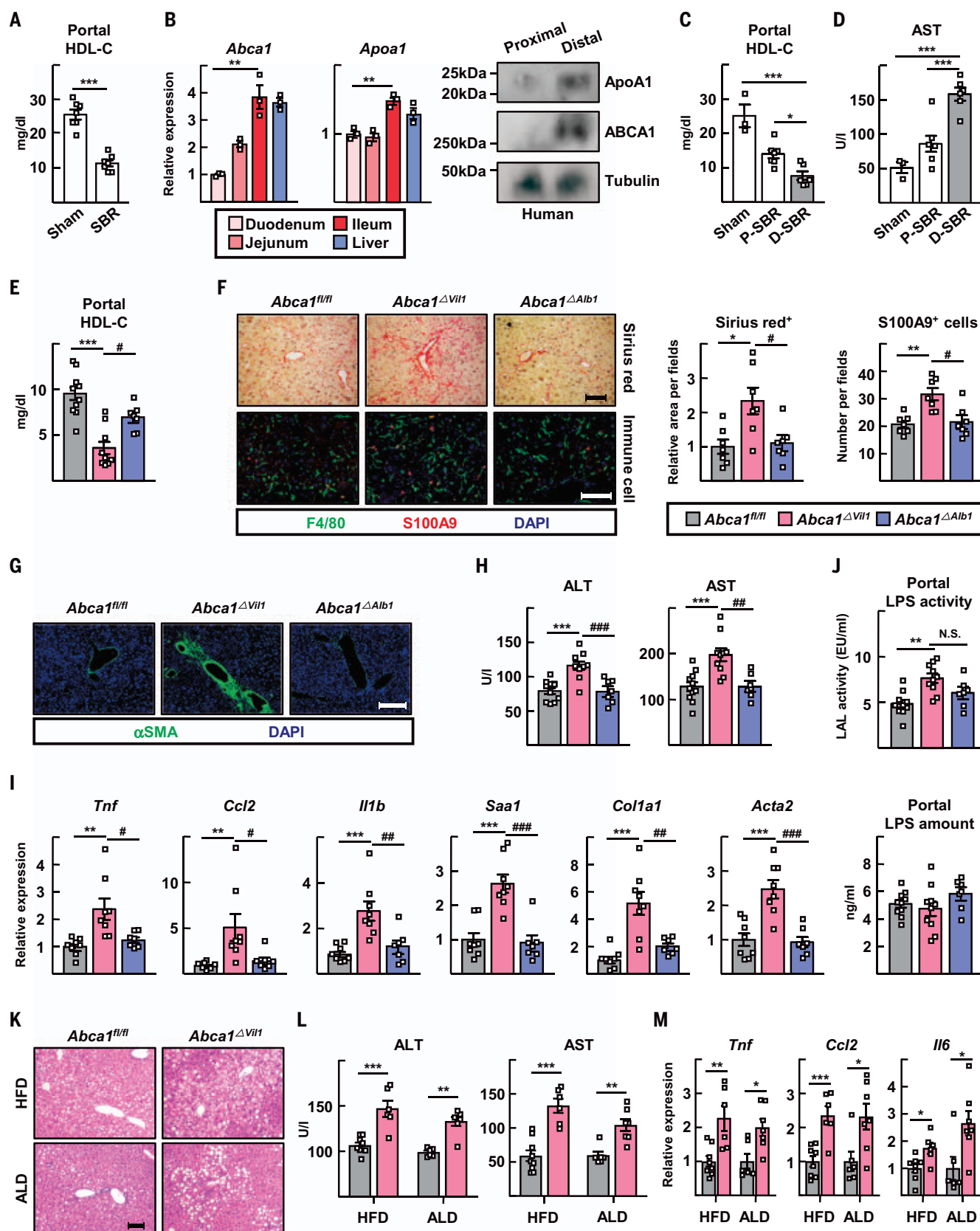


Fig. 6. Disruption of enteric HDL production worsens small bowel resection-induced liver injuries. (A) Portal HDL cholesterol levels in mice receiving sham or 75% SBR operations. (B) RT-PCR for *Abca1* and *Apoa1* in

mouse duodenum, jejunum, ileum, and liver (left). Protein expression of apoA1 and ABCA1 in human proximal and distal gut were analyzed by immunoblotting (right). (C and D) SBR operations were conducted for WT mice. Mice receiving

sham ($n = 3$), proximal 50% SBR (P-SBR) ($n = 7$), and distal 50% SBR (D-SBR) ($n = 6$) operations were euthanized 10 weeks later. (C) Portal HDL cholesterol levels. (D) Plasma AST levels. (E to J) 75% SBR operations were performed for *Abca1^{fl/fl}* ($n = 10$), *Abca1^{ΔVil1}* ($n = 10$), and *Abca1^{ΔVil1}* ($n = 7$) mice, which were euthanized 8 weeks later. (E) Portal HDL cholesterol levels. (F) Representative sirius red staining of liver sections (left, top row) and relative area per field (near right). F4/80⁺ macrophages and S100A9⁺ neutrophils were visualized (left, bottom row); cell numbers per field are shown (far right). Scale bars, 100 μ m. (G) Representative anti-SMA immunostaining of liver sections. Scale bar, 200 μ m. (H) Plasma ALT and AST levels. (I) Hepatic RT-PCR for inflammatory genes. (J) Portal LAL LPS activity (top) and LPS quantification

by ELISA (bottom). (K to M) *Abca1^{fl/fl}* and *Abca1^{ΔVil1}* male mice were fed HFD or ALD. (K) Representative H&E-stained liver sections. Scale bar, 100 μ m. (L) Plasma ALT and AST levels. (M) RT-PCR for inflammatory genes. (A) to (C) show results from $n = 3$ to 7 mice per condition using WT mice (one experiment each panel). (E) to (I) combine data from two experiments using 28 mice ($n = 7$ to 11 mice per genotype). (L) and (M) depict two experiments (one HFD, one ALD) from $n = 15$ mice on HFD ($n = 9$ *Abca1^{fl/fl}* mice, 6 *Abca1^{ΔVil1}* mice) or $n = 13$ mice on ALD ($n = 6$ *Abca1^{fl/fl}* mice, 7 *Abca1^{ΔVil1}* mice). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, # $P < 0.05$, ## $P < 0.01$, ### $P < 0.001$. Statistical analysis on data presented as bar graphs used one-way ANOVA except for (A), for which a t test was used.

sham control operation (bowel transection with reanastomosis alone), as previously described (28). In brief, through a midline laparotomy, the small bowel was exteriorized and transected 1–2 cm distal from the ligament of Treitz and ~12 cm (for 50% resection) or 6 cm (for 75% resection) proximal to the ileocecal junction. For sham operations, a transection 12 cm proximal to the ileocecal junction with immediate reanastomosis was performed. For distal 50% small bowel resection, the ileum (last 12 cm of small bowel) was removed with an anastomosis of the jejunum to a small cuff of small bowel on the cecum. All anastomoses were hand sewn end to end with interrupted 9-0 nylon sutures. Postoperative care included housing in an incubator for temperature stability and 24 hour fasting before starting a liquid diet (PMI Micro-Stabilized Rodent Liquid Diet LD 101; TestDiet), on which the mice were maintained for 8–12 weeks until euthanasia.

For bone marrow transplants, WT recipient mice received whole-body irradiation at a dose of 11 Gy at 8 weeks of age, and then were intravenously injected with 5×10^6 bone marrow cells from WT or *Tlr4^{-/-}* donor mice. After 4 weeks, short bowel resections were conducted.

For LXR agonist treatment, GW3965 (Sigma-Aldrich, #G6295) was suspended in 0.5% carboxymethyl cellulose and orally administered twice weekly at 1 mg/kg body weight per day for the last 5 weeks in the 10-week period after intestinal resection. The different experimental groups of mice maintained a similar body weight during liquid diet feeding and/or drug treatment.

Dietary challenge models

Where specified, mice were given HFD or ALD to induce liver inflammation and injury (48, 49). The HFD study, containing 60% kcal from fat (Research Diets, #D12492), was conducted for 12 weeks. For ethanol feeding, the mice were acclimated to increasing alcohol concentration of 2.1, 4.2, and 6.4% v/v (ethanol and liquid diet) over 3 days, respectively. After alcohol adaptation, a 6.4% ethanol-enriched diet was supplied in the same liquid diet used after small bowel resection surgeries (diet changed daily) for 4 weeks. In supplemental experiments, mice were fasted (with ad libitum

access to water) for up to 20 hours, or challenged for 3 weeks with an atherogenic diet containing 42% kcal from fat (Harlan Teklad, #TD.88137).

Immunostaining and confocal microscopy

Left lobes of liver tissues and small intestines were excised and fixed in 4% paraformaldehyde (Santa Cruz Biotechnology) overnight at 4°C. Ten-micrometer paraffin-embedded sections were prepared and slides were boiled in Diva Decloaker solution (Biocare Medical, #DV2004) in a pressurized chamber for 15 min. Sections were blocked in phosphate-buffered saline (PBS) containing 5% donkey serum, 1% bovine serum albumin (BSA) (Sigma-Aldrich), and 0.03% Triton X-100 (Plusone, #I7-1315-01) for 1 hour; then incubated with rat anti-F4/80 (Abcam, #ab6640), goat anti-S100A9 (R&D Systems, #AF2065), rabbit anti-von Willebrand Factor (DAKO, #a0082), rat anti-PV1 (BD pharmingen, #550563), or rabbit anti-ABCA1 (Novus Biologicals, NB400-105) at 4°C overnight. Primary antibodies were detected using Cy3- or Cy5-conjugated secondary antibodies (Jackson ImmunoResearch). The stained sections were imaged using an SP8 confocal microscope (Leica) equipped with nine lasers and four tunable detectors (two hybrid, two tunable) and a 20× HC PL Apo CS 2 multi-immersion objective, numerical aperture 0.75. Images were processed with Imaris software (Bitplane). Ten fields were quantified and averaged for each sample, with cell counts per image quantified using Image J software (NIH). All slides were analyzed in a blinded and randomized fashion.

Quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA from tissues or cells was isolated by using RNeasy Mini or Micro kits according to the manufacturer's protocol (Qiagen). cDNA was synthesized using the high-capacity cDNA reverse transcription kit (Applied Biosystems, #4368814). qRT-PCR experiments were performed using ABI StepOnePlus Real-Time PCR machine with specific primers (Applied Biosystems). Primer sequences are available upon request. The relative transcriptional expression of target genes was evaluated by the Eq. $2^{-\Delta Ct}$ ($\Delta Ct = Ct$ of target gene minus Ct of

18S rRNA). Relative transcription, where plotted, was calculated with the mean of the control group set as 1.

Intestinal permeability assay

After the mice had fasted for 4 hours, 200 mg/kg body weight of 4-kDa fluorescein isothiocyanate (FITC)-dextran (Sigma-Aldrich) was administered by gavage. After 0.5, 1, 2, or 4 hours, blood for the preparation of plasma was collected from the tail vein, and the fluorescence intensity of the samples and standards was read at excitation 485 nm/emission 525 nm using the Cytation 5 Cell Imaging Multi-Mode Reader (BioTek).

Photoactivation of *PGAI^{KI/+}* mice

For phototagging HDL from skin, 8- to 10-week-old *PGAI^{KI/+}* mice were anesthetized and a region of shaved skin was photoconverted using a SOKY, Violet 405 nm 500 mW (FDA), PL-405-500B laser, as described previously (15). For photoactivation of the lumen of the small intestine of anesthetized mice, we stretched the mesentery and intestine over the solid surface of a Petri dish, located the region of interest, and surgically clipped the bowel just enough so that we could thread into the lumen a fiberoptic endoscopic laser (Laserland, Violet 405 nm 100 mW) to photoactivate enterocytes. Unless otherwise specified, three areas were activated for one data point, with the laser being held on for 10 s and off for 20 s, cycling for 1 min 10 s to achieve three exposures per location. For photoactivation of the exterior of the small intestine, the Laserland Violet 405 nm 100 mW 5V laser was used to activate area of 14.668 mm² of gut with a similar on-off cycle as for the intestinal lumen.

Collection of blood and lymph

The portal blood was collected using a 33-gauge needle to a volume of 40 μ l per mouse. Systemic blood was collected from the inferior vena cava using a 26-gauge needle in EDTA-containing tubes. Mesenteric lymph fluid was also collected, and the fluorescence intensity of plasma or lymph fluid was measured using the Cytation 5 Cell Imaging Multi-Mode Reader (BioTek). Mesenteric lymphatic cannulations were accomplished under general anesthesia using an

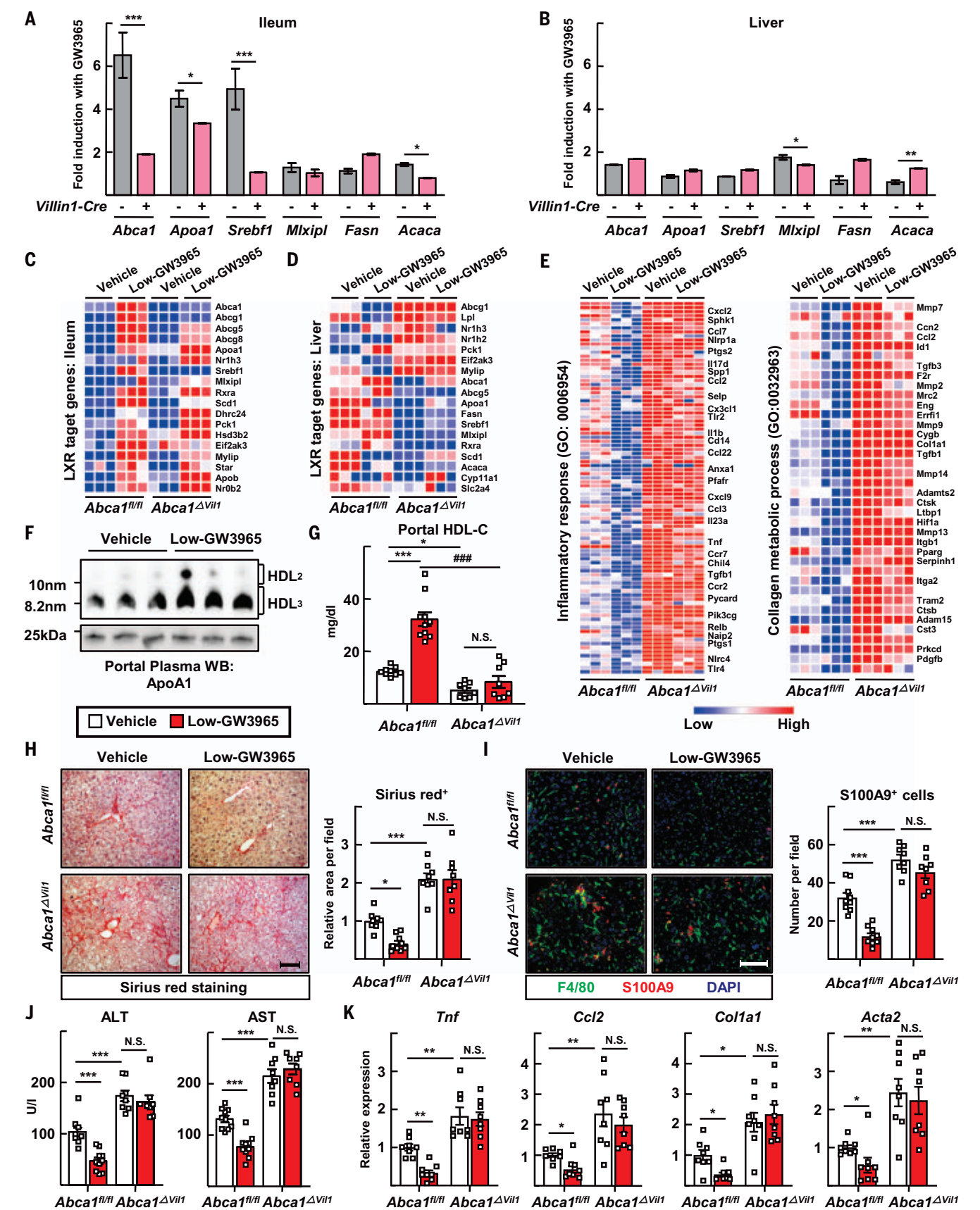


Fig. 7. Intestine-restricted activation of LXR ameliorates liver injury in a manner dependent upon enteric HDL production. (A to E) *Abca1*^{fl/fl} and *Abca1*^{ΔVil1} mice received vehicle ($n = 10$ *Abca1*^{fl/fl} mice, $n = 8$ *Abca1*^{ΔVil1} mice) or 1 mg/kg/day GW3965 ($n = 10$ *Abca1*^{fl/fl} mice, 8 *Abca1*^{ΔVil1} mice) by gavage twice weekly in the last 5 weeks of a 10-week feeding after SBR (Low-GW3965). (A and B) RT-PCR analysis of select target genes in the ileum and liver tissues was used to determine fold induction in response to GW3965 treatment. (C and D) Heatmap of mRNA transcripts of LXR-regulated genes in the ileum and liver tissues. (E) Heatmap of mRNA transcripts encoding genes associated with the inflammatory response or collagen metabolism within the liver according to GW3965 treatment and genotype. (F) Immunoblots for apoA1

to detect HDL after portal vein plasma were run on a nondenaturing gel. (G) Portal plasma HDL cholesterol levels. (H) Sirius red staining of liver sections (left) and relative area per field (right). Scale bar, 100 μ m. (I) F4/80⁺ macrophages and S100A9⁺ neutrophils visualized by immunofluorescence in liver sections (left) (scale bar, 100 μ m) and cell numbers per field counted (right). (J) Plasma ALT and AST levels. (K) RT-PCR measuring inflammation- or fibrosis-associated mediators. Plots show mean \pm SEM in male mice, with each symbol on the bar graphs representing a single mouse with data combined from two experiments; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, #### $P < 0.001$. t tests were used for statistical comparisons in (A) and (B); one-way ANOVA was used for statistical comparisons in (G) to (K).

operating microscope. A midline laparotomy with an extension to a left subcostal incision was made and the intestine was mobilized to expose the mesenteric lymphatic duct proximal to the cisterna chyli. A small incision was made on the mesenteric lymphatic duct and gently cannulated using polyethylene tubing (inner diameter, 0.28 mm; outer diameter, 0.61 mm; Intramedic, Sparks, MD). At the completion of the collection, the mouse was euthanized.

Preparation of HDL fractions

Human and mouse plasma collected from the portal vein or peripheral vein (inferior vena cava for mouse, antecubital vein for human) was collected and ultracentrifuged overnight using standard methodology at 100,000g at 4°C in sequential steps, adjusting the solution to 1.063 g/ml using KBr (Sigma-Aldrich, #221864) to remove LDL. Then the HDL fraction was prepared by another centrifugation at 100,000g after further adjusting density of the solution to 1.21 g/ml with KBr. Isolated HDL fractions were dialyzed using the Slide-A-Lyzer Dialysis cassette kit (Thermo Fisher Scientific) with PBS solutions containing NaCl, Tris, and EDTA for 4 hours at 4°C to remove KBr. Human HDL₂ (1.063–1.125 g/ml) and HDL₃ (1.125–1.21 g/ml) fractions were obtained from GenWay Biotech.

FPLC and measurement of HDL cholesterol

Fifty microliters of blood from mice fasted for 4 hours was collected in microcentrifuge tubes containing 10 μ l of 0.5 mM EDTA and then centrifuged at 500g to collect plasma. For cholesterol distribution of total lipoproteins, plasma was prepared and 100–200 μ l was flowed over a Superose 6 10/300GL gel filtration column (GE Healthcare) to separate the different classes of lipoproteins. Cholesterol in each fraction was measured by an enzymatic assay kit (Wako Diagnostics Cholesterol E, #439-17501). An HDL cholesterol assay kit (Cell Biolabs, #STA-394) was used to measure HDL-C levels.

Immunoblots

Protein immunoblotting was performed using rabbit anti-mouse apoA1 (Meridian Life Sciences), rabbit anti-human apoA1 (Millipore,

#MAB011), rabbit anti-human ABCA1 (Novus Biologicals, #NB400-105), mouse anti-PON1 (Abcam, #ab24261), rabbit anti-ApoB (Proteintech, #20578-1), rabbit anti-LBP (Abcam, #ab233524), rabbit anti-AOAH (Proteintech, #12911-1), rabbit anti-SERPINA1 (Thermo Fisher Scientific, #PA5-16661), mouse anti-ApoE (kindly provided by D. M. Holtzmann, Washington University), or rabbit anti-albumin (Proteintech, #16475-1) antibodies. Table S2 specifies the dilutions of the antibodies used. The HDL fractions were loaded to achieve the same protein concentration per lane, and plasma was loaded without dilution. For native gels, the samples were diluted in 2 \times native sample buffer (Bio-Rad) and run on 4 to 20% Mini-PROTEAN Tris-glycine gels (Bio-Rad) with Tris-glycine running buffer. For denaturing gels, the samples were diluted in 2 \times Tris-Glycine-SDS sample buffer (EZ Bioresarch) and heated at 95°C for 10 min. The samples were loaded onto 4 to 20% Mini-PROTEAN gels and run with Tris-glycine-SDS running buffer. The separated proteins were transferred to 0.45- μ m polyvinylidene fluoride membrane (Millipore, #IPVH00010) with Tris-glycine transfer buffer for 2 hours at 20 V. Membranes were blocked with 5% nonfat dry skim milk (Bio-Rad, #170-6404) for 1 hour, and primary antibodies were incubated overnight at 4°C. After incubation with horseradish peroxidase-conjugated secondary antibodies, signal was detected using Clarity Western ECL solution (Bio-Rad).

Electron microscopy of HDL particles

The isolated HDL fractions were diluted to 15 μ g/ml of total protein and negatively stained with 1% uranium acetate. The samples were deposited on carbon-coated, 200-mesh copper grids (Electron Microscopy Sciences). Images were acquired with a transmission electron microscope (TEM; JEOL, #JEM-1400^{Plus}) at 120 KeV and 80,000 \times or 150,000 \times magnifications. The diameter of HDL particles was measured using ImageJ software.

Isolation or culture of liver immune cells and macrophages

For quantification of neutrophils, monocyte-derived macrophages, and KCs in livers of mice subjected to short bowel resection or sham

surgery, livers were collected and homogenized in Hank's buffered saline solution containing 1.49 mg/ml of collagenase type IV (Sigma-Aldrich, #C5139) and dissociated using the gentleMACS Octo Dissociator (Miltenyi Biotec). After centrifugation at 50g, the supernatant containing nonparenchymal cells was separated using 33% Percoll (GE Healthcare).

For KC isolation and culture, the livers of 7- to 10-week-old male C57BL/6 mice were perfused through the inferior vena cava with collagenase type IV solution as described previously (50). The cell suspension was centrifuged. Cell suspensions collected in the supernatant were collected and again centrifuged in 50%/25% Percoll (GE Healthcare). The layer containing liver macrophages was plated in RPMI-1640 (Hyclone) containing 10% fetal bovine serum (FBS). After 2 hours of culture to allow for cell attachment, the cell medium was changed to “vehicle” medium, which was serum-free RPMI-1640 containing 1 μ g/ml of recombinant LBP (R&D Systems, #6635-LP) for 3 hours of culture. HDL preparations were added, or not, to these cultures with 20 ng/ml of LPS (Sigma-Aldrich, #L2630), 10 μ g/ml lipoteichoic acid (LTA) (Sigma-Aldrich, #L2515), 10 μ g/ml CpG DNA (Invivogen, #tlrl-1826), 100 ng/ml TNF (Sigma-Aldrich, #T7539), or 100 ng/ml IL-1 β (Sigma-Aldrich, #I5271).

Flow cytometry

Isolated liver immune cells and cultured liver macrophages were collected and counted in an automated cell counters (Cellometer Auto X4; Nexelcom Bioscience) after staining for acridine orange (Sigma-Aldrich). Antibodies (details in table S2) including BUV396-anti-CD45 (BD Biosciences, #563791), FITC-anti-Ly6G (BioLegend, #127605), APC/Cy7-anti-F4/80 (BioLegend, #123117), PerCP/Cy5.5-anti-Ly6C (BioLegend, #128011), PE/Cy7-anti-CD31 (BioLegend, #102417), Alexa Fluor 488-anti-iNOS (Thermo Fisher Scientific, #53-5920-82), APC-anti-CD11b (Thermo Fisher Scientific, #17-0112-82), PE-anti-Tim4 (Thermo Fisher Scientific, #12-5866-82), or goat anti-Clec4f (R&D Systems, #AF2784) were incubated with FACS buffer (2% FBS, 2 mM EDTA, and sodium azide in PBS) on ice for 30 min. In some experiments, the primary cultures of KCs were

incubated with 100 ng/ml of biotinylated LPS (Invivogen, #tlrl-lpsbiot) for 2 hours. Then, biotin was detected using PE/Cy7-streptavidin (BioLegend, #405206). After surface staining of biotin, in some experiments, internalized biotin-LPS was stained using BV605-streptavidin (BioLegend, #405229) in cells permeabilized using the Intracellular Fixation & Permeabilization Buffer Set (Thermo eBioscience, #88-8824). After washing and resuspension, cells were analyzed on a BD Biosciences FACS Symphony machine and analyzed by FlowJo software (BD Biosciences).

Enzyme-Linked Immunosorbent Assay

Enzyme-linked immunosorbent assay (ELISA) kits were used according to the manufacturer's protocol and included the AST Activity Assay Kit (Sigma-Aldrich, MAK055), TNF- α (Sigma-Aldrich, RAB0477), or CCL2 ELISA (Sigma-Aldrich, #RAB0055). Limulus amoebocyte lysate (LAL) endotoxin activity was measured using the Pierce LAL Chromogenic Endotoxin Quantitation Kit (Thermo Fisher Scientific, #88282). LPS quantification by ELISA used the LPS ELISA kit from MyBiosource (#MBS700021). For sandwich ELISA to analyze LPS-HDL binding, we used high-binding clear polystyrene microtiter plates (R&D Systems, #DY990), and purified HDL was immobilized for 2 hours on these plates at 37°C at 10 μ g/ml. After washing, plates were blocked with 1% BSA for 1 hour. Then biotinylated LPS was preincubated with or without 1 μ g/ml of recombinant LBP (R&D Systems) for 1 hour and incubated in plates for 30 min. For ELISA to quantify biotin, samples including biotinylated LPS were incubated in high-binding clear polystyrene microtiter plates overnight at 37°C. Streptavidin peroxidase (R&D Systems) was added, followed by diaminobenzidine substrate (Abcam), for colorimetric reactions. Colorimetric or fluorometric absorbance was detected using the Cytation 5 Cell Imaging Multi-Mode Reader (BioTek).

IP of HDL, LDL, and biotinylated LPS

For proteomic analysis of systemic and portal plasma, human plasma samples were separated using size-exclusion FPLC and then plasma albumin and IgG were depleted using the Pierce Albumin/IgG Removal kit (Thermo Fisher Scientific, #89875). Immunoprecipitation (IP) on the separated HDL fractions was conducted using Pierce MS-Compatible Streptavidin Magnetic IP Kit (Thermo Fisher Scientific, #90408). The anti-human apoA1 antibody (Proteintech, #14427-1) was biotinylated using Pierce Antibody Biotinylation kit for IP (Thermo Fisher Scientific, #90407) for later capture by streptavidin IP. The samples were eluted at low pH and neutralized to 100 mM Tris, pH 8.0.

Biotinylated LPS in cell media was mixed with 1% SDS to disrupt lipoproteins and then

purified using Streptavidin Magnetic IP kit. Endotoxin activity of the immunopurified biotinylated LPS was measured using the LAL kit and normalized by relative amount of biotin as detected in ELISA.

For binding studies, 50 μ g/ml of protein in HDL₃, HDL₂, or LDL fractions was mixed with 1 μ g/ml of LBP and 20 μ g/ml of biotinylated LPS in 100 mM NaCl, 100 mM Tris, and 1 mM EDTA buffer for 2 hours. The mixtures were incubated with anti-apoA1 antibody (Proteintech) or anti-apoB antibody (Thermo Fisher Scientific, #M1A1605) overnight and immunoprecipitated using Protein A agarose beads (Abcam, #ab193254). The bead-bound samples were eluted using 0.1 M glycine buffer (pH 2.5) and neutralized using 100 mM Tris, pH 8.0. The distributions of LBP and biotinylated LPS were measured by immunoblotting and streptavidin-peroxidase ELISA, respectively.

Generation of reconstituted HDL (rHDL)

Human apoA1 isolation and purification from fresh human plasma was performed as reported previously (51). The lyophilized protein was solubilized and denatured in STB (10 mM Tris, 0.15 M NaCl, 1 mM EDTA, and 0.2% NaN₃) containing 3 M guanidine HCl, followed by refolding at 4°C by dialyzing against three changes of 4 liters of STB for a minimum of 3 hours each. rHDL particles were generated with a modified sodium cholate dialysis as previously described (52). A molar ratio of 80:1 POPC (Avanti Polar Lipids) and apoA1, respectively, was used to generate rHDL.

Depletion of HDL or AOA

The HDL in mouse plasma was removed with the HDL Depletion Column IgY Kit (Genway Biotech, #GWB-HDLIGY). AOA in cell medium was depleted through IP by anti-AOA antibody (Proteintech) and Protein A agarose beads. After IP, the supernatant was harvested and depletion of HDL or AOA confirmed by immunoblotting.

Chromatographic analysis of biotinylated LPS transfer between lipoprotein species

HDL₃ or HDL₂ (100 μ g protein/ml) was incubated with or without 2 μ g/ml of LBP and 1 μ g/ml of biotinylated LPS in Tris buffer for 30 min and then dialyzed in a Slide-A-Lyzer Dialysis cassette kit. Isolated VLDL (10 μ g/ml of protein) and LDL (50 μ g/ml of protein) were subsequently added to the HDL-LPS mixture and incubated for 2 hours at 37°C. The 200- μ l mixture was then subjected to FPLC separation using a Superose 6 10/300GL gel filtration column. The amount of biotin was measured through streptavidin peroxidase ELISA, and lipoprotein-associated LPS was thereby determined.

Portal vein injection of biotinylated LPS and lipoprotein complex

A mixture containing 5 μ g of biotinylated LPS and 2 μ g of LBP was incubated with 0.1 mg of HDL₃, HDL₂, or LDL (concentration determined by protein not lipid) in saline buffer for 3 hours. Mice were anaesthetized by injection of a mixture of ketamine (50 mg/kg body weight) and xylazine (8 mg/kg body weight) intraperitoneally. The LPS-loaded lipoprotein mixtures described above were then injected through the portal vein in a 100- μ l volume using an ultrasmall 33-gauge needle. VETSPON Absorbable Hemostatic Gelatin Sponges (Novartis) were used to stop bleeding of the portal vein. Systemic blood and livers were harvested 30 min later.

Peptide preparation and nano-liquid chromatography–tandem mass spectrometry

Peptides were prepared as previously described (53). Then modification of a previous method (54) was followed. First, the column was equilibrated to 0.1% formic acid (FA) for a total of 11 μ l at 700 bar pressure. The samples in FA (1%) were loaded on an EASY nanoLC (Thermo Fisher), with sample (2.5 μ l) applied onto a 75- μ m inner diameter \times 50-cm Acclaim PepMap 100 C18 RSLC column (Thermo Fisher). A constant pressure of 700 bar was maintained at 0.1% FA. Peptide chromatography was performed using mobile phase A (1% FA) containing 2% B (100% MeCN, 1%FA) for 5 min, then increased to 20% B over 100 min, to 32% B over 20 min, to 95% B over 1 min, and held at 95% B for 29 min. The flow rate was 250 nl/min. Data were acquired in data-dependent acquisition (DDA) mode. Full-scan mass spectra were acquired with the Orbitrap mass analyzer with a scan range of m/z = 350 to 1500 and a mass resolving power set to 70,000. Ten data-dependent, high-energy collisional dissociations were performed with a mass-resolving power set to 17,500, a fixed lower value of m/z 100, an isolation width of 2 Da, and a normalized collision energy setting of 27. The maximum injection time was 60 ms for parent ion analysis and product ion analysis. The target ions that were selected for tandem mass spectrometry (MS/MS) were dynamically excluded for 20 s. The automatic gain control was set at a target value of 3×10^{-6} ions for full MS scans and 1×10^{-5} ions for MS2. Peptide ions with charge states of 1 or >8 were excluded for HCD acquisition.

Protein identification

The MS unprocessed data from the mass spectrometer were converted to peak lists using Proteome Discoverer (Thermo Fisher Scientific). The MS2 spectra with charges +2, +3, and +4 were analyzed using Mascot software (Matrix Science). Mascot was set up to search against a UniProt (July 2019) database of human

proteins (20,667 entries), using trypsin cleavage specificity (trypsin/P) with four missed cleavages allowed. The searches were performed using Mascot software according to previously described parameters for peptide identification (53). Peptides and proteins were filtered at 1% false discovery rate (FDR) by searching against a reversed protein sequence database. The ontology of HDL signature proteins were acquired from DAVID bioinformatics functional annotation (<https://david.ncifcrf.gov/>).

RNA sequencing

Three individual mice per experimental group were used for the generation of whole intestinal or whole liver RNA sequencing. Total RNA integrity was determined using an Agilent Technologies Bioanalyzer or 4200 TapeStation. Library preparation was performed with 500 ng to 1 µg of total RNA. Ribosomal RNA was removed by an RNase-H method using RiboErase kits (Kapa Biosystems), and mRNA was then fragmented in reverse transcriptase buffer with heating to 94°C for 8 min. Then mRNA was reverse transcribed to yield cDNA using SuperScript III RT enzyme (Life Technologies) and random hexamers. A second strand reaction was performed to yield ds-cDNA. cDNA was blunt-ended, had an A base added to the 3' ends, and then had Illumina sequencing adapters ligated to the ends. Ligated fragments were then amplified for 12 to 15 cycles using primers incorporating unique dual index tags. Fragments were sequenced on an Illumina NovaSeq-6000 using paired-end reads extending 150 bases. The gene counts were quantified with CPM transformations added with custom R scripting. Normalized Log2 CPM values were visualized as heatmaps through the web interface Phantasus (<https://genome.ifmo.ru/phantasus>). Gene ontology pathway analyses were acquired from DAVID bioinformatics functional annotation. The sequencing and expression data have been deposited in the Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Information with the accession number GSE167983.

Human studies

Human portal and peripheral systemic blood was acquired from adult patients undergoing open surgical procedures in which the surgical team deemed that the portal vein was safely accessible. Blood was collected in EDTA-containing tubes and centrifuged for immediate analysis of plasma for HDL. The population of patients from which plasma was acquired was composed of three males and three females ranging in age from 54 to 80. Four underwent Whipple pancreatic surgery, one underwent gastric bypass, and one underwent an orthotopic liver transplant. Immunoblots were performed using tissue specimens from other patients under-

going surgical resection for proximal or distal gut after traumatic injury. These were collected and fixed in formalin by the Department of Pathology and Immunology for a routine surgical pathology workup and were shared for research after sign out of the clinical case. After dissecting tissue enriched in epithelium, dissected tissues were stabilized in PAXgene Tissue Stabilization buffer (PreAnalytiX, #765512) for at least 3 hours. Protein lysates were then prepared by homogenization in Extraction EXB buffer (Qiagen, #37623). All human studies were approved by the Human Research Protection Office at Washington University (institutional review board protocols #201111038 and #2019101009, PI G. J. Randolph).

Statistics

All graphs are plotted to depict mean ± SEM. A paired or unpaired two-tailed Student's *t* test was used for simple comparisons, a one-way ANOVA with Tukey's post hoc test for multiple comparisons for three or more groups with one variable, or a two-way ANOVA with Sidak post hoc test for three or more groups with two variables. Statistical differences were analyzed and graphs were prepared using GraphPad Prism software version 8.0. *P* < 0.05 was considered to be a significant difference. Replicates in the bar graphs are shown by distinct symbols. Replicates in biochemical and/or binding experiments were independent technical replicates using the same reagents (technical replicates; Figs. 3 and 4, A to F). In other experiments, each symbol in a graph represents experiments where independent cell preparations of plasma (Fig. 1, A and F, and Fig. 2, A and B; paired samples) or primary cells (Fig. 3, A to E, and Fig. 4, G to L) were generated and studied within the experimental design. Finally, in all in vivo experiments, each symbol in a graph represents data generated from an individual mouse (Fig. 1, C to E; Fig. 4M; Fig. 5, A, C, and E to H; Fig. 5, J to L; Fig. 6, A to F, H to J, and L and M; and Fig. 7, A to K). All data using mice contained at least two mice in each experimental group from the same litter, so that the data in a given experimental cohort contained littermate controls in the other groups of the same experiment. Some studies combined mice from three or four litters to generate sufficient numbers. If these mice were subjected to experimental manipulations such as short bowel surgery on different dates, then we refer to the combination of those data into one graph as combining different experiments, with each distinct start date of the experimental manipulation considered an independent experiment. Experimental manipulations that began on the same day with age-matched litters were not classified as independent experiments. The distribution of data was not affected by combining independent experiments. Where possible, data collected

from different experiments were subjected to assays simultaneously to minimize batch effects.

REFERENCES AND NOTES

1. A. Albillos, A. de Gottardi, M. Rescigno, The gut-liver axis in liver disease: Pathophysiological basis for therapy. *J. Hepatol.* **72**, 558–577 (2020). doi: [10.1016/j.jhep.2019.10.003](https://doi.org/10.1016/j.jhep.2019.10.003); pmid: [31622696](https://pubmed.ncbi.nlm.nih.gov/31622696/)
2. Y. Duan et al., Bacteriophage targeting of gut bacterium attenuates alcoholic liver disease. *Nature* **575**, 505–511 (2019). doi: [10.1038/s41586-019-1742-x](https://doi.org/10.1038/s41586-019-1742-x); pmid: [31723265](https://pubmed.ncbi.nlm.nih.gov/31723265/)
3. R. F. Schwabe, E. Seki, D. A. Brenner, Toll-like receptor signaling in the liver. *Gastroenterology* **130**, 1886–1900 (2006). doi: [10.1053/j.gastro.2006.01.038](https://doi.org/10.1053/j.gastro.2006.01.038); pmid: [16697751](https://pubmed.ncbi.nlm.nih.gov/16697751/)
4. J. Todoric et al., Fructose stimulated de novo lipogenesis is promoted by inflammation. *Nat. Metab.* **2**, 1034–1045 (2020). doi: [10.1038/s42255-020-0261-2](https://doi.org/10.1038/s42255-020-0261-2); pmid: [32839596](https://pubmed.ncbi.nlm.nih.gov/32839596/)
5. I. Hritz et al., The critical role of toll-like receptor (TLR) 4 in alcoholic liver disease is independent of the common TLR adapter MyD88. *Hepatology* **48**, 1224–1231 (2008). doi: [10.1002/hep.22470](https://doi.org/10.1002/hep.22470); pmid: [18792393](https://pubmed.ncbi.nlm.nih.gov/18792393/)
6. L. K. Barron et al., Toll-like receptor 4 is critical for the development of resection-associated hepatic steatosis. *J. Pediatr. Surg.* **52**, 1014–1019 (2017). doi: [10.1016/j.jpedsurg.2017.03.026](https://doi.org/10.1016/j.jpedsurg.2017.03.026); pmid: [28351520](https://pubmed.ncbi.nlm.nih.gov/28351520/)
7. S. Inokuchi et al., Toll-like receptor 4 mediates alcohol-induced steatohepatitis through bone marrow-derived and endogenous liver cells in mice. *Alcohol. Clin. Exp. Res.* **35**, 1509–1518 (2011). doi: [10.1111/j.1530-0277.2011.01487.x](https://doi.org/10.1111/j.1530-0277.2011.01487.x); pmid: [21463341](https://pubmed.ncbi.nlm.nih.gov/21463341/)
8. D. M. Levine, T. S. Parker, T. M. Donnelly, A. Walsh, A. L. Rubin, In vivo protection against endotoxin by plasma high density lipoprotein. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 12040–12044 (1993). doi: [10.1073/pnas.90.24.12040](https://doi.org/10.1073/pnas.90.24.12040); pmid: [8265667](https://pubmed.ncbi.nlm.nih.gov/8265667/)
9. M. M. Wurfel, E. Hailman, S. D. Wright, Soluble CD14 acts as a shuttle in the neutralization of lipopolysaccharide (LPS) by LPS-binding protein and reconstituted high density lipoprotein. *J. Exp. Med.* **181**, 1743–1754 (1995). doi: [10.1084/jem.181.5.1743](https://doi.org/10.1084/jem.181.5.1743); pmid: [7536794](https://pubmed.ncbi.nlm.nih.gov/7536794/)
10. O. Meilhac, S. Tanaka, D. Couret, High-density lipoproteins are bug scavengers. *Biomolecules* **10**, 598 (2020). doi: [10.3390/biom10040598](https://doi.org/10.3390/biom10040598); pmid: [32290632](https://pubmed.ncbi.nlm.nih.gov/32290632/)
11. H. G. Windmueller, P. N. Herbert, R. I. Levy, Biosynthesis of lymph and plasma lipoprotein apoproteins by isolated perfused rat liver and intestine. *J. Lipid Res.* **14**, 215–223 (1973). doi: [10.1016/S0022-2275\(20\)36909-1](https://doi.org/10.1016/S0022-2275(20)36909-1); pmid: [4349053](https://pubmed.ncbi.nlm.nih.gov/4349053/)
12. L. R. Brunham et al., Intestinal ABCA1 directly contributes to HDL biogenesis in vivo. *J. Clin. Invest.* **116**, 1052–1062 (2006). doi: [10.1172/JCI27352](https://doi.org/10.1172/JCI27352); pmid: [16543947](https://pubmed.ncbi.nlm.nih.gov/16543947/)
13. C. Martel et al., Lymphatic vasculature mediates macrophage reverse cholesterol transport in mice. *J. Clin. Invest.* **123**, 1571–1579 (2013). doi: [10.1172/JCI63685](https://doi.org/10.1172/JCI63685); pmid: [23524964](https://pubmed.ncbi.nlm.nih.gov/23524964/)
14. H. Y. Lim et al., Lymphatic vessels are essential for the removal of cholesterol from peripheral tissues by SR-BI-mediated transport of HDL. *Cell Metab.* **17**, 671–684 (2013). doi: [10.1016/j.cmet.2013.04.002](https://doi.org/10.1016/j.cmet.2013.04.002); pmid: [23663736](https://pubmed.ncbi.nlm.nih.gov/23663736/)
15. L. H. Huang et al., Interleukin-17 drives interstitial entrapment of tissue lipoproteins in experimental psoriasis. *Cell Metab.* **29**, 475–487.e7 (2019). doi: [10.1016/j.cmet.2018.10.006](https://doi.org/10.1016/j.cmet.2018.10.006); pmid: [30415924](https://pubmed.ncbi.nlm.nih.gov/30415924/)
16. G. J. Randolph, N. E. Miller, Lymphatic transport of high-density lipoproteins and chylomicrons. *J. Clin. Invest.* **124**, 929–935 (2014). doi: [10.1172/JCI71610](https://doi.org/10.1172/JCI71610); pmid: [24590278](https://pubmed.ncbi.nlm.nih.gov/24590278/)
17. R. M. Glickman, P. H. R. Green, The intestine as a source of apolipoprotein A1. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2569–2573 (1977). doi: [10.1073/pnas.74.6.2569](https://doi.org/10.1073/pnas.74.6.2569); pmid: [196292](https://pubmed.ncbi.nlm.nih.gov/196292/)
18. W. S. Davidson et al., Proteomic analysis of defined HDL subpopulations reveals particle-specific protein clusters: Relevance to antioxidative function. *Arterioscler. Thromb. Vasc. Biol.* **29**, 870–876 (2009). doi: [10.1161/ATVBAHA.109.186031](https://doi.org/10.1161/ATVBAHA.109.186031); pmid: [19325143](https://pubmed.ncbi.nlm.nih.gov/19325143/)
19. A. N. Hoofnagle, J. W. Heinecke, Lipoproteomics: Using mass spectrometry-based proteomics to explore the assembly, structure, and function of lipoproteins. *J. Lipid Res.* **50**, 1967–1975 (2009). doi: [10.1194/jlr.R900015-JLR200](https://doi.org/10.1194/jlr.R900015-JLR200); pmid: [19738003](https://pubmed.ncbi.nlm.nih.gov/19738003/)
20. J. K. Ryu et al., Reconstruction of LPS transfer cascade reveals structural determinants within LBP, CD14, and TLR4-MD2 for efficient LPS recognition and transfer. *Immunity* **46**, 38–50 (2017). doi: [10.1016/j.immuni.2016.11.007](https://doi.org/10.1016/j.immuni.2016.11.007); pmid: [27986454](https://pubmed.ncbi.nlm.nih.gov/27986454/)
21. C. L. Scott et al., Bone marrow-derived monocytes give rise to self-renewing and fully differentiated Kupffer cells. *Nat. Commun.* **7**, 10321 (2016). doi: [10.1038/ncomms10321](https://doi.org/10.1038/ncomms10321); pmid: [26813785](https://pubmed.ncbi.nlm.nih.gov/26813785/)

22. J. S. Seidman *et al.*, Niche-specific reprogramming of epigenetic landscapes drives myeloid cell diversity in nonalcoholic steatohepatitis. *Immunity* **52**, 1057–1074 (2020). doi: [10.1016/j.immuni.2020.04.001](https://doi.org/10.1016/j.immuni.2020.04.001); pmid: [32362324](https://pubmed.ncbi.nlm.nih.gov/32362324/)
23. C. O. Odeyale, Y. H. Kang, Biotinylation of bacterial lipopolysaccharide and its applications to electron microscopy. *J. Histochem. Cytochem.* **36**, 1131–1137 (1988). doi: [10.1177/36.9.3136207](https://doi.org/10.1177/36.9.3136207); pmid: [3136207](https://pubmed.ncbi.nlm.nih.gov/3136207/)
24. J. M. Luk, A. Kumar, R. Tsang, D. Staunton, Biotinylated lipopolysaccharide binds to endotoxin receptor in endothelial and monocytic cells. *Anal. Biochem.* **232**, 217–224 (1995). doi: [10.1006/abio.1995.0010](https://doi.org/10.1006/abio.1995.0010); pmid: [8747478](https://pubmed.ncbi.nlm.nih.gov/8747478/)
25. T. Massamiri, P. S. Tobias, L. K. Curtiss, Structural determinants for the interaction of lipopolysaccharide binding protein with purified high density lipoproteins: Role of apolipoprotein A-I. *J. Lipid Res.* **38**, 516–525 (1997). doi: [10.1016/S0022-2275\(20\)37259-X](https://doi.org/10.1016/S0022-2275(20)37259-X); pmid: [9101432](https://pubmed.ncbi.nlm.nih.gov/9101432/)
26. R. S. Munford, J. P. Weiss, M. Lu, Biochemical transformation of bacterial lipopolysaccharides by acyloxyacyl hydrolase reduces host injury and promotes recovery. *J. Biol. Chem.* **295**, 17842–17851 (2020). doi: [10.1074/jbc.REV120.015254](https://doi.org/10.1074/jbc.REV120.015254); pmid: [33454018](https://pubmed.ncbi.nlm.nih.gov/33454018/)
27. C. M. Courtney, B. W. Warner, Pediatric intestinal failure-associated liver disease. *Curr. Opin. Pediatr.* **29**, 363–370 (2017). doi: [10.1097/MOP.0000000000000484](https://doi.org/10.1097/MOP.0000000000000484); pmid: [28333693](https://pubmed.ncbi.nlm.nih.gov/28333693/)
28. E. J. Onufer *et al.*, Effects of high-fat diet on liver injury after small bowel resection. *J. Pediatr. Surg.* **55**, 1099–1106 (2020). doi: [10.1016/j.jpedsurg.2020.02.037](https://doi.org/10.1016/j.jpedsurg.2020.02.037); pmid: [32164985](https://pubmed.ncbi.nlm.nih.gov/32164985/)
29. I. Spadoni *et al.*, A gut-vascular barrier controls the systemic dissemination of bacteria. *Science* **350**, 830–834 (2015). doi: [10.1126/science.1240135](https://doi.org/10.1126/science.1240135); pmid: [26564856](https://pubmed.ncbi.nlm.nih.gov/26564856/)
30. M. Sakai *et al.*, Liver-derived signals sequentially reprogram myeloid enhancers to initiate and maintain Kupffer cell identity. *Immunity* **51**, 655–670.e8 (2019). doi: [10.1016/j.immuni.2019.09.002](https://doi.org/10.1016/j.immuni.2019.09.002); pmid: [31587991](https://pubmed.ncbi.nlm.nih.gov/31587991/)
31. L. E. Cho, L. R. Yu, M. A. Abdelmegeed, S. H. Yoo, B. J. Song, Apoptosis of enterocytes and nitration of junctional complex proteins promote alcohol-induced gut leakiness and liver injury. *J. Hepatol.* **69**, 142–153 (2018). doi: [10.1016/j.jhep.2018.02.005](https://doi.org/10.1016/j.jhep.2018.02.005); pmid: [29458168](https://pubmed.ncbi.nlm.nih.gov/29458168/)
32. J. Mouries *et al.*, Microbiota-driven gut vascular barrier disruption is a prerequisite for non-alcoholic steatohepatitis development. *J. Hepatol.* **71**, 1216–1228 (2019). doi: [10.1016/j.jhep.2019.08.005](https://doi.org/10.1016/j.jhep.2019.08.005); pmid: [31419514](https://pubmed.ncbi.nlm.nih.gov/31419514/)
33. L. R. Brunham *et al.*, Tissue-specific induction of intestinal ABCA1 expression with a liver X receptor agonist raises plasma HDL cholesterol levels. *Circ. Res.* **99**, 672–674 (2006). doi: [10.1161/01.RES.0000244014.19589.8e](https://doi.org/10.1161/01.RES.0000244014.19589.8e); pmid: [16946132](https://pubmed.ncbi.nlm.nih.gov/16946132/)
34. G. Lo Sasso *et al.*, Intestinal specific LXR activation stimulates reverse cholesterol transport and protects from atherosclerosis. *Cell Metab.* **12**, 187–193 (2010). doi: [10.1016/j.cmet.2010.07.002](https://doi.org/10.1016/j.cmet.2010.07.002); pmid: [20674863](https://pubmed.ncbi.nlm.nih.gov/20674863/)
35. B. Shao, R. S. Munford, R. Kitchens, A. W. Varley, Hepatic uptake and deacylation of the LPS in bloodborne LPS-lipoprotein complexes. *Innate Immun.* **18**, 825–833 (2012). doi: [10.1177/1753425912442431](https://doi.org/10.1177/1753425912442431); pmid: [22441700](https://pubmed.ncbi.nlm.nih.gov/22441700/)
36. C. Grunfeld *et al.*, Lipoproteins inhibit macrophage activation by lipoteichoic acid. *J. Lipid Res.* **40**, 245–252 (1999). doi: [10.1016/S0022-2275\(20\)33363-0](https://doi.org/10.1016/S0022-2275(20)33363-0); pmid: [9925653](https://pubmed.ncbi.nlm.nih.gov/9925653/)
37. Y. Haberman *et al.*, Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *J. Clin. Invest.* **124**, 3617–3633 (2014). doi: [10.1172/JCI75436](https://doi.org/10.1172/JCI75436); pmid: [25003194](https://pubmed.ncbi.nlm.nih.gov/25003194/)
38. S. S. Martin *et al.*, HDL cholesterol subclasses, myocardial infarction, and mortality in secondary prevention: The Lipoprotein Investigators Collaborative. *Eur. Heart J.* **36**, 22–30 (2015). doi: [10.1093/eurheartj/ehu264](https://doi.org/10.1093/eurheartj/ehu264); pmid: [24980493](https://pubmed.ncbi.nlm.nih.gov/24980493/)
39. J. J. Albers, A. Slee, J. L. Fleg, K. D. O'Brien, S. M. Marcovina, Relationship of baseline HDL subclasses, small dense LDL and LDL triglyceride to cardiovascular events in the AIM-HIGH clinical trial. *Atherosclerosis* **251**, 454–459 (2016). doi: [10.1016/j.atherosclerosis.2016.06.019](https://doi.org/10.1016/j.atherosclerosis.2016.06.019); pmid: [27320173](https://pubmed.ncbi.nlm.nih.gov/27320173/)
40. G. Gracia, E. Cao, A. P. R. Johnston, C. J. H. Porter, N. L. Trevasakis, Organ-specific lymphatics play distinct roles in regulating HDL trafficking and composition. *Am. J. Physiol. Gastrointest. Liver Physiol.* **318**, G725–G735 (2020). doi: [10.1152/ajpgi.00340.2019](https://doi.org/10.1152/ajpgi.00340.2019); pmid: [32068443](https://pubmed.ncbi.nlm.nih.gov/32068443/)
41. G. P. Forester, A. R. Tall, C. L. Bisgaier, R. M. Glickman, Rat intestine secretes spherical high density lipoproteins. *J. Biol. Chem.* **258**, 5938–5943 (1983). doi: [10.1016/S0021-9258\(20\)81987-5](https://doi.org/10.1016/S0021-9258(20)81987-5); pmid: [6853560](https://pubmed.ncbi.nlm.nih.gov/6853560/)
42. H. R. Beartot, R. M. Glickman, L. Weinberg, P. H. Green, A. R. Tall, Effect of biliary diversion on rat mesenteric lymph apolipoprotein-I and high density lipoprotein. *J. Clin. Invest.* **69**, 210–217 (1982). doi: [10.1172/JCI10432](https://doi.org/10.1172/JCI10432); pmid: [6798073](https://pubmed.ncbi.nlm.nih.gov/6798073/)
43. H. C. Oliveira, K. Nilansen, H. Meinertz, E. C. Quintão, Cholesteryl esters in lymph chylomicrons: Contribution from high density lipoprotein transferred from plasma into intestinal lymph. *J. Lipid Res.* **34**, 1729–1736 (1993). doi: [10.1016/S0022-2275\(20\)35735-7](https://doi.org/10.1016/S0022-2275(20)35735-7); pmid: [8245723](https://pubmed.ncbi.nlm.nih.gov/8245723/)
44. E. C. Quintão, A. Drewiacki, K. Stechhaln, E. C. de Faria, A. M. Sipahi, Origin of cholesterol transported in intestinal lymph: Studies in patients with filarial chyluria. *J. Lipid Res.* **20**, 941–951 (1979). doi: [10.1016/S0022-2275\(20\)39995-8](https://doi.org/10.1016/S0022-2275(20)39995-8); pmid: [533829](https://pubmed.ncbi.nlm.nih.gov/533829)
45. I. Pierantonelli *et al.*, HDL cholesterol protects from liver injury in mice with intestinal specific LXRα activation. *Liver Int.* **40**, 3127–3139 (2020). doi: [10.1111/liv.14712](https://doi.org/10.1111/liv.14712); pmid: [33098723](https://pubmed.ncbi.nlm.nih.gov/33098723/)
46. T. Yasuda *et al.*, Tissue-specific liver X receptor activation promotes macrophage reverse cholesterol transport in vivo. *Arterioscler. Thromb. Vasc. Biol.* **30**, 781–786 (2010). doi: [10.1161/ATVBAHA.109.195693](https://doi.org/10.1161/ATVBAHA.109.195693); pmid: [20110577](https://pubmed.ncbi.nlm.nih.gov/20110577/)
47. B. Wang, P. Tontonoz, Liver X receptors in lipid signalling and membrane homeostasis. *Nat. Rev. Endocrinol.* **14**, 452–463 (2018). doi: [10.1038/s41574-018-0037-x](https://doi.org/10.1038/s41574-018-0037-x); pmid: [29904174](https://pubmed.ncbi.nlm.nih.gov/29904174/)
48. S. Daemen *et al.*, Dynamic shifts in the composition of resident and recruited macrophages influence tissue remodeling in NASH. *Cell Rep.* **34**, 108626 (2021). doi: [10.1016/j.celrep.2020.108626](https://doi.org/10.1016/j.celrep.2020.108626); pmid: [33440159](https://pubmed.ncbi.nlm.nih.gov/33440159/)
49. F. Guo, K. Zheng, R. Benedé-Ubieto, F. J. Cubero, Y. A. Nevzorova, The Lieber-DeCarli diet: A flagship model for experimental alcoholic liver disease. *Alcohol. Clin. Exp. Res.* **42**, 1828–1840 (2018). doi: [10.1111/acer.13840](https://doi.org/10.1111/acer.13840); pmid: [30025151](https://pubmed.ncbi.nlm.nih.gov/30025151/)
50. Y. H. Han *et al.*, RORα induces KLF4-mediated M2 polarization in the liver macrophages that protect against nonalcoholic steatohepatitis. *Cell Rep.* **20**, 124–135 (2017). doi: [10.1016/j.celrep.2017.06.017](https://doi.org/10.1016/j.celrep.2017.06.017); pmid: [28683306](https://pubmed.ncbi.nlm.nih.gov/28683306/)
51. W. S. Davidson, G. M. Hilliard, The spatial organization of apolipoprotein A-I on the edge of discoidal high density lipoprotein particles: A mass spectrometry study. *J. Biol. Chem.* **278**, 27199–27207 (2003). doi: [10.1074/jbc.M302764200](https://doi.org/10.1074/jbc.M302764200); pmid: [12724319](https://pubmed.ncbi.nlm.nih.gov/12724319/)
52. E. A. Bonomo, J. B. Swaney, A rapid method for the synthesis of protein-lipid complexes using adsorption chromatography. *J. Lipid Res.* **29**, 380–384 (1988). doi: [10.1016/S0022-2275\(20\)38530-8](https://doi.org/10.1016/S0022-2275(20)38530-8); pmid: [3132520](https://pubmed.ncbi.nlm.nih.gov/3132520)
53. J. Rustenhoven *et al.*, Functional characterization of the dural sinuses as a neuroimmune interface. *Cell* **184**, 1000–1016.e27 (2021). doi: [10.1016/j.cell.2020.12.040](https://doi.org/10.1016/j.cell.2020.12.040); pmid: [33508229](https://pubmed.ncbi.nlm.nih.gov/33508229/)
54. D. Patra, J. Kim, Q. Zhang, E. Tykseen, L. J. Sandell, Site-1 protease ablation in the osterix-lineage in mice results in bone marrow neutrophilia and hematopoietic stem cell alterations. *Biol. Open* **9**, bio052993 (2020). doi: [10.1242/bio.052993](https://doi.org/10.1242/bio.052993); pmid: [32576566](https://pubmed.ncbi.nlm.nih.gov/32576566/)

ACKNOWLEDGMENTS

We thank R. Apte for *Lys2-Cre×Abca1^{fl/fl} × Abcg1^{fl/fl}* mice; N. Abumrad for *Cd36^{-/-}* mice; E.B. Cho for assistance with figures; M. Wallendorf for statistics consultation; and P. Erdmann-Gilmore, Y. Mi, and R. Connors for assistance with proteomics. **Funding:** This work was supported by the National Institutes of Health (NIH grants R01DK119147 to G.J.R. and B.W.W. and AIO499653 to G.J.R., including a Primary Caregiver Supplement to E.J.O. Y.-H.H. was supported in part by the National Research Foundation (NRF) of Korea (2021R1C1C1004023). E.J.O. also received support from NIH T32DK077653. L.-H.H. was supported by an AHA Career Development Award (AHA-18CDA34110273). R.S.C. was supported by a Lawrence C. Pakula, MD IBD Research Fellowship. M.G.S.-T. was supported by NIH R01HL127649 and HL138908. Further core facility support was to the Digestive Diseases Research Core Center of Washington University (P30 DK052574); to Washington University Center for Cellular Imaging (WUCCI) by the Children's Discovery Institute of Washington University (CDI-CORE-2015-505 and CDI-CORE-2019-813) and the Foundation for Barnes-Jewish Hospital (3770); to Washington University Proteomics Shared Resource, directed by R. R. Townsend) by the WU Institute of Clinical and Translational Sciences (NCATS UL1 TR000448), the Mass Spectrometry Research Resource (NIGMS P41 GM103422; R24GM136766) and the Sitman Comprehensive Cancer Center (NCI P30 CA091842); and to the Genome Technology Access Center by NCI Cancer Center grant no. P30 CA91842 and ICTS/CTSA grant no. UL1TR002345 from the National Center for Research Resources (NCRR). **Author contributions:** Y.-H.H., E.J.O., M.G.S.-T., and G.J.R. designed the experiments. Y.-H.H., E.J.O., L.-H.H., R.S.C., R.W.S., and M.W. conducted experiments and/or developed techniques. W.S.D. generated materials. Y.-H.H. and G.J.R. analyzed data. Y.-H.H. and G.J.R. wrote the manuscript. L.-H.H., R.S.C., B.W.W., W.S.D., and M.G.S.-T. edited the manuscript. G.J.R., B.W.W., and M.G.S.-T. obtained compliance approvals. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Gene expression raw data have been deposited in the Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Information under accession number GSE167983. Proteomics data have been deposited at ProteomeXchange (no. PXD025648; the massIVE dataset name is MSV000079070). All other data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/eabe6729/suppl/DC1
Figs. S1 to S10
Tables S1 and S2
MDAR Reproducibility Checklist

Accepted 3 June 2021

6 September 2020; resubmitted 16 March 2021
10.1126/science.abe6729

RESEARCH ARTICLE SUMMARY

ENZYMOLOGY

Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics

C. J. Markin[†], D. A. Mokhtari[†], F. Sunden, M. J. Appel, E. Akiva, S. A. Longwell, C. Sabatti, D. Herschlag^{*}, P. M. Fordyce^{*}

INTRODUCTION: Enzymes possess extraordinary catalytic proficiency and specificity. These properties ultimately derive from interactions not just between the active-site residues and the substrate but from functional interactions throughout a folded enzyme. Therefore, understanding the origins of catalytic proficiency and specificity will require the ability to make mutations throughout the protein. Traditionally, enzyme active sites have been characterized by means of site-directed mutagenesis (SDM), revealing much about the catalytic functions of these residues; nevertheless, SDM is low-throughput, costly, and labor intensive. By contrast, recently developed high-throughput mutational scanning techniques assay large numbers of sequences but provide only coarse estimates of function, such as the amount of product generated at a particular time under a particular set of conditions or overall organismal fitness.

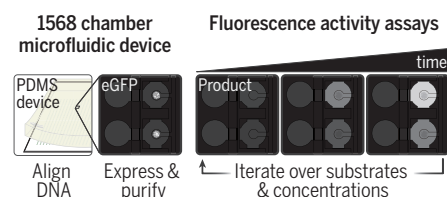
RATIONALE: New technologies are needed to overcome the limitations of current approaches and allow deep characterization of many enzyme variants in a cost- and time-efficient manner. To meet this challenge, we developed a high-throughput microfluidic platform that allows the simultaneous expression and purification of more than 1500 rationally chosen enzyme mutants in hours and allows their quantitative functional characterization in days. HT-MEK (High-Throughput Microfluidic Enzyme Kinetics) can be used with any enzyme system that can be tagged and expressed *in vitro* and has a direct or coupled fluorogenic assay.

RESULTS: As a first application of HT-MEK, we functionally characterized 1036 single-site mutants that contain either a glycine or valine substitution at each position within PafA (phosphate-irrepressible alkaline phosphatase

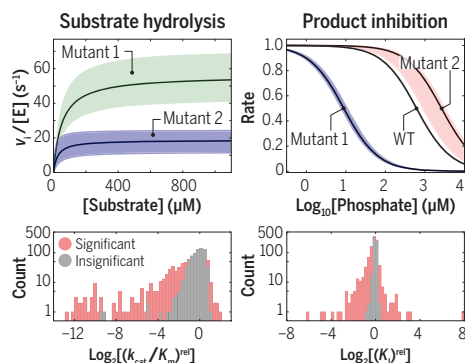
of *Flavobacterium*), a well-studied enzyme from the alkaline phosphatase superfamily. For each mutant, we measured Michaelis-Menten kinetics [apparent unimolecular rate constant (k_{cat}), Michaelis constant (K_m), and k_{cat}/K_m] for multiple substrates, inhibition constants, and effects on folding, obtaining more than 5000 kinetic and thermodynamic constants from more than 670,000 total reactions. We found that most mutations (702 of 1036) yielded statistically significant effects on some aspect of catalysis. By systematically and independently varying expression and assay conditions, we determined that 232 of these mutations reduced catalysis by promoting the formation of a long-lived, catalytically inactive misfolded state, whereas none did so through equilibrium unfolding under our assay conditions. Combining these functional measurements with prior mechanistic knowledge allowed us to systematically assess the effect of each mutation. Different groups of residues affected different aspects of function, with residues that affect a particular function forming large, spatially contiguous regions that spanned from the active site up to 20 Å from the active site and to the enzyme surface.

CONCLUSION: HT-MEK has allowed us to uncover functional effects throughout PafA and to identify the catalytic features affected by different groups of residues. Some of these effects are readily rationalized through inspection of structural interconnections to the active site, whereas others were nonobvious, including large distal and surface effects and the discovery of a long-lived misfolded state. These results underscore the need to measure the effects of mutations on multiple kinetic and thermodynamic parameters across multiple reaction conditions and thus the need for this new technology. Because HT-MEK is applicable to any enzyme with a direct or coupled fluorescent readout and provides an in-depth and quantitative analysis of mutant space rapidly and at modest cost, it may be the method of choice to characterize new enzymes. In future applications, HT-MEK can be used to dissect potential evolutionary trajectories, determine the functional consequences of human disease-associated allelic variants, identify surfaces with nascent allosteric potential for rational control of catalysis, and direct the adaptation of natural and designed enzymes for new functions and roles. ■

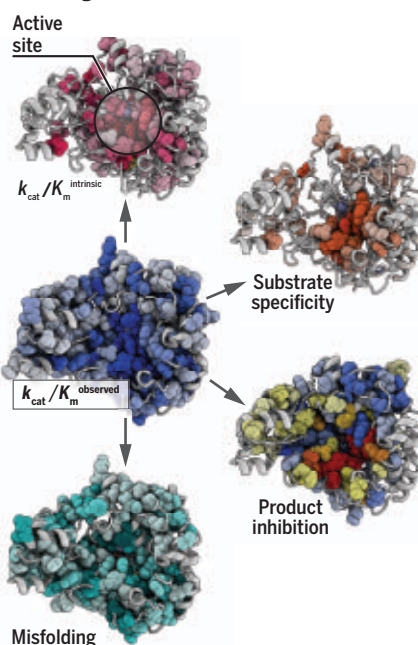
High-throughput quantitative enzymology



Kinetic and thermodynamic parameters for 1036 PafA phosphatase variants



Systematic, multidimensional measurements revealing PafA functional architecture



Simultaneous expression, purification, and biochemical characterization of enzyme variants in a microfluidic device makes it possible to measure Michaelis-Menten parameters and inhibition constants for more than 1500 variants in days. Mutational effects across multiple assays reveal an extensive functional architecture in which physically contiguous residue regions extending to the enzyme surface control or alter particular aspects of catalysis.

The list of author affiliations is available in the full article online.

^{*}Corresponding author. Email: herschlag@stanford.edu

(D.H.); pfordyce@stanford.edu (P.M.F.)

[†]These authors contributed equally to this work.

Cite this article as C. J. Markin *et al.*, *Science* 373, eabf8761 (2021). DOI: 10.1126/science.abf8761

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abf8761>

RESEARCH ARTICLE

ENZYMOLOGY

Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics

C. J. Markin^{1†}, D. A. Mokhtari^{1†}, F. Sunden¹, M. J. Appel¹, E. Akiva², S. A. Longwell³, C. Sabatti^{4,5}, D. Herschlag^{1,6,7*}, P. M. Fordyce^{3,7,8,9*}

Systematic and extensive investigation of enzymes is needed to understand their extraordinary efficiency and meet current challenges in medicine and engineering. We present HT-MEK (High-Throughput Microfluidic Enzyme Kinetics), a microfluidic platform for high-throughput expression, purification, and characterization of more than 1500 enzyme variants per experiment. For 1036 mutants of the alkaline phosphatase PafA (phosphate-irrepressible alkaline phosphatase of *Flavobacterium*), we performed more than 670,000 reactions and determined more than 5000 kinetic and physical constants for multiple substrates and inhibitors. We uncovered extensive kinetic partitioning to a misfolded state and isolated catalytic effects, revealing spatially contiguous regions of residues linked to particular aspects of function. Regions included active-site proximal residues but extended to the enzyme surface, providing a map of underlying architecture not possible to derive from existing approaches. HT-MEK has applications that range from understanding molecular mechanisms to medicine, engineering, and design.

Understanding how sequence encodes function remains a fundamental challenge in biology. Linear chains of amino acids fold into three-dimensional protein structures that carry out the physical and chemical tasks needed for life, such as highly efficient and specific catalysis. Sequence variations across organisms and individuals confer beneficial and deleterious effects: Variation throughout evolution creates proteins with improved or new functions, but variation among individuals can also compromise function and cause disease (1–3). An enhanced predictive understanding of the sequence-function landscape could have profound impacts across biology, from enabling efficient protein design to improving detection of rare allelic variants that drive disease (4–7), but new approaches and data are needed to attain this goal.

Understanding sequence-function relationships within enzymes poses a particular challenge. Structural and biochemical studies of enzymes have revealed the sites of substrate binding and catalytic transformation, the residues directly involved in catalysis, and roles

for these residues. Nevertheless, residues outside the active site are needed for the active site to assemble and function and for control of function by allosteric ligands and covalent modifications (8–10). Despite their importance, the roles played by residues outside the active site, which comprise the majority of amino acids in an enzyme, remain largely unexplored.

This dearth of knowledge stems from the nature of experimental approaches currently available. Site-directed mutagenesis (SDM) has traditionally been used to assess function by means of in-depth biochemical assays that yield kinetic and thermodynamic constants. However, SDM is time-, resource-, and labor-intensive, limiting investigation to a small number of residues. By contrast, deep mutational scanning (DMS) provides the ability to assay the effects of all 20 amino acids at every position within an enzyme (5, 11, 12). However, DMS lacks the depth and dimensionality of traditional SDM studies, typically providing a scalar readout with an uncertain relationship to the multiple fundamental physical constants needed to describe an enzyme's function.

Marrying the strengths of traditional SDM and emerging DMS is needed to usher in a new era of mechanistic enzymology. Here, we present HT-MEK (High-Throughput Microfluidic Enzyme Kinetics), a platform capable of simultaneously expressing, purifying, and characterizing more than 1500 rationally chosen enzyme mutants in parallel with the depth and precision of traditional SDM. Each HT-MEK experiment provides 1000s of measurements and multiple kinetic and thermodynamic constants [such as the apparent unimolecular rate constant (k_{cat}), Michaelis constant (K_m), $k_{cat}/$

K_m , and inhibition constant (K_i)] in days and at low cost.

To guide HT-MEK development and demonstrate its capabilities, we carried out a comprehensive mechanistic investigation of the effects of mutations to every residue within the alkaline phosphatase superfamily member PafA (phosphate-irrepressible alkaline phosphatase of *Flavobacterium*) (Fig. 1A and fig. S1). PafA and related phosphomonoesterases are among the most prodigious catalysts known, with rate enhancements of up to $\sim 10^{27}$ -fold, providing a large dynamic range to explore (13). We also anticipated that PafA, a secreted enzyme, would be highly stable, potentially allowing us to more deeply probe catalysis without obfuscation from global unfolding. We found that 702 of the 1036 mutants investigated have substantial functional consequences, with none arising from equilibrium unfolding. Additional experiments revealed that many mutations promote the formation of a long-lived, catalytically incompetent misfolded state both in vitro and in cells. The multidimensional measurements provided by HT-MEK allowed us to decouple this misfolding from catalytic effects and quantify mutational effects on particular aspects of catalysis and mechanism using an approach we call Functional Component Analysis. This approach revealed spatially contiguous regions of residues that extend from the active site to the enzyme surface responsible for optimizing particular catalytic strategies. Surface residues with functional effects upon mutation may compose candidate allosteric regions useful for engineering rational control of catalytic activity. The HT-MEK platform and the quantitative multidimensional datasets it can provide will have broad utility for future efforts to understand catalytic mechanisms, natural variation, and evolutionary trajectories and to design enzymes with new functions.

Results

HT-MEK device and experimental pipeline

HT-MEK is built around a two-layer polydimethylsiloxane (PDMS) microfluidic device with 1568 chambers and integrated pneumatic valves (Fig. 1B and fig. S2) (14, 15). Each chamber is composed of two compartments (DNA and Reaction) separated by a valve (Neck), with adjacent chambers isolated from one another by a second valve (Sandwich). A third valve (Button) reversibly excludes or exposes a circular patch of the reaction compartment surface, enabling surface patterning for on-chip protein immobilization and purification (Fig. 1C) and subsequent simultaneous initiation of successive on-chip reactions across the device (Fig. 1D). Each DNA compartment of each chamber is programmed with a specified enzyme variant by aligning the device to a spotted array of DNA plasmids that encode for the expression of C-terminally enhanced green

¹Department of Biochemistry, Stanford University, Stanford, CA 94305, USA. ²Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA 94158, USA. ³Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. ⁴Department of Biomedical Data Science, Stanford University, Stanford, CA 94305, USA. ⁵Department of Statistics, Stanford University, Stanford, CA 94305, USA. ⁶Department of Chemical Engineering, Stanford University, Stanford, CA 94305, USA. ⁷ChEM-H Institute, Stanford University, Stanford, CA 94305, USA. ⁸Department of Genetics, Stanford University, Stanford, CA 94305, USA. ⁹Chan Zuckerberg Biohub, San Francisco, CA 94110, USA.

*Corresponding author. Email: herschla@stanford.edu (D.H.); pfordyce@stanford.edu (P.M.F.)

†These authors contributed equally to this work.

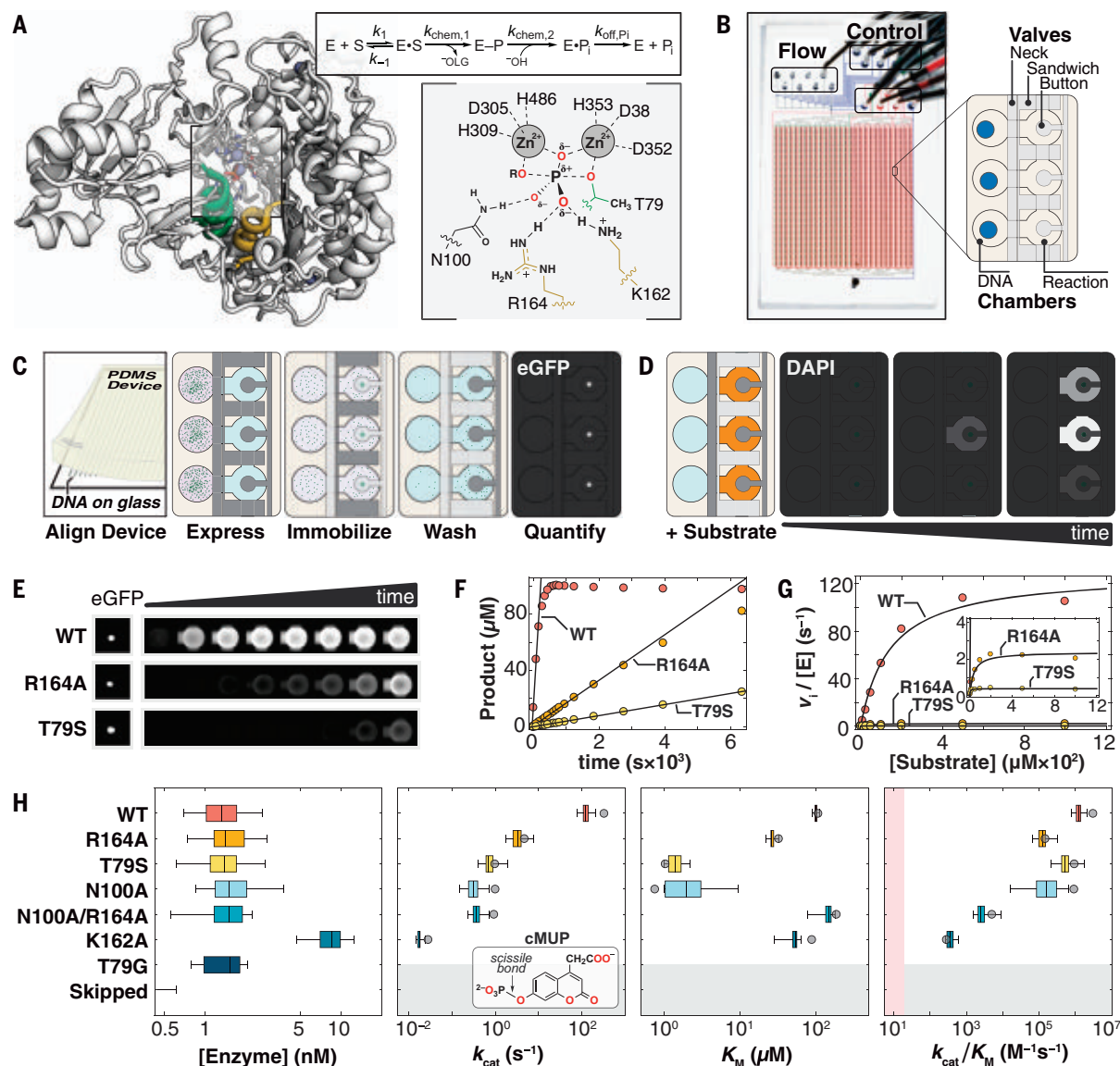


Fig. 1. Overview of HT-MEK and PafA. (A) (Left) Crystal structure of WT PafA [Protein Data Bank (PDB) ID: 5JTJ] highlighting “nucleophile” (residues 77 to 89, green) and “monoesterase” (residues 161 to 171, yellow) helices. (Right) PafA catalytic cycle (S, phosphate monoester substrate; E–P, covalent phospho-threonine intermediate; $^{\text{OLG}}$, alkoxide leaving group) and active site in the transition state of monoester hydrolysis. (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. In the mutants, other amino acids were substituted at certain locations; for example, T79S indicates that threonine at position 79 was replaced by serine.) (B) HT-MEK microfluidic device image and schematic showing solution (Flow) and pneumatic manifold (Control) input ports, device valves, and chambers. (C) Schematic of on-chip enzyme expression pipeline. Dark and light gray valves are pressurized (closed) and depressurized (open), respectively. (D) Schematic of

on-chip activity assays by using fluorogenic substrate in reaction chambers (orange). (E) Sample images of (left) immobilized enzyme and (right) fluorogenic product over time for WT PafA and two active-site mutants (R164A and T79S) with the substrate cMUP. (F) Example cMUP progress curves for chambers containing wild type and two active-site mutants and initial rate fits to these data. (G) Michaelis-Menten fits to initial rates yield k_{cat} , K_{M} , and $k_{\text{cat}}/K_{\text{M}}$ for cMUP. (H) (Left) On-chip expressed concentrations for WT PafA and six active-site mutants calculated by using eGFP calibration curves. (Middle left to right) Comparisons of on-chip (box plots) and off-chip (gray circle) values of k_{cat} , K_{M} , and $k_{\text{cat}}/K_{\text{M}}$ values (log₁₀ axis) for cMUP for seven PafA variants (table S1) and (inset) cMUP structure. K162A was expressed at higher concentration in a later experimental tier (with higher $[E]$); “skipped” refers to chambers without plasmid DNA; pink shaded region denotes 10-fold above the median apparent second-order rate constant for T79G control chambers (supplementary materials, materials and methods).

fluorescent protein (eGFP)-tagged variants (fig. S3). After alignment, device surfaces are patterned with antibodies to eGFP beneath the Button valve and passivated with bovine serum albumin elsewhere. All enzymes are then expressed in parallel through the introduction

of an *Escherichia coli* in vitro transcription-translation system and purified by means of capture with surface-patterned immobilized antibody and washing (fig. S4). Production of up to 1568 different purified enzymes takes ~10 hours, with most steps automated. Enzymes

are immobilized under the Button valves that protect against flow-induced loss of enzyme during solution exchange and allow repeated synchronous initiation of reactions.

To obtain catalytic rate parameters, we quantify (i) the concentration of immobilized

enzyme in each chamber, using an eGFP calibration curve (fig. S5), and (ii) the amount of product formed as a function of reaction time, using a chamber-specific product calibration curve (fig. S6). We then fit reaction progress curves in each chamber to obtain initial rates (v_i) for each substrate concentration using a custom image-processing pipeline and convert observed rates (v_i) to enzyme-normalized rate constants according to the eGFP intensity in each chamber and calibration curve (Fig. 1, E and F, and figs. S7 and S8). This process—repeated on a single device for multiple substrate concentrations, multiple substrates, and multiple inhibitors—provides the data necessary to obtain Michaelis-Menten parameters and other kinetic and thermodynamic constants (Fig. 1G and fig. S7).

HT-MEK reproduces kinetic constants previously measured with traditional assays

To demonstrate the technical capabilities of HT-MEK, we applied it to study seven previously characterized PafA variants: wild type, five active site mutants (T79S, N100A, R164A, K162A, and N100A/R164A), and one mutant lacking detectable activity (T79G, negative control) (16). Activities of wild-type (WT) PafA and the six mutants span a broad range in k_{cat} ($>10^4$ -fold), k_{cat}/K_m ($>10^4$ -fold), and K_m ($>10^2$ -fold) for aryl phosphate monoester hydrolysis, providing a stringent initial test of HT-MEK dynamic range (table S1). Nearly all DNA-containing chambers expressed enzyme ($>90\%$), and all mutants expressed at similar levels as determined from eGFP fluorescence (K162A was deliberately expressed at higher concentrations in a later experimental tier, described below) (Fig. 1H and fig. S9).

Although fluorogenic phosphate ester substrates permit kinetic assays of phosphatase activity with a high dynamic range, microfluidic assays that use the commercial 4-methylumbelliferyl phosphate ester (MUP) were complicated by partitioning of the hydrophobic fluorescent product into the hydrophobic PDMS, which increases background and distorts kinetic measurements. To address this limitation, we synthesized MUP derivatives of similar reactivity (cMUP and a corresponding methyl phosphodiester, MecMUP) that bear a charged moiety on the leaving group to eliminate PDMS absorption (figs. S10 to S12 and table S1).

Accurately resolving enzymatic rates spanning many orders of magnitude poses technical challenges because different acquisition times are needed at catalytic extremes (fig. S13), and even a small concentration of contaminating fast enzyme introduced into nearby chambers during fluid exchanges can obscure the true rates for the most catalytically compromised mutants. To address the first challenge, we expressed enzymes at two concentrations: ~ 1.5 nM

(for accurate measurement of fast enzymes) and ~ 15 nM (to speed reactions for efficient detection of slow enzymes) (Fig. 1H and fig. S14). To identify regions of the device with contaminating enzymes from other chambers, we interspersed chambers that were empty or contained the inactive T79G mutant and also measured their apparent activity (figs. S15 and S16 and table S2). Per-device normalizations (0.4 and 1.4-fold) were used to account for small variations in apparent activity owing to nonspecific adsorption of mutant enzymes to chamber walls (supplementary materials, materials and methods). This normalization increased precision across replicates but did not affect conclusions (fig. S17). HT-MEK assays recapitulated cMUP kinetic parameters (k_{cat} , K_m , and k_{cat}/K_m) accurately and over a wide dynamic range ($>10^4$ -fold in k_{cat}/K_m), indicating that the device and added eGFP tag did not alter activity (Fig. 1H and table S3).

Many mutations throughout PafA affect phosphate monoester hydrolysis

To explore functional effects of mutations throughout PafA, we created mutant libraries in which we introduced two residues with widely differing side-chain properties at each position: (i) glycine, to ablate side chain interactions and increase backbone flexibility, and (ii) valine, to introduce a branched chain hydrophobe of average volume. Native valine and glycine residues were mutated to alanine. Nearly all of these 1052 possible mutants (1036; 98%) were successfully cloned, sequenced, expressed, and assayed with HT-MEK (fig. S18).

We first measured the catalytic effect of each substitution on the steady-state kinetic parameters for cMUP hydrolysis (k_{cat} , K_m , and k_{cat}/K_m) (Fig. 2A). To facilitate efficient measurement, we performed experiments in three tiers based on reaction rates (table S2). In tier 1, we assayed devices that contained all variants at low enzyme concentration ($[E] \sim 1.5$ nM), with enzymes printed in duplicate and the above-noted active site mutants distributed throughout as fiducial controls. Tier 2 and 3 measurements, using higher enzyme concentration ($[E] \approx 15$ nM), focused successively on the slowest variants with increasing assay times and increasing numbers of replicates to provide high-precision measurements of these slower mutants. Each device was used to measure tens of cMUP progress curves, and all expressed variants were stable over days, facilitating high-throughput data collection. Per-experiment data reports contain all data collected for each chamber, including initial rate plots and fit Michaelis-Menten curves (an example is provided in fig. S19, and full data sets are available as data files S1 and S2); per-mutant summaries combine data from all experiments and include estimates of statistical significance (Fig. 2, B and C, and fig.

S20). In total, we acquired a median of nine and seven replicates for valine and glycine mutants, respectively, over 16 experiments (figs. S21 to 24 and table S2). The wealth and precision of these data allowed us to resolve differences across ranges of 10^4 , 10^2 , and 10^5 -fold for k_{cat} , K_m , and k_{cat}/K_m , respectively.

As expected, mutations of active-site residues and catalytic Zn^{2+} ligands were highly deleterious, and positional effects varied for valine and glycine (Fig. 2, D to F). Nevertheless, a surprisingly large number of mutations throughout the enzyme were deleterious, with decreases in k_{cat}/K_m observed for 267 of the 1036 mutants ($P < 0.01$). We also observed 35 mutants with increased activity (fig. S21). These measurements provide a quantitative survey of Michaelis-Menten kinetic constants for mutations throughout a large enzyme but do not tell us why so many mutations alter activity.

The most obvious explanation for these widespread effects would be destabilization leading to a significant fraction of unfolded enzyme. Beyond destabilization, mutations can have other repercussions, altering the catalytic effectiveness of particular active-site residues, reducing Zn^{2+} affinity at the bimetallo active site or altering enzymatic protonation states. The ability to efficiently measure catalytic activity for all mutants under different assay and expression conditions as afforded by HT-MEK allowed us to test each of these possibilities.

Widespread mutational effects do not arise from equilibrium unfolding

Reflecting its role as a secreted phosphatase designed to function in harsh and variable environments, WT PafA is highly stable and remains folded, as inferred from circular dichroism (CD) spectra, even after exposure to 4 M urea for 14 days (fig. S25). This stability suggests that any individual mutation is unlikely to substantially unfold the enzyme. To directly test this expectation, we measured cMUP activities in the presence of increasing concentrations of urea. If a variant were already partially unfolded in the absence of urea, then even low concentrations of added urea would cause substantial additional unfolding (17), proportionally lowering activity (Fig. 2G and supplementary text S1). By contrast, we observed only minor rate effects for all mutants (six \pm onefold decrease at 3 M compared with 0 M urea, \pm SEM), which is considerably less than the $>10^9$ -fold decrease predicted for an unfolding effect and consistent with inhibition by urea, with $K_i = 2.6$ M for all variants (Fig. 2H and fig. S26).

Equilibrium unfolding would also predict a Zn^{2+} concentration dependence for the observed rate because added or removed Zn^{2+} would pull the equilibrium between folded (and Zn^{2+} -bound) and unfolded states (Fig. 2G). Consistent with an absence of unfolding

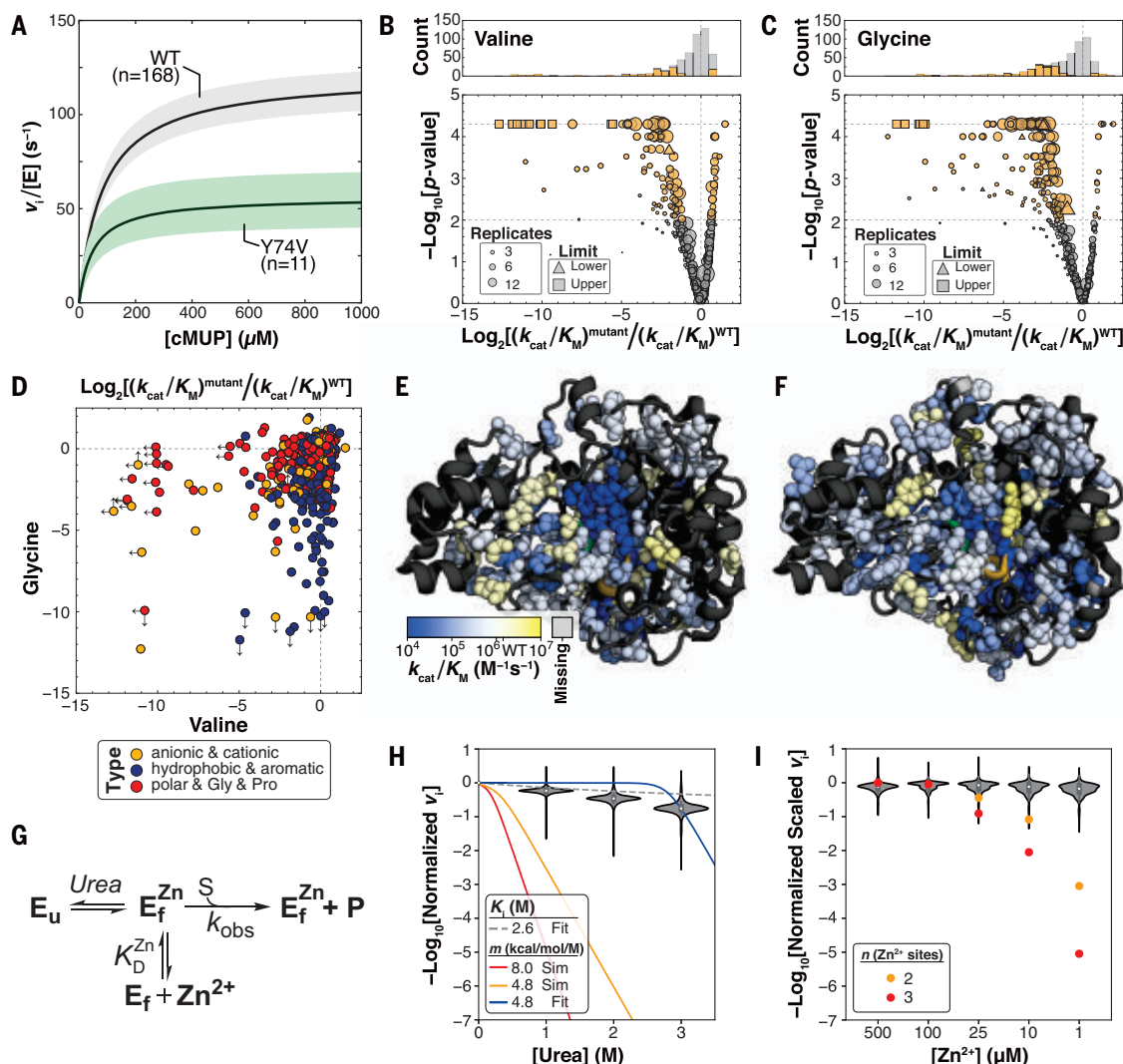


Fig. 2. HT-MEK measurements of aryl phosphate monoester (cMUP) hydrolysis for valine and glycine scans of PafA. (A) Median Michaelis-Menten curves for WT PafA and an example mutant with k_{cat} and K_{M} effects. Colored regions denote 99% confidence intervals (CIs) on medians of parameters from replicates. (B) Valine substitutions at 126 positions alter cMUP $k_{\text{cat}}/K_{\text{M}}$ at $P < 0.01$ (105 slower, 21 faster, gold markers; gray, $P \geq 0.01$). (C) Glycine substitutions at 176 positions alter cMUP $k_{\text{cat}}/K_{\text{M}}$ at $P < 0.01$ (162 slower, 14 faster, gold; gray, $P \geq 0.01$). (D) Effect of glycine versus valine substitution on cMUP $k_{\text{cat}}/K_{\text{M}}$ at each position, colored by identity type of the native residue. Arrows denote the presence and direction of measurement limits. (E and F) Valine (E) and glycine (F) substitution effects on cMUP hydrolysis on the PafA structure. $P < 0.01$ sites shown as spheres, $P \geq 0.01$ and missing

sites shown as ribbons. Alternate views are provided in fig. S21. (G) Model of equilibrium unfolding with varying urea and Zn^{2+} . (H) Activity of mutants relative to wild type at 50 μM cMUP as a function of urea. Red and yellow lines show predicted dependencies if mutants were 10% unfolded in the absence of urea, assuming m values of 4.8 and 8 kcal/mol/M (supplementary text S1). The blue curve is the best fit of the data for $m = 4.8$ kcal/mol/M with fraction unfolded (f_u) at 0 M [urea] as a free parameter (fit $f_u = 1 \times 10^{-10}$), and the dashed gray curve is the fit competitive inhibition of WT PafA by urea (fig. S26). (I) Normalized activity of mutants relative to wild type as a function of $[\text{Zn}^{2+}]$. Orange and red points show expected activities for Zn^{2+} -dependent unfolding events if mutants were 10% unfolded at 100 μM Zn^{2+} assuming two or three Zn^{2+} binding events, respectively.

effects, we observed no dependence on Zn^{2+} concentration over a $\sim 10^3$ -fold range (Fig. 2I and supplementary text S1). This observation also establishes that observed rate effects for mutants with measurable rates do not arise from loss of bound Zn^{2+} owing to lowered Zn^{2+} affinity. Last, mutants were unaltered in their pH dependencies (fig. S27 and supplementary text S1), ruling out altered protonation states as responsible for observed kinetic effects.

A general high-throughput assay for phosphate release and additional mutational effects

Although fluorogenic probes provide a sensitive and convenient method for directly visualizing enzyme activity in kinetic assays, many reactions lack a direct fluorogenic readout. To allow future application of HT-MEK to a much broader range of enzymes, we developed an on-chip coupled assay in which inorganic phosphate (P_i) is detected through fluorescence em-

itted upon binding to a modified phosphate binding protein (PBP) (Fig. 3A) (18). Calibration curves for P_i and PBP and control measurements using the PafA substrate methyl phosphate (MeP) established that this coupled on-chip assay can detect submicromolar P_i formation and accurately reproduce off-chip kinetic constants (figs. S28 to S30).

Beyond their limited availability, fluorogenic substrates are often more reactive than

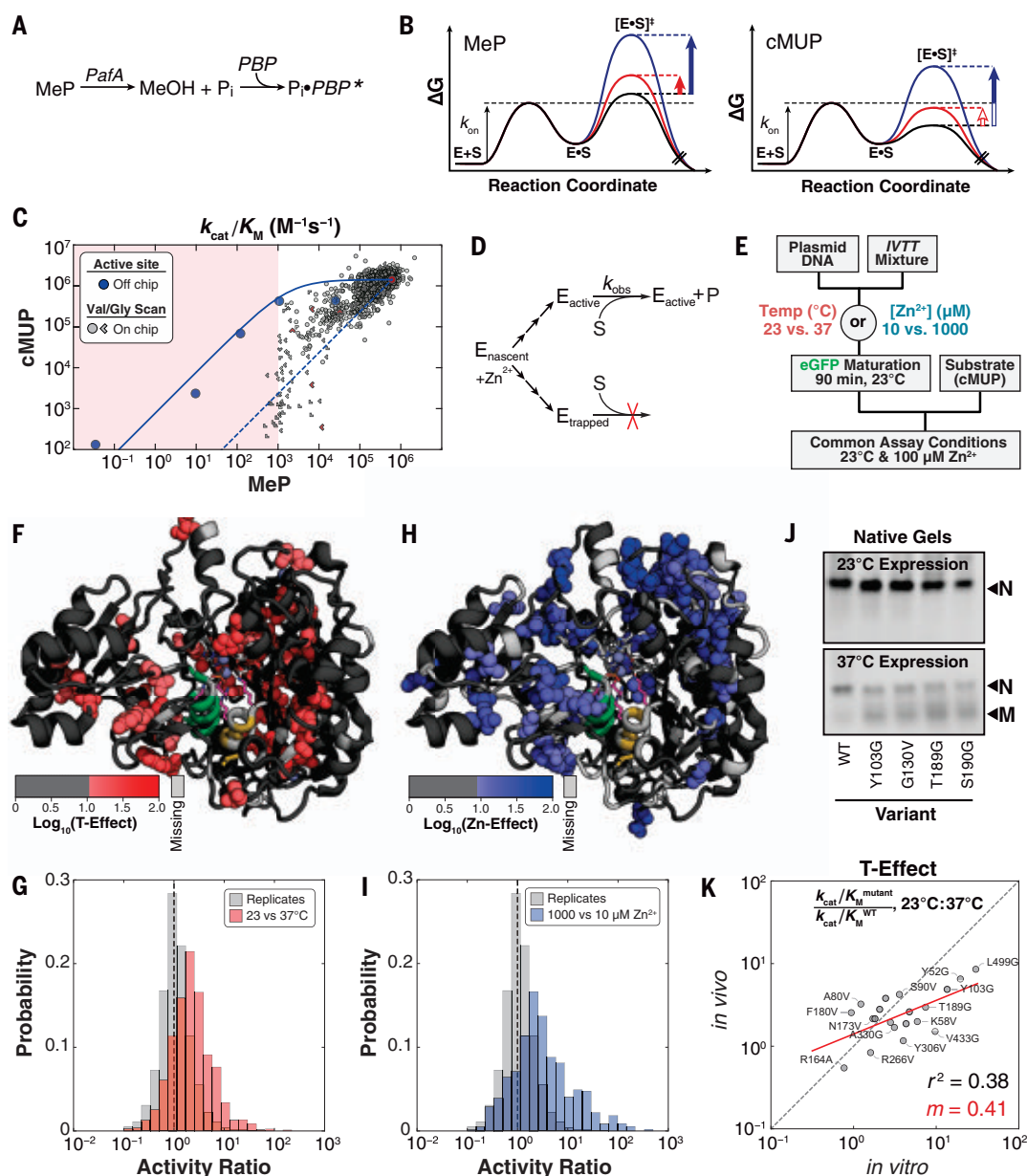


Fig. 3. Substrates with different intrinsic reactivity reveal kinetic folding traps.

(A) Schematic illustrating real-time detection of P_i produced by MeP hydrolysis on-chip using a fluorophore-conjugated PBP (18). (B) Reaction coordinate diagrams schematically illustrating substrates with either chemistry (MeP) or binding (cMUP) as the rate-limiting step. Solid arrows indicate observable mutational effects on the chemical step, and open arrows indicate effects that are not observable. (C) Measured k_{cat}/K_M for MeP versus k_{cat}/K_M for cMUP for active-site mutants off-chip (blue points) and for valine and glycine mutants. Limits in one or both directions are indicated with chevrons pointing in the direction or quadrant of the limit. The red shaded region denotes lower limit of detection for MeP, the solid blue line indicates predicted relationship between k_{cat}/K_M for cMUP and MeP hydrolysis, and the dashed blue line shows predicted relationship for enzymes with WT activity but a varying fraction of inactive enzyme (supplementary text S2). (D) Scheme showing kinetic partitioning between folded (active) and misfolded (kinetically trapped) enzyme during expression. (E) Pulse-chase experiment to test for expression-dependent misfolding by varying expression conditions while keeping reaction conditions constant. (F) Residues with temperature-dependent activity changes shown

as spheres on the PafA structure, with active-site residues N100, K162, R164, and T79 shown as purple sticks. (G) Histograms of the ratio of measured activities after expression at 23° and 37°C [k_{cat}/K_M ($k_{\text{cat}}/K_M^{\text{mutant}}/k_{\text{cat}}/K_M^{\text{WT}}$), 23°C/37°C] (red) compared with the ratios of randomly paired replicates expressed under standard conditions (37°C and 100 μM Zn^{2+} ; gray). (H) Residues with expression [Zn^{2+}]-dependent activity changes shown as spheres on the PafA structure. (I) Histograms of the ratio of measured activities across Zn^{2+} concentrations [k_{cat}/K_M ($k_{\text{cat}}/K_M^{\text{mutant}}/k_{\text{cat}}/K_M^{\text{WT}}$), 1000 μM /10 μM Zn^{2+}] (blue) compared with the ratios of randomly paired replicates expressed under standard conditions (37°C and 100 μM Zn^{2+} ; gray). (J) (Top) Native gels for enzymes expressed by means of in vitro transcription-translation off-chip at 23°C. All constructs appear as a single band corresponding to the natively folded species ("N"). (Bottom) At 37°C, non-WT variants have an additional band corresponding to the putatively misfolded ("M") species. (K) Expression temperature effects on-chip (in vitro) versus in *E. coli* (in vivo). The red line indicates linear regression of log-transformed data with the indicated slope (values provided in table S5). In vivo protein levels were determined from eGFP fluorescence to estimate total expressed protein concentration.

naturally occurring substrates, potentially rendering binding rate-limiting and obscuring mutational effects on the chemical step of catalysis, and there is evidence for this behavior with PafA (16). On-chip MeP hydrolysis assays, therefore, complement the wide dynamic range of cMUP assays by identifying modest effects on the chemical step. Off-chip measurements of several active-site mutants revealed a decrease in observed MeP activity (owing to effects on the chemical step) without a concomitant change in cMUP activity, as expected for rate-limiting cMUP binding (Fig. 3B, red vertical arrows). Once transition-state destabilization was sufficiently large, MeP and cMUP reactions both slowed (Fig. 3B, blue vertical arrows). The solid blue line in Fig. 3C is a fit to the rate model derived from the free energy-reaction profiles in Fig. 3B for a series of active-site mutants (Fig. 3C, blue points; figs. S31; and supplementary text S2) (19) and predicts the kinetic behavior expected for the PafA glycine- and valine-scanning library mutants.

HT-MEK kinetic measurements revealed k_{cat}/K_m effects for almost half of the mutants (498 of 1035 with $P < 0.01$) (fig. S32), but few exhibited the predicted behavior (Fig. 3C, solid blue line versus gray symbols). Instead, the mutants tended to fall between the predicted line and a diagonal line representing equally deleterious effects on the reactions of both substrates (Fig. 3C, blue solid and dashed lines, respectively). Equally deleterious effects are expected for enzymes with WT activity but only a fraction of the enzyme in the active configuration, with less correctly folded variants further down the diagonal. Thus, observed intermediate effects could represent combinations of effects on the chemical step and on the fraction of the mutant population that is active.

Many mutations reduce catalysis by altering folding

The urea, Zn^{2+} , and pH data presented above provided strong evidence against equilibrium unfolding for any of the variants (Fig. 2, H and I). We therefore considered and tested an alternate model in which inactive enzyme resulted from a nonequilibrium process: the formation of long-lived misfolded proteins during expression (Fig. 3D) (20). Because temperature is known to affect folding efficiency (21, 22), and because PafA binds multiple Zn^{2+} ions during folding, we varied the expression temperature (23° and 37°C) and Zn^{2+} concentration (10 and 1000 μM) from our standard expression conditions (37°C and 100 μM Zn^{2+}). We then measured reaction rates under identical assay conditions (23°C and 100 μM Zn^{2+} , our standard assay conditions), so that any observed rate changes must arise from differences during folding that persisted over time (Fig. 3E). Many mutations had differential ef-

fects on observed catalytic activity when expressed at 23° versus 37°C (“T-Effect”) or with different concentrations of Zn^{2+} (“Zn-Effect”). Whereas T-Effects were found predominantly in the enzyme core, Zn-Effects were concentrated around the distal Zn^{2+} (Fig. 3, F to I, and figs. S33 and S34). These results strongly support the presence of persistent nonequilibrium folding effects, which have been anecdotally observed in other systems (20, 21, 23–25). A second prediction of the misfolding model was also met: Variants with T- and Zn-Effects were not systematically altered in K_m under varied expression conditions (fig. S35), as expected under a model of nonequilibrating states. PafA folding thus apparently involves one or more branchpoints sensitive to temperature and Zn^{2+} that lead to active PafA or one or more long-lived inactive states (Fig. 3D).

Altered folding pathways promote a long-lived inactive state in vitro and in vivo

Misfolding could be an artifact of high-throughput on-chip expression or could also arise during standard expression in vitro and possibly in vivo. To test for chip-induced misfolding effects, we selected 19 variants with varying predicted amounts of misfolding for off-chip expression by means of in vitro transcription-translation and kinetic characterization (fig. S36 and table S4). Activities were similar off- and on-chip (fig. S37), suggesting that the chip is not responsible for the observed misfolding.

Native gels and kinetic assays provided additional support for the misfolded state. Mutants predicted to misfold had an additional band of distinct mobility when expressed at high temperature that was not present or diminished when expressed at lower temperature (Fig. 3J, misfolded state “M”). Transient treatment with thermolysin, a protease that cleaves within exposed hydrophobic regions that occur in unfolded or misfolded proteins, resulted in loss of M but not the native state (native state “N”) (fig. S38A) (26). Nevertheless, despite degradation of the majority of the protein (present as M), the total observed enzyme activity for each mutant was unchanged before and after degradation (fig. S38B), indicating that M lacked substantial activity and that N and M did not equilibrate over the hours taken to carry out these experiments. The observation that different temperature-sensitive mutants gave the same mobility on a native gel and similar protease sensitivity suggests the formation of a single misfolded state or set of states with similar properties.

Our observation of a long-lived inactive state raised the question of whether analogous misfolding occurs in cells, where cellular machinery can assist folding. We recombinantly expressed 21 variants in *E. coli* that did and did not undergo temperature-dependent misfold-

ing in vitro (fig. S36 and table S5). Expression in vivo was also temperature dependent, with changes in apparent k_{cat}/K_m values for the in vivo-expressed PafA mutants that correlated with the change observed in vitro [coefficient of determination (R^2) = 0.38] (Fig. 3K). Temperature effects were generally smaller for in vivo-expressed mutants, suggesting that cellular factors may partially rescue or preferentially degrade the misfolded state, analogous to cellular degradation of thermodynamically destabilized mutants (27).

To test whether misfolding was a consequence of the eGFP tag, we overexpressed a non-eGFP-tagged variant (Y103G). We found wild type-like activity when it was expressed at 23°C but 260-fold less activity when it was expressed at 37°C (table S6 and fig. S39), suggesting that >99% of the 37°C-expressed enzyme was in the M state. CD spectra of the purified Y103G mutant provided independent evidence for a structural alteration that accompanied misfolding in vivo. WT PafA exhibits identical CD spectra when expressed at 37° or 23°C; by contrast, the CD spectrum for Y103G matches that of wild type when the mutant is expressed at low temperature (23°C) but not at higher temperature (37°C) (fig. S40), with the observed difference at 37°C suggesting loss of about one-third of PafA’s α -helical character in M (table S7).

Together, these results suggest that cellular folding conditions and chaperones are insufficient to prevent mutations from causing PafA to misfold and form long-lived inactive states in an *E. coli* expression strain [BL21 (DE3)]. However, the degree to which misfolding is rescued likely varies between strains and is sensitive to growth conditions and expression levels (28). A tendency to form kinetically stable misfolded states may therefore exert a selective pressure and influence the fitness landscape of proteins in cells (29–33).

Dissecting the origins of observed catalytic effects

HT-MEK assays allow us to quantify and dissect the degree to which observed changes in activity arise from changes in the amount of expressed protein, the amount that is correctly folded, and the catalytic efficiency of the correctly folded enzyme. Below, we isolate the catalytic effects for our PafA variants. We then take advantage of HT-MEK’s ability to provide quantitative kinetic and thermodynamic constants for multiple substrates and inhibitors and use these data to probe PafA’s functional architecture at a global level.

To remove folding effects, we compared k_{cat}/K_m values measured for substrates with different rate-limiting steps (cMUP and MeP) (Fig. 3C and supplementary text S3 and S4), represented the datapoint for each PafA variant in Fig. 3C as a superposition of a catalytic

effect (Fig. 3C, blue solid line, defined by PafA active site variants that do not have misfolding effects; fig. S41) and a misfolding effect (Fig. 3C, diagonal blue dashed line), and solved for both (Fig. 4A and fig. S42). With this approach, we were able to quantify catalytic effects ($k_{\text{cat}}/K_{\text{m}}^{\text{chem}}$) for 946 of the 1036 variants and obtain upper limits for an additional 65 (fig. S43). Deleterious catalytic effects were found for mutations at 161 of PafA's 526 positions (Fig. 4B, figs. S44 and S45, and table S8). Mutations at an even larger number of positions, 232, gave folding effects. The magnitude of folding effects did not correlate with observed eGFP fluorescence intensities, establishing that enzyme misfolding does not influence eGFP folding (fig. S46).

The largest catalytic effects cluster in and around the active site, with the fraction of residues exhibiting effects diminishing with distance from the active site (Fig. 4C and table S8). Although positions that give catalytic effects tend to cluster, the pattern is asymmetric

and complex (Fig. 4D). Many large distal effects are found surrounding the bound Zn^{2+} that is $>15 \text{ \AA}$ from the active site ("distal Zn^{2+} "), suggesting that its coordination may restrain the conformational mobility of surrounding residues and propagate to the active site. Nevertheless, it is difficult to rationalize catalytic effects, especially more distal effects. To better understand these patterns and to relate effects to the specific mechanisms used in catalysis, we defined functional components (FCs) from prior mechanistic studies of PafA and other alkaline phosphatase (AP) superfamily members (16, 34, 35) and systematically assessed the effect of each mutation on each FC.

FC1: Mutations disrupting O2 phosphoryl oxygen atom interactions

Our first FC (FC1) is derived from the observation that removal of two active-site side chains, K162 and R164, reduces monoester hydrolysis by 10^5 -fold but has no effect on diester hydrolysis, rendering PafA an equally

potent phosphate mono- and diesterase (fig. S47) (16). Because K162 and R164 interact with the phosphoryl oxygen that is anionic in monoesters but esterified in diesters (Fig. 5A, O2) (34), we expect other residues that support the formation of these interactions to similarly disrupt monoester but not diester hydrolysis. We define these effects as $\text{FC1} = \Delta^{\text{diester}} / \Delta^{\text{monoester}}$, where $\Delta = (k_{\text{cat}}/K_{\text{m}})^{\text{mutant}} / (k_{\text{cat}}/K_{\text{m}})^{\text{WT}}$. Although the simplest expectation is that mutations to residues neighboring K162 and R164 will have FC1 effects, we cannot predict how large and how varied these effects are, how far they extend, or whether there are remote regions or surface sites that have large effects. In addition, we cannot predict whether residues that affect FC1 also contribute to other catalytic mechanisms, represented as other FCs below.

To measure diester activity on-chip and determine FC1 effects, we synthesized a fluorogenic diester substrate suitable for HT-MEK (MecMUP) and measured $k_{\text{cat}}/K_{\text{m}}$ for the PafA mutant libraries (high K_{m} values for the non-cognate diesterase activities preclude estimation of k_{cat} and K_{m} separately) (fig. S48) (16). We obtained $k_{\text{cat}}/K_{\text{m}}$ values for 857 of the 1036 mutants and upper limits for an additional 178 (fig. S49, A and B), and on-chip diester rate constants matched off-chip measurements (fig. S49C). Because folding affects hydrolysis of both substrates equally, we compared measured rates directly, without correcting for folding effects (fig. S50A), allowing stronger statistical inference (fig. S50B).

Many mutants had FC1 effects: 88 Val and 93 Gly mutations (fig. S50C), corresponding to 156 of the 494 measurable positions (of a total of 526 positions) (Fig. 5, B to D). Seven of the 10 measurable non-active-site residues directly contacting K162 or R164 (second-shell residues) exhibited FC1 effects (Fig. 5E; fig. S51, A and B; and table S9A), which is consistent with frequent second-shell effects observed in directed evolution experiments (36–38). Of the three active-site Zn^{2+} ligands with measurable effects upon mutation, we observed an FC1 effect for D38G (which accepts a hydrogen bond from K162) but no FC1 effect for D352G or H353V (which do not interact with K162 or R164). The active-site variants T79S and N100G also had FC1 effects (tables S9B and S10), which is consistent with coupling between active-site residues attributable to shared contacts with K162 and R164 (Fig. 5E).

Although the largest effects were observed for active-site residues and next-largest for the second-shell residues, there was no additional drop in effect size after the third shell; the majority of residues with >10 -fold FC1 effects were found in the third shell and beyond (15 of 23) (fig. S51 and table S11). Four of these lie at the enzyme surface (Fig. 5D and fig. S51), which is consistent with the hypothesis that

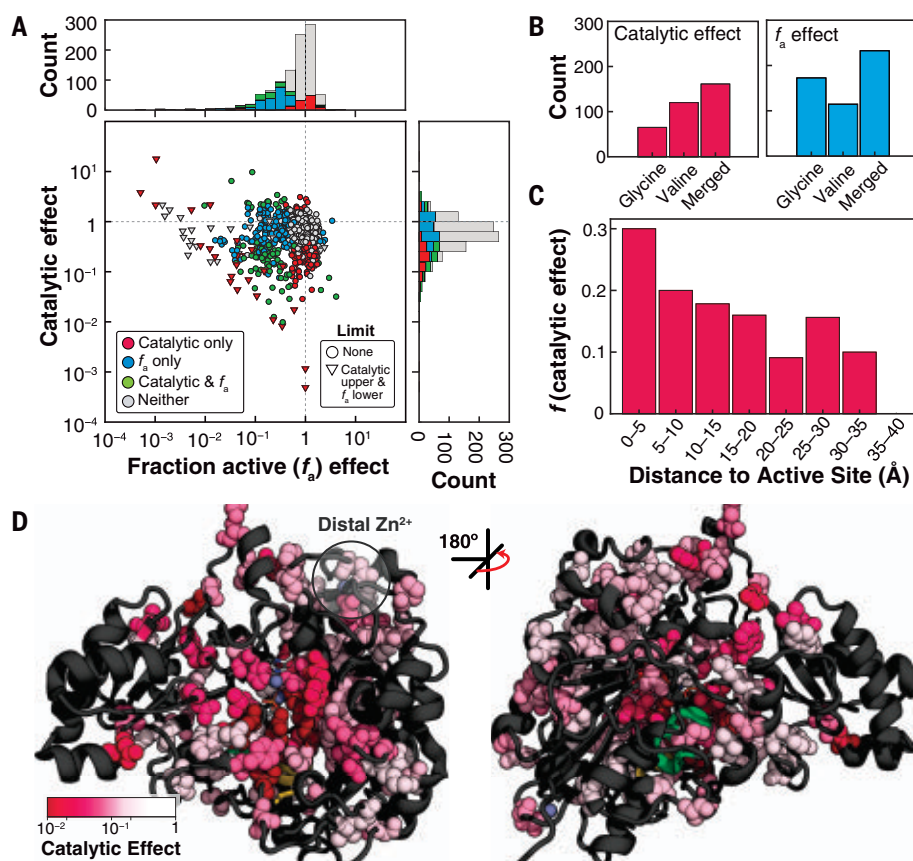


Fig. 4. Separating mutational effects on catalysis from changes in the fraction of active enzyme (f_a). (A) Catalytic and f_a effects (relative to wild type) for all variants colored by effect. Distributions of effects are projected along axes (stacked). (B) Counts of mutants with catalytic ($P < 0.05$) and f_a effects ($P < 0.01$) for the valine and glycine libraries, and total number of positions with effects (out of 525, "Merged") (fig. S45). (C) Fraction of non-active-site mutants within each distance bin with deleterious catalytic effects. $P < 0.05$; 5-Å bins (table S8). (D) Positions with deleterious catalytic effects on the PafA structure [16; "Merged" set in (B)]. Residues with deleterious effects (greater than fivefold down from wild type) are colored according to the most deleterious effect and shown as spheres.

enzymes possess a reservoir of allosteric potential and suggests that HT-MEK can be used to identify regions that are potential sites for allosteric inhibitors and drugs (39–43).

PafA has three of four nonterminal auxiliary domains (ADs) found within the AP superfamily, referred to as ADs 2 to 4 (supplementary text S5), that sit around the universally conserved Rossmann fold (Fig. 5F and fig. S52). ADs 2 and 4 are present in both AP superfamily phosphate mono- and diesterases, whereas AD 3 contains K162 and R164 and is considerably

more extensive in the monoesterases (figs. S52 and S53 and table S12). Despite these apparent functional and evolutionary differences, FC1 effects are found in all three ADs to a similar extent and with similar magnitudes (fig. S54 and table S13). The largest FC1 effect outside of the active site or second shell comes from a solvent-exposed surface residue, D473, within AD 4 (Fig. 5F, yellow); a dramatically larger effect for the valine substitution (>60-fold versus less than twofold for D473G) suggests that a change in local folding may allosteri-

cally disrupt the O2 site 20 Å away. Two nearby residues (L434 and I451) also exhibit >10-fold FC1 effects, and all three lie along a three-way interface between ADs 3 and 4 and the Rossmann core domain (Fig. 5F). Future multimutant cycle experiments (44–46) using HT-MEK may allow testing of whether mutations at interfaces result in more extensive conformational rearrangements than mutations within a single domain and dissect the underlying structural properties accounting for long-range effects.

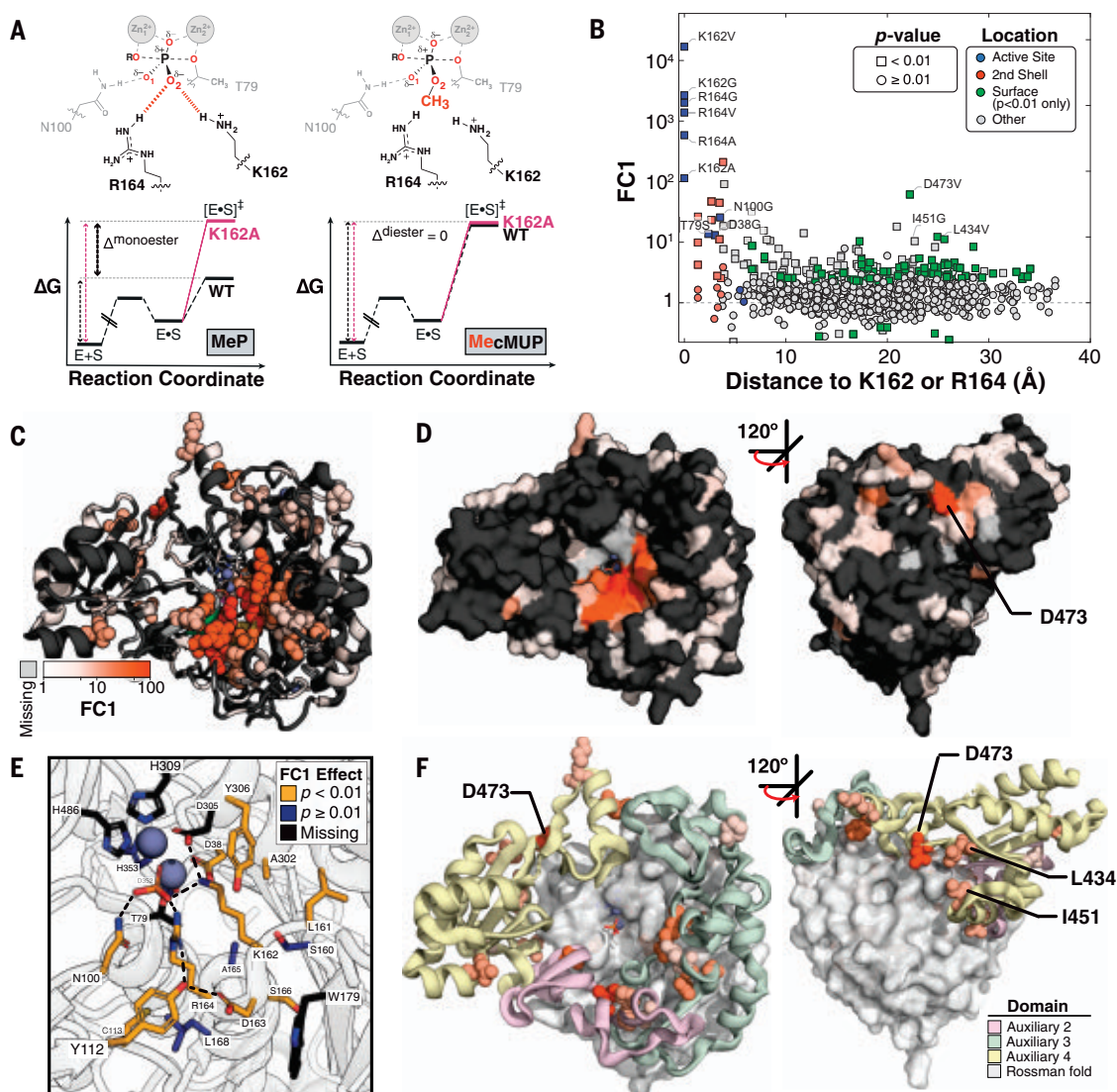


Fig. 5. FC1: Catalytic effects through the O₂ phosphoryl oxygen atom.

(A) Schematic of the transition states for reaction of (left) MeP and (right) MecMUP highlighting the methyl group on the O₂ phosphoryl atom of MecMUP (orange) and the O₂ interactions with K162 and R164. Thin black and pink arrows indicate the reaction barrier for wild type and a mutant PafA, respectively, for each reaction; the thick black arrow denotes their difference for the two reactions ($\Delta_{\text{monoester}}$ and Δ_{diester}). The ratio of k_{cat}/K_m effects (the difference between $\Delta_{\text{monoester}}$ and Δ_{diester} in energy space) gives FC1. (B) FC1 values as a function of the minimum distance to K162 or R164, with active site, second shell, and significantly affected surface residues colored. (C) PafA positions with FC1

effects ($P < 0.01$) when mutated to valine or glycine, colored by FC1 magnitude. Effects greater than fivefold are shown as spheres, with ribbon coloring for positions with effects less than or equal to fivefold. (D) PafA surface representation with FC1 effects colored as in (C). (E) FC1 effects of active site and second-shell residues indicating residues with significant FC1 effects (yellow), without FC1 effects (blue), and without measurements (black) (table S9). (F) Distribution of FC1 effects within the Rossmann core and Auxiliary Domains. Distal (greater than or equal to third shell) positions with effects greater than fivefold shown as spheres colored corresponding to their auxiliary domain. Only effects in auxiliary domains are shown.

FC2 and FC3: Effects on phosphate affinity

To provide catalysis, enzymes must bind their transition states more strongly than their substrates; otherwise, the energetic barrier for the reaction and reaction rate would remain the same as in solution (47, 48). Enzymes must also limit the binding of substrates and products to allow sufficient turnover in the presence of higher substrate and product concentrations (49, 50). Ground-state destabilization has been considered a possible mechanism for achieving this, and there is evidence for ground-state destabilization by PafA and other AP superfamily members through electrostatic repulsion between the anionic nucleophile, T79 in

PafA, and the negatively charged phosphoryl oxygens (Fig. 6A) (16, 35, 51). For PafA, mutating T79 to serine increases the affinity for P_i , the reaction's product and a ground-state analog, by 100-fold, whereas in *E. coli* AP, the nucleophile S102G mutation increases affinity more than 1000-fold (16, 35); the S102G mutation ablates the nucleophile entirely, and the chemically conservative T79S substitution presumably allows greater mobility, reducing electrostatic repulsion (16). We therefore expect that mutations that perturb this mechanism will increase P_i affinity and define the second FC, FC2, as strengthened P_i binding ($K_i^{\text{mutant}}/K_i^{\text{WT}} < 1$).

Conversely, active-site residues typically make both ground-state and transition-state interactions, so their removal weakens binding and diminishes catalysis, in some instances to a similar extent ("uniform binding") and in others preferentially destabilizing the transition state (52–55). As expected, mutations to the PafA active-site residues that interact with the phosphoryl O1 and O2 oxygen atoms (N100, K162, and R164) weaken P_i binding, but N100 and K162 mutations have much larger effects on catalysis, indicating preferential transition-state stabilization (fig. S55) (16). We therefore define the third FC (FC3) as weakened P_i binding ($K_i^{\text{mutant}}/K_i^{\text{WT}} > 1$).

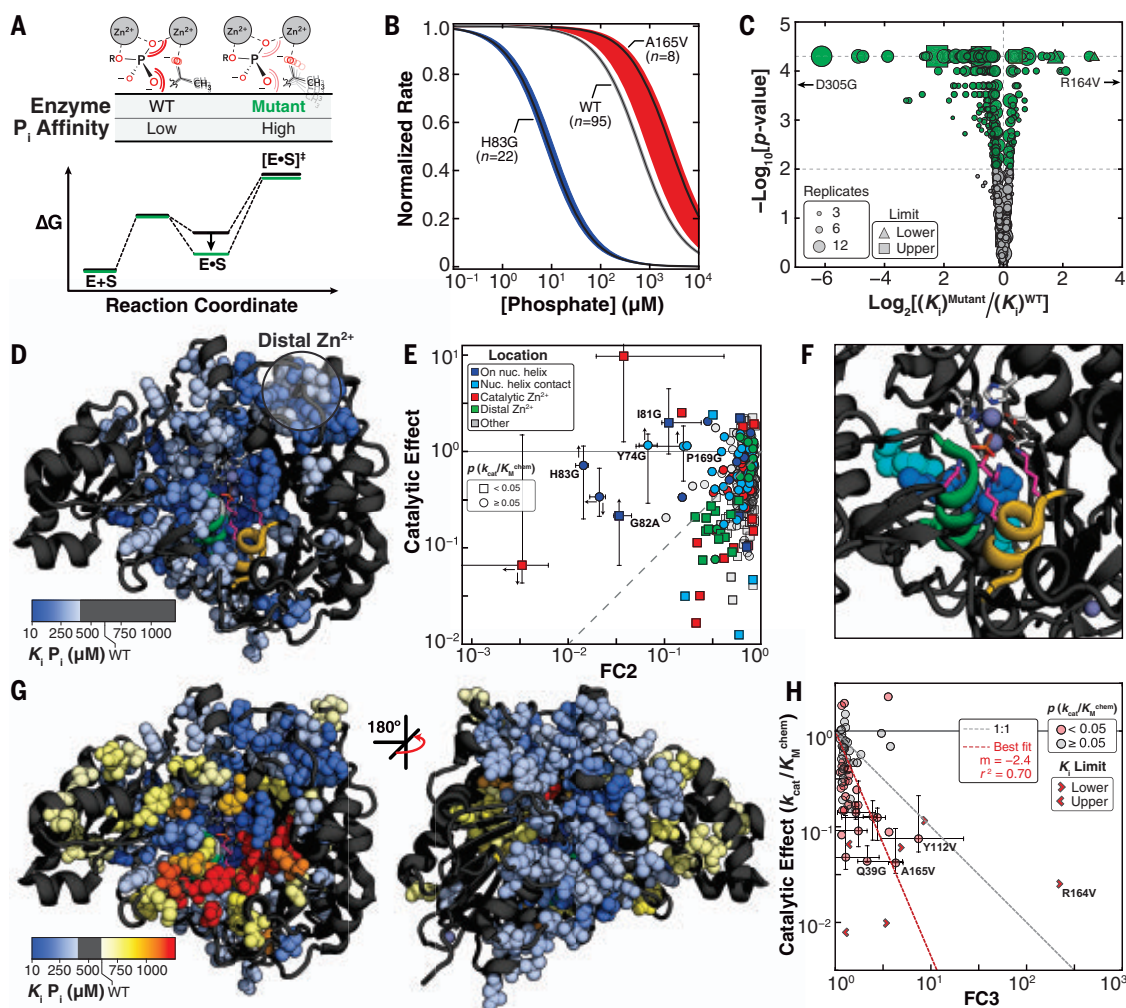


Fig. 6. FC2 and FC3: Mutational effects on affinity for P_i . (A) Reaction coordinate diagram for hypothetical mutant (green) with diminished destabilization relative to wild type (black) and schematic illustrating a model for an FC2 effect through increased flexibility of the T79 nucleophile. (B) Median P_i inhibition curves for wild type and two mutants with FC2 (H83G) and FC3 (A165V) effects. Colored regions denote 99% CIs on the medians of replicate measurements. (C) Volcano plot of $K_i P_i$ effects for glycine and valine scan mutants. $P < 0.01$, green markers; $P \geq 0.01$, gray markers. (D) PafA structure with positions for which $K_i P_i$ is >1.5-fold tighter than that of wild type (at $P < 0.01$) shown as spheres colored by the magnitude of the larger effect. The “nucleophile” helix (green) is largely obscured. (E) Scatter plot of catalytic and FC2 effects ($P < 0.01$) colored by location. Error bars correspond

to 95% CIs determined from bootstrapping; up and down arrows indicate lower and upper $k_{\text{cat}}/K_m^{\text{chem}}$ limits, respectively; left and right arrows indicate upper and lower $K_i P_i$ limits. (F) PafA active-site (pink residues) and the five mutants possessing large FC2 effects without $k_{\text{cat}}/K_m^{\text{chem}}$ effects shown as spheres, corresponding to the labeled mutants in (E). (G) Front and back views of FC2 (blue spheres) and FC3 effects (red spheres). FC2 effects are those shown in (D), and FC3 effects are shown for $P < 0.01$. (H) Scatter plot of catalytic ($k_{\text{cat}}/K_m^{\text{chem}}$) and FC3 effects for mutants with FC3 effects ($P < 0.01$). Gray and red points indicate significant and insignificant $k_{\text{cat}}/K_m^{\text{chem}}$ effects, respectively. Error bars correspond to 95% CIs, as in (E); red dashed line indicates the best fit line to mutants with significant FC3 and catalytic ($k_{\text{cat}}/K_m^{\text{chem}}$) effects, excluding active-site mutants.

To measure inhibition constants, we quantified rates of cMUP hydrolysis as a function of P_i concentration and fit observed initial rates to a competitive inhibition model (Fig. 6B and fig. S56). HT-MEK-determined inhibition constants agreed with previous off-chip measurements (fig. S57) and are of higher precision than kinetic constants because they are not sensitive to uncertainties in total enzyme concentration or active fraction (Fig. 6C; reports of inhibition measurements are available in data files S1 and S2).

We uncovered 331 mutants that increased P_i affinity (FC2) and 73 that decreased P_i affinity (FC3) (Fig. 6C and fig. S58). Thus, about one-third of all mutants measurably altered affinity, and four times as many mutations enhanced binding as weakened it. Because it is highly unusual to enhance function through random variation, this observation suggests that residues at many positions are evolutionarily selected to prevent tight P_i binding. Mutations with ground-state destabilization effects (FC2) were located in an extended yet spatially contiguous region that included the helix containing the T79 oxyanion (“nucleophile helix”), the catalytic Zn^{2+} ions, and the distal Zn^{2+} site (Fig. 6D and tables S14 and S15).

Catalytic effects (k_{cat}/K_m^{chem}) for mutants with FC2 effects ranged from insignificant to 100-fold reductions (Fig. 6E and fig. S59). For most, the catalytic effect was greater than the FC2 effect, which is consistent with functional roles beyond the prevention of too-tight binding. Nevertheless, five mutants of residues that form a spatially contiguous subregion on and adjacent to the nucleophile helix have the largest FC2 effects but little or no catalytic effect (Fig. 6E and F), suggesting that these mutations allow rearrangements to reduce electrostatic repulsion in the presence of bound substrate.

Mutants that weakened P_i binding (FC3 effects) were also located in a contiguous region (Fig. 6G). The largest FC3 effects were near active-site residues K162 and R164 and had larger effects on catalysis (k_{cat}/K_m^{chem}) than P_i binding (Fig. 6H, red, and figs. S60 to S62), presumably reflecting the greater constraints present in the transition state than in the ground state (56, 57).

Comparisons of the mutational effects across FC1 to FC3 reveal that many mutations outside the active site preferentially affect either FC1 or FC2. Further, several preferentially alter FC2 without dramatically altering FC1, reducing the fraction of active enzyme, or reducing overall catalysis (fig. S63). A small number of residues at the interface of FC2 and FC3 effects yield either an FC2 or FC3 effect depending on the substitution (fig. S64). The ability to selectively alter particular properties through specific mutations provides a potential starting point for attempts to engineer enzymes with

desired kinetic and thermodynamic constants and behaviors.

FC4: Rates of phosphoenzyme hydrolysis

Linking observed mutational rate effects to their physical and chemical origins requires knowledge of the step that is being observed: the rate-limiting step. Pre-steady-state approaches (such as stopped flow and rapid quench) are the gold standard for determining rates of individual reaction steps but do not readily scale to large libraries. For PafA, extensive prior mechanistic knowledge allowed us to determine mutational effects on individual PafA reaction steps for 992 PafA variants without requiring pre-steady-state measurements.

The steady-state kinetic constant k_{cat}/K_m can be limited by substrate binding or chemical cleavage of the substrate to form the covalent enzyme-phosphate (E-P) species (Fig. 7A, k_1 and $k_{chem,1}$ steps), and k_{cat} can be limited by hydrolysis of E-P ($k_{chem,2}$) or dissociation of product P_i ($k_{off,Pi}$) (Fig. 7A). We deconvoluted mutational effects on $k_{chem,2}$ using measurements of k_{cat} for cMUP and K_1 for P_i , and these $k_{chem,2}$ effects define FC4 (supplementary text S6). Seven mutants changed the rate-limiting step from E-P hydrolysis to P_i release ($k_{off,Pi} < k_{chem,2}$) (fig. S65), which is consistent with the observation of naturally occurring alkaline phosphatases of the AP superfamily with either of these steps rate limiting (58).

Overall, we found 18 Val and 36 Gly mutants that reduce $k_{chem,2}$ (Fig. 7B and fig. S66), and these overlap substantially with those affecting k_{cat}/K_m^{chem} (table S16). In the simplest scenario, mutations reducing $k_{chem,2}$ would be a subset of those reducing k_{cat}/K_m^{chem} because k_{cat}/K_m^{chem} includes both a phosphoryl transfer chemical step and an additional step (binding). Consistent with this expectation, mutations affect either k_{cat}/K_m^{chem} alone or k_{cat}/K_m^{chem} and $k_{chem,2}$ to a similar or smaller extent (Fig. 7C).

Discussion

HT-MEK uses automatic valved microfluidics to carry out high-throughput expression, purification, and comprehensive biochemical characterization of enzymes at unprecedented scale and mechanistic depth. In the future, HT-MEK can be applied to the vast number of enzymes whose activity can be monitored through fluorescence directly or with a coupled assay. In particular, the ability to sensitively detect P_i (18) renders HT-MEK immediately applicable to adenosine triphosphatases, guanosine triphosphatases, helicases, protein chaperones, polymerases (with pyrophosphatase present), and many others.

Deleterious effects throughout the enzyme and at the surface, with many of the largest effects distant from the active site, mirror results obtained from DMS studies of other enzymes that often reveal distal effects on product

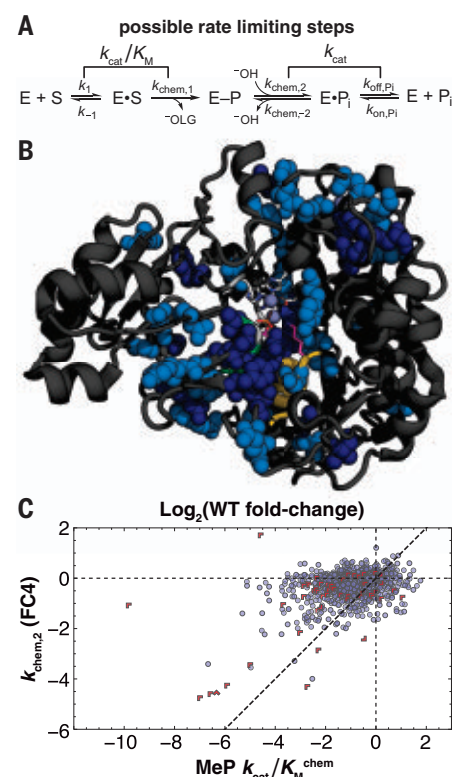


Fig. 7. FC4: Mutational effects on phosphoenzyme intermediate hydrolysis. (A) Schematic of PafA catalytic cycle with possible rate-limiting steps under saturating (k_{cat}) and subsaturating (k_{cat}/K_m) conditions. (B) Structure showing positions with decreased $k_{chem,2}$ ($P < 0.05$ and $P < 0.1$ in dark and light blue, respectively) upon mutation (spheres). Positions lacking an estimate of $k_{chem,2}$ for both substitutions are light gray ribbons, and positions without significant effects are black ribbons. (C) Scatter plot comparing mutational effects on phosphoenzyme hydrolysis ($k_{chem,2}$, or FC4) and k_{cat}/K_m^{chem} . Limits are indicated with red chevrons pointing in limit directions (supplementary text S6).

formation or organismal fitness (12, 59–61). However, the ability of HT-MEK to provide quantitative measurements of a battery of kinetic and thermodynamic parameters yields an atlas of PafA functional “architecture” with unprecedented detail (table S17). Regions of residues with similar catalytic signatures extending throughout the enzyme affirm that the enzyme beyond the active site is not a passive, monolithic scaffold but rather contributes to function in multiple ways.

Comparing patterns of mutational effects between FCs reveals that these architectural solutions are idiosyncratic. For example, K162/R164, and the T79 nucleophile, sit at the N-termini of their respective helices, but the largest FC2 effects were on and around the nucleophile helix, whereas FC1 effects were not similarly distributed around the monoesterase helix. This

difference perhaps reflects a need for more interactions to secure the nucleophile helix against ground-state electrostatic repulsive forces. Twice as many glycine as valine mutations increased P_i binding (fig. S58), potentially because side-chain ablation allows structural rearrangements to reduce electrostatic repulsion without disrupting favorable binding interactions.

The need for interactions throughout the enzyme for optimal catalysis highlights the challenges inherent in de novo enzyme design because it suggests that computational efforts must consider in detail the interactions of a prohibitively large number of residues. The detailed architectural maps provided by HT-MEK can focus computational and experimental mutagenesis by distinguishing catalytic from folding effects and by identifying residues and regions that affect particular aspects of catalysis.

The presence of a long-lived inactive population of many PafA variants emphasizes how different underlying (and unexpected) biophysical mechanisms can compromise an enzyme's activity and thus the need to explicitly decouple effects on folding and catalysis to understand function. Analogous behavior observed for these mutants in *E. coli* is consistent with selective pressure to avoid misfolding in vivo and adds to growing evidence that kinetic factors affect stable protein expression in cells (20, 30, 62–64). We speculate that highly stable proteins such as PafA and other secreted enzymes may be more prone to forming kinetically trapped states.

Although PafA functional regions have superficial similarities to sectors and other measures of evolutionary covariation and coconservation (65–69), our data report on sequence-function relationships in ways that sequence analyses alone cannot. On the basis of a metagenomic alignment of 14,505 PafA-like AP superfamily sequences (fig. S67 and supplementary text S7), we found that observed k_{cat}/K_m values correlate only modestly with conservation at a given position (figs. S68 and S69). Positions at which mutations promote misfolding tend to be most conserved, suggesting that the largest selective pressure at the residue level may be related to ensuring correct folding, whereas correlations between individual FCs and conservation are weak, with many large effects at nonconserved residues (figs. S69 and S70). For PafA, selective pressures will differ with the available P_i at the organism's physical location, whether available P_i varies temporally, and whether there is competition for P_i from other organisms in the same ecological niche. Residues with critical roles in responding to these adaptive pressures would be poorly conserved, changing frequently despite being tightly linked to survival. These results underscore the need for in-depth functional studies to uncover the

relationships between the evolution of particular functions and conservation.

Enzymes are the targets of many therapeutics, are altered in genetic diseases, serve as tools for molecular biology, and play critical roles in industrial processes. Its rapidity and low cost make HT-MEK a powerful tool for future applications across all these areas. In basic research, the large, highly quantitative datasets provided by HT-MEK can greatly extend and even supplant traditional SDM approaches for the initial characterization of new enzymes and for in-depth mechanistic investigation. Combined with recent advances in gene synthesis, HT-MEK can rapidly functionally characterize metagenomic variants, providing a critically needed dimension to phylogenetic analyses. In medicine, we anticipate that HT-MEK will rapidly determine the functional effects of human enzyme allelic variants of unknown relevance identified from sequence data and systematically identify candidate allosteric surfaces within currently “undruggable” therapeutic target enzymes. We anticipate HT-MEK contributing to these and still more areas of basic and applied biology, medicine, and engineering.

Materials and methods summary

A full description of materials and methods is provided in the supplementary materials. Briefly, we fabricated microfluidic devices and aligned them to plasmid DNA arrays as previously described (70–72), then connected devices to a custom pneumatic manifold (73) and imaged using a fully automated fluorescence microscope. The plasmid DNAs coded for a set of specified PafA variants, each with a C-terminal eGFP tag. The PafA-eGFP variants were expressed in vitro on-chip and immobilized to Button valves by means of antibodies to eGFP (72) in parallel, before quantification of the enzyme concentration within each chamber by using an eGFP standard curve. This entire process takes ~10 hours to complete. Before measuring kinetics of substrate hydrolysis for cMUP, MeP, and MecMUP, we determined product standard curves for each chamber by introducing a series of product concentrations [cMU for cMUP and MecMUP hydrolysis and P_i in the presence of 30 μ M of modified PBP (18) for MeP hydrolysis], imaging the device after each addition. To measure rates of hydrolysis, we flowed a single concentration of substrate into the device with the Button valves closed and then started reactions by (i) closing the Sandwich valves to isolate individual reaction chambers and (ii) opening the Button valves to expose the immobilized enzyme within each chamber to substrate. We quantified initial rates of product formation for each chamber by imaging over time across the entire device and converting measured intensities to product con-

centrations using standard curves. To determine k_{cat} , K_m , and k_{cat}/K_m for cMUP hydrolysis and k_{cat}/K_m for MeP and MecMUP hydrolysis, we quantified initial rates of product formation for a series of increasing substrate concentrations, scaled rates by measured enzyme concentration, and fit to a simple Michaelis-Menten model. For the nonfluorogenic MeP substrate, we detected formation of P_i by performing reactions in the presence of 30 μ M PBP. To determine the K_i for P_i inhibition, we assayed rates of cMUP hydrolysis at 10 or 50 μ M cMUP in the presence of increasing P_i concentration, fit initial rates as described above, and then fit these rates to a competitive inhibition model. For each assay type, we carried out multiple replicate experiments for each parameter for each mutant, allowing us to perform extensive quality control measurements and determine the statistical significance of effects.

REFERENCES AND NOTES

- Q. Wang *et al.*, Landscape of multi-nucleotide variants in 125,748 human exomes and 15,708 genomes. *Nat. Commun.* **11**, 2539 (2020). doi: [10.1038/s41467-019-12438-5](https://doi.org/10.1038/s41467-019-12438-5); pmid: 32461613
- M. Claussnitzer *et al.*, A brief history of human disease genetics. *Nature* **577**, 179–189 (2020). doi: [10.1038/s41586-019-1879-7](https://doi.org/10.1038/s41586-019-1879-7); pmid: 31915397
- M. Lek *et al.*, Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016). doi: [10.1038/nature19057](https://doi.org/10.1038/nature19057); pmid: 27535533
- A. Stein, D. M. Fowler, R. Hartmann-Petersen, K. Lindorff-Larsen, Biophysical and mechanistic models for disease-causing protein variants. *Trends Biochem. Sci.* **44**, 575–588 (2019). doi: [10.1016/j.tibs.2019.01.003](https://doi.org/10.1016/j.tibs.2019.01.003); pmid: 30712981
- J. B. Kinney, D. M. McCandlish, Massively parallel assays and quantitative sequence-function relationships. *Annu. Rev. Genomics Hum. Genet.* **20**, 99–127 (2019). doi: [10.1146/annurev-genom-083118-014845](https://doi.org/10.1146/annurev-genom-083118-014845); pmid: 31091417
- C. Blanco, E. Janzen, A. Pressman, R. Saha, I. A. Chen, Molecular fitness landscapes from high-coverage sequence profiling. *Annu. Rev. Biophys.* **48**, 1–18 (2019). doi: [10.1146/annurev-biophys-052118-115333](https://doi.org/10.1146/annurev-biophys-052118-115333); pmid: 30601678
- J. Weile, F. P. Roth, Multiplexed assays of variant effects contribute to a growing genotype-phenotype atlas. *Hum. Genet.* **137**, 665–678 (2018). doi: [10.1007/s00439-018-1916-x](https://doi.org/10.1007/s00439-018-1916-x); pmid: 30073413
- G. P. Lisi, J. P. Loria, Allosteric in enzyme catalysis. *Curr. Opin. Struct. Biol.* **47**, 123–130 (2017). doi: [10.1016/j.sbsi.2017.08.002](https://doi.org/10.1016/j.sbsi.2017.08.002); pmid: 28865247
- K. A. Reynolds, R. N. McLaughlin, R. Ranganathan, Hot spots for allosteric regulation on protein surfaces. *Cell* **147**, 1564–1575 (2011). doi: [10.1016/j.cell.2011.10.049](https://doi.org/10.1016/j.cell.2011.10.049); pmid: 22196731
- D. A. Keedy *et al.*, An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *eLife* **7**, e36307 (2018). doi: [10.7554/eLife.36307](https://doi.org/10.7554/eLife.36307); pmid: 29877794
- D. M. Fowler *et al.*, High-resolution mapping of protein sequence-function relationships. *Nat. Methods* **7**, 741–746 (2010). doi: [10.1038/nmeth.1492](https://doi.org/10.1038/nmeth.1492); pmid: 20711194
- C. L. Araya, D. M. Fowler, Deep mutational scanning: Assessing protein function on a massive scale. *Trends Biotechnol.* **29**, 435–442 (2011). doi: [10.1016/j.tibtech.2011.04.003](https://doi.org/10.1016/j.tibtech.2011.04.003); pmid: 21561674
- J. K. Lassila, J. G. Zalatan, D. Herschlag, Biological phosphoryl-transfer reactions: Understanding mechanism and catalysis. *Annu. Rev. Biochem.* **80**, 669–702 (2011). doi: [10.1146/annurev-biochem-060409-092741](https://doi.org/10.1146/annurev-biochem-060409-092741); pmid: 21513457
- P. M. Fordyce *et al.*, De novo identification and biophysical characterization of transcription-factor binding sites with microfluidic affinity analysis. *Nat. Biotechnol.* **28**, 970–975 (2010). doi: [10.1038/nbt.1675](https://doi.org/10.1038/nbt.1675); pmid: 20802496

15. S. J. Maerkl, S. R. Quake, A systems approach to measuring the binding energy landscapes of transcription factors. *Science* **315**, 233–237 (2007). doi: [10.1126/science.1131007](https://doi.org/10.1126/science.1131007); pmid: [17218526](https://pubmed.ncbi.nlm.nih.gov/17218526/)
16. F. Sundén *et al.*, Mechanistic and evolutionary insights from comparative enzymology of phosphomonoesterases and phosphodiesterases across the alkaline phosphatase superfamily. *J. Am. Chem. Soc.* **138**, 14273–14287 (2016). doi: [10.1021/jacs.6b06186](https://doi.org/10.1021/jacs.6b06186); pmid: [27670607](https://pubmed.ncbi.nlm.nih.gov/27670607/)
17. M. M. Santoro, D. W. Bolen, Unfolding free energy changes determined by the linear extrapolation method. I. Unfolding of phenylmethanesulfonyl α -chymotrypsin using different denaturants. *Biochemistry* **27**, 8063–8068 (1988). doi: [10.1021/bi00421a014](https://doi.org/10.1021/bi00421a014); pmid: [3233195](https://pubmed.ncbi.nlm.nih.gov/3233195/)
18. M. Brune, J. L. Hunter, J. E. T. Corrie, M. R. Webb, Direct, real-time measurement of rapid inorganic phosphate release using a novel fluorescent probe and its application to actomyosin subfragment 1 ATPase. *Biochemistry* **33**, 8262–8271 (1994). doi: [10.1021/bi00193a013](https://doi.org/10.1021/bi00193a013); pmid: [8031761](https://pubmed.ncbi.nlm.nih.gov/8031761/)
19. P. J. O'Brien, D. Herschlag, Alkaline phosphatase revisited: Hydrolysis of alkyl phosphates. *Biochemistry* **41**, 3207–3225 (2002). doi: [10.1021/bi012166y](https://doi.org/10.1021/bi012166y); pmid: [11863460](https://pubmed.ncbi.nlm.nih.gov/11863460/)
20. R. M. Addabbo *et al.*, Complementary role of co- and post-translational events in the *de novo* protein biogenesis. *J. Phys. Chem. B* **124**, 6488–6507 (2020). doi: [10.1021/acs.jpcb.0c03039](https://doi.org/10.1021/acs.jpcb.0c03039); pmid: [32456434](https://pubmed.ncbi.nlm.nih.gov/32456434/)
21. J. M. Sturtevant, M. H. Yu, C. Haase-Pettingell, J. King, Thermostability of temperature-sensitive folding mutants of the P22 tailspike protein. *J. Biol. Chem.* **264**, 10693–10698 (1989). doi: [10.1016/S0021-9258\(18\)81678-7](https://doi.org/10.1016/S0021-9258(18)81678-7); pmid: [2525128](https://pubmed.ncbi.nlm.nih.gov/2525128/)
22. M. Guo, Y. Xu, M. Gruebele, Temperature dependence of protein folding kinetics in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17863–17867 (2012). doi: [10.1073/pnas.1201797109](https://doi.org/10.1073/pnas.1201797109); pmid: [22665776](https://pubmed.ncbi.nlm.nih.gov/22665776/)
23. D. H. Smith, J. King, Temperature-sensitive mutants blocked in the folding or subunit assembly of the bacteriophage P22 tail spike protein. III. Intensive polypeptide chains synthesized at 39°C. *J. Mol. Biol.* **145**, 653–676 (1981). doi: [10.1016/0022-2836\(81\)90308-9](https://doi.org/10.1016/0022-2836(81)90308-9); pmid: [7265218](https://pubmed.ncbi.nlm.nih.gov/7265218/)
24. D. Baker, J. L. Sohl, D. A. Agard, A protein-folding reaction under kinetic control. *Nature* **356**, 263–265 (1992). doi: [10.1038/356263a0](https://doi.org/10.1038/356263a0); pmid: [1552947](https://pubmed.ncbi.nlm.nih.gov/1552947/)
25. H. Im, M.-S. Woo, K. Y. Hwang, M.-H. Yu, Interactions causing the kinetic trap in serpin protein folding. *J. Biol. Chem.* **277**, 46347–46354 (2002). doi: [10.1074/jbc.M207682200](https://doi.org/10.1074/jbc.M207682200); pmid: [12244055](https://pubmed.ncbi.nlm.nih.gov/12244055/)
26. C. Park, S. Marqusee, Probing the high energy states in proteins by proteolysis. *J. Mol. Biol.* **343**, 1467–1476 (2004). doi: [10.1016/j.jmb.2004.08.085](https://doi.org/10.1016/j.jmb.2004.08.085); pmid: [15491624](https://pubmed.ncbi.nlm.nih.gov/15491624/)
27. S. Thompson, Y. Zhang, C. Ingle, K. A. Reynolds, T. Kortemme, Altered expression of a quality control protease in *E. coli* reshapes the in vivo mutational landscape of a model enzyme. *eLife* **9**, e53476 (2020). doi: [10.7554/eLife.53476](https://doi.org/10.7554/eLife.53476); pmid: [32701056](https://pubmed.ncbi.nlm.nih.gov/32701056/)
28. T. Gidalevitz, V. Prahla, R. I. Morimoto, The stress of protein misfolding: From single cells to multicellular organisms. *Cold Spring Harb. Perspect. Biol.* **3**, a009704–a009704 (2011). doi: [10.1101/cshperspect.a009704](https://doi.org/10.1101/cshperspect.a009704); pmid: [21536706](https://pubmed.ncbi.nlm.nih.gov/21536706/)
29. J. L. Sohl, S. S. Jaswal, D. A. Agard, Unfolded conformations of α -lytic protease are more stable than its native state. *Nature* **395**, 817–819 (1998). doi: [10.1038/27470](https://doi.org/10.1038/27470); pmid: [9796818](https://pubmed.ncbi.nlm.nih.gov/9796818/)
30. S. S. Jaswal, J. L. Sohl, J. H. Davis, D. A. Agard, Energetic landscape of α -lytic protease optimizes longevity through kinetic stability. *Nature* **415**, 343–346 (2002). doi: [10.1038/415343a](https://doi.org/10.1038/415343a); pmid: [11797014](https://pubmed.ncbi.nlm.nih.gov/11797014/)
31. A. L. Watters *et al.*, The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell* **128**, 613–624 (2007). doi: [10.1016/j.cell.2006.12.042](https://doi.org/10.1016/j.cell.2006.12.042); pmid: [17289578](https://pubmed.ncbi.nlm.nih.gov/17289578/)
32. F. Bemporad *et al.*, Biological function in a non-native partially folded state of a protein. *EMBO J.* **27**, 1525–1535 (2008). doi: [10.1038/emboj.2008.82](https://doi.org/10.1038/emboj.2008.82); pmid: [18451804](https://pubmed.ncbi.nlm.nih.gov/18451804/)
33. R. E. Jefferson, T. M. Blois, J. U. Bowie, Membrane proteins can have high kinetic stability. *J. Am. Chem. Soc.* **135**, 15183–15190 (2013). doi: [10.1021/ja407232b](https://doi.org/10.1021/ja407232b); pmid: [24032628](https://pubmed.ncbi.nlm.nih.gov/24032628/)
34. J. G. Zalatan, T. D. Fenn, D. Herschlag, Comparative enzymology in the alkaline phosphatase superfamily to determine the catalytic role of an active-site metal ion. *J. Mol. Biol.* **384**, 1174–1189 (2008). doi: [10.1016/j.jmb.2008.09.059](https://doi.org/10.1016/j.jmb.2008.09.059); pmid: [18851975](https://pubmed.ncbi.nlm.nih.gov/18851975/)
35. L. D. Andrews, T. D. Fenn, D. Herschlag, Ground state destabilization by anionic nucleophiles contributes to the activity of phosphoryl transfer enzymes. *PLOS Biol.* **11**, e1001599 (2013). doi: [10.1371/journal.pbio.1001599](https://doi.org/10.1371/journal.pbio.1001599); pmid: [23843744](https://pubmed.ncbi.nlm.nih.gov/23843744/)
36. D. Hilvert, Design of protein catalysts. *Annu. Rev. Biochem.* **82**, 447–470 (2013). doi: [10.1146/annurev-biochem-072611-101825](https://doi.org/10.1146/annurev-biochem-072611-101825); pmid: [23746259](https://pubmed.ncbi.nlm.nih.gov/23746259/)
37. M. I. Freiberger, A. B. Guzovsky, P. G. Wolynes, R. G. Parra, D. U. Ferreira, Local frustration around enzyme active sites. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4037–4043 (2019). doi: [10.1073/pnas.1819859116](https://doi.org/10.1073/pnas.1819859116); pmid: [30765513](https://pubmed.ncbi.nlm.nih.gov/30765513/)
38. K. L. Morley, R. J. Kazlauskas, Improving enzyme properties: When are closer mutations better? *Trends Biotechnol.* **23**, 231–237 (2005). doi: [10.1016/j.tibtech.2005.03.005](https://doi.org/10.1016/j.tibtech.2005.03.005); pmid: [15866000](https://pubmed.ncbi.nlm.nih.gov/15866000/)
39. O. Dagliyan, N. V. Dokholyan, K. M. Hahn, Engineering proteins for allosteric control by light or ligands. *Nat. Protoc.* **14**, 1863–1883 (2019). doi: [10.1038/s41596-019-0165-3](https://doi.org/10.1038/s41596-019-0165-3); pmid: [31076662](https://pubmed.ncbi.nlm.nih.gov/31076662/)
40. J. W. McCormick, D. Pincus, O. Resnekov, K. A. Reynolds, Strategies for engineering and rewiring kinase regulation. *Trends Biochem. Sci.* **45**, 259–271 (2020). doi: [10.1016/j.tibs.2019.11.005](https://doi.org/10.1016/j.tibs.2019.11.005); pmid: [31866305](https://pubmed.ncbi.nlm.nih.gov/31866305/)
41. K. Gunasekaran, B. Ma, R. Nussinov, Is allostery an intrinsic property of all dynamic proteins? *Proteins* **57**, 433–443 (2004). doi: [10.1002/prot.20232](https://doi.org/10.1002/prot.20232); pmid: [15382234](https://pubmed.ncbi.nlm.nih.gov/15382234/)
42. J. A. Hardy, J. A. Wells, Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* **14**, 706–715 (2004). doi: [10.1016/j.sbi.2004.10.009](https://doi.org/10.1016/j.sbi.2004.10.009); pmid: [15582395](https://pubmed.ncbi.nlm.nih.gov/15582395/)
43. G. R. Bowman, E. R. Bolin, K. M. Hart, B. C. Maguire, S. Marqusee, Discovery of multiple hidden allosteric sites by combining Markov state models and experiments. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2734–2739 (2015). doi: [10.1073/pnas.1417811112](https://doi.org/10.1073/pnas.1417811112); pmid: [25730859](https://pubmed.ncbi.nlm.nih.gov/25730859/)
44. F. Sundén, A. Peck, J. Salzman, S. Ressler, D. Herschlag, Extensive site-directed mutagenesis reveals interconnected functional units in the alkaline phosphatase active site. *eLife* **4**, e06181 (2015). doi: [10.7554/eLife.06181](https://doi.org/10.7554/eLife.06181); pmid: [25902402](https://pubmed.ncbi.nlm.nih.gov/25902402/)
45. P. J. Carter, G. Winter, A. J. Wilkinson, A. R. Fersht, The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (*Bacillus stearothermophilus*). *Cell* **38**, 835–840 (1984). doi: [10.1016/0092-8674\(84\)90278-2](https://doi.org/10.1016/0092-8674(84)90278-2); pmid: [6488318](https://pubmed.ncbi.nlm.nih.gov/6488318/)
46. P. Carter, J. A. Wells, Dissecting the catalytic triad of a serine protease. *Nature* **332**, 564–568 (1988). doi: [10.1038/332564a0](https://doi.org/10.1038/332564a0); pmid: [3282170](https://pubmed.ncbi.nlm.nih.gov/3282170/)
47. D. Herschlag, A. Natarajan, Fundamental challenges in mechanistic enzymology: Progress toward understanding the rate enhancements of enzymes. *Biochemistry* **52**, 2050–2067 (2013). doi: [10.1021/bi4000113](https://doi.org/10.1021/bi4000113); pmid: [23488725](https://pubmed.ncbi.nlm.nih.gov/23488725/)
48. W. P. Jencks, Mechanism of enzyme action. *Annu. Rev. Biochem.* **32**, 639–676 (1963). doi: [10.1146/annurev-bi.32.070163.003231](https://doi.org/10.1146/annurev-bi.32.070163.003231); pmid: [14140708](https://pubmed.ncbi.nlm.nih.gov/14140708/)
49. A. R. Fersht, *Enzyme Structure and Mechanism* (W.H. Freeman and Co., ed. 2, 1985).
50. W. P. Jencks, Binding energy, specificity, and enzymic catalysis: The Circe effect. *Adv. Enzymol. Relat. Areas Mol. Biol.* **43**, 219–410 (1975). doi: [10.1002/9780470122884.ch4](https://doi.org/10.1002/9780470122884.ch4); pmid: [892](https://pubmed.ncbi.nlm.nih.gov/892/)
51. L. D. Andrews, H. Deng, D. Herschlag, Isotope-edited FTIR of alkaline phosphatase resolves paradoxical ligand binding properties and suggests a role for ground-state destabilization. *J. Am. Chem. Soc.* **133**, 11621–11631 (2011). doi: [10.1021/ja203370b](https://doi.org/10.1021/ja203370b); pmid: [21692505](https://pubmed.ncbi.nlm.nih.gov/21692505/)
52. W. J. Albery, J. R. Knowles, Evolution of enzyme function and the development of catalytic efficiency. *Biochemistry* **15**, 5631–5640 (1976). doi: [10.1021/bi00670a032](https://doi.org/10.1021/bi00670a032); pmid: [999839](https://pubmed.ncbi.nlm.nih.gov/999839/)
53. F. M. Menger, Analysis of ground-state and transition-state effects in enzyme catalysis. *Biochemistry* **31**, 5368–5373 (1992). doi: [10.1021/bi00138a018](https://doi.org/10.1021/bi00138a018); pmid: [1606161](https://pubmed.ncbi.nlm.nih.gov/1606161/)
54. J. D. McCarter, M. J. Adam, S. G. Withers, Binding energy and catalysis. Fluorinated and deoxygenated glycosides as mechanistic probes of Escherichia coli (lacZ) β -galactosidase. *Biochem. J.* **286**, 721–727 (1992). doi: [10.1042/bj2860721](https://doi.org/10.1042/bj2860721); pmid: [1417731](https://pubmed.ncbi.nlm.nih.gov/1417731/)
55. S. A. Moore, W. P. Jencks, Formation of active site thiol esters of CoA transferase and the dependence of catalysis on specific binding interactions. *J. Biol. Chem.* **257**, 10893–10907 (1982). doi: [10.1016/S0021-9258\(18\)33908-5](https://doi.org/10.1016/S0021-9258(18)33908-5); pmid: [6955308](https://pubmed.ncbi.nlm.nih.gov/6955308/)
56. A. Whitty, C. A. Fierke, W. P. Jencks, Role of binding energy with coenzyme A in catalysis by 3-oxoacid coenzyme A transferase. *Biochemistry* **34**, 11678–11689 (1995). doi: [10.1021/bi00037a005](https://doi.org/10.1021/bi00037a005); pmid: [7547900](https://pubmed.ncbi.nlm.nih.gov/7547900/)
57. Y.-L. Zhang *et al.*, Impaired transition state complementarity in the hydrolysis of O-arylphosphorothioates by protein-tyrosine phosphatases. *Biochemistry* **38**, 12111–12123 (1999). doi: [10.1021/bi990836i](https://doi.org/10.1021/bi990836i); pmid: [10508416](https://pubmed.ncbi.nlm.nih.gov/10508416/)
58. W. E. Hull, S. E. Halford, H. Gutfreund, B. D. Sykes, Phosphorus-31 nuclear magnetic resonance study of alkaline phosphatase: The role of inorganic phosphate in limiting the enzyme turnover rate at alkaline pH. *Biochemistry* **15**, 1547–1561 (1976). doi: [10.1021/bi00652a028](https://doi.org/10.1021/bi00652a028); pmid: [4092](https://pubmed.ncbi.nlm.nih.gov/4092/)
59. E. E. Wrenbeck, L. R. Azouz, T. A. Whitehead, Single-mutation fitness landscapes for an enzyme on multiple substrates reveal specificity is globally encoded. *Nat. Commun.* **8**, 15695 (2017). doi: [10.1038/ncomms15695](https://doi.org/10.1038/ncomms15695); pmid: [28585537](https://pubmed.ncbi.nlm.nih.gov/28585537/)
60. J. R. Klesmith, J.-P. Bacik, E. E. Wrenbeck, R. Michalczyk, T. A. Whitehead, Trade-offs between enzyme fitness and solubility illuminated by deep mutational scanning. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2265–2270 (2017). doi: [10.1073/pnas.1614437114](https://doi.org/10.1073/pnas.1614437114); pmid: [28196882](https://pubmed.ncbi.nlm.nih.gov/28196882/)
61. P. Bandaru *et al.*, Deconstruction of the Ras switching cycle through saturation mutagenesis. *eLife* **6**, e27810 (2017). doi: [10.7554/eLife.27810](https://doi.org/10.7554/eLife.27810); pmid: [28686159](https://pubmed.ncbi.nlm.nih.gov/28686159/)
62. C. M. Dobson, Protein folding and misfolding. *Nature* **426**, 884–890 (2003). doi: [10.1038/nature02261](https://doi.org/10.1038/nature02261); pmid: [14685248](https://pubmed.ncbi.nlm.nih.gov/14685248/)
63. M. S. Faber, E. E. Wrenbeck, L. R. Azouz, P. J. Steiner, T. A. Whitehead, Impact of in vivo protein folding probability on local fitness landscapes. *Mol. Biol. Evol.* **36**, 2764–2777 (2019). doi: [10.1093/molbev/msz184](https://doi.org/10.1093/molbev/msz184); pmid: [31400199](https://pubmed.ncbi.nlm.nih.gov/31400199/)
64. V. Zhao, W. M. Jacobs, E. I. Shakhnovich, Effect of protein structure on evolution of cotranslational folding. *Biophys. J.* **119**, 1123–1134 (2020). doi: [10.1016/j.bpj.2020.06.037](https://doi.org/10.1016/j.bpj.2020.06.037); pmid: [32857962](https://pubmed.ncbi.nlm.nih.gov/32857962/)
65. N. Halabi, O. Rivoire, S. Leibler, R. Ranganathan, Protein sectors: Evolutionary units of three-dimensional structure. *Cell* **138**, 774–786 (2009). doi: [10.1016/j.cell.2009.07.038](https://doi.org/10.1016/j.cell.2009.07.038); pmid: [19703402](https://pubmed.ncbi.nlm.nih.gov/19703402/)
66. O. Rivoire, K. A. Reynolds, R. Ranganathan, Evolution-based functional decomposition of proteins. *PLOS Comput. Biol.* **12**, e1004817 (2016). doi: [10.1371/journal.pcbi.1004817](https://doi.org/10.1371/journal.pcbi.1004817); pmid: [27254668](https://pubmed.ncbi.nlm.nih.gov/27254668/)
67. V. H. Salinas, R. Ranganathan, Coevolution-based inference of amino acid interactions underlying protein function. *eLife* **7**, e34300 (2018). doi: [10.7554/eLife.34300](https://doi.org/10.7554/eLife.34300); pmid: [30024376](https://pubmed.ncbi.nlm.nih.gov/30024376/)
68. T. Teşileanu, L. J. Colwell, S. Leibler, Protein sectors: Statistical coupling analysis versus conservation. *PLOS Comput. Biol.* **11**, e1004091 (2015). doi: [10.1371/journal.pcbi.1004091](https://doi.org/10.1371/journal.pcbi.1004091); pmid: [25723535](https://pubmed.ncbi.nlm.nih.gov/25723535/)
69. P. Creixell *et al.*, Hierarchical organization endows the kinase domain with regulatory plasticity. *Cell Syst.* **7**, 371–383.e4 (2018). doi: [10.1016/j.cels.2018.08.008](https://doi.org/10.1016/j.cels.2018.08.008); pmid: [30243563](https://pubmed.ncbi.nlm.nih.gov/30243563/)
70. P. M. Fordyce *et al.*, Basic leucine zipper transcription factor Hac1 binds DNA in two distinct modes as revealed by microfluidic analyses. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E3084–E3093 (2012). doi: [10.1073/pnas.1212457109](https://doi.org/10.1073/pnas.1212457109); pmid: [23054834](https://pubmed.ncbi.nlm.nih.gov/23054834/)
71. D. D. Le *et al.*, Comprehensive, high-resolution binding energy landscapes reveal context dependencies of transcription factor binding. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E3702–E3711 (2018). doi: [10.1073/pnas.1715888115](https://doi.org/10.1073/pnas.1715888115); pmid: [29588420](https://pubmed.ncbi.nlm.nih.gov/29588420/)
72. A. K. Adithan, C. J. Markin, D. A. Mokhtari, N. DelRosso, P. M. Fordyce, High-throughput affinity measurements of transcription factor and DNA mutations reveal affinity and specificity determinants. *Cell Syst.* **12**, 112–127.e11 (2021). doi: [10.1016/j.cels.2020.11.012](https://doi.org/10.1016/j.cels.2020.11.012); pmid: [33340452](https://pubmed.ncbi.nlm.nih.gov/33340452/)
73. K. Brower *et al.*, An open-source, programmable pneumatic setup for operation and automated control of single- and multi-layer microfluidic devices. *HardwareX* **3**, 117–134 (2018). doi: [10.1016/j.ohx.2017.10.001](https://doi.org/10.1016/j.ohx.2017.10.001); pmid: [30221210](https://pubmed.ncbi.nlm.nih.gov/30221210/)
74. C. J. Markin *et al.*, Revealing enzyme functional architecture via high-throughput microfluidic enzyme kinetics. Open Science Repository (2021); doi: [10.17605/OSF.IO/QRN3C](https://doi.org/10.17605/OSF.IO/QRN3C)

ACKNOWLEDGMENTS

We thank members of the Herschlag and Fordyce laboratories for discussions and review of the manuscript. M. Madsen for technical assistance, K. Brower for microfluidic mold fabrication and optics assistance, and S. Costello and S. Marqusee for helpful discussions. **Funding:** This research was supported by the NIH grant R01 (GM064798) awarded to D.H. and P.M.F., a Joint Initiative for Metrology in Biology (JIMB) seed grant, a Stanford Bio-X Interdisciplinary Initiative Seed Grant, an Ono Pharma Foundation Breakthrough Innovation Prize, and the Gordon and Betty Moore Foundation (grant 8415). P.M.F. acknowledges the support of an Alfred P. Sloan Foundation fellowship and is a Chan Zuckerberg Biohub Investigator. C.J.M. acknowledges the support of a Canadian Institutes of Health Research (CIHR) Postdoctoral Fellowship. D.A.M. acknowledges support from the Stanford Medical Scientist Training Program

and a Stanford Interdisciplinary Graduate Fellowship (Anonymous Donor) affiliated with Stanford ChEM-H. E.A. acknowledges support from the NIH (GM0595). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a US Department of Energy (DOE) Office of Science User Facility operated under contract DE-AC02-05CH11231, as well as resources obtained from the Facilities Integrating Collaborations for User Science (FICUS) program associated with the Joint Genome Institute of the DOE, proposal 503369. **Author contributions:** C.J.M., D.A.M., and F.S. designed experiments and collected and analyzed data. M.J.A. collected and analyzed data. C.S. provided assistance with statistical analyses. D.A.M. and E.A. generated the sequence alignments and performed phylogenetic analyses. S.A.L. contributed to hardware automation and software

development. D.H. and P.M.F. conceived and supervised the project and acquired funding. C.J.M., D.A.M., D.H., and P.M.F. wrote and revised the manuscript, with input from all authors. **Competing interests:** The authors declare no competing financial interests. **Data and materials availability:** Summary tables of all kinetic and thermodynamic parameters measured for each mutant are included in the supplementary materials. All data acquired in this study, all code used to obtain and process images and fit kinetic and thermodynamic parameters, and PyMOL files corresponding to all figures containing structural information are available in a registered Open Science Foundation Repository (DOI: 10.17605/OSF.IO/QRN3C) (74). Additional per-experiment and per-mutant summary PDFs are available on the Fordyce Lab website (www.fordycelab.com/publication-data).

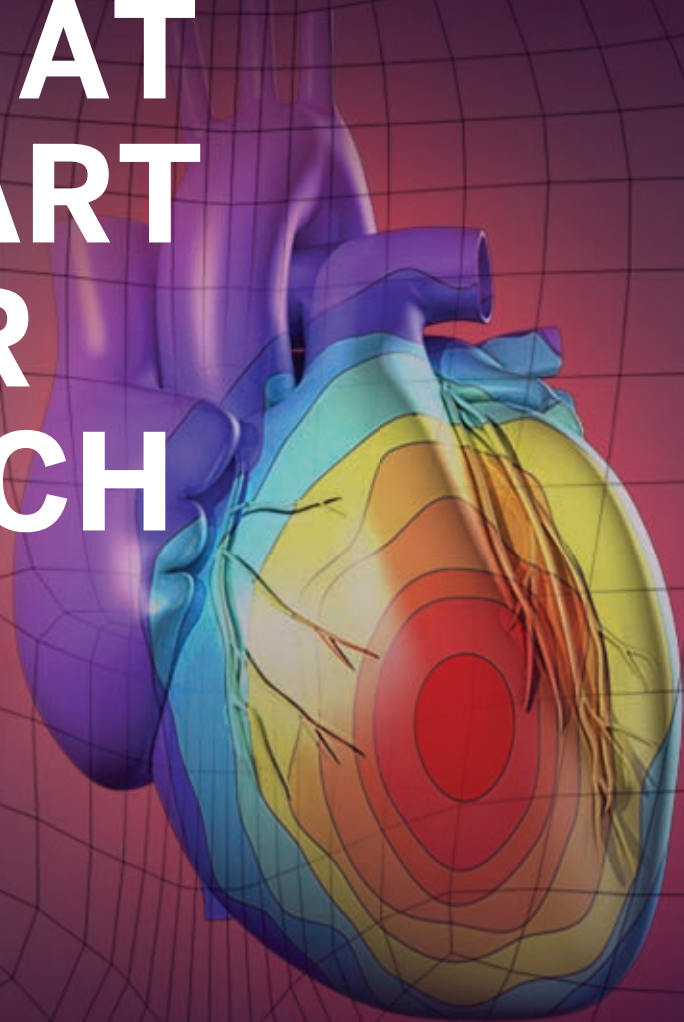
SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/eabf8761/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S70
Tables S1 to S17
References (75–118)
MDAR Reproducibility Checklist
Data files S1 and S2

25 November 2020; accepted 24 May 2021
10.1126/science.abf8761

PUT HUMAN HEALTH AT THE HEART OF YOUR RESEARCH

Submit your research:
cts.ScienceMag.org



Science
Translational
Medicine
 AAAS



Twitter: @ScienceTM



Facebook: @ScienceTranslationalMedicine

RESEARCH ARTICLE SUMMARY

NEURODEVELOPMENT

Retinal waves prime visual motion detection by simulating future optic flow

Xinxin Ge[†], Kathy Zhang[†], Alexandra Gribizis, Ali S. Hamodi, Aude Martinez Sabino, Michael C. Crair^{*}

INTRODUCTION: Fundamental circuit features of the mouse visual system emerge before the onset of vision, allowing the mouse to perceive objects and detect visual motion immediately upon eye opening. How the mouse visual system achieves self-organization by the time of eye opening without structured external sensory input is not well understood. In the absence of sensory drive, the developing retina generates spontaneous activity in the form of propagating waves. Past work has shown that spontaneous retinal waves provide the correlated activity necessary to refine the development of gross topographic maps in downstream visual areas, such as retinotopy and eye-specific segregation, but it is unclear whether waves also convey information that

instructs the development of higher-order visual response properties, such as direction selectivity, at eye opening.

RATIONALE: Spontaneous retinal waves exhibit stereotyped changing spatiotemporal patterns throughout development. To characterize the spatiotemporal properties of waves during development, we used one-photon wide-field calcium imaging of retinal axons projecting to the superior colliculus in awake neonatal mice. We identified a consistent propagation bias that occurred during a transient developmental window shortly before eye opening. Using quantitative analysis, we investigated whether the directionally biased retinal waves conveyed ethological information relevant to future vis-

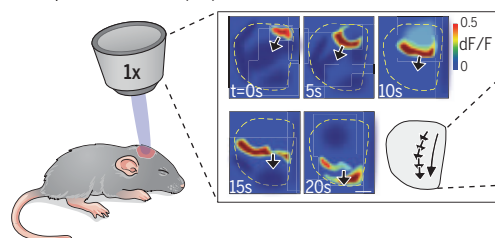
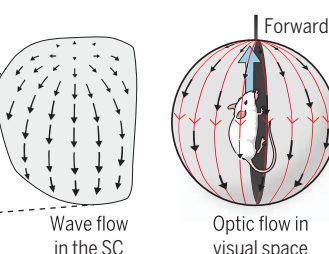
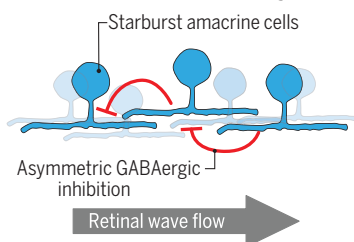
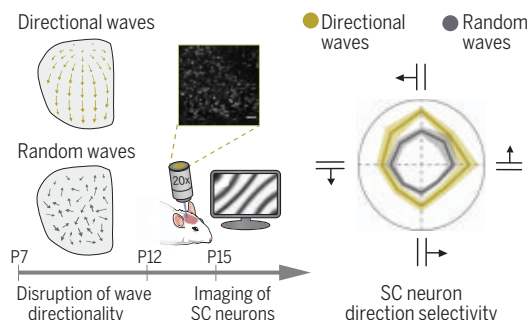
ual inputs. To understand the origin of directional retinal waves, we used pharmacological, optogenetic, and genetic strategies to identify the retinal circuitry underlying the propagation bias. Finally, to evaluate the role of directional retinal waves in visual system development, we used pharmacological and genetic strategies to chronically manipulate wave directionality and used two-photon calcium imaging to measure responses to visual motion in the midbrain superior colliculus immediately after eye opening.

RESULTS: We found that spontaneous retinal waves in mice exhibit a distinct propagation bias in the temporal-to-nasal direction during a transient window of development (postnatal day 8 to day 11). The spatial geometry of directional wave flow aligns strongly with the optic flow pattern generated by forward self-motion, a dominant natural optic flow pattern after eye opening. We identified an intrinsic asymmetry in the retinal circuit that enforced the wave propagation bias involving the same circuit elements necessary for motion detection in the adult retina, specifically asymmetric inhibition from starburst amacrine cells through γ -aminobutyric acid type A (GABA_A) receptors. Finally, manipulation of directional retinal waves, through either the chronic delivery of gabazine to block GABAergic inhibition or the starburst amacrine cell-specific mutation of the *FRMD7* gene, impaired the development of responses to visual motion in superior colliculus neurons downstream of the retina.

CONCLUSION: Our results show that spontaneous activity in the developing retina prior to vision onset is structured to convey essential information for the development of visual response properties before the onset of visual experience. Spontaneous retinal waves simulate future optic flow patterns produced by forward motion through space, due to an asymmetric retinal circuit that has an evolutionarily conserved link with motion detection circuitry in the mature retina. Furthermore, the ethologically relevant information relayed by directional retinal waves enhances the development of higher-order visual function in the downstream visual system prior to eye opening. These findings provide insight into the activity-dependent mechanisms that regulate the self-organization of brain circuits before sensory experience begins. ■

A Spontaneous retinal waves are directional

Imaging of retinal axon activity in the superior colliculus (SC)

**B Directional retinal waves mimic optic flow from forward motion****C Asymmetric retinal inhibition mediates wave directionality****D Disruption of wave directionality impairs motion detection in SC**

Origin and function of directional retinal waves. (A) Imaging of retinal axon activity reveals a propagation bias in spontaneous retinal waves (scale bar, 500 μ m). (B) Cartoon depiction of wave flow vectors projected onto visual space. Vectors (black arrows) align with the optic flow pattern (red arrows) generated by forward self-motion. (C) Asymmetric GABAergic inhibition in the retina mediates wave directionality. (D) Developmental manipulation of wave directionality disrupts direction-selective responses in downstream superior colliculus neurons at eye opening.

Department of Neuroscience, Yale University School of Medicine, New Haven, CT 06510, USA.

^{*}Corresponding author. Email: michael.crair@yale.edu[†]These authors contributed equally to this work.Cite this article as X. Ge *et al.*, *Science* 373, eabd0830 (2021).

DOI: 10.1126/science.abd0830

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.abd0830>

RESEARCH ARTICLE

NEURODEVELOPMENT

Retinal waves prime visual motion detection by simulating future optic flow

Xinxin Ge^{†‡}, Kathy Zhang[†], Alexandra Gribizis[§], Ali S. Hamodi, Aude Martinez Sabino[¶], Michael C. Crair^{*}

The ability to perceive and respond to environmental stimuli emerges in the absence of sensory experience. Spontaneous retinal activity prior to eye opening guides the refinement of retinotopy and eye-specific segregation in mammals, but its role in the development of higher-order visual response properties remains unclear. Here, we describe a transient window in neonatal mouse development during which the spatial propagation of spontaneous retinal waves resembles the optic flow pattern generated by forward self-motion. We show that wave directionality requires the same circuit components that form the adult direction-selective retinal circuit and that chronic disruption of wave directionality alters the development of direction-selective responses of superior colliculus neurons. These data demonstrate how the developing visual system patterns spontaneous activity to simulate ethologically relevant features of the external world and thereby instruct self-organization.

Essential functions of the mammalian visual system, including the ability to locate objects and detect motion, are present even at the first onset of vision (1–7). How the visual system organizes its functional maturation before the start of sensory experience remains unclear. Prior to vision, spontaneous retinal activity occurs in the form of propagating waves that spread throughout the ascending visual system (8–10). Spontaneous retinal waves correlate the activity of neighboring cells in the retina and, through a Hebb-based mechanism, are critical for the development of visual maps (such as retinotopy and eye-specific segregation) that reflect near-neighbor relations in visual space (1, 2, 11–21). Whether and how the spatiotemporal pattern of spontaneous retinal waves conveys additional information beyond the gross correlations between topographically neighboring cells and contributes to the development of higher-order visual response properties remains unclear.

Spontaneous retinal waves in mice are divided into three stages according to their requisite mode of interneuronal communication (fig. S1A) (22). The spatiotemporal pattern of waves differs across these stages (10, 23, 24), but relatively little is known about whether or which of these spatiotemporal features might be relevant for visual system development. Previous studies have shown that retinal waves

exhibit an early bias in propagation direction of unknown origin and function (9, 10, 18). In amphibian larvae, whose development takes place largely externally and in the presence of visual stimuli capable of driving retinal activity, the direction of visual optic flow as the animal moves through space instructs visual system development (25). We sought to examine whether the directional propagation bias of spontaneous retinal waves prior to eye opening might also convey critical information for the functional maturation of the visual system in mammals.

Directional retinal waves resemble future optic flow

To determine the developmental time course of the directional bias in retinal waves, we examined the spontaneous activity of retinal ganglion cells in young mice at multiple ages throughout development *in vivo*. Spontaneous retinal activity was imaged in topographically mapped retinal ganglion cell axon arbors in the superior colliculus (Fig. 1A). To gain full access to retinal activity, we used *Emx1-Cre; Tra2 $\beta^{f/f}$* mice (cortexless mice) (26), which have a completely exposed superior colliculus due to suppressed cortical neurogenesis. Retinal ganglion cells were labeled with the genetically encoded calcium indicator GCaMP6s by intra-ocular injections of AAV2/1-hSyn-GCaMP6s virus at postnatal day 0 to 1 (P0–P1) for calcium imaging of activity in retinal axon arbors in the superior colliculus at later ages (Fig. 1A). Spontaneous retinal activity was observed in all age groups between P3 and P29 in head-fixed, awake mice (movie S1).

During the first two postnatal weeks, wave directionality changed together with other spatiotemporal properties of waves, such as

wave interval, duration, size, and propagation distance (fig. S1). In retinal coordinates, used throughout for data presentation and description, retinal waves exhibited a temporal-to-nasal directional bias, corresponding to forward-to-backward motion in visual space. The wave directional bias spanned the end of stage 2 (P8–P9) and persisted into stage 3 (P10–P11) (Fig. 1C and fig. S1F), even though waves at P8–P9 and P10–P11 differed in other spatiotemporal properties (Fig. 1B, fig. S1, B to E, and table S1). To exclude the possibility that the wave directional bias was due to the loss of cortex in *Emx1-Cre; Tra2 $\beta^{f/f}$* mice, we also examined retinal activity in mice with an intact cortex. A similar pattern of wave directionality was observed in mice with an intact cortex and in *Emx1-Cre; Tra2 $\beta^{f/f}$* (cortexless) mice with a restricted field of view to resemble cortex-intact mice (fig. S2 and table S1). These results reveal the presence of a developmentally transient temporal-to-nasal bias in wave propagation direction that spanned both stage 2 and stage 3 waves.

To examine whether locations across the superior colliculus experienced a consistent directional bias, we quantified the average wave propagation across each pixel separately. The pixel-based directionality analysis confirmed the temporal-to-nasal wave-based analysis across development (fig. S3). However, the pixel-based flow patterns were inhomogeneous and stereotypically deviated from the dominant temporal-to-nasal wave direction, depending on location in the superior colliculus (Fig. 1D and fig. S4A). We investigated the relationship between this stereotyped wave flow pattern and optic flow by first measuring retinotopy in the superior colliculus of cortexless mice after eye opening (~P21) (fig. S4B). We then compared the wave flow patterns with translatory or rotatory optic flow fields. To quantify the alignment between wave flow and translatory or rotatory optic flow fields, we calculated a concordance index (27) as the percentage of pixels with wave flow within 10° of local optic flow. We found that wave flow vectors were significantly better aligned with a translatory optic flow field than with a rotatory optic flow field (Fig. 1, E and F). The translatory optic flow alignment was best fit with a center of expansion in the temporal retina, consistent with the flow patterns produced by forward self-motion through space (fig. S5). These results demonstrate that spontaneous retinal waves prior to eye opening resemble forward translatory optic flow, the dominant future optic flow pattern received by animals as they move through space after eye opening.

Asymmetric GABAergic inhibition mediates wave directionality

To investigate whether directional retinal waves affect visual system development, we

Department of Neuroscience, Kavli Institute for Neuroscience, Yale University School of Medicine, New Haven, CT 06510, USA.

*Corresponding author. Email: michael.crair@yale.edu

†These authors contributed equally to this work. ‡Present address:

Department of Physiology, University of California, San Francisco, CA 94158, USA. §Present address: Department of Functional Architecture and Development of Cerebral Cortex, Max Planck Florida Institute for Neuroscience, Jupiter, FL 33458, USA.

¶Present address: Bioengineering Department, University of Technology of Compiègne, 60200 Compiègne, France.

sought to first understand the origin of wave directionality. We considered two nonexclusive mechanisms for wave directionality (Fig. 2A). The first is an intrinsic propagation bias: an asymmetric retinal circuit that intrinsically

favors wave propagation in the temporal-to-nasal direction. The second potential mechanism is an initiation bias: a tendency for waves to initiate at the temporal edge of the retina that could cause an overall temporal-to-

nasal directionality as waves travel away from the temporal edge.

We first examined the possibility of a propagation bias mechanism by assessing the directionality of waves initiated from the nasal part

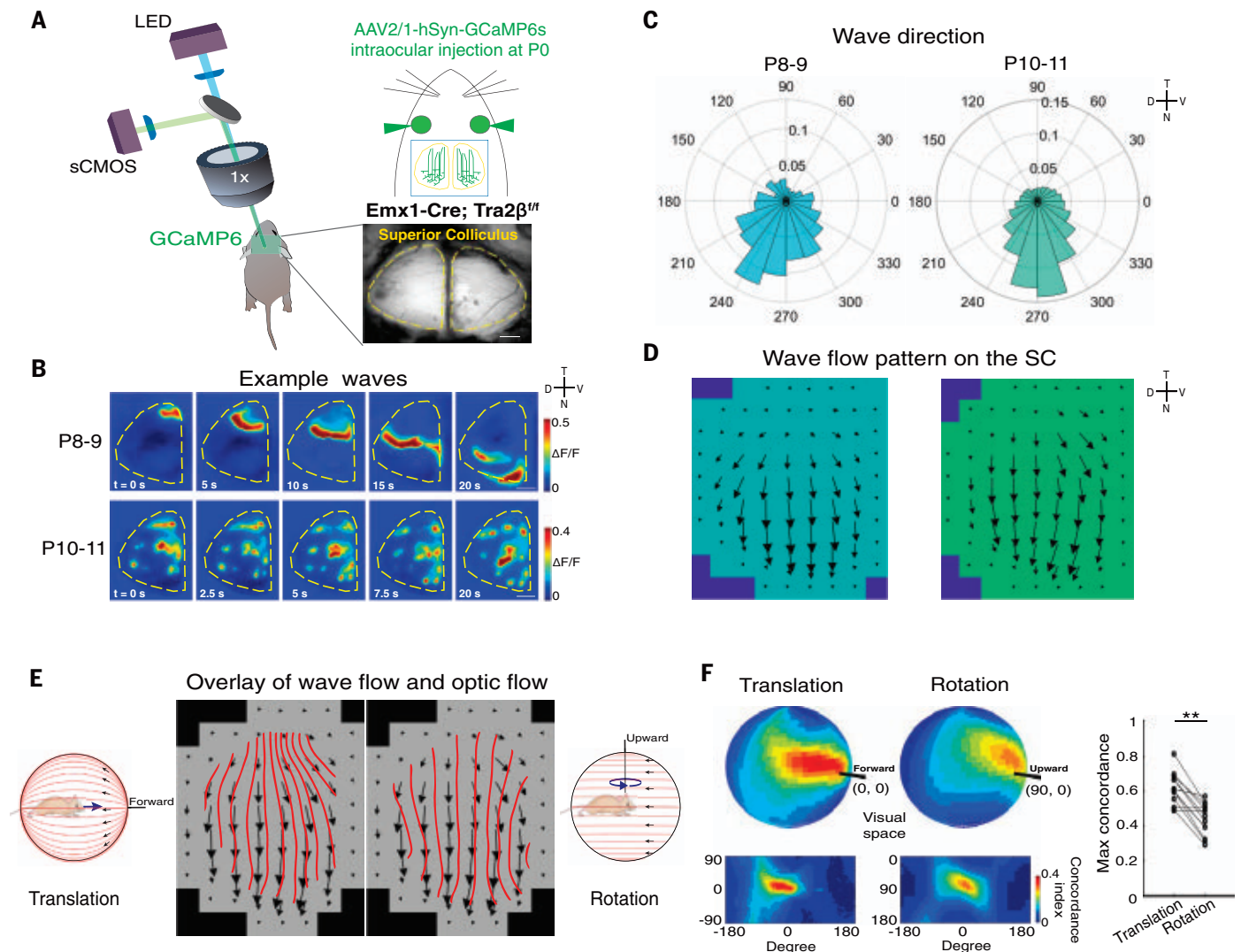


Fig. 1. Spontaneous retinal waves are directional and resemble translatory optic flow during a transient developmental window. (A) Schematic showing the wide-field single-photon microscopy used to perform Ca^{2+} imaging of retinal ganglion cell axon activity in the superior colliculus (SC). GCaMP6s expression in retinal ganglion cells was achieved by intraocular injections of AAV2/1-hSyn-GCaMP6s in *Emx1-Cre; Tra2 β^{fl}* mice. A cranial window above the SC was surgically prepared before imaging. Mice were unanesthetized and head-fixed during the imaging. Scale bar, 500 μm . (B) $\Delta F/F$ montages of example waves at P8-P9 (top) and P10-P11 (bottom). (C) Frequency distributions of propagation directions for all waves at P8-P9 [$n = 5$ hemispheres, 831 waves, total circular variance (CV) = 0.6087] and at P10-P11 ($n = 8$ hemispheres, 7937 waves, total CV = 0.5611). Inset at right: Directions T, N, D, and V correspond to temporal, nasal, dorsal, and ventral in the retina unless otherwise stated. (D) Averaged pixel-based wave flow fields on the SC at P8-P9 (left) and at P10-P11 (right). (E) Left: Optic flow pattern produced by an animal's forward translation through space, and an overlay between average wave flow (black arrows) and the translatory optic flow that best matched with wave flow (red lines) on the left SC. Right: Optic flow induced by an animal's rotation in space,

and an overlay between average wave flow and rotatory optic flow that best matched with wave flow on the SC. Wave flow was averaged from all recordings at P8-P11 ($n = 13$ hemispheres). Optic flow fields were generated from retinotopic maps in the SC (see supplementary materials for details). Forward and upward indicate directions in visual space. (F) Concordance index between wave and translatory or rotatory optic flow as a function of the axis of the animal's translation or rotation. All possible axes of translation or rotation were tested. Concordance index is the percentage of pixels on the SC that have the same direction ($<10^\circ$ difference) as the local optic flow. Top: Spherical maps of average concordance index ($n = 13$ hemispheres) between wave and translatory (top left) and rotatory (top right) optic flow in global visual space. The best match between wave flow and optic flow due to translation deviated slightly from forward motion (10° elevation and 15° azimuth). Bottom: Flattened concordance maps. Black bars indicate the forward and upward directions in visual space. Right: Maximum concordance index between individual wave flow fields and translatory optic flow was much higher than rotatory optic flow (translatory, 0.611 ± 0.026 ; rotatory, 0.438 ± 0.025 ; $n = 13$ hemispheres; $**P = 0.0012$, Wilcoxon signed-rank test). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. Data are means \pm SEM.

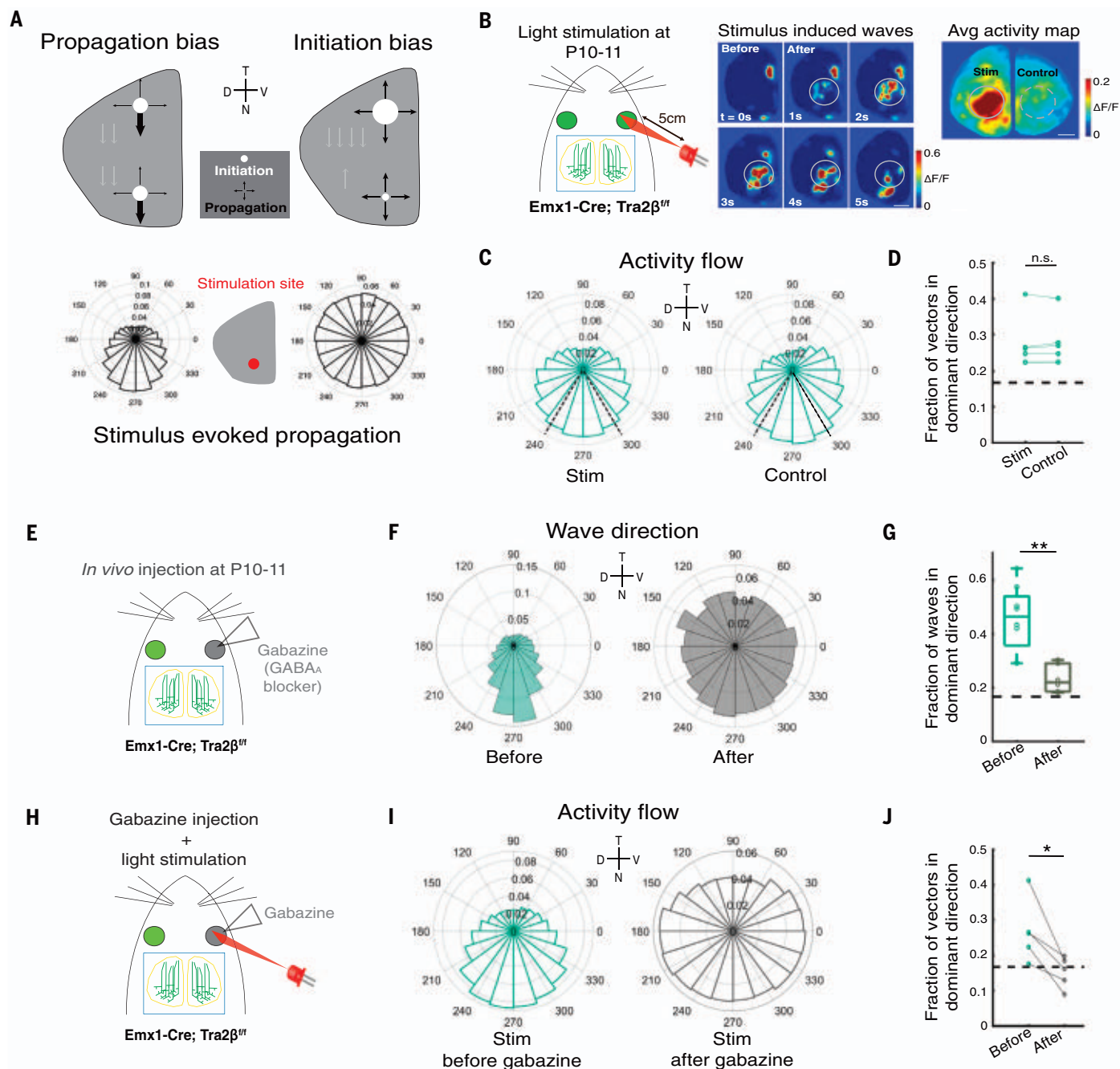


Fig. 2. Wave directionality requires a propagation bias mediated by asymmetric GABAergic inhibition. (A) Possible mechanisms of wave directionality. Top left: Wave bias may be generated by a propagation bias in the retinal circuitry that favors wave propagation in the temporal-to-nasal direction. Top right: Biased initiation in the temporal retina can induce wave directional bias. White circles, likelihood of initiation; black arrows, likelihood of propagation in certain directions. Bottom: Hypothetical propagation directions for waves stimulated in the nasal retina for the two proposed mechanisms. (B) Left: At P10–P11, light stimulation was performed in *Emx1-Cre; Tra2 $\beta^{fl/fl}$* mice that had received intraocular injections of AAV2/1-hSyn-GCaMP6s at P0–P1. The 615-nm LED strobe was placed ~5 cm away from the stimulated eye and was oriented toward the nasal retina. Light stimuli had a duration of 500 ms and were repeated every 30 s. Center: Example montages of waves induced by light stimulation at P10–P11. Right: Averaged activity 0 to 10 s after each stimulation in the same example animal. Solid circles (stim, surrounding the stimulation site)

and dashed circles [control, mirrored region of interest (ROI) in the opposite hemisphere] indicate the ROIs that were selected for the analyses of wave propagation directions. (C) Pooled pixel-based activity flow directions within stim (left) and control (right) ROIs 0 to 10 s after each stimulation ($n = 5$ animals; stim, total CV = 0.680; control, total CV = 0.682). Dashed black lines indicate the 60° around the dominant direction that was used for further analyses. (D) Proportion of pixel-based activity flow vectors within 60° around the dominant direction. Each data point represents one hemisphere; data points from the same animal were paired. Dashed black line, fraction in the dominant direction if wave propagation directions are random (stim, 0.282 ± 0.033 ; control, 0.284 ± 0.030 ; $P = 0.625$, Wilcoxon signed-rank test). (E) Gabazine (50 μM) was applied in vivo at P10 in *Emx1-Cre; Tra2 $\beta^{fl/fl}$* mice that received intraocular injections of AAV2/1-hSyn-GCaMP6s at P0–P1. (F) Frequency distributions of wave propagation directions in animals without gabazine injection (before, $n = 8$ hemispheres, total CV = 0.561) or after gabazine injection

(after, $n = 6$ hemispheres, total CV = 0.931) at P10. **(G)** Gabazine reduced the proportion of waves propagating within 60° of the dominant temporal-to-nasal direction (control, 0.456 ± 0.044 ; gabazine, 0.234 ± 0.021 ; $**P = 0.003$, Wilcoxon rank sum test). **(H)** Light stimuli were applied in *Emx1-Cre; Tra2 β ^{fl/fl}* mice before and after gabazine injection to the eye, following the same protocol described in (A). **(I)** Pixel-based activity flow directions within stim ROIs 0 to 10 s

after each stimulation before ($n = 5$ animals, total CV = 0.682) and after gabazine injection ($n = 5$ animals, total CV = 0.938). **(J)** Gabazine injections reduced the proportion of pixel-based activity flow vectors within 60° surrounding the dominant direction. Data points from the same hemisphere were paired (before, 0.267 ± 0.040 ; after, 0.152 ± 0.020 ; $*P = 0.031$, one-tailed Wilcoxon signed-rank test). All boxplots report medians with the 75% confidence interval unless otherwise noted.

of the retina at P10–P11 (Fig. 2A). To control the site of wave initiation, we made use of exogenous stimuli. Starting around P9, photoreceptors are weakly but functionally integrated into the retinal circuit (28–31), which enabled us to evoke waves using focal light stimuli (Fig. 2B and movie S2). Waves stimulated in the nasal retina at P10–P11 exhibited the same temporal-to-nasal bias as spontaneous retinal waves (Fig. 2, C and D), suggesting an inherent anisotropy in retinal circuitry that favors retinal wave propagation in the temporal-to-nasal direction at this age. Consistent with P10–P11, P8–P9 waves initiated in the nasal retina exhibited the same temporal-to-nasal bias (fig. S6, A to C). To ensure that the propagation bias was not an artifact of stimulation, we examined waves initiated in the nasal retina at P6–P7 when spontaneous retinal waves are not biased in the temporal-to-nasal direction. At this age, waves were initiated by optogenetic stimulation of starburst amacrine cells, interneurons that are responsible for stage 2 wave initiation and propagation (fig. S6A and movie S3). In contrast to P8–P9 and P10–P11, waves initiated in the nasal retina at P6–P7 were not biased in the temporal-to-nasal direction (fig. S6, B and C). We next quantified the distribution of wave initiation locations at P8–P9 and P10–P11. Although waves were equally likely to initiate from any location in the retina at P10–P11, there was an initiation bias at P8–P9 (fig. S6D). The initiation bias at P8–P9 could be due to an extra source of excitatory drive in the temporal retina at this age, or could be a consequence of the slow afterhyperpolarizations in the recurrent network of starburst amacrine cells that entrains an initiation zone (22, 32–34). Regardless, neonatal mice at both P8–P9 and P10–P11 displayed a consistent retinal circuit propagation bias, which suggests a generic role of intrinsic propagation anisotropy in mediating retinal wave directionality that is absent at P6–P7.

We next investigated the origin of the propagation anisotropy described above. In the adult mouse retina, direction selectivity in retinal ganglion cells is achieved through asymmetric γ -aminobutyric acid-mediated (GABAergic) inhibition (35, 36). We reasoned that asymmetric circuit wiring might also be present during development and provide imbalanced inhibition to waves propagating in different directions to generate the observed propagation anisotropy. To determine whether asymmetric inhibition is involved in wave di-

rectionality, we examined wave directionality after the application of inhibitory antagonists to the retina in vivo at P10 (Fig. 2E). Application of gabazine, a GABA_A receptor (GABA_AR) antagonist, significantly reduced the wave directional bias (Fig. 2, F and G, and movie S4), whereas applications of TPMPA, a GABA_C receptor (GABA_CR) antagonist, or strychnine, an antagonist of glycine, had little effect (fig. S7, A to C). As a result of disrupted wave directionality, the concordance between translatory optic flow and wave flow was significantly reduced after gabazine injection (fig. S7H). No other spatiotemporal feature was significantly altered by gabazine application, suggesting the specificity of GABA_ARs at this age in mediating wave directionality (fig. S7, D to G, and table S1). We further examined the directionality of exogenously stimulated waves after retinal application of gabazine (Fig. 2H). Consistent with the significant reduction of directional bias in spontaneous waves, stimulated waves were similarly disrupted after gabazine application (Fig. 2, I and J). These data demonstrate that asymmetric inhibition mediated by retinal GABA_ARs is crucial for wave directionality.

Wave directionality requires starburst amacrine cells

We next sought to identify the underlying cellular components that produce the asymmetric inhibition necessary for wave directionality. Starburst amacrine cells are known to confer motion direction-selective responses on direction-selective ganglion cells in the adult mouse retina (35, 37). This property suggests starburst amacrine cells as a candidate for providing asymmetric inhibition in wave propagation. However, in the developing retina, starburst amacrine cells are only thought to generate stage-2 retinal waves via their recurrent excitatory network (33), and little is known about the role of inhibition from starburst amacrine cells in wave propagation. To examine the role of starburst amacrine cells in directional waves, we selectively ablated starburst amacrine cells through daily retinal application of diphtheria toxin (DT) from P5 to P7 in *ChAT-Cre; LSL-DTR* mice, which specifically express DT receptors in starburst amacrine cells (Fig. 3, A and B), and examined wave directionality at P10. Retinal waves in P10 starburst amacrine cell-ablated animals maintained a normal peak wave area and propagation distance relative to controls (fig. S8, B to E, table S1, and movie S5) but had sig-

nificantly disrupted wave directionality (Fig. 3, D and E). Wave directionality in starburst amacrine cell-ablated animals was not further altered after gabazine injection (fig. S8A), consistent with GABA_AR-dependent wave directionality acting through starburst amacrine cells. These results reveal that generation of the temporal-to-nasal wave directional bias observed at P10 requires starburst amacrine cells.

To further identify the features of starburst amacrine cells crucial for wave directionality, we examined retinal waves in transgenic animals that have a starburst amacrine cell-specific mutation of the *FRMD7* gene (38, 39). In the adult retina, mutation of *FRMD7* abolishes temporal and nasal retinal ganglion cell direction selectivity by eliminating the asymmetry of inhibition from starburst amacrine cells onto temporal- or nasal-preferring direction-selective ganglion cells (38). During development, spontaneous retinal waves in *FRMD7*tm mice displayed a significant disruption in wave directionality similar to that observed in starburst amacrine cell-ablated animals (Fig. 3, F and G). The disruption in wave directionality in *FRMD7*tm mice was not further altered by gabazine (fig. S6F), and other wave properties remained unchanged (fig. S6, G to J, table S1, and movie S5). These results show that asymmetric inhibition from starburst amacrine cells, which confer direction selectivity in the adult retina, is also necessary for mediating temporal-to-nasal wave directionality during development. Taken together with the finding that gabazine blocks wave directionality, our data suggest that asymmetric GABA_AR-mediated inhibition by starburst amacrine cells is necessary for the transient temporal-to-nasal directional bias in spontaneous retinal waves during development.

Because starburst amacrine cells in the adult retina provide asymmetric inhibition to downstream direction-selective ganglion cells as well as to neighboring starburst amacrine cells or other amacrine cells (27, 37, 40, 41), it remains unclear which asymmetric inhibitory connection mediates wave directionality. To determine whether starburst amacrine cells modulate wave directionality by providing asymmetric inhibition onto direction-selective ganglion cells, we examined wave directionality in two different subtypes of direction-selective ganglion cells: *Cdh6*-positive ganglion cells that receive asymmetric inhibition from starburst amacrine cells along the dorsal-ventral axis

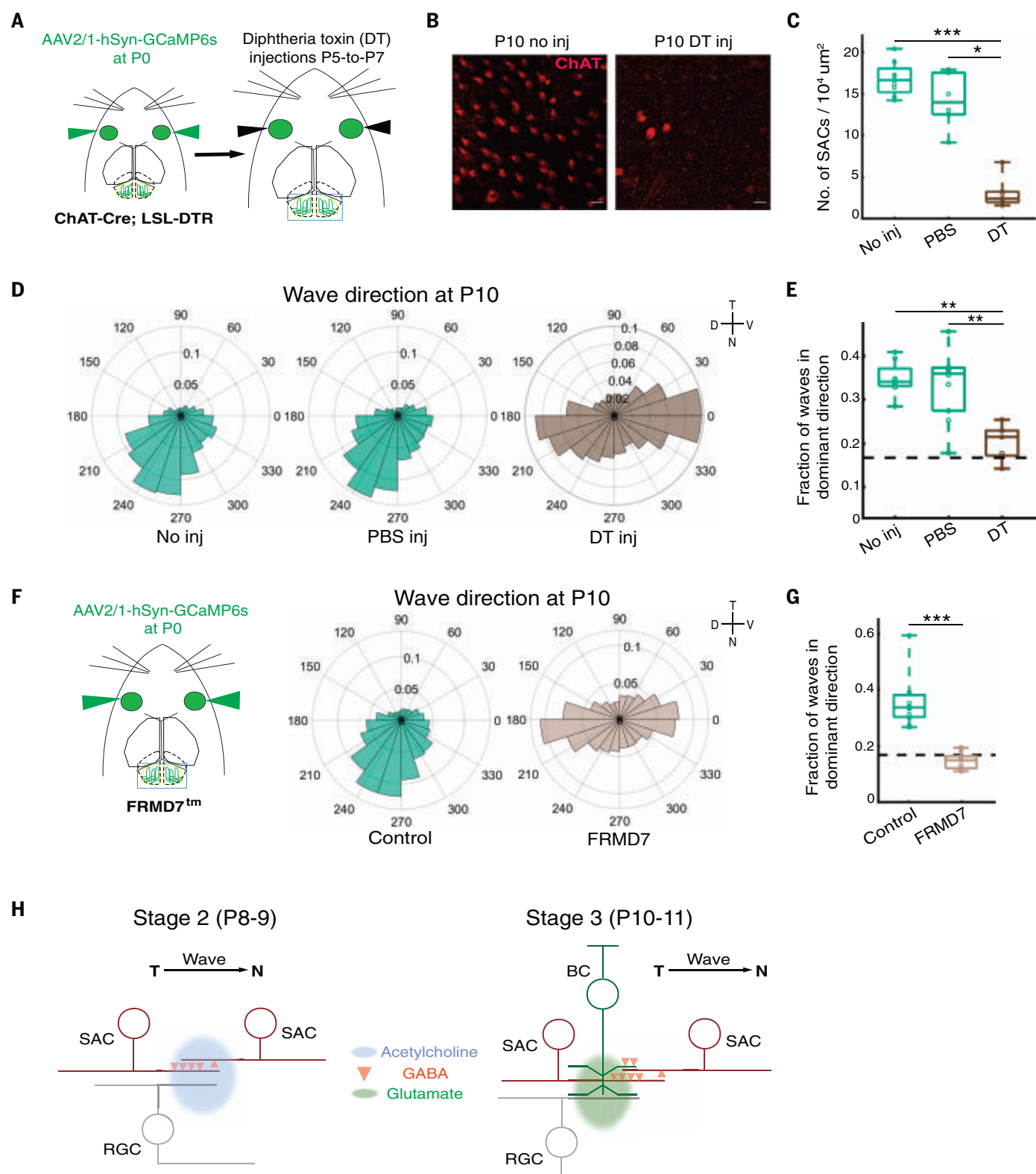


Fig. 3. Starburst amacrine cells (SACs) mediate wave directionality. (A) AAV2/1-hSyn-GCaMP6s was intraocularly injected in ChAT-Cre; LSL-DTR mouse pups at P0. Intraocular injections of diphtheria toxin (DT, 0.8 ng/ μl , or PBS as controls) were performed every 24 hours between P5 and P7. (B) Immunostaining of ChAT in the whole-mount retina at P10 with (left) or without (right) DT treatment. Scale bar, 20 μm . Images were background-

subtracted. (C) Number of SACs per unit area ($1 \times 10^4 \mu\text{m}^2$). Each data point represents the averaged quantification from one retina (no inj: $n = 8$ retinas, 16.79 ± 0.74 ; PBS: $n = 6$ retinas, 14.17 ± 1.36 ; DT: $n = 7$ retinas, 2.98 ± 0.67 ; *** $P = 0.0009$, Kruskal-Wallis test; Tukey's post hoc test, no injection versus PBS $P = 0.659$, no injection versus DT *** $P = 7.405 \times 10^{-4}$, PBS versus DT * $P = 0.028$). (D) Frequency distributions of wave propagation directions

show that SAC-depleted retinas lack temporal-to-nasal propagation bias. Left to right, no injection ($n = 8$ hemispheres, total CV = 0.5167), PBS-injected ($n = 10$ hemispheres, total CV = 0.5475), and DT-injected ($n = 7$ hemispheres, total CV = 0.8373). **(E)** Proportion of wave propagation directions within 60° surrounding the dominant direction (no inj = 0.347 ± 0.014 ; PBS = 0.332 ± 0.027 ; DT = 0.146 ± 0.021 ; $**P = 0.0011$, Kruskal-Wallis test; Tukey's post hoc test, no injection versus DT-injected $**P = 0.0043$, PBS injection versus DT-injected $**P = 0.0024$, no injection versus PBS-injected $P = 1$). **(F)** Left: FRMD7tm mice received intraocular injections of AAV2/1-hSyn-GCaMP6s at P0. Middle right: FRMD7tm mice lack temporal-to-nasal wave propagation bias as revealed by the frequency distribution of wave propagation directions in littermate controls ($n = 11$ hemispheres, total CV = 0.559) and in FRMD7tm mice ($n = 9$ hemispheres, total CV = 0.9214) at P10. **(G)** Proportion of wave propagation directions within

60° surrounding the dominant direction (control, 0.353 ± 0.027 ; FRMD7tm, 0.145 ± 0.010 ; $***P = 1.971 \times 10^{-6}$, Wilcoxon rank sum test). Dashed black lines in **(E)** and **(G)** indicate the fraction in the dominant direction if wave propagation directions are random. **(H)** Schematic models of retinal circuits that generate temporal-to-nasal directional waves during stage 2 (left) and stage 3 (right). Left: Stage 2 retinal waves propagate via SAC release of acetylcholine (blue) onto other cells, wave directionality is mediated by asymmetric GABAergic inhibition (red) among the recurrent SAC network. Greater inhibition on the nasal SAC dendrite prevents temporal propagation of waves. Right: Stage 3 retinal waves propagate via spillover of glutamate (green) from bipolar cells (BCs) onto other BCs, wave directionality is mediated by asymmetric GABAergic inhibition among SACs, and their feedback onto BCs. Feedback from SACs onto BCs may also be relayed by other types of intermediate amacrine cells in the retina (41).

(42) and JamB-positive ganglion cells that lack strong asymmetric inhibitory input from starburst amacrine cells in the adult retina (43). If wave directionality arises at the synapse between starburst amacrine cells and direction-selective ganglion cells, which emerges around P8 (44, 45), we should observe a dorsal-ventral propagation bias in Cdh6-positive ganglion cells and little propagation bias in JamB-positive ganglion cells at P10. However, waves in both JamB-positive and Cdh6-positive ganglion cells propagated in the temporal-to-nasal direction (fig. S9 and movie S6). These data demonstrate that retinal ganglion cells inherit biased propagating activity from upstream circuits in the retina, not from asymmetric inhibitory inputs from starburst amacrine cells onto direction-selective ganglion cells. Our results also suggest that starburst amacrine cells modulate wave directionality by providing inhibition onto other cell types involved in wave propagation, perhaps onto starburst amacrine cells themselves during stage 2 and directly onto bipolar cells or indirectly onto bipolar cells via other amacrine cells (41) during stage 3 (Fig. 3H). Further experiments are required to test these hypotheses.

Directional retinal waves enhance downstream visual function

To directly probe the role of directional waves in visual system development, we chronically manipulated the directionality of spontaneous retinal waves and measured the visual function of neurons in the superior colliculus at eye opening. We used two-photon imaging to record the activity of neurons expressing nucleus-targeted GCaMP6s in the superior colliculus of Emx1-Cre; Tra2 $\beta^{f/f}$ mice (Fig. 4A). To specifically disrupt wave directionality, we performed intraocular injections of gabazine (or saline) every 24 hours from P7 to P12 (Fig. 4A). We collected cellular responses in the superficial layers (50 to 60 μm) at the center of the superior colliculus 2 to 4 days after the last gabazine treatment (at P14–P16) in awake head-fixed mice as they viewed drifting grating visual stimuli (Fig. 4A). Drifting gratings elicited responses from similar proportions of

neurons in saline-treated and chronic gabazine-treated mice (fig. S10A). Cells that preferentially responded to certain directions or orientations were found in both groups (Fig. 4C and supplementary materials). For all responsive cells, we measured the strength of direction and orientation tuning with a global direction selectivity index (gDSI) and a global orientation selectivity index (gOSI). We observed a significant reduction overall in the direction selectivity of cells in chronic gabazine-treated mice relative to saline controls (Fig. 4D). In contrast, gabazine treatment had little effect on the orientation selectivity of neurons in the superior colliculus (fig. S11A).

To determine whether disruption of wave directionality affected the selectivity of neurons for specific stimulus directions, we divided the responsive cells into eight groups according to their preferred direction and compared the tuning strength in each group. We observed a significant change in the distribution of gDSI across directions in chronic gabazine-treated mice in comparison to saline-treated mice, but the strength of direction selectivity was not significantly altered in any single direction (Fig. 4E and fig. S10B). We next examined whether the disruption of wave directionality altered the proportion of direction-selective cells preferring different directions. To do so, we computed the proportion of direction-selective cells preferring each direction among all responding cells. Consistent with the overall reduction in the strength of direction selectivity, we observed that the overall proportion of direction-selective cells was also significantly reduced in chronic gabazine-treated mice relative to saline-treated mice (Fig. 4F). We also observed that the distribution of direction preference across the eight different directions was significantly altered in chronic gabazine-treated mice relative to saline-treated mice, with a significant reduction in the dorsal direction (Fig. 4G and fig. S10C). Finally, we examined whether orientation selectivity was impaired in an orientation-dependent way, even though gabazine treatment had no impact on overall orientation tuning. No difference between saline- and gabazine-treated

groups was observed in the tuning strength for any orientation (fig. S11, A and B), and the proportion of orientation-selective cells preferring each orientation among all responding cells was not significantly different between saline and gabazine groups (fig. S11C).

Because the effects of gabazine had a restricted duration in vivo (fig. S13), the manipulation of wave directionality using chronic intraocular injections of gabazine was likely limited. To more consistently disrupt wave directionality, we examined the direction-selective responses of superior colliculus neurons in Emx1-Cre; Tra2 $\beta^{f/f}$; FRMD7tm mice, which lack a temporal-to-nasal wave directional bias, as described earlier (Fig. 3, F and G). Like our findings in chronic gabazine-treated mice, we observed a significant impact on direction selectivity, but not orientation selectivity, of neurons in the superior colliculus of FRMD7tm mice at eye opening (P14–P16) relative to littermate controls (Fig. 4, J to M, and fig. S11, D to F). In particular, there was a significant decrease overall in the direction selectivity of responding neurons in FRMD7tm mice relative to littermate controls (Fig. 4J) and a significant change in the distribution of gDSI across the eight directions, with significant reductions in the temporal quadrant of directions in retinal coordinates (Fig. 4K and fig. S10E). Like the results of chronic gabazine treatment, we also observed a significant reduction in the overall proportion of direction-selective neurons among all responding neurons in the superior colliculus of FRMD7tm mice (Fig. 4L). When divided according to direction preference, we also observed a significant change in the distribution of direction preference in FRMD7tm mice relative to control mice, with a significant reduction in the temporal quadrant (Fig. 4M and fig. S10F). Direction selectivity is reduced in temporal- and nasal-preferring direction-selective retinal ganglion cells in FRMD7tm mice (38), which may have contributed to the changes observed here in the superior colliculus. However, although there was an overall reduction in both the tuning and proportion of direction-selective cells in the superior colliculus of FRMD7tm mice, only cells preferring the temporal

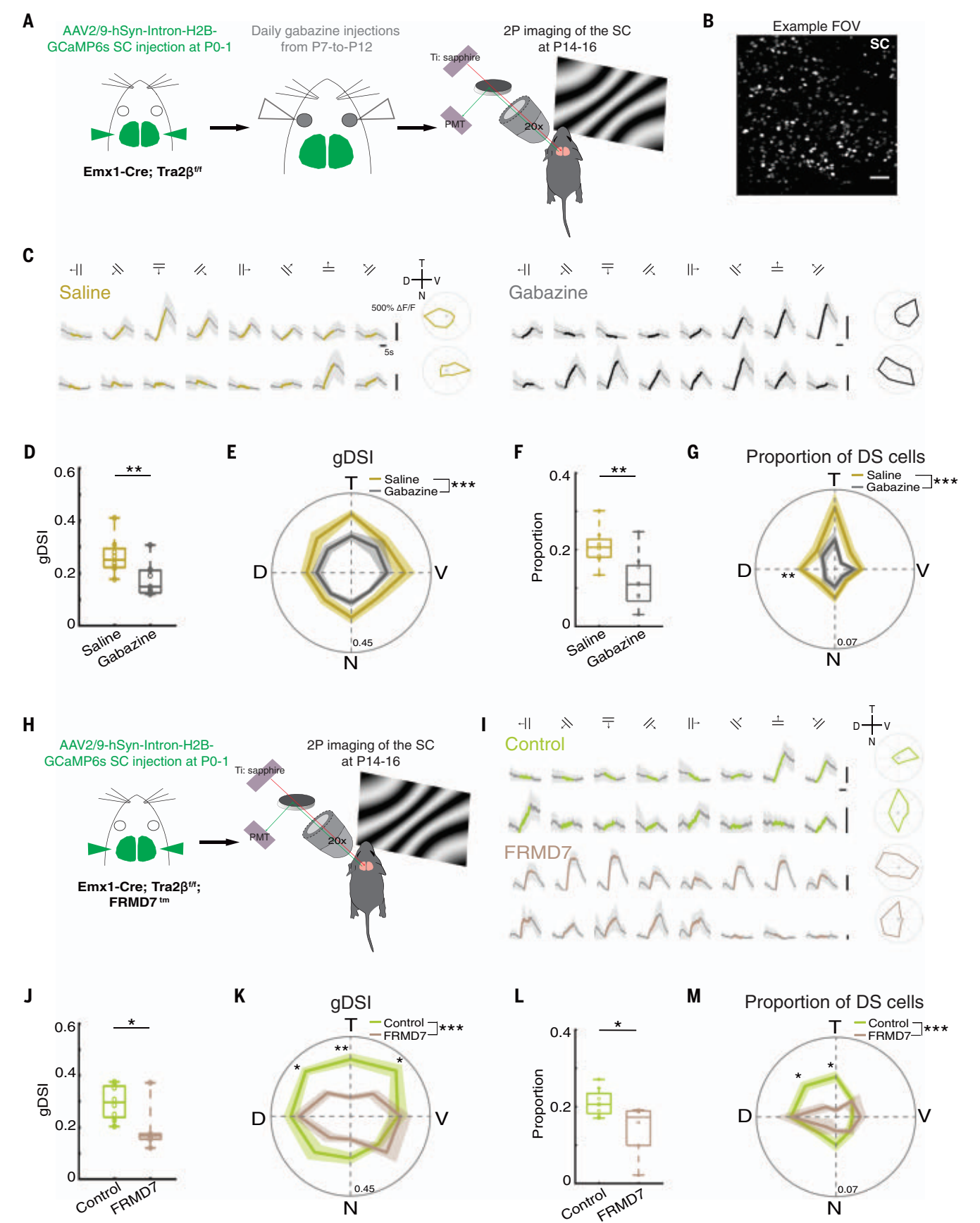


Fig. 4. Wave directionality instructs development of visual response properties.

(A) AAV2/9-Syn-Intron-H2B-GCaMP6s was injected into both hemispheres of the SC in *Emx1-Cre; Tra2 $\beta^{f/f}$* mice at P0–P1. Mouse pups then received daily (chronic) intraocular injections of gabazine (200 μ M) from P7 to P12. Two-photon imaging of SC neurons in the center of the left SC in response to visual stimuli presented to the right eye was performed in P14–P16 mice. **(B)** Example two-photon field of view in the SC \sim 50 μ m from the surface. Scale bar, 50 μ m. **(C)** Average fluorescence signal ($\Delta F/F \pm$ SD) of individual neurons in the left SC of an example saline-treated P15 mouse (left) and an example chronic gabazine-treated P15 mouse (right) in response to 20 repetitions each of drifting gratings in eight directions. For each trial, gratings were static for 8 s, drifting for 8 s, and then static for another 8 s before switching to another direction. Solid lines, averaged $\Delta F/F$ signal where colored part indicates the $\Delta F/F$ during drifting grating period; gray shadow, SD of $\Delta F/F$. Top, direction of drifting grating stimuli. The left most direction is nasal-to-temporal (backward-to-forward in visual space), the third from the left is ventral-to-dorsal (upward-to-downward in visual space). (Right) Polar plots of the averaged responses to each direction, with inset showing temporal (T), nasal (N), dorsal (D) and ventral (V) directions in retinal coordinates. **(D)** Median gDSI of all responding cells in chronic gabazine-treated mice ($n = 9$) was significantly lower than in saline-treated mice ($n = 8$) (saline, 0.265 ± 0.025 ; gabazine, 0.175 ± 0.021 ; $^{**}P = 0.008$, Wilcoxon rank sum test). **(E)** Polar plot showing the median gDSI of responding cells in saline- and gabazine-treated mice. Cells were divided into eight groups according to the directions for which they had the largest responses (see supplementary materials). The distribution of gDSI across directions was significantly altered in chronic gabazine-treated mice in comparison to saline-treated mice ($^{***}P = 2.724 \times 10^{-4}$, Mack-Skillings test; $P > 0.05$ for individual directions, Wilcoxon rank sum test with Benjamini-Hochberg false discovery rate adjustment). Solid line and shadow, mean \pm SEM of median gDSI across all animals. Polar plot r range, [0, 0.45]. **(F)** The proportion of direction-selective (DS) cells (gDSI > 0.25) among all responding cells was significantly reduced in chronic gabazine-treated mice relative to saline-treated mice (saline = 0.208 ± 0.017 , gabazine = $0.113 \pm$

0.023 ; $^{**}P = 0.008$, Wilcoxon rank sum test). **(G)** The distribution of direction preference of DS cells among all responding cells was significantly altered in chronic gabazine-treated mice relative to saline-treated mice ($^{***}P = 4.002 \times 10^{-6}$, Mack-Skillings test), with a significant reduction in the dorsal direction ($^{**}P = 0.005$, Wilcoxon rank sum test with Benjamini-Hochberg false discovery rate adjustment). Solid line and shadow, mean \pm SEM of proportions across all mice. Polar plot r range, [0, 0.07]. **(H)** *Emx1-Cre; Tra2 $\beta^{f/f}$; FRMD7 tm* and littermate control mice received AAV2/9-Syn-Intron-H2B-GCaMP6s injections into the SC at P0–P1 and then visual response properties of SC neurons were measured using two-photon imaging of the center of left SC at P14–P16. **(I)** Average fluorescence signals ($\Delta F/F \pm$ SD) of individual neurons in the left SC of an example littermate control P15 mouse (upper) and an example *Emx1-Cre; Tra2 $\beta^{f/f}$; FRMD7 tm* P15 mouse (lower) in response to 20 repetitions of drifting gratings in each of eight directions. **(J)** The median gDSI of all responding cells was significantly reduced in *FRMD7 tm* mice ($n = 6$ animals) relative to controls ($n = 8$ animals) (control, 0.182 ± 0.035 ; *FRMD7 tm* , 0.144 ± 0.021 ; $^{*}P = 0.029$, Wilcoxon rank sum test). **(K)** The distribution of gDSI across directions was significantly altered in *FRMD7 tm* mice relative to control mice ($^{***}P = 2.361 \times 10^{-13}$, Mack-Skillings test), with a significant reduction in the ventral-temporal ($^{*}P = 0.021$), temporal ($^{**}P = 0.005$), and dorsal-temporal ($^{*}P = 0.011$) directions (Wilcoxon rank sum test with Benjamini-Hochberg false discovery rate adjustment). Solid line and shadow, mean \pm SEM of median gDSI across all animals. Polar plot r range, [0, 0.45]. **(L)** The proportion of DS cells (gDSI > 0.25) among all responding cells was significantly reduced in *FRMD7 tm* mice relative to control mice (control = 0.211 ± 0.012 , *FRMD7 tm* = 0.141 ± 0.028 ; $^{*}P = 0.029$, Wilcoxon rank sum test). **(M)** The distribution of direction preference of DS cells among all responding cells was significantly altered in *FRMD7 tm* mice relative to control mice ($^{***}P = 2.569 \times 10^{-8}$, Mack-Skillings test), with a significant reduction in the temporal ($^{*}P = 0.011$) and dorsal-temporal ($^{*}P = 0.019$) directions (Wilcoxon rank sum test with Benjamini-Hochberg false discovery rate adjustment). Solid line and shadow, mean \pm SEM of proportions across all animals. Polar plot r range, [0, 0.07].

and not the nasal quadrant of directions were significantly disrupted. This finding suggests that both the retina and colliculus may contribute to the disrupted development of direction-selective responses in superior colliculus neurons of *FRMD7 tm* mice.

To further examine whether impaired direction selectivity in the superior colliculus in chronic gabazine-treated and *FRMD7 tm* mice was a product of abnormal direction selectivity derived from the retina or alterations emerging in the superior colliculus, we treated mice with a single (acute) application of intraocular gabazine at eye opening immediately before imaging (fig. S14A). Gabazine application to the retina is known to disrupt direction-selective responses in retinal ganglion cells (44). In contrast to the effects of chronic gabazine well before eye opening (Fig. 4 and fig. S10), superior colliculus neurons in mice that received an acute gabazine injection showed no significant difference in the strength of direction selectivity or the proportion of direction-selective neurons (fig. S14, B to F), whereas direction-selective responses in the nasal but not the temporal quadrant were modestly degraded in the superior colliculus relative to controls (fig. S14, B to H). The differential effects of acute and chronic gabazine treatment

on direction selectivity and the proportion of direction-selective cells suggests that disruptions in the development of circuits in the superior colliculus likely contribute to changes in the selectivity of superior colliculus neurons in chronic gabazine-treated mice and *FRMD7 tm* mice (Fig. 4 and fig. S10), although disruptions in the retina may also contribute. Overall, these results demonstrate that interfering with the directional bias of spontaneous retinal waves during development compromises the emergence of direction-selective responses in the superior colliculus at eye opening, highlighting the role of directional retinal waves in the emergence of functional response properties in mice.

Prior studies show that spontaneous retinal waves are necessary for the refinement of retinotopy in the superior colliculus (11, 12, 20, 21). To investigate whether the impairment in direction selectivity in the superior colliculus can be explained by altered retinotopy, we quantified the refinement of retinotopy in the superior colliculus in mice with disrupted wave directionality. No difference was observed between the sizes of retinal ganglion cell arbors in the superior colliculus in *FRMD7 tm* mice and littermate controls (fig. S15). Furthermore, retinotopy measured at eye opening was similar between chronic gabazine-treated and

chronic saline-treated mice (fig. S16). These data suggest that retinotopic refinement remained normal in mice with disrupted wave directionality, and exclude the possibility that retinotopic impairment contributes to the reduction of direction selectivity in the superior colliculus after the manipulation of wave directionality.

Discussion

Our data show that spontaneous retinal waves during development exhibit a distinct and strong propagation bias in the temporal-to-nasal direction during a transient window of development (P8 to P11) in mice. Directional waves mimic the translatory optic flow pattern that would be produced by an animal's own forward motion through space, a dominant optic flow pattern that is ethologically relevant after eye opening. Wave directionality during development is enforced by asymmetric inhibition through starburst amacrine cells mediated by GABA_ARs, which is the same circuit element necessary for the computation of direction selectivity by retinal ganglion cells in the adult (35, 37). Disrupting the directionality of spontaneous retinal waves compromises the emergence of direction-selective neurons in the superior colliculus at eye opening, causing

a general degradation in direction-selective responses and linking visual motion detection to retinal wave patterns during development.

The results of in vitro experiments suggest that starburst amacrine cells are responsible for the generation of stage 2 (P1–P9) retinal waves via their recurrent cholinergic neurotransmission (22, 32–34, 46). Here, we showed that starburst amacrine cells also enforce biased propagation of waves, even at P10–P11 when stage 3 waves are dependent on glutamatergic, not cholinergic, neurotransmission. Specifically, we found that asymmetric inhibitory outputs from starburst amacrine cells, which confer motion direction selectivity in direction-selective ganglion cells after the onset of vision (35–37), also mediate the temporal-to-nasal bias in wave directionality. This reveals an evolutionarily conserved link between the retinal circuitry responsible for biased wave propagation during development and the motion detection circuitry in mature retinas. There are also some differences between the developmental starburst amacrine cell-dependent wave propagation circuitry and the adult motion detection circuitry. Retinal waves in JamB-positive and Cdh6-positive ganglion cells, which prefer dorsal or ventral directions of motion in the adult retina, exhibit the same temporal-to-nasal wave propagation bias as the overall population of retinal ganglion cells during development. Because JamB-positive ganglion cells receive little inhibition from starburst amacrine cells (43), and because Cdh6-positive ganglion cells receive asymmetric inhibition from starburst amacrine cells along the dorsal-ventral axis (42), these results suggest that wave directionality is not generated by asymmetric inhibition from starburst amacrine cells onto direction-selective ganglion cells. Instead, direct or indirect asymmetric inhibition from starburst amacrine cells onto cells that are required for wave propagation—that is, starburst amacrine cells during stage 2 and bipolar cells during stage 3—mediates wave propagation bias.

Notably, the spatial geometry of the directional wave flow field resembles translatory optic flow from forward motion, which is the same optic flow pattern preferred by nasal-preferring direction-selective ganglion cells in the adult retina (27). This suggests that the asymmetric dendrites of starburst amacrine cells along the temporal-nasal axis that provide inhibition to nasal-preferring direction-selective ganglion cells—perhaps because they develop precociously or are more numerous (27)—may contribute the most to directional wave propagation. Further experiments are needed to examine this hypothesis.

Chronic blockade of GABA_A transmission does not alter the development of direction selectivity in the retina, at least for a subset of direction-selective retinal ganglion cells for

which this has been tested (44, 47). In our experiments, we observed that blocking retinal direction selectivity with gabazine injections at eye opening, but leaving the directional waves intact earlier during development, resulted in little disruption in the direction selectivity of superior colliculus neurons. Moreover, FRMD7tm mice that lack a temporal-to-nasal wave directional bias exhibited a general disruption in the strength and proportion of direction-selective neurons in the superior colliculus, with a more specific effect on direction-selective cells preferring the temporal quadrant of directions. These results suggest that impaired direction selectivity of neurons in the superior colliculus after the disruption of wave directionality is not fully explained by disruptions in retinal ganglion cell directional selectivity, and likely involves miswiring between the retina and the superior colliculus or within the superior colliculus. Our data also suggest that direction selectivity in superior colliculus neurons at eye opening, unlike the adult (48), is not wholly dependent on direction-selective inputs from the retina. Further experiments are needed to determine the origin and mechanism underlying the computation of direction selectivity in the developing superior colliculus.

We used Emx1-Cre; Tra2 $\beta^{f/f}$ (cortexless) mice because they have a fully accessible superior colliculus due to cortical agenesis. Although the absence of cortical feedback to the superior colliculus may compromise some visual functions (26, 49), retinal waves in cortexless mice were similar to those in mice with an intact cortex, which suggests that waves were not altered in the absence of cortex during development. Previous studies also show that Emx1-Cre; Tra2 $\beta^{f/f}$ mice have normal direction selectivity, orientation selectivity, and ON-OFF responses in the superior colliculus (26, 49), indicating that the impairment of direction selectivity we observed in Emx1-Cre; Tra2 $\beta^{f/f}$ mice after the disruption of wave directionality is not attributable to cortical agenesis.

Our experiments demonstrate that spontaneous retinal activity is self-structured to convey information essential for the development of visual response properties before the onset of visual experience. By simulating ethologically relevant features of the external environment, spontaneous retinal activity specifically and causally enhances higher-order visual function at eye opening. This is analogous to amphibian larvae, in which visual optic flow as the animal moves forward through space contributes to the development of retinotopy (25), as well as preplay in the developing mammalian hippocampus, where preconfigured sequences of neuronal firing foreshadow the development of hippocampal place cell memory function (50). Future experiments are necessary to determine the detailed mechanism by which wave directionality instructs the development

of direction selectivity of neurons in the superior colliculus. Spontaneous retinal waves are transmitted through the entire developing visual system (9, 10). Whether wave directionality contributes to the functional development of other downstream visual areas remains to be established.

Methods summary

All animal procedures followed the Yale Institutional Animal Care and Use Committee (IACUC), the US Department of Health and Human Services, and institution guidelines. Mouse pups received intraocular or intracollicular injections of virus encoding GCaMP6s at P0–P1 and were prepared with a craniotomy over the superior colliculus at later ages for wide-field calcium imaging of retinal ganglion cell axons in the superior colliculus or two-photon calcium imaging of superior colliculus neurons, respectively. Light stimulation or optogenetic stimulation of the retina was applied to initiate retinal waves in vivo. Retinal waves after acute intraocular injections of gabazine, TPMPA, or strychnine in mice at P10 were measured and analyzed to assess the role of inhibition in wave directionality. To ablate starburst amacrine cells, intraocular injections of diphtheria toxin (0.8 ng/ μ l in PBS) or PBS in littermate controls were performed daily from P5 to P7 in ChAT-Cre; iDTR mice. Daily gabazine (200 μ M) injections (or saline injections in control mice) were performed from P7 to P12 to chronically disrupt wave directionality during development. Procedures for the experiments and data analyses are described in the supplementary materials.

REFERENCES AND NOTES

1. A. D. Huberman, M. B. Feller, B. Chapman, Mechanisms underlying development of visual maps and receptive fields. *Annu. Rev. Neurosci.* **31**, 479–509 (2008). doi: [10.1146/annurev.neuro.31.060407.125533](https://doi.org/10.1146/annurev.neuro.31.060407.125533); pmid: 18558864
2. T. A. Seabrook, T. J. Burbridge, M. C. Crair, A. D. Huberman, Architecture, Function, and Assembly of the Mouse Visual System. *Annu. Rev. Neurosci.* **40**, 499–538 (2017). doi: [10.1146/annurev-neuro-071714-033842](https://doi.org/10.1146/annurev-neuro-071714-033842); pmid: 28772103
3. L. E. White, D. Fitzpatrick, Vision and cortical map development. *Neuron* **56**, 327–338 (2007). doi: [10.1016/j.neuron.2007.10.011](https://doi.org/10.1016/j.neuron.2007.10.011); pmid: 17964249
4. M. C. Crair, D. C. Gillespie, M. P. Stryker, The role of visual experience in the development of columns in cat visual cortex. *Science* **279**, 566–570 (1998). doi: [10.1126/science.279.5350.566](https://doi.org/10.1126/science.279.5350.566); pmid: 9438851
5. N. L. Rochefort et al., Development of direction selectivity in mouse cortical neurons. *Neuron* **71**, 425–432 (2011). doi: [10.1016/j.neuron.2011.06.013](https://doi.org/10.1016/j.neuron.2011.06.013); pmid: 21835340
6. J. C. Horton, D. R. Hocking, An adult-like pattern of ocular dominance columns in striate cortex of newborn monkeys prior to visual experience. *J. Neurosci.* **16**, 1791–1807 (1996). doi: [10.1523/JNEUROSCI.16-05-01791.1996](https://doi.org/10.1523/JNEUROSCI.16-05-01791.1996); pmid: 8774447
7. L. M. Chalupa, R. W. Rhoades, Directional selectivity in hamster superior colliculus is modified by strobe-rearing but not by dark-rearing. *Science* **199**, 998–1001 (1978). doi: [10.1126/science.622583](https://doi.org/10.1126/science.622583); pmid: 622583
8. M. Weliky, L. C. Katz, Correlational structure of spontaneous neuronal activity in the developing lateral geniculate nucleus in vivo. *Science* **285**, 599–604 (1999). doi: [10.1126/science.285.5427.599](https://doi.org/10.1126/science.285.5427.599); pmid: 10417392
9. J. B. Ackman, T. J. Burbridge, M. C. Crair, Retinal waves coordinate patterned activity throughout the developing visual

- system. *Nature* **490**, 219–225 (2012). doi: [10.1038/nature11529](https://doi.org/10.1038/nature11529); pmid: 23060192
10. A. Gribizis *et al.*, Visual Cortex Gains Independence from Peripheral Drive before Eye Opening. *Neuron* **104**, 711–723.e3 (2019). doi: [10.1016/j.neuron.2019.08.015](https://doi.org/10.1016/j.neuron.2019.08.015); pmid: 31561919
 11. J. Cang, D. A. Feldheim, Developmental mechanisms of topographic map formation and alignment. *Annu. Rev. Neurosci.* **36**, 51–77 (2013). doi: [10.1146/annurev-neuro-062012-170341](https://doi.org/10.1146/annurev-neuro-062012-170341); pmid: 23642132
 12. T. J. Burbridge *et al.*, Visual circuit development requires patterned activity mediated by retinal acetylcholine receptors. *Neuron* **84**, 1049–1064 (2014). doi: [10.1016/j.neuron.2014.10.051](https://doi.org/10.1016/j.neuron.2014.10.051); pmid: 25466916
 13. H. P. Xu *et al.*, Spatial pattern of spontaneous retinal waves instructs retinotopic map refinement more than activity frequency. *Dev. Neurobiol.* **75**, 621–640 (2015). doi: [10.1002/dneu.22288](https://doi.org/10.1002/dneu.22288); pmid: 25787992
 14. D. A. Butts, P. O. Kanold, C. J. Shatz, A burst-based “Hebbian” learning rule at retinogeniculate synapses links retinal waves to activity-dependent refinement. *PLoS Biol.* **5**, e61 (2007). doi: [10.1371/journal.pbio.0050061](https://doi.org/10.1371/journal.pbio.0050061); pmid: 17341130
 15. A. D. Huberman, C. M. Speer, B. Chapman, Spontaneous retinal activity mediates development of ocular dominance columns and binocular receptive fields in v1. *Neuron* **52**, 247–254 (2006). doi: [10.1016/j.neuron.2006.07.028](https://doi.org/10.1016/j.neuron.2006.07.028); pmid: 17046688
 16. C. L. Torborg, K. A. Hansen, M. B. Feller, High frequency, synchronized bursting drives eye-specific segregation of retinogeniculate projections. *Nat. Neurosci.* **8**, 72–78 (2005). doi: [10.1038/nn1376](https://doi.org/10.1038/nn1376); pmid: 15608630
 17. C. Pfeifenberger, J. Yamada, D. A. Feldheim, Ephrin-As and patterned retinal activity act together in the development of topographic maps in the primary visual system. *J. Neurosci.* **26**, 12873–12884 (2006). doi: [10.1523/JNEUROSCI.3595-06.2006](https://doi.org/10.1523/JNEUROSCI.3595-06.2006); pmid: 17167078
 18. B. K. Stafford, A. Sher, A. M. Litke, D. A. Feldheim, Spatial-temporal patterns of retinal waves underlying activity-dependent refinement of retinofugal projections. *Neuron* **64**, 200–212 (2009). doi: [10.1016/j.neuron.2009.09.021](https://doi.org/10.1016/j.neuron.2009.09.021); pmid: 19874788
 19. B. Chapman, Necessity for afferent activity to maintain eye-specific segregation in ferret lateral geniculate nucleus. *Science* **287**, 2479–2482 (2000). doi: [10.1126/science.287.5462.2479](https://doi.org/10.1126/science.287.5462.2479); pmid: 10741966
 20. J. Cang, L. Wang, M. P. Stryker, D. A. Feldheim, Roles of ephrin-as and structured activity in the development of functional maps in the superior colliculus. *J. Neurosci.* **28**, 11015–11023 (2008). doi: [10.1523/JNEUROSCI.2478-08.2008](https://doi.org/10.1523/JNEUROSCI.2478-08.2008); pmid: 18945909
 21. T. D. Mrsic-Flogel *et al.*, Altered map of visual space in the superior colliculus of mice lacking early retinal waves. *J. Neurosci.* **25**, 6921–6928 (2005). doi: [10.1523/JNEUROSCI.1555-05.2005](https://doi.org/10.1523/JNEUROSCI.1555-05.2005); pmid: 16033902
 22. A. G. Blankenship, M. B. Feller, Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nat. Rev. Neurosci.* **11**, 18–29 (2010). doi: [10.1038/nrn2759](https://doi.org/10.1038/nrn2759); pmid: 19953103
 23. A. Akrouh, D. Kerschensztein, Intersecting circuits generate precisely patterned retinal waves. *Neuron* **79**, 322–334 (2013). doi: [10.1016/j.neuron.2013.05.012](https://doi.org/10.1016/j.neuron.2013.05.012); pmid: 23830830
 24. A. Maccione *et al.*, Following the ontogeny of retinal waves: Pan-retinal recordings of population dynamics in the neonatal mouse. *J. Physiol.* **592**, 1545–1563 (2014). doi: [10.1113/jphysiol.2013.262840](https://doi.org/10.1113/jphysiol.2013.262840); pmid: 24366261
 25. M. Hiramoto, H. T. Cline, Optic flow instructs retinotopic map formation through a spatial to temporal to spatial transformation of visual information. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E5105–E5113 (2014). doi: [10.1073/pnas.1416953111](https://doi.org/10.1073/pnas.1416953111); pmid: 25385606
 26. J. A. Shanks *et al.*, Corticothalamic Axons Are Essential for Retinal Ganglion Cell Axon Targeting to the Mouse Dorsal Lateral Geniculate Nucleus. *J. Neurosci.* **36**, 5252–5263 (2016). doi: [10.1523/JNEUROSCI.4599-15.2016](https://doi.org/10.1523/JNEUROSCI.4599-15.2016); pmid: 27170123
 27. S. Sabbah *et al.*, A retinal code for motion along the gravitational and body axes. *Nature* **546**, 492–497 (2017). doi: [10.1038/nature22818](https://doi.org/10.1038/nature22818); pmid: 28607486
 28. M. Chen, S. Weng, Q. Deng, Z. Xu, S. He, Physiological properties of direction-selective ganglion cells in early postnatal and adult mouse retina. *J. Physiol.* **587**, 819–828 (2009). doi: [10.1113/jphysiol.2008.161240](https://doi.org/10.1113/jphysiol.2008.161240); pmid: 19103682
 29. N. Tian, D. R. Copenhagen, Visual stimulation is required for refinement of ON and OFF pathways in postnatal retina. *Neuron* **39**, 85–96 (2003). doi: [10.1016/S0896-6273\(03\)00389-1](https://doi.org/10.1016/S0896-6273(03)00389-1); pmid: 12848934
 30. J. M. Rosa, R. D. Morrie, H. C. Baertsch, M. B. Feller, Contributions of Rod and Cone Pathways to Retinal Direction Selectivity Through Development. *J. Neurosci.* **36**, 9683–9695 (2016). doi: [10.1523/JNEUROSCI.3824-15.2016](https://doi.org/10.1523/JNEUROSCI.3824-15.2016); pmid: 27629718
 31. A. Tiriac, B. E. Smith, M. B. Feller, Light Prior to Eye Opening Promotes Retinal Waves and Eye-Specific Segregation. *Neuron* **100**, 1059–1065.e4 (2018). doi: [10.1016/j.neuron.2018.10.011](https://doi.org/10.1016/j.neuron.2018.10.011); pmid: 30392793
 32. J. Zheng, S. Lee, Z. J. Zhou, A transient network of intrinsically bursting starburst cells underlies the generation of retinal waves. *Nat. Neurosci.* **9**, 363–371 (2006). doi: [10.1038/nn1644](https://doi.org/10.1038/nn1644); pmid: 16462736
 33. K. J. Ford, A. L. Félix, M. B. Feller, Cellular mechanisms underlying spatiotemporal features of cholinergic retinal waves. *J. Neurosci.* **32**, 850–863 (2012). doi: [10.1523/JNEUROSCI.5309-12.2012](https://doi.org/10.1523/JNEUROSCI.5309-12.2012); pmid: 22262883
 34. H. P. Xu *et al.*, Retinal Wave Patterns Are Governed by Mutual Excitation among Starburst Amacrine Cells and Drive the Refinement and Maintenance of Visual Circuits. *J. Neurosci.* **36**, 3871–3886 (2016). doi: [10.1523/JNEUROSCI.3549-15.2016](https://doi.org/10.1523/JNEUROSCI.3549-15.2016); pmid: 27030771
 35. S. I. Fried, T. A. Münch, F. S. Werblin, Mechanisms and circuitry underlying directional selectivity in the retina. *Nature* **420**, 411–414 (2002). doi: [10.1038/nature01179](https://doi.org/10.1038/nature01179); pmid: 12459782
 36. A. S. Mauss, A. Vlasits, A. Borst, M. Feller, Visual Circuits for Direction Selectivity. *Annu. Rev. Neurosci.* **40**, 211–230 (2017). doi: [10.1146/annurev-neuro-072116-031335](https://doi.org/10.1146/annurev-neuro-072116-031335); pmid: 28418757
 37. W. R. Taylor, R. G. Smith, The role of starburst amacrine cells in visual signal processing. *Vis. Neurosci.* **29**, 73–81 (2012). doi: [10.1017/S0952523811000393](https://doi.org/10.1017/S0952523811000393); pmid: 22310373
 38. K. Yonehara *et al.*, Congenital Nystagmus Gene FRMD17 Is Necessary for Establishing a Neuronal Circuit Asymmetry for Direction Selectivity. *Neuron* **89**, 177–193 (2016). doi: [10.1016/j.neuron.2015.11.032](https://doi.org/10.1016/j.neuron.2015.11.032); pmid: 26711119
 39. D. Hillier *et al.*, Causal evidence for retina-dependent and -independent visual motion computations in mouse cortex. *Nat. Neurosci.* **20**, 960–968 (2017). doi: [10.1038/nn.4566](https://doi.org/10.1038/nn.4566); pmid: 28530661
 40. J. R. Sanes, R. H. Masland, The types of retinal ganglion cells: Current status and implications for neuronal classification. *Annu. Rev. Neurosci.* **38**, 221–246 (2015). doi: [10.1146/annurev-neuro-071714-034120](https://doi.org/10.1146/annurev-neuro-071714-034120); pmid: 25897874
 41. A. Matsumoto *et al.*, Synapse-specific direction selectivity in retinal bipolar cell axon terminals. *bioRxiv* <https://doi.org/10.1101/2020.10.12.335810v1> [preprint]. 12 October 2020.
 42. J. N. Kay *et al.*, Retinal ganglion cells with distinct directional preferences differ in molecular identity, structure, and central projections. *J. Neurosci.* **31**, 7753–7762 (2011). doi: [10.1523/JNEUROSCI.0907-11.2011](https://doi.org/10.1523/JNEUROSCI.0907-11.2011); pmid: 21613488
 43. I. J. Kim, Y. Zhang, M. Yamagata, M. Meister, J. R. Sanes, Molecular identification of a retinal cell type that responds to upward motion. *Nature* **452**, 478–482 (2008). doi: [10.1038/nature06739](https://doi.org/10.1038/nature06739); pmid: 18368118
 44. W. Wei, A. M. Hamby, K. Zhou, M. B. Feller, Development of asymmetric inhibition underlying direction selectivity in the retina. *Nature* **469**, 402–406 (2011). doi: [10.1038/nature09600](https://doi.org/10.1038/nature09600); pmid: 21131947
 45. K. Yonehara *et al.*, Spatially asymmetric reorganization of inhibition establishes a motion-sensitive circuit. *Nature* **469**, 407–410 (2011). doi: [10.1038/nature09711](https://doi.org/10.1038/nature09711); pmid: 21170022
 46. M. B. Feller, D. P. Wellis, D. Stellwagen, F. S. Werblin, C. J. Shatz, Requirement for cholinergic synaptic transmission in the propagation of spontaneous retinal waves. *Science* **272**, 1182–1187 (1996). doi: [10.1126/science.272.5265.1182](https://doi.org/10.1126/science.272.5265.1182); pmid: 8638165
 47. L. Sun, X. Han, S. He, Direction-selective circuitry in rat retina develops independently of GABAergic, cholinergic and action potential activity. *PLoS ONE* **6**, e19477 (2011). doi: [10.1371/journal.pone.0019477](https://doi.org/10.1371/journal.pone.0019477); pmid: 21573361
 48. X. Shi *et al.*, Retinal origin of direction selectivity in the superior colliculus. *Nat. Neurosci.* **20**, 550–558 (2017). doi: [10.1038/nn.4498](https://doi.org/10.1038/nn.4498); pmid: 28192394
 49. K. H. Lee, A. Tran, Z. Turan, M. Meister, The sifting of visual information in the superior colliculus. *eLife* **9**, e50678 (2020). doi: [10.7554/eLife.50678](https://doi.org/10.7554/eLife.50678); pmid: 32286224
 50. U. Farooq, G. Dragoi, Emergence of preconfigured and plastic time-compressed sequences in early postnatal development. *Science* **363**, 168–173 (2019). doi: [10.1126/science.aav0502](https://doi.org/10.1126/science.aav0502); pmid: 30630930

ACKNOWLEDGMENTS

We thank D. Clark, J. Demb, D. Lee, X. Papademetris, J. Zhou, and all members of the Crair lab for their helpful comments on this project; I.-J. Kim, X. Duan, and B. Roska for providing mouse lines; the Yale Vision Core for viral constructs; and especially Y. Zhang for help with mouse husbandry and genotyping. **Funding:** Supported by NIH grants R01EY015788, U01NS094358, P30EY026878, R01MH111424 (M.C.C.). M.C.C. also thanks the family of William Ziegler III for their support. **Author contributions:** X.G. and M.C.C. conceived and designed the study. X.G. performed and analyzed wide-field calcium imaging experiments. X.G. and K.Z. performed and analyzed two-photon calcium imaging experiments. A.G. helped with pharmacology experiments. A.S.H. helped with two-photon imaging experiments and A.M.S. helped with immunostaining experiments. X.G., K.Z., A.M.S., and M.C.C. analyzed and interpreted the results. X.G. and M.C.C. wrote the manuscript with input from all authors. **Competing interests:** Authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/eabd0830/suppl/DC1
Materials and Methods
Figs. S1 to S16
Table S1
Movies S1 to S6
References (51–58)

28 July 2020; accepted 26 May 2021
10.1126/science.abd0830

RESEARCH ARTICLES

STRUCTURAL BIOLOGY

Structure of an AMPK complex in an inactive, ATP-bound state

Yan Yan¹, Somnath Mukherjee^{2†}, Kaleeckal G. Harikumar^{3†}, Timothy S. Strutzenberg^{4†}, X. Edward Zhou^{1†}, Kelly Suino-Powell¹, Ting-Hai Xu^{1,5}, Ryan D. Sheldon⁶, Jared Lamp⁷, Joseph S. Brunzelle⁸, Katarzyna Radziwon², Abigail Ellis⁶, Scott J. Novick⁴, Irving E. Vega⁷, Russell G. Jones⁶, Laurence J. Miller³, H. Eric Xu⁹, Patrick R. Griffin⁴, Anthony A. Kossiakoff^{2,10}, Karsten Melcher^{1*}

Adenosine monophosphate (AMP)-activated protein kinase (AMPK) regulates metabolism in response to the cellular energy states. Under energy stress, AMP stabilizes the active AMPK conformation, in which the kinase activation loop (AL) is protected from protein phosphatases, thus keeping the AL in its active, phosphorylated state. At low AMP:ATP (adenosine triphosphate) ratios, ATP inhibits AMPK by increasing AL dynamics and accessibility. We developed conformation-specific antibodies to trap ATP-bound AMPK in a fully inactive, dynamic state and determined its structure at 3.5-angstrom resolution using cryo-electron microscopy. A 180° rotation and 100-angstrom displacement of the kinase domain fully exposes the AL. On the basis of the structure and supporting biophysical data, we propose a multistep mechanism explaining how adenine nucleotides and pharmacological agonists modulate AMPK activity by altering AL phosphorylation and accessibility.

Adenosine monophosphate-activated protein kinase (AMPK) is the primary energy sensor and regulator of energy homeostasis in eukaryotes (1–5). Upon activation by energy stress, AMPK reprograms metabolism by phosphorylating and modulating the activities of metabolic enzymes and key regulators of metabolism, growth, and proliferation (6–9). Deregulation of AMPK is associated with metabolic diseases, and AMPK is a pharmacological target for the treatment of diabetes, obesity, cancer, and cardiometabolic disease (10–12).

AMPK is a heterotrimeric protein kinase that is assembled from an α subunit ($\alpha 1$ or $\alpha 2$) containing the kinase domain (KD) and regulatory β subunits ($\beta 1$ or $\beta 2$) and γ subunits ($\gamma 1$, $\gamma 2$, or $\gamma 3$) (13) (Fig. 1A). AMPK senses the

cellular energy state by competitive binding of adenosine monophosphate (AMP), adenosine diphosphate (ADP), and adenosine triphosphate (ATP) to three sites in its γ subunit, of which site 3 (CBS3) is the primary sensor (14, 15) (Fig. 1A). In addition, pharmacological compounds including Abbot A769662 and Merck 991 bind to a distant site called the allosteric drug and metabolite (ADaM) site, which is at the interface between the KD and the carbohydrate-binding module (CBM) of the β subunit (16, 17) (Fig. 1A). AMPK is activated about 100-fold by phosphorylation of a residue (Thr¹⁷⁴ in human $\alpha 1$ and Thr¹⁷² in human $\alpha 2$) in its kinase activation loop (AL) and up to 10-fold by direct allosteric kinase activation (18–22). In reconstituted systems, AMP, ADP, and ADaM agonists all increase net AL phosphorylation by stabilizing a conformation that limits phosphatase access to the AL (14, 19, 20, 23, 24), although AL protection by ADaM agonists has been less clear in a cellular context (23, 25). In addition, binding of AMP, but not ADP or ADaM ligands, further increases AL phosphorylation by stabilizing an interaction with AMPK's main upstream kinase, the tumor suppressor LKB1 (22, 26–28), and possibly with other upstream kinases (27). Under energy excess, ATP competes with AMP and ADP at one or more of the allosteric sites and inhibits AMPK.

Modulation of AL accessibility by adenine nucleotides was identified as a major mechanism of AMPK regulation more than 20 years ago (19), but how adenine nucleotides and ADaM agonists induce conformational changes that modulate AL accessibility is unknown because the structure of the AMPK complex in

its highly dynamic, ATP-bound state has yet to be resolved. Equally important, how agonists can restrict access to protein phosphatases without impeding access to AL kinases has remained enigmatic.

Conformational trapping of inactive-state AMPK by protein engineering and conformation-selective Fabs

AMP binding stabilizes, and ATP binding destabilizes, an interaction between the γ subunit and the α linker that is required for AMPK activation (29). To determine the structure of ATP-bound AMPK in the phosphorylated, more-AL-accessible state, we stabilized the phosphorylated AMPK (pAMPK) construct shown in Fig. 1A through (i) adenosine-5'-(γ -thio)-triphosphate (ATP γ S), which we found to more potently inhibit the interaction between the γ subunit and the α linker than ATP (fig. S1A); (ii) a nonflexibly linked maltose-binding protein (MBP) tag; (iii) the AMPK inhibitor compound C (CpdC) (30); and (iv) an E199A/E200A surface entropy reduction mutation in the β subunit (see materials and methods in the supplementary materials). These modifications did not change AMPK modulation by AMP and ATP (fig. S1, B to D) and allowed us to solve its crystal structure at a modest resolution of 5.4 Å (fig. S1E and table S1). Although the resolution was insufficient to gain detailed mechanistic insight, it showed that the KD was associated with the β subunit as in active AMPK and that all three binding-competent CBS sites were bound by adenine nucleotides. A difference density map with the phosphate group omitted from the model suggested that AL residue T174 was phosphorylated (fig. S1E).

The resolution could not be improved through further optimization, indicating the need for more potent trapping of the dynamic complex in a distinct, inhibitory conformation. To do so, we generated Fabs specific for the inactive conformation of the AMPK crystallization construct (Fig. 1B and fig. S2). We also included a Fab-binding nanobody (31) to eliminate flexibility between the Fab's constant and variable domains. Together, these measures allowed us to determine a cryo-electron microscopy (cryo-EM) structure at 3.48-Å nominal resolution from 286,895 particles used in the three-dimensional reconstruction (Fig. 1, C to E; fig. S3; and table S2). Because CpdC is a nonphysiological AMPK inhibitor, we also determined the structure in its absence at a nominal resolution of 3.92 Å (fig. S4 and table S2). The two cryo-EM structures were very similar [root mean square deviation (RMSD) of 1.48 Å; fig. S6], indicating that CpdC did not change the overall domain organization or complex conformation of the inhibited AMPK.

Structures of Fab-stabilized, ATP-bound AMPK

The cryo-EM structures allowed an unambiguous assignment of the KD, core domain, and

¹Department of Structural Biology, Van Andel Institute, Grand Rapids, MI 49503, USA. ²Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA. ³Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Scottsdale, AZ 85259, USA. ⁴Department of Molecular Medicine, The Scripps Research Institute, Scripps Florida, Jupiter, FL 33458, USA. ⁵Center for Epigenetics, Van Andel Institute, Grand Rapids, MI 49503, USA. ⁶Metabolic and Nutritional Programming, Center for Cancer and Cell Biology, Van Andel Institute, Grand Rapids, MI 49503, USA. ⁷Integrated Mass Spectrometry Unit, Department of Translational Neuroscience, Michigan State University College of Human Medicine, Grand Rapids Research Center, Grand Rapids, MI 49503, USA. ⁸Life Sciences Collaborative Access Team, Northwestern University Synchrotron Research Center, Northwestern University, Argonne, IL 60439, USA. ⁹Center for Structure and Function of Drug Targets, The CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences (CAS), Shanghai 201203, China. ¹⁰Institute of Biophysical Dynamics, University of Chicago, Chicago, IL 60637, USA.

*Corresponding author. Email: karsten.melcher@vai.org

†These authors contributed equally to this work.

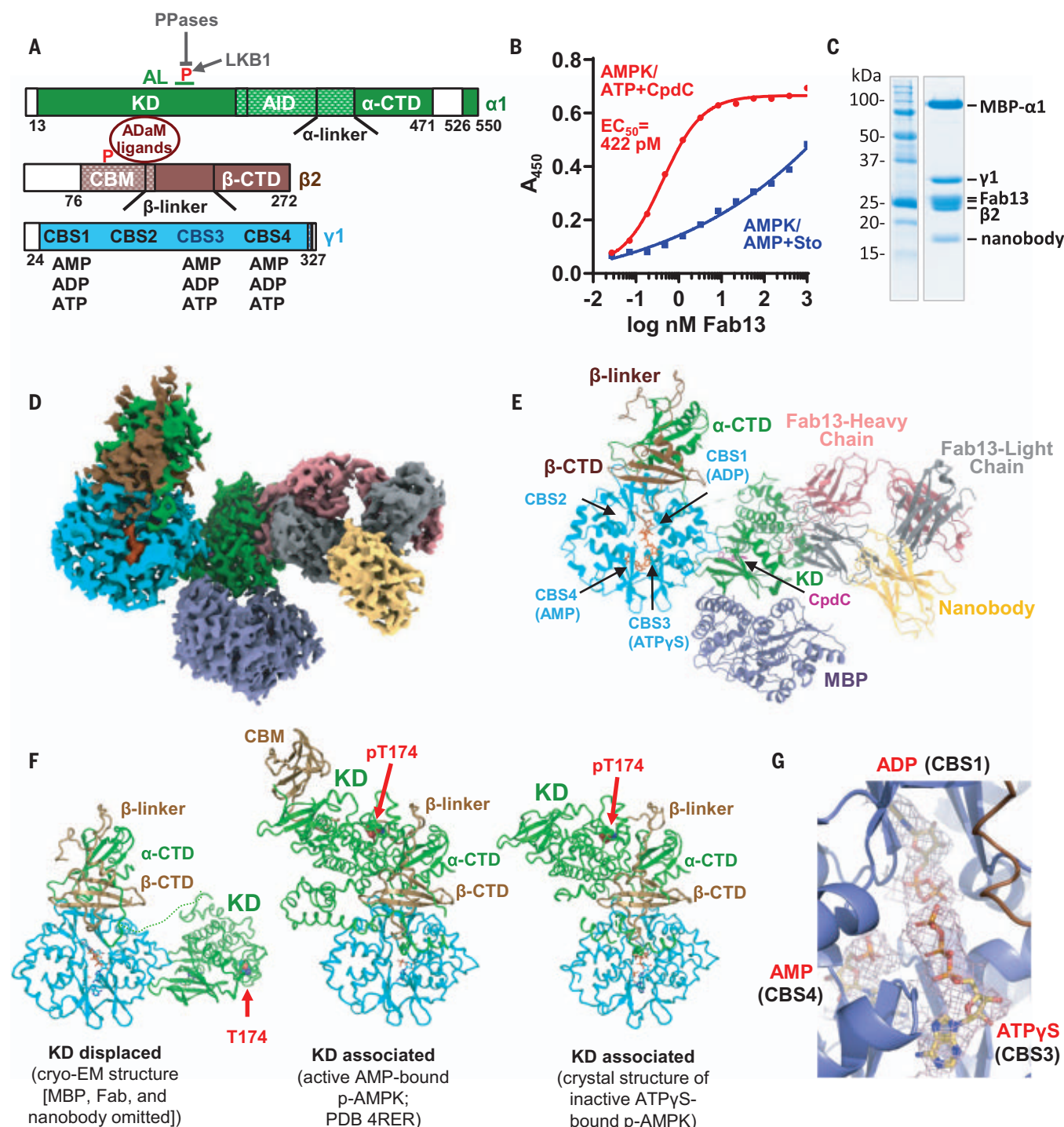


Fig. 1. Overall structure of the ATP γ S-bound AMPK/CpdC/Fab/nanobody complex. (A) Diagram of the AMPK construct and its domain architecture, using the same color code as in (D) to (G). It lacks flexible termini and the ST loop (α 1[477–528]) (white). These loops impede structure determination, are unresolved in all AMPK structures, and are not required for AMPK structural integrity or regulation by adenine nucleotides or ADaM site ligands (14, 17, 29, 35). Segments with patterned shading are not resolved. ADaM, allosteric drug and metabolite site (brown oval); AID, autoinhibitory domain; AL, activation loop; CBM, carbohydrate-binding module; CBS1, CBS2, and CBS4, adenine nucleotide binding sites; CTD, C-terminal domain; KD, kinase domain; PPases, protein phosphatases; p, phosphorylation site in the AL (T174) and the CBM (S108). Single-letter abbreviations for the amino acid residues

are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (B) Fab13 binding to active (AMP+staurosporine) and inactive (ATP+CpdC) AMPK. A_{450} , absorbance at 450-nm wavelength. (C) SDS-polyacrylamide gel electrophoresis analysis of the purified AMPK $\alpha 1\beta 2\gamma 1$ complex bound to Fab and nanobody. (D) Cryo-EM density map. (E) Atomic model. (F) Side-by-side structural alignment of the inactive, Fab-stabilized AMPK cryo-EM structure (left) with the crystal structures of active pAMPK [Protein Data Bank (PDB) ID 4RER; middle] and Fab-free ATP γ S- and CpdC-bound pAMPK (right). MBP, Fab, and nanobody were omitted. (G) Nucleotide occupancy at CBS1, CBS3, and CBS4 overlaid with EM density (shown as mesh) contoured at 7.5 σ (PyMOL).

amino acid residues 180 through 203 of the β linker, as well as of the MBP tag, Fab, and nanobody in the EM density map (Fig. 1, D and E). We observed a conformation that is very different from both active-state pAMPK and the crystal structure of inactive ATP γ S/CpdC-bound pAMPK (Fig. 1F). Although the core complex remained largely superimposable with that of active-state AMPK (RMSD = 2.1 Å), the KD was completely dissociated from the β subunit by a rotation of $\sim 180^\circ$ and a translation of ~ 100 Å (Fig. 1, D to F, and movie S1). In this conformation, the KD associated weakly with the γ subunit (see fig. S7 for details and mutational analysis) in an orientation in which the AL became fully accessible to phosphatases and upstream kinases.

AL residues 169 through 177 showed relatively weak density compared with the core domains in both structures, but the lack of the characteristic strong phosphate density [as seen for AMP in the same structure (compare fig. S6 with Fig. 1G)] suggested that residue T174 was not phosphorylated. This conclusion was supported by the lack of countercharges around the buried T174 to balance a heavily negatively charged phosphate group. To confirm this interpretation, we also determined the structure of AMPK in complex with Fab and CpdC from a protein preparation that was not phosphorylated (figs. S5 and S8). As seen in fig. S8, the structure from the nonphosphorylated sample was superimposable with the other two cryo-EM structures, with the same interfaces between subunits, including an identical interface between KD and the γ subunit. Using mass spectrometry, we confirmed that T174 in AMPK used for Fab selection and structure determination was incompletely phosphorylated (5.3% nonphosphorylated; table S3). We expect that preclearing with phosphorylated active-state AMPK during Fab selection enriched for Fabs binding to the nonphosphorylated target.

The KD in the cryo-EM structures adopted an inhibitory conformation showing the hallmarks of inactive kinase domains (fig. S9). Moreover, ATP destabilizes the interaction between the α linker and CBS3 (29). Consistently, the α linker was disordered in the structures, and the position of the resolved flanking sequences of the α linker and preceding autoinhibitory domain (AID) suggests that the α linker was indeed displaced (fig. S10). Further, we saw no density for the CBM, consistent with double electron-electron resonance spectra indicating that the CBM can dissociate from the KD in the absence of ADaM ligands and adopt multiple alternative positions, especially in nonphosphorylated AMPK (32).

In this conformation, the CBM, AID, and α linker are unresolved and completely or partially dissociated, and the KD makes only limited interactions with the γ subunit [515-Å²

buried interface compared with buried interfaces of 1677 Å² with β subunit and α -CTD in the KD-associated conformation (4RER)]. This state is in equilibrium with the more stable active AMPK conformation. The Fab shifts this equilibrium by binding a surface in nonphosphorylated AMPK that largely overlaps the AL-containing surface in pAMPK that is protected by β linker and core AMPK (fig. S11). This mutually exclusive binding is the basis for the Fab's conformational selectivity and stabilization of the inactive conformation.

An important question in the AMPK field is which adenine nucleotides bind to which individual CBS sites. Unexpectedly, although AMP and ADP were not added for the ATP-complex formation, we observed AMP-occupied CBS4 and ADP-occupied CBS1 (Fig. 1G; see detailed discussion in the supplementary materials and fig. S12). Only CBS3, the nucleotide sensor site, was occupied by ATP γ S.

In contrast to the cryo-EM structures, in the crystal structure in the absence of Fab, the AL was phosphorylated, as expected, and the KD remained associated with the β subunit (fig. S1E). On the basis of these results, we hypothesized that allosteric ATP binding of phosphorylated AMPK may partially destabilize AL protection in a KD-associated conformation, whereas AL dephosphorylation may lead to KD displacement and full inhibition.

Dephosphorylation is a switch for KD displacement

We tested whether AL dephosphorylation indeed leads to KD displacement in the context of full-length AMPK in solution and in the absence of Fab and nanobody. First, we scrutinized our previously published hydrogen-deuterium exchange mass spectrometry (HDX-MS) data (33). We overlaid all resolved peptides (Fig. 2A, colored gray) having phosphorylation-induced reduction in deuterium exchange (i.e., that have increased interaction and/or stability in pAMPK) onto the structures of KD-associated and KD-displaced AMPK. These peptides were found in three regions: (i) the AL-containing surface of the KD that is protected in KD-associated conformation but is solvent-accessible in the KD-displaced conformation; (ii) the β -CTD surface that interacts with the AL in the KD-associated conformation; and (iii) the γ subunit surface that interacts with the α linker, which is dissociated from the γ subunit in KD-displaced AMPK. In contrast, no such changes occurred in purified pAMPK upon the shift from energy stress to energy excess nucleotide levels (fig. S13) (32). These data therefore support that the KD-displaced conformation is primarily adopted by fully inactive, dephosphorylated AMPK.

Next we probed the AMPK conformation using AlphaScreen luminescence proximity sensors (29). This assay measures the proxim-

ity between a hexahistidine (His6) tag introduced at the N terminus of the KD and a biotin tag introduced at the N terminus of the γ subunit (Fig. 2B). AMP, ATP, and ATP γ S had only small effects on tag distances in purified phosphorylated AMPK. In contrast, the luminescence signal was more than threefold stronger (i.e., the distance between the tags shorter) in nonphosphorylated than in phosphorylated AMPK (Fig. 2B). This is in agreement with the difference in distances seen in the structures of KD-associated versus KD-displaced AMPK and further supports dephosphorylation-induced KD displacement.

The assay also provided insight into AL rephosphorylation. AMP stimulates AL rephosphorylation by inducing an interaction of AMPK with LKB1-bound axin (22). How AMP can stimulate a conformation in which the AL is protected against dephosphorylation and yet is accessible to upstream kinases had been puzzling. Our data indicate that the AL is protected only in phosphorylated AMPK, the substrate for dephosphorylation; it is accessible in nonphosphorylated AMPK. As in our previous study (29), incubation of nonphosphorylated AMPK with AMP caused a strong and highly significant increase in the AlphaScreen luminescence proximity signal indicative of an AMP-induced movement of the N termini of the KD and the γ subunit toward each other (Fig. 2B). While the details of this induced conformational change are unclear, we speculate that they mediate the effects of AMP-induced binding of AMPK to LKB1-bound axin.

We further probed the protein structures of phosphorylated, AMP-bound AMPK (fully active) and ATP γ S-bound, nonphosphorylated AMPK (fully inactive) by cross-linking coupled with mass spectrometry (XL-MS), using the MS-cleavable cross-linker disuccinimidyl sulfoxide (DSSO) (34). The maximum C α -C α distance between any two stably positioned lysine residues cross-linked by DSSO is ~ 30 Å (35). Because of the distance constraints, cross-linked residues are often located in flexible loops, whose movements allow extended cross-linking distances.

We identified seven cross-links that (i) could be reliably identified (cross-linking scores > 50); (ii) were statistically significant ($P < 0.05$); (iii) could be mapped onto the structures of both pAMPK/AMP and non-pAMPK/ATP γ S (i.e., the lysine residues of both peptides were resolved or bordered resolved regions); and (iv) had C α -C α distances that differed in both structures (Fig. 2, C and D; see data S1 for full data). The two cross-links enriched in pAMPK/AMP were between residues that are ~ 19 Å ($\alpha 1$ -K71- $\alpha 1$ -K476) and 30.1 Å ($\alpha 1$ -K53- $\beta 2$ -K203) apart in the structure of pAMPK/ (AMP+staurosporine) but that would be ~ 65 and 87.4 Å apart, respectively, in the structure of AMPK/(ATP γ S+CpdC) (Fig. 2C). Similarly, the cross-links enriched in AMPK/ATP γ S

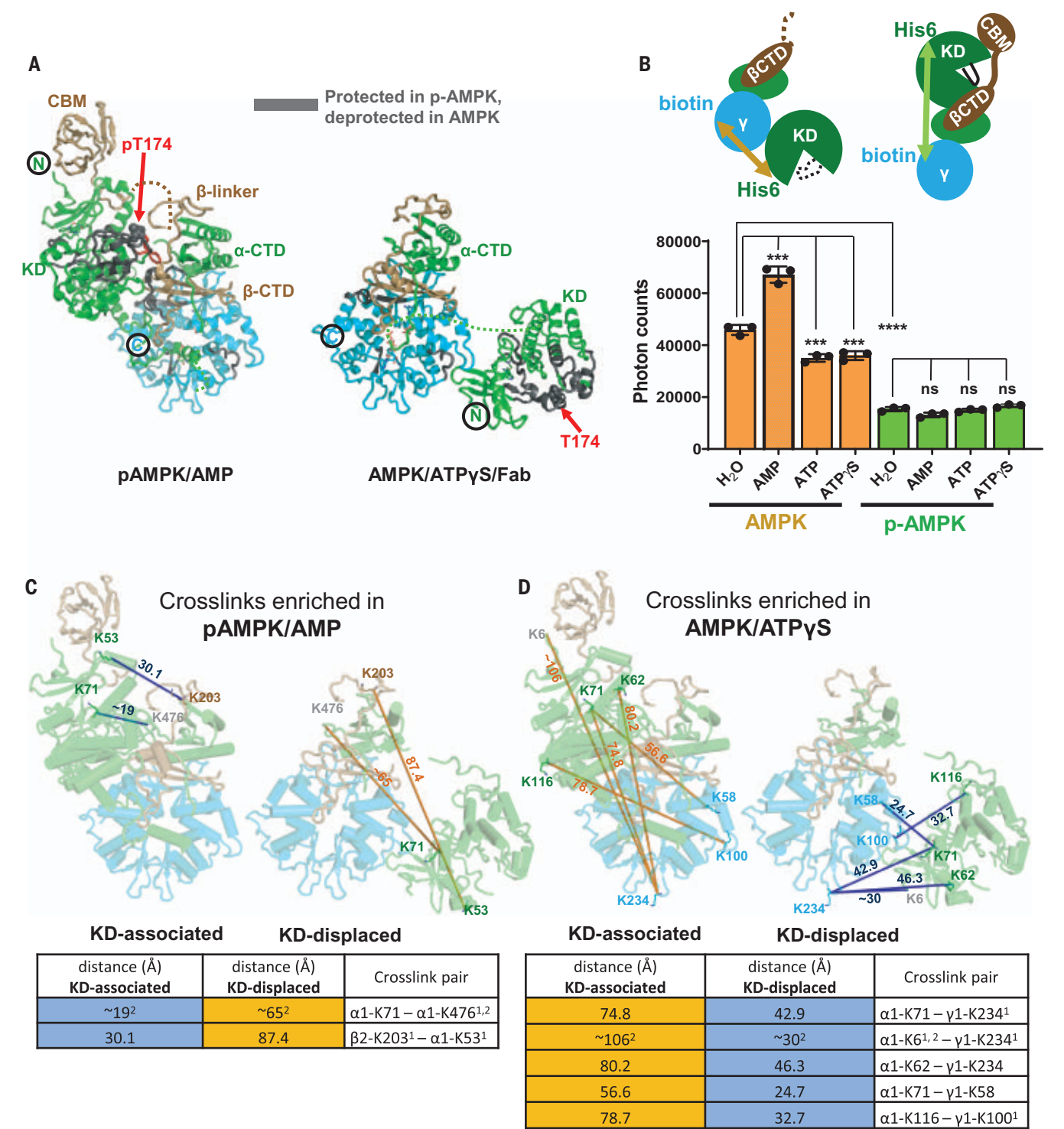


Fig. 2. In vitro analysis of AMPK conformations. (A) Phosphorylation-induced HDX-MS changes overlaid on the structures of pAMPK/AMP/staurosporine (PDB ID 4RER) (left) and AMPK/ATP/CpdC/Fab (right). Peptides with reduced deuterium exchange in phosphorylated AMPK relative to non-phosphorylated AMPK are overlaid in gray. T174 is shown as spheres. (B) (Top) Cartoons of His6- and biotin-tagged AMPK. (Bottom) AlphaScreen assay ($n = 3$ technical replicates, error bars= SEM; one-way analysis of variance (ANOVA): *** $P < 0.001$; **** $P < 0.0001$; ns, not significant). (C) and

D) XL-MS analysis. Cross-linking distances that are compatible with the indicated structure are shown in blue, those that are incompatible, in orange. (C) Cross-links that are enriched in phosphorylated AMPK/AMP overlaid onto the KD-associated (PDB ID 4RER) and KD-displaced (cryo-EM) structures. (D) Cross-links that are enriched in nonphosphorylated AMPK/ATP γ S overlaid onto the structures. Superscript number 1 indicates a loop residue. Superscript number 2 indicates a residue that borders a resolved region, so the residue–residue distance is only estimated.

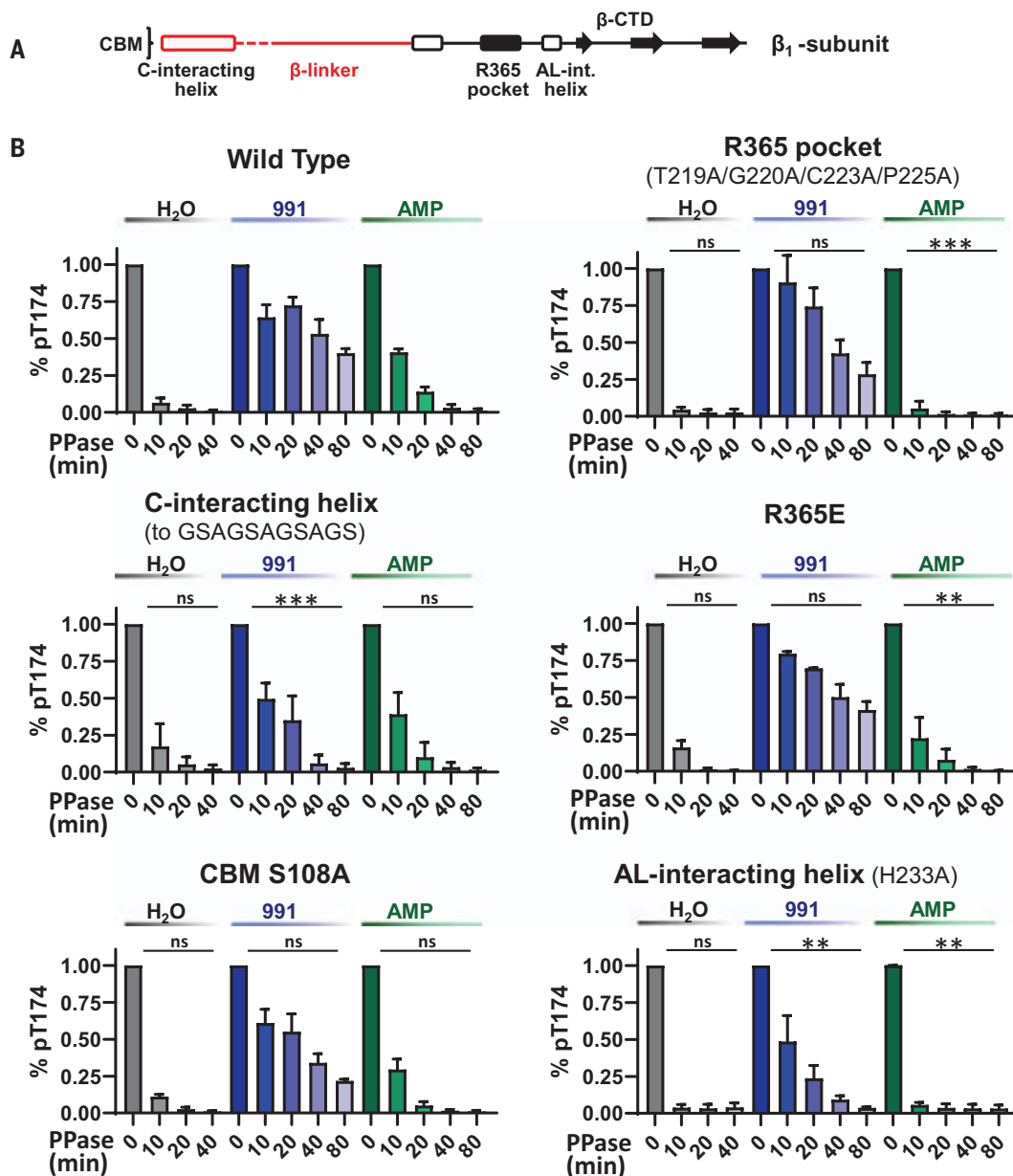


Fig. 3. Elements at the termini of the β linker mediate the effects of compound 991 and AMP on AL accessibility.

(A) Schematic presentation of the AMPK β linker with flanking elements. AL-int., AL-interacting. (B) PP2C α protection assays. Full-length wild-type and mutant AMPK α 1 β 1 γ 1 were incubated with PP2C for the indicated amount of time, and T174 phosphorylation determined by immunoblotting (fig. S17B); $n = 3$ biological repeats, error bars = SEM; ** $P < 0.01$, *** $P < 0.001$ (cumulative robust linear regression of each time course relative to the corresponding wild-type time course; with ordinary regression, the P values for R365E/AMP and H233A/AMP change from ** to *, while all other values remain the same).

(α 1-K71- γ 1-K234, α 1-K6- γ 1-K234, α 1-K62- γ 1-K234, α 1-K71- γ 1-K58, and α 1-K116- γ 1-K100) were only compatible with the KD-displaced structure (distances between 24.7 and 46.3 Å, with distances >30 Å involving loop residues); these would be between 56.6 and >100 Å apart in the KD-associated structure (Fig. 2D). Together, the XL-MS data confirmed that AMP-bound phosphorylated AMPK preferentially adopted the KD-associated conformation and that ATP γ S-bound, nonphosphorylated AMPK preferentially adopted the KD-dissociated conformation. The XL-MS analysis also provided insight on the positions of the unresolved N terminus of the β subunit and the position(s) of the CBM (fig. S14 and additional data analysis section of the supplementary materials).

ATP binding stimulates AL dephosphorylation by destabilizing an interaction between AL and the β linker

To understand how AMPK ligands change AL accessibility in pAMPK, we first identified the AL-protecting elements. The AMPK core consisting of the γ subunit and the α - and β -CTDs was previously implicated in AL protection (14), and only the core and the β linker are in the vicinity of the AL and have the potential to shield it (fig. S15, A and B) (14, 29). Whereas pT174 was rapidly dephosphorylated by only 100 nM human protein phosphatase 2C α (PP2C) in the isolated KD, similar rates of dephosphorylation required about threefold higher concentration of PP2C for a truncated AMPK complex that contains the core but lacks the CBM and β linker (Δ CBM- β linker)

and an ~ 260 -fold higher concentration for full-length AMPK (fig. S16). This result indicated that the core complex contributes to restricting AL access but that most of the protection is due to the β linker. Energy excess nucleotide levels (4.5 mM ATP, 400 μ M ADP, 40 μ M AMP) slightly reduced AL protection, although the decrease was not statistically significant, whereas energy stress levels (3.8 mM ATP, 1 mM ADP, 0.3 mM AMP) moderately, and compound 991 strongly, increased AL protection for full-length AMPK. In contrast, none of the ligands affected AL protection in the isolated KD or in AMPK(Δ CBM- β linker), consistent with a previous finding (16) (fig. S16). Therefore, both adenine nucleotides and ADaM site ligands likely regulate pT172 or pT174 access through modulation of the β linker-AL interaction.

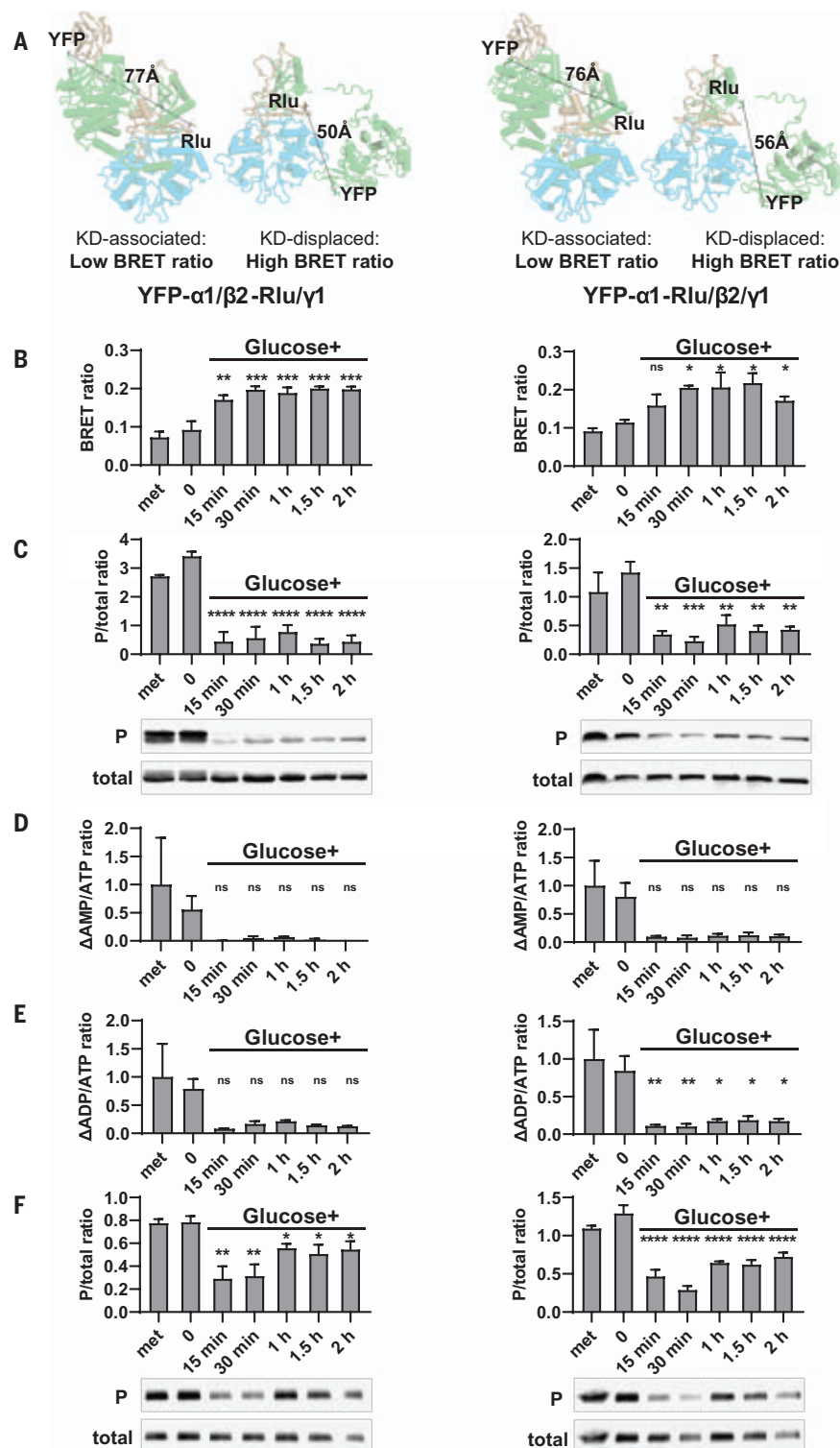


Fig. 4. Cell-based BRET-immunoblot-MS analysis. (Left) Construct combination 1: YFP- α 1/ β 2-Rlu/ γ 1. (Right) Construct combination 2: YFP- α 1-Rlu/ β 2/ γ 1. **(A)** BRET constructs in the context of the KD-associated and KD-displaced AMPK structures. **(B)** BRET ratios. **(C)** AL phosphorylation ratios determined by immunoblotting. pT174 phosphorylation (P) and total expression (total) were quantified with untreated sample set as 1.0. Representative immunoblots are shown below the bar graphs. **(D and E)** AMP/ATP (D) and ADP/ATP (E) ratios relative to ratios under energy stress (met = 1.0). **(F)** ACC phosphorylation ratios determined by immunoblotting and quantified with untreated sample set as 1.0. Representative immunoblots are shown below the bar graphs. Met: 5 mM metformin. $n = 3$ biological replicates, error bars = SEM, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ (against 0 min glucose; one-way ANOVA). Uncropped blots are shown in fig. S22.

Next we sought to identify the connections between ligand binding and the AL- β linker interaction. Both AMP and 991 allosterically stabilize the active KD conformation (3, 33) thereby ordering the inherently flexible AL (36) and its interaction with the β linker. To test whether the active KD conformation contributes to pT174 protection, we stabilized the conformation by an unrelated approach: binding of the catalytic cosubstrate Mg^{2+} -ATP or the high-affinity, ATP-competitive kinase inhibitor staurosporine (fig. S17A). Both reduced AL dephosphorylation, suggesting that the allosteric kinase activation by AMP and 991 contributed to the increase in AL protection.

In addition, we hypothesized that both ligands stabilize the β linker by stabilizing its flanking domains, the CBM and the β -CTD (Fig. 3A). ADaM ligands, and to a lesser degree CBM S108 phosphorylation, stabilize the CBM conformation and position and induce formation of a helix, called the C-interacting helix, at the N terminus of the β linker (Fig. 3A) (16, 17, 29). Confirming our hypothesis, replacement of the C-interacting helix with an equally long, flexible linker (GSAGSAGSAGS), and possibly to a lesser degree the CBM S108A mutation, reduced AL protection by 991 while having little effect on protection in the absence of ligands or in the presence of AMP (Fig. 3B and fig. S17, B and C).

AMP is sensed through formation of a CBS3-AMP- α linker interaction through α -E364 (fig. S15B, inset) that is required for both direct AMPK activation and protection against AL dephosphorylation (14, 29, 37). The E364-adjacent R365 binds an inducible pocket (R365 pocket) in the β -CTD C terminal to the β linker (fig. S15B, inset) (37). Mutation of either R365 or the R365 pocket reduced AL protection by AMP yet had little effect on protection in the absence of ligands or in the presence of 991 (Fig. 3B and fig. S17, B and C). This result strongly supports an important role for the R365 pocket in AMP-mediated AL protection. In contrast, mutation of the AL-interacting core residue β 2-H233 reduced AL protection, as previously reported (14), regardless of the presence or absence of AMP or 991 (Fig. 3B and fig. S17, B and C). Stabilization by the H233 interaction is thus not regulated by AMPK agonists. These results are confirmed in the context of full-length wild-type AMPK by analyzing our previously published HDX-MS data (33, 38, 39), which showed that AMP, ADaM ligand A769662, and AMPK phosphorylation induced stabilization of the β linker flanking regions (fig. S18).

Showing further agreement with our previously reported results, the α linker is dissociated from its position at the γ subunit in ATP-bound AMPK (fig. S19), and the R365 pocket moved up to 6 Å relative to its position in AMP-bound AMPK (fig. S19A). In addition, the lack of density for the CBM in all three structures of ATP-bound AMPK, together with our XL-MS data

(see supplementary materials) and previous double electron-electron resonance spectra (32), indicates that the CBM can dissociate from the KD in the absence of ADaM ligands and adopt multiple alternative positions, especially in nonphosphorylated AMPK. Finally, the resolved part of the β linker moved outward by up to 9 Å (fig. S19B). Collectively, our results thus indicate that a series of connected conformational changes link ATP binding and the absence of ADaM agonists to an α linker movement, CBM dissociation, and destabilization of the β linker, which then promotes dephosphorylation. Dephosphorylation in turn destabilizes KD association, which depends on phosphate-mediated charge interactions that stabilize the AL and the CBM–KD interaction (fig. S20, A and B).

Live-cell time course of energy excess-induced conformational changes

Our luminescence proximity, HDX-MS, and XL-MS assays provided three independent lines of support for the KD-displaced conformation seen in our cryo-EM structures. Because these experiments used recombinant, purified proteins, we developed a bioluminescence resonance energy transfer (BRET) conformational sensor to monitor changes in the position of the KD in cells. We genetically fused yellow fluorescent protein (YFP) to the N terminus of the KD and *Renilla* luciferase (Rlu) to either the C terminus of β 2 or the C terminus of the α -CTD. For both combinations, Rlu and YFP move closer together when the KD adopts the displaced conformation (50 and 56 Å, respectively) relative to the KD-associated conformation (Fig. 4A).

To determine the time course of AMPK inactivation, we first pretreated COS-1 cells expressing these two sets of BRET conformational sensors with 5 mM metformin, which activates AMPK by producing energy stress (30, 40). Because COS-1 cells are highly dependent on glycolysis to generate ATP, we induced the transition to energy excess by replacing metformin with 25 mM glucose. We followed AMPK inactivation over 2 hours by monitoring, in parallel, AMPK conformation by BRET (Fig. 4B); AL dephosphorylation by immunoblotting (Fig. 4C and fig. S22); changes in the AMP/ATP and ADP/ATP ratios by liquid chromatography-mass spectrometry (Fig. 4, D and E); and AMPK activity by immunoblotting of the AMPK substrate acetyl-coenzyme A carboxylase (ACC) (Fig. 4F). At 15 min after glucose addition, the AMP/ATP and ADP/ATP ratios, AL phosphorylation, and ACC phosphorylation had already dropped drastically and remained at low levels (Fig. 4, C to F). Simultaneously, BRET ratios for both BRET combinations approximately doubled and reached a maximum of about a 2.5-fold increase after 30 min of glucose treatment (Fig. 4B), indicating a closer distance

between KD and core AMPK upon AMPK inactivation by energy excess. This indicated that energy excess and AL dephosphorylation were coupled to the conformational change of the KD–AMPK core interaction. The delay in reaching maximum BRET ratios relative to AMPK dephosphorylation, while in part due to a small delay of pelleting cells for immunoblotting (see methods), is consistent with our data suggesting that, predominantly, phosphorylation precedes KD displacement in cells. Together, this illustrates a multistep model of AMPK inactivation and reactivation (fig. S21). In the fully active state, AMPK is bound by AMP and/or an ADaM agonist and both the KD and CBM are phosphorylated. AMP stabilizes the β linker through the CBS3/ α linker/R365 pocket network and ADaM ligands through the KD/CBM/C-interacting helix network. Under conditions of energy excess, ATP replaces AMP at CBS3. ATP weakens the interaction network, thereby destabilizing the β linker and the KD active conformation, resulting in increased AL access to phosphatases. Partial destabilization of AL– β linker and CBM–KD interactions increases AL and CBM dephosphorylation leading to KD and CBM displacement and full inhibition. Upon energy stress, AMP or ADP replaces ATP at CBS3 again. This induces a conformational change in AMPK that facilitates AL rephosphorylation by LKB1 and CBM phosphorylation by autophosphorylation. The detailed mechanism of AL accessibility regulation revealed in this study provides a structural framework for the rational development of therapeutics that selectively modulate AL accessibility, and it may serve as a paradigm for other signaling kinases that are regulated through changes in AL protection.

REFERENCES AND NOTES

- D. Garcia, R. J. Shaw, *Mol. Cell* **66**, 789–800 (2017).
- D. G. Hardie, *Genes Dev.* **25**, 1895–1908 (2011).
- D. G. Hardie, *J. R. Soc. Interface* **15**, 20170774 (2018).
- G. R. Steinberg, B. E. Kemp, *Physiol. Rev.* **89**, 1025–1078 (2009).
- H. X. Yuan, Y. Xiong, K. L. Guan, *Mol. Cell* **49**, 379–387 (2013).
- N. Dzamko et al., *J. Biol. Chem.* **285**, 115–122 (2010).
- D. F. Egan et al., *Science* **331**, 456–461 (2011).
- H. O. Akman et al., *Pediatr. Res.* **62**, 499–504 (2007).
- T. L. Marin et al., *BMC Syst. Biol.* **9**, 13 (2015).
- S. Fogarty, D. G. Hardie, *Biochim. Biophys. Acta* **1804**, 581–591 (2010).
- A. K. Wong, J. R. Petrie, C. C. Lang, *Clin. Sci.* **116**, 607–620 (2009).
- G. A. Rutter, I. Leclerc, *Mol. Cell. Endocrinol.* **297**, 41–49 (2009).
- F. A. Ross, C. MacKintosh, D. G. Hardie, *FEBS J.* **283**, 2987–3001 (2016).
- B. Xiao et al., *Nature* **472**, 230–233 (2011).
- F. J. Xin, J. Wang, R. Q. Zhao, Z. X. Wang, J. W. Wu, *Cell Res.* **23**, 1237–1240 (2013).
- B. Xiao et al., *Nat. Commun.* **4**, 3017 (2013).
- M. F. Calabrese et al., *Structure* **22**, 1161–1172 (2014).
- D. Carling, P. R. Clarke, V. A. Zammit, D. G. Hardie, *Eur. J. Biochem.* **186**, 129–136 (1989).
- S. P. Davies, N. R. Helps, P. T. W. Cohen, D. G. Hardie, *FEBS Lett.* **377**, 421–425 (1995).

- G. J. Gowan, S. A. Hawley, F. A. Ross, D. G. Hardie, *Cell Metab.* **18**, 556–566 (2013).
- J. S. Oakhill et al., *Science* **332**, 1433–1435 (2011).
- Y. L. Zhang et al., *Cell Metab.* **18**, 546–555 (2013).
- O. Göransson et al., *J. Biol. Chem.* **282**, 32549–32560 (2007).
- M. J. Sanders et al., *J. Biol. Chem.* **282**, 32539–32548 (2007).
- J. W. Scott et al., *Chem. Biol.* **21**, 619–627 (2014).
- Chen et al., *eLife* **6**, e31268 (2017).
- P. Puustinen et al., *Autophagy* **16**, 1871–1888 (2020).
- J. Jia et al., *Mol. Cell* **77**, 951–969.e9 (2020).
- X. Li et al., *Cell Res.* **25**, 50–66 (2015).
- G. Zhou et al., *J. Clin. Invest.* **108**, 1167–1174 (2001).
- J. Ereño-Orbea et al., *J. Mol. Biol.* **430**, 322–336 (2018).
- X. Gu et al., *J. Biol. Chem.* **293**, 16994–17007 (2018).
- Y. Yan et al., *J. Biol. Chem.* **294**, 953–967 (2019).
- A. Kao et al., *Mol. Cell. Proteomics* **10**, M110.002212 (2011).
- X. Wang et al., *J. Biol. Chem.* **292**, 16310–16320 (2017).
- B. Nolen, S. Taylor, G. Ghosh, *Mol. Cell* **15**, 661–675 (2004).
- L. Chen et al., *Nature* **498**, E8–E10 (2013).
- X. Gu et al., *J. Biol. Chem.* **292**, 12653–12666 (2017).
- R. R. Landgraf et al., *Structure* **21**, 1942–1953 (2013).
- R. J. Shaw et al., *Science* **310**, 1642–1646 (2005).

ACKNOWLEDGMENTS

We thank D. Wu (Guangzhou Institute for Biomedicine and Health, China) for the CaMKK β expression plasmid described in (29). Cryo-EM data were collected at the David Van Andel Advanced Cryo-Electron Microscopy Suite at the Van Andel Institute. We are grateful to X. Meng for cryo-EM data collection and EM technical support and to D. Nadziejka for language editing of the manuscript. We also thank staff members of the Life Science Collaborative Access Team of the Advanced Photon Source (APS) for assistance in data collection at the beam lines of sector 21. **Funding:** This work was supported in part by NIH grants R01 GM117372 (A.A.K.) and R01 GM129436 (K.M. and P.R.G.) and by the Van Andel Institute (R.G.J. and K.M.). Use of the Advanced Photon Source (APS), an Office of Science User Facility operated for the US Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the US DOE under contract DE-AC02-06CH11357. **Author contributions:** K.M. conceived of the project. I.E.V., R.G.J., L.J.M., H.E.X., P.R.G., A.A.K., and K.M. designed the experiments. Y.Y., S.M., K.G.H., T.S.S., X.E.Z., K.S.-P., T.-H.X., R.D.S., J.L., J.S.B., K.R., A.E., and S.J.N. performed and/or interpreted the experiments. Y.Y. and K.M. wrote the paper, with support from all authors. **Competing interests:** R.G.J. holds an adjunct professorship at McGill University, serves on the scientific advisory board of Immunomet Therapeutics, and is a scientific consultant for Agios Pharmaceuticals. **Data and materials availability:** Crystal structure data were deposited in the Protein Data Bank (PDB) under ID 7JJJ. Cryo-EM data were deposited in the Electron Microscopy Data Bank (EMD) under accession number EMD-22336 and in PDB under ID 7JHG (AMPK-CpdC-ATP γ S-Fab-nanobody); EMD-22337 and PDB ID 7JHH (AMPK-ATP γ S-Fab-nanobody); and EMD-23708 and PDB ID 7M74 (AMPK-CpdC-ATP γ S-Fab-nanobody, nonphosphorylated sample). The XL-MS raw data and search results have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD019017. All other data needed to evaluate the conclusions in the paper are presented in the main text or supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/413/suppl/DC1
Additional Data Analysis
Materials and Methods
Figs. S1 to S22
Tables S1 to S3
References (41–67)
MDAR Reproducibility Checklist
Data S1
Movie S1

12 September 2020; resubmitted 31 March 2021
Accepted 7 June 2021
10.1126/science.abe7565

PLANT SCIENCE

Plant “helper” immune receptors are Ca^{2+} -permeable nonselective cation channels

Pierre Jacob^{1†}, Nak Hyun Kim^{1†}, Feihua Wu^{2,3}, Farid El-Kasmi⁴, Yuan Chi², William G. Walton⁵, Oliver J. Furzer¹, Adam D. Lietzan⁶, Sruthi Sunil⁴, Korina Kempthorn¹, Matthew R. Redinbo⁵, Zhen-Ming Pei^{2*}, Li Wan^{1,7*}, Jeffery L. Dangl^{1*}

Plant nucleotide-binding leucine-rich repeat receptors (NLRs) regulate immunity and cell death. In *Arabidopsis*, a subfamily of “helper” NLRs is required by many “sensor” NLRs. Active NRG1.1 oligomerized, was enriched in plasma membrane puncta, and conferred cytoplasmic calcium ion (Ca^{2+}) influx in plant and human cells. NRG1.1-dependent Ca^{2+} influx and cell death were sensitive to Ca^{2+} channel blockers and were suppressed by mutations affecting oligomerization or plasma membrane enrichment. Ca^{2+} influx and cell death mediated by NRG1.1 and ACTIVATED DISEASE RESISTANCE 1 (ADR1), another helper NLR, required conserved negatively charged N-terminal residues. Whole-cell voltage-clamp recordings demonstrated that *Arabidopsis* helper NLRs form Ca^{2+} -permeable cation channels to directly regulate cytoplasmic Ca^{2+} levels and consequent cell death. Thus, helper NLRs transduce cell death signals directly.

In plants, successful pathogens inject effectors into the host cell to dampen the immune response. Plants evolved a surveillance system consisting of intracellular nucleotide-binding leucine-rich repeat receptors (NLRs) capable of triggering immunity in response to effector activity or by direct effector recognition (1). Effector-triggered immunity leads to activation of pathogen defense and culminates in the death of the host cell, the “hypersensitive response,” which can further limit pathogen growth (2). NLR activation is sufficient to determine the outcome of a plant-pathogen interaction (1). In plants, NLRs are divided into two major classes based on their N-terminal domains: Toll/interleukin-1 receptor/Resistance (TIR)-NLRs, (hereafter, TNL) and the coiled-coil (CC)-NLRs (hereafter, CNL) (3, 4). All tested TNLs require ENHANCED DISEASE SUSCEPTIBILITY 1 (EDS1), as well as a subfamily of five redundant “helper” NLRs also known as RNLs because of their CC-R [RPW8 (Resistance to Powdery Mildew 8)-like CC] domain (5–7). There are two subfamilies of RNLs, NRG1 (N REQUIREMENT GENE 1) and

ADR1 (ACTIVATED DISEASE RESISTANCE 1), in nearly all flowering plants (8).

Activation of the *Arabidopsis* RNL NRG1.1 can be mimicked using an autoactive allele (D485V; hereafter DV), mutated in the conserved methionine-histidine-aspartate motif. This is an accepted proxy for pathogen effector-mediated NLR activation (9–12). NRG1.1 DV-induced cell death is independent of the native RNLs NRG1 and ADR1 and of the EDS1-signaling module in the heterologous host, *Nicotiana benthamiana* [(4, 7, 13, 14); fig. S1].

Results

Structure of NRG1.1-signaling domain

We obtained x-ray crystal structures of two mutant NRG1.1 CC-R domains (residues 1 to 124), K94E/K96E/R99E/K100E/R103E/K106E/K110E [7K(R)/E] and K94E/K96E (2K/E), which diffracted to 3.0 and 2.5 Å, respectively (table S1). These putative surface mutations were required to achieve monodispersity of the protein. Structural homology modeling suggested that the CC-R domains share an N-terminal four-helical bundle with cell death pore forming mammalian mixed-lineage kinase domain-like (MLKL) proteins and fungal HET-s/HELO domain proteins (6, 15, 16). Similar to predictions, the two mutant structures superimposed well with the four-helical bundles of the resting-state CC domain of ZAR1 (17) and the cation channel-forming domain of MLKL (18) (Fig. 1, A to D, and table S2).

The four-helical bundle of MLKL is sufficient to cause cell death, and this requires two hydrophobic residues, L58 and I76, which maintain the four-helical bundle hydrophobic core (19). Structural overlay predicted that the two hydrophobic residues are conserved in NRG1.1 as L69 and F86 (Fig. 1E); mutating them abolished cell death activity of an NRG1.1 fragment from residues 1 to 180 (hereafter NRG1.1-180) (Fig. 1F). In ZAR1, a hydrophobic groove

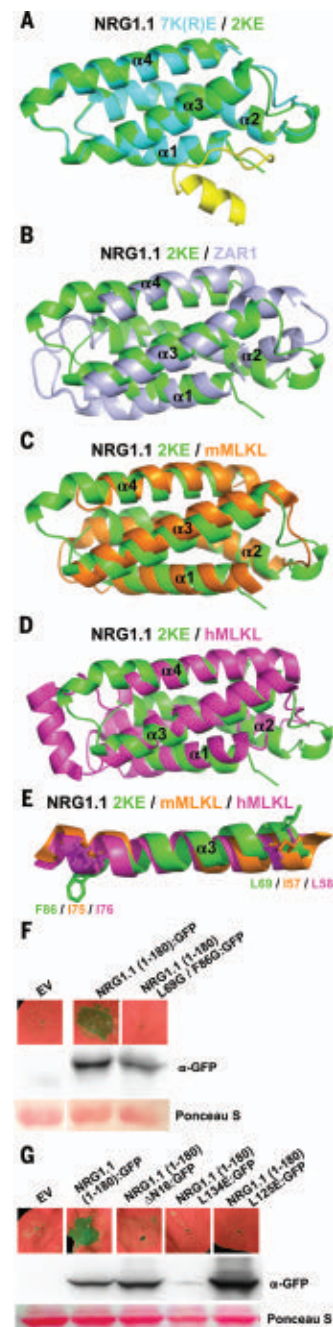


Fig. 1. NRG1.1 CC-R resembles the ZAR1 and MLKL four-helical bundle. (A) Structural overlay of 7K(R)/E (PDB: 7L7V) and 2K/E (7L7W) with 2K/E in green, 7K(R)/E four-helical bundle (4HB) in cyan, and 7K(R)/E N-terminal region in yellow. (B) Structural overlay of 2K/E in green and the ZAR1 4HB (6J5V) structure in light purple. (C) Structural overlay of 2K/E in green and mMLKL 4HB (4BTF) in orange. (D) Structural overlay of 2K/E in green and hMLKL 4HB (2MSV) in magenta. (E) Conserved F86/I75/I76 and L69/I57/L58 shown in sticks in the α 3 helix of superimposed 2K/E, mMLKL and hMLKL 4HB structures, respectively. (F and G) In planta (*N. benthamiana*) phenotypes at 2 days after induction and protein accumulation of WT NRG1.1 CC-R and structure-derived mutants. EV, empty vector.

¹Department of Biology and Howard Hughes Medical Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. ²Department of Biology, Duke University, Durham, NC 27708, USA. ³Department of Horticulture, Foshan University, Foshan, China. ⁴Department of Plant Physiology, Centre of Plant Molecular Biology (ZMBP), University of Tübingen, Tübingen, Germany. ⁵Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. ⁶Division of Oral and Craniofacial Health Sciences, Adams School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. ⁷National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China. *Corresponding author. Email: zpei@duke.edu (Z.-M.P.); lwan@cemps.ac.cn (L.W.); dangl@email.unc.edu (J.D.) †These authors contributed equally to this work and are listed alphabetically.

made by $\alpha 2$ and $\alpha 4B$ is important for oligomerization and function (17). We found two hydrophobic residues, L134 and L125, in the homologous region in NRG1.1; mutating them to glutamic acid abolished cell death in NRG1.1 1-180 (Fig. 1G). These data validate the NRG1.1 four-helical bundle structure.

Active ZAR1 oligomerizes into a pentamer in which the $\alpha 1$ helix of the four-helical bundle rearranges, flips out, and forms a plasma membrane (PM)-penetrating, funnel-like structure (17, 20, 21), consistent with models proposed for fungal Het-s/HELO activation (22). This first helix was essential for ZAR1 function at

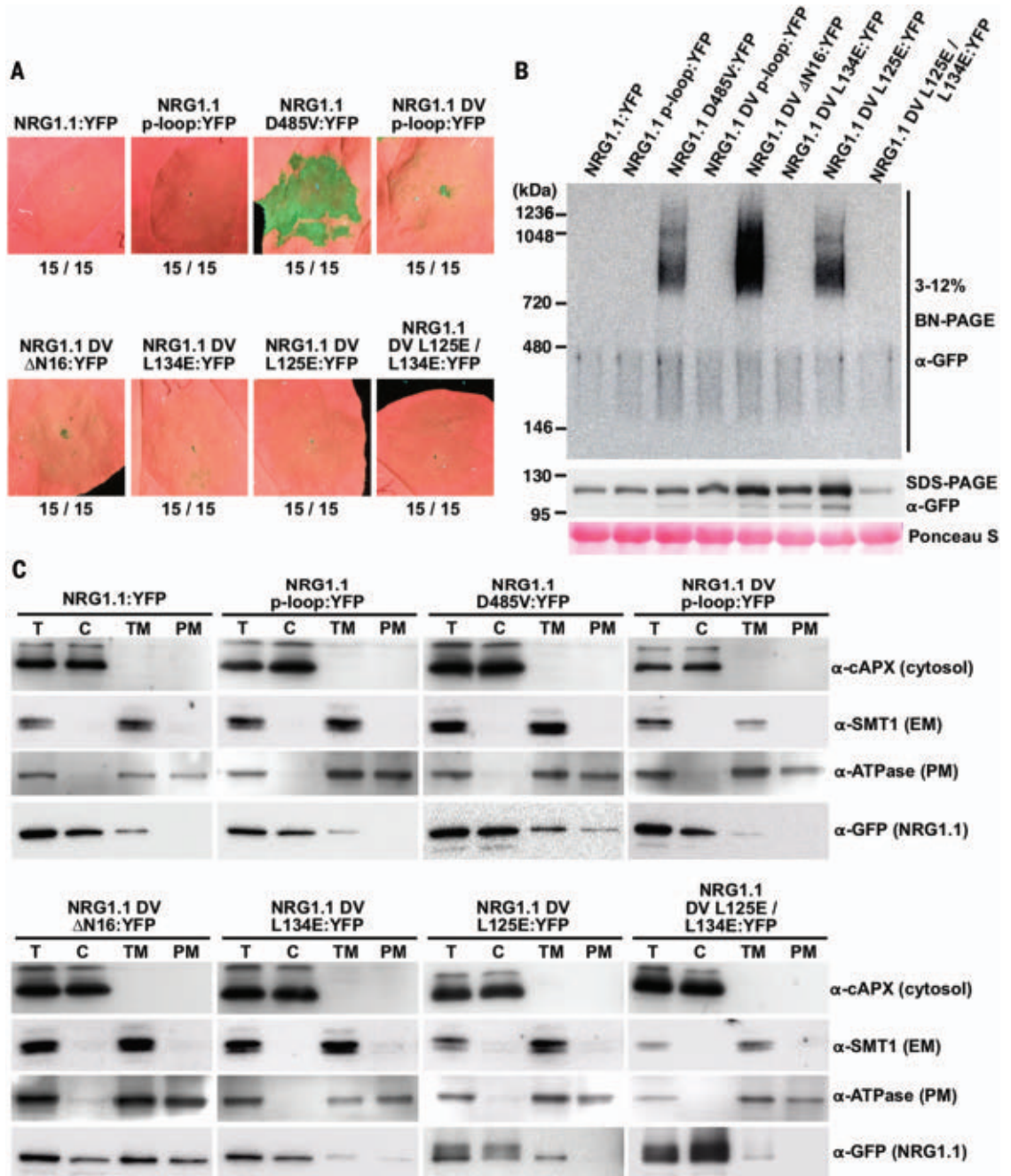
the PM (17). One of the NRG1.1 CC-R structures [7(K)R/E] revealed a potentially flexible N terminus (residues 1 to 16; absent in ZAR1; Fig. 1A) that could extend the four-helical bundle $\alpha 1$ helix of a putative funnel in the active NRG1.1 protein. This was disordered in the 2K/E structure. This region was required for NRG1.1 1-180 cell death induction (Δ N16; Fig. 1G).

Active NRG1.1 oligomerizes and forms puncta

We introduced several of the structure-derived mutations into the full-length NRG1.1 in *cis* with the autoactive DV allele because the wild-type (WT) resting NRG1.1 is inactive in the absence

of sensor NLR activation (Fig. 2A). We found that the Δ N16, L134E, and L125E *cis* mutations each abolished the cell death function of NRG1.1 DV (Fig. 2A). Blue native-PAGE analyses revealed that active NRG1.1 DV formed high-molecular-weight complexes, whereas the inactive WT, catalytic p-loop (G199A/K200A), and DV p-loop *cis* double mutants did not (Fig. 2B). Mutation of L134E, but not of L125E, in *cis* also abolished NRG1.1 DV oligomer formation (Fig. 2B). Unlike the WT or inactive p-loop mutants, NRG1.1 DV was enriched in the PM fraction, whereas DV L125E showed substantially reduced PM enrichment (Fig. 2C). Although it

Fig. 2. Active NRG1.1 (D485V) oligomerizes on the PM and triggers cell death. (A) In planta phenotypes of NRG1.1 variants in *N. benthamiana* (Nb) at 2 dpi. YFP, yellow fluorescent protein. (B) Oligomerization of NRG1.1 DV variants. Accumulation of YFP-tagged NRG1.1 variants was verified by SDS-polyacrylamide gel electrophoresis (PAGE) immunoblotting, and their oligomeric states were checked by blue native-PAGE immunoblotting in *Nb adr1 nrg1*. (C) Activation promotes PM enrichment. Total proteins (T) extracted from *Nb*-expressing NRG1.1:YFP variants were fractionated into cytosolic (C), total membrane (TM), and PM fractions and verified by marker proteins: Cytosol, cytosolic ascorbic peroxidase; endoplasmic reticulum membrane (EM), sterol methyltransferase 1; PM, H⁺ATPase.



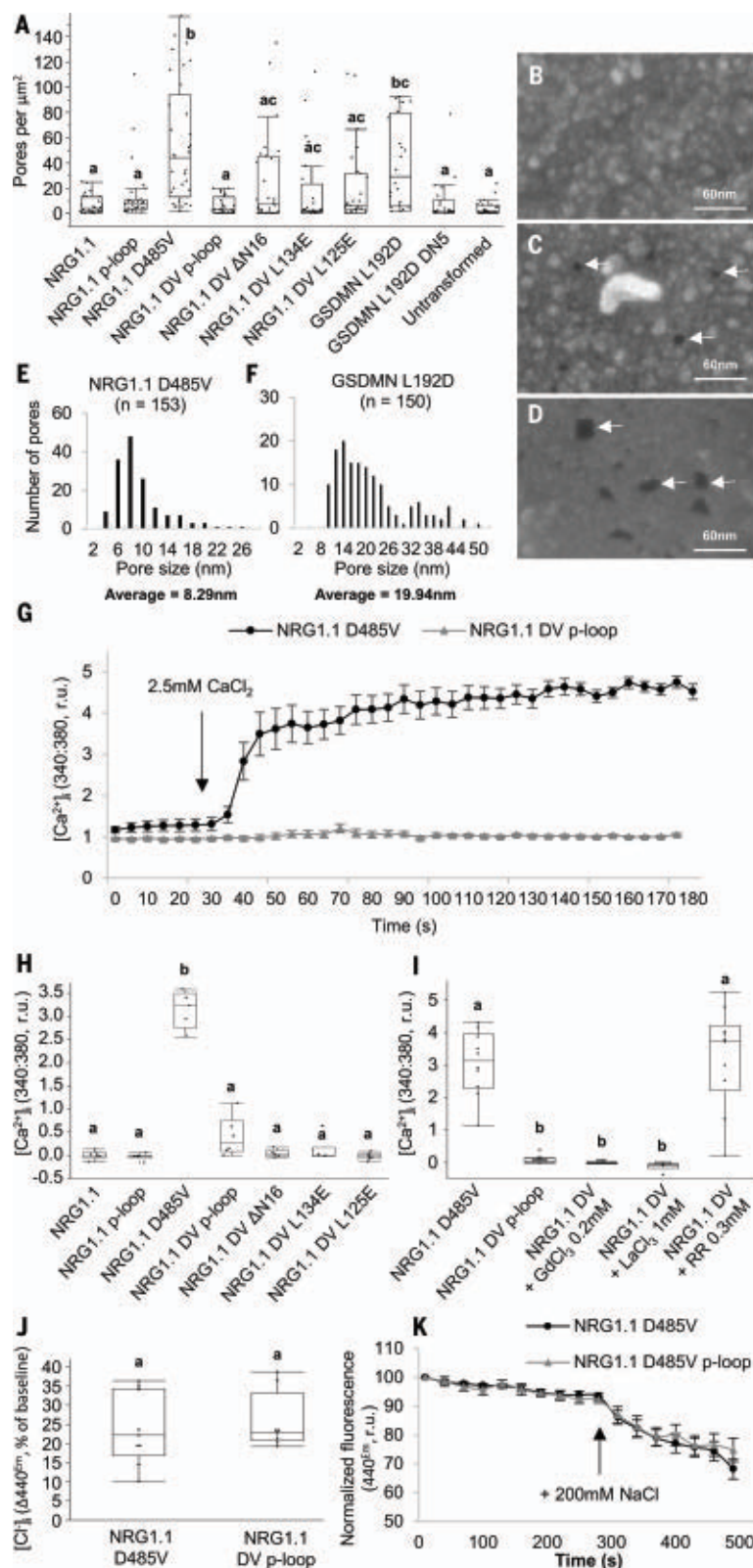


Fig. 3. NRG1.1 forms ion channels permeable to Ca^{2+} and not Cl^- . (A) Quantification of apparent PM pores in HeLa cells expressing NRG1.1 variants. (B to D) Representative scanning electron micrographs of cells exhibiting apparent PM pores in cell lines expressing the NRG1.1 DV p-loop (B), NRG1.1 D485V (C) or the pore-forming protein GSDMN 1-275 L192D (D). White arrows indicate the visible PM pores. (E and F) Distribution of the diameter of PM pores visible after NRG1.1 D485V (E) or GSDMN L192D (F) expression. The average diameters are significantly different (t test, $P < 0.0001$). (G) $[\text{Ca}^{2+}]_i$ in NRG1.1 D485V or NRG1.1 DV p-loop-expressing HeLa cells. Black arrow indicates CaCl_2 addition. (H) Ca^{2+} influx in HeLa cells expressing NRG1.1 variants. (I) Effect of the Ca^{2+} channel blockers LaCl_3 , GdCl_3 , and ruthenium red (RR) on NRG1.1 D485V-induced Ca^{2+} influx. (J) Intracellular $[\text{Cl}^-]$ accumulation in cells expressing the NRG1.1 D485V or NRG1.1 DV p-loop. (K) Representative time course experiment showing variation of $[\text{Cl}^-]$. Black arrow indicates 200 mM NaCl addition. Letters indicate statistical significance (ANOVA with post hoc Tukey, $P < 0.01$).

the NRG1.1 DV ΔN16 double mutant but were less common for the missense loss-of-function alleles, which colocalized more with the endoplasmic reticulum marker than the PM marker (fig. S3). These results show that the NRG1.1 N-terminal 16 residues may extend a funnel-like structure similar to that of ZARI.

NRG1.1 is functional in human HeLa cells

We investigated the possibility that NRG1.1 forms PM pores and functions as a channel. We expressed NRG1.1 variants in HeLa cells. If NRG1.1 caused cell death, formed pores, and exhibited channel activity in this evolutionarily distant cellular background, then the most parsimonious conclusion would be that it did so autonomously. The alternative hypothesis would require a signaling partner conserved between plants and humans. NRG1.1 DV induced significant cell death at 6 hours after induction (fig. S5). This did not occur when WT NRG1.1 was expressed or when NRG1.1 DV activity was suppressed by p-loop, ΔN16 , L125E, or L134E mutations in *cis* (fig. S5). Oligomerization and PM localization of the NRG1.1 variants in HeLa cells were similar to the in planta results (fig. S6). Thus, the genetic requirements for NRG1.1 DV cell death initiation, (p-loop, N16, L125, and L134) are similar in HeLa cells and plant cells.

We observed the morphology of the dying HeLa cells using scanning electron microscopy. We found that the number of PM pores, appearing as dark spots (23), correlated with the cell death activity of NRG1.1 variants (fig. S7 and Fig. 3, A to C). NRG1.1 DV-expressing cells exhibited 8-nm (average) pores in these processed samples (Fig. 3, E and F), although

oligomerized and was enriched at the PM (figs. S2 to S4), NRG1.1 DV ΔN16 failed to induce cell death (Fig. 2, A to C).

Confocal microscopy demonstrated that NRG1.1 DV exhibits increased PM localization

compared with inactive alleles (figs. S2 to S4). Active NRG1.1 DV exhibited numerous puncta on the PM, whereas the NRG1.1 DV p-loop double mutant exhibited many fewer (figs. S2 to S4). These puncta were also observed for

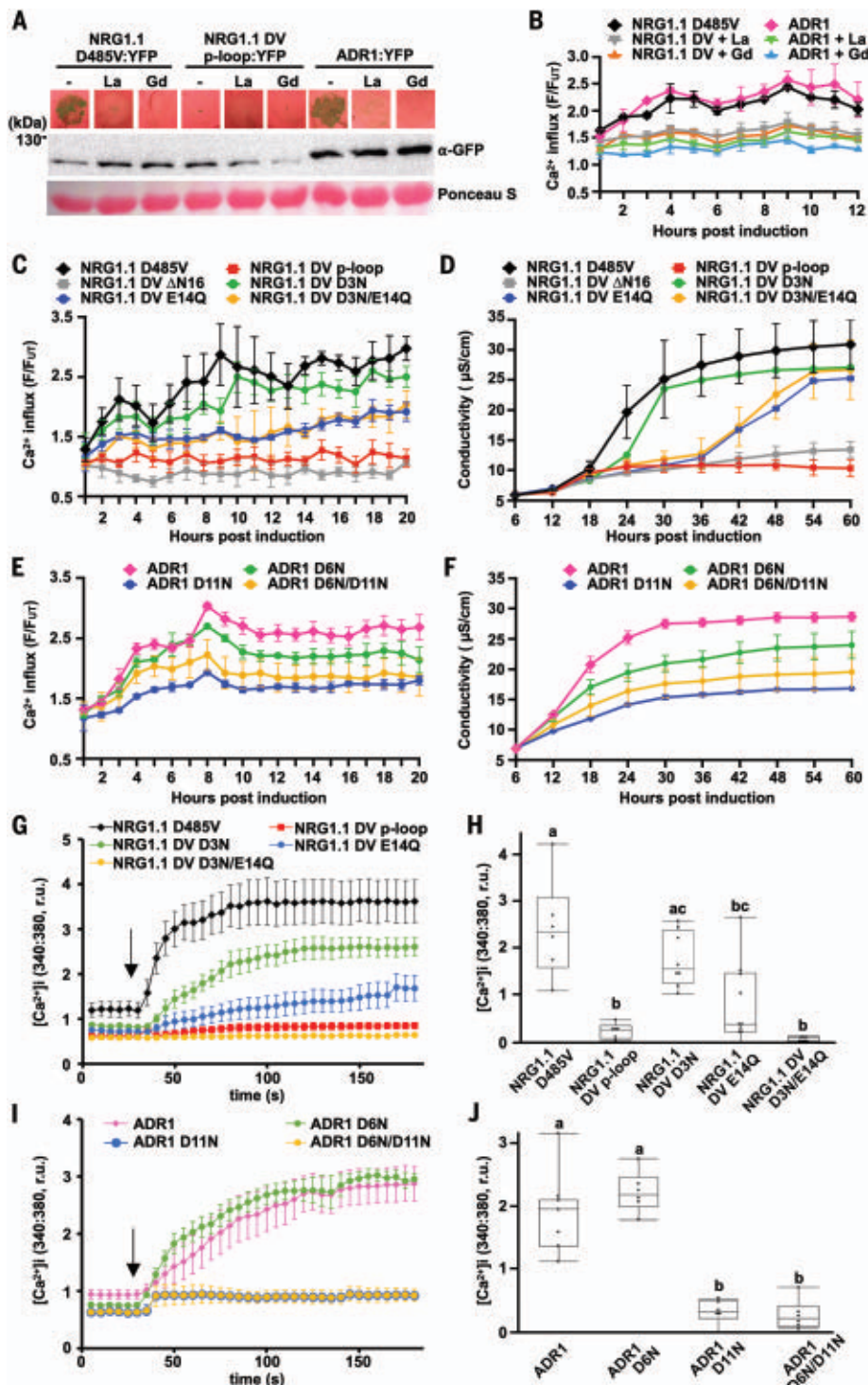


Fig. 4. Negatively charged residues in the conserved RNL motif are required for Ca²⁺ influx and cell death. (A) In planta phenotypes (2 dpi) of plant RNLs in the presence of the Ca²⁺ channel blockers LaCl₃ (La) and GdCl₃ (Gd). Accumulation of YFP-tagged RNL variants was verified by SDS-PAGE and immunoblotting, in *Nb adr1 nrg1*. (B, C, and E) Time course RNL-induced Ca²⁺ influx measurements with GCaMP3-transgenic *Nb*. (D and F) Time course conductivity measurement depicting RNL-triggered cell death in *Nb*. (G to J) [Ca²⁺]_i in HeLa cells expressing NRG1.1 DV and variants (G and H) or ADR1 and variants (I and J). Black arrows indicate the addition of CaCl₂. Letters indicate statistical significance (ANOVA with post hoc Tukey, *P* < 0.05). YFP-tagged RNL protein expression in *Nb* and HeLa cells was verified (fig. S16).

(25, 26) and is a requirement for NLR signaling (25, 27). Using Fura-2-based Ca²⁺ imaging (28, 29), we measured the cytosolic free Ca²⁺ concentration ([Ca²⁺]_i) in HeLa cells expressing NRG1.1 DV or the DV p-loop double mutant to determine whether NRG1.1 DV affected cytoplasmic Ca²⁺ homeostasis. We observed sustained [Ca²⁺]_i increases specifically in the NRG1.1 DV-expressing cells seconds after CaCl₂ addition (Fig. 3G and fig. S8). Loss of the cell death *cis* mutations ΔN16, L125E, or L134E all suppressed NRG1.1 DV-driven Ca²⁺ influx (Fig. 3H and fig. S8). The general Ca²⁺ influx channel blockers LaCl₃ and GdCl₃, but not the Ca²⁺ release channel blocker ruthenium red (30–32), blocked NRG1.1 DV-driven Ca²⁺ influx (Fig. 3I and fig. S9). These observations are consistent with NRG1.1 DV directly or indirectly facilitating Ca²⁺ influx. We investigated the specificity of NRG1.1 DV-driven ion flux by measuring cytosolic [Cl[−]] using 6-methoxy-N-ethylquinolinium iodide (MEQ)/dihydroMEQ (33). There was no difference between the NRG1.1 DV and NRG1.1 DV p-loop-expressing cells (Fig. 3, J and K, and fig. S10), indicating that NRG1.1 DV channels do not drive Cl[−] influx. Thus, we conclude that NRG1.1 might form Ca²⁺-permeable channels, or facilitate their formation, in the PM of HeLa cells.

Conserved negatively charged RNL residues are required for Ca²⁺ influx

We focused on the first 16 amino acids of NRG1.1 across 334 plant RNL sequences because NRG1.1 DV ΔN16 retained oligomerization and PM enrichment but lost Ca²⁺ influx. We observed a pattern of glycine or negatively charged or polar residues separated by two to three hydrophobic residues (table S3 and fig. S11). This motif was conserved in the ADR1 clade of RNLs, partially degenerated in the NRG1 clade, and further degenerated in CNLs (fig. S12). Although different from an

this may not represent the actual pore size. This apparent pore size was significantly different from the size of larger pores formed by GASDERMIN-D 1-275 L192D, a partial loss-of-function mutant of the pore-forming protein GASDERMIN-D that induces cell death and allows high protein accumulation [(17, 24); Fig. 3, D to F, and fig. S5B]. Cell fractionation experiments indicated that NRG1.1 DV was

enriched in the PM fraction in HeLa cells, whereas the NRG1.1 DV p-loop *cis* double mutant was not (fig. S6B). Overall, active NRG1.1 DV localized to the PM and its cell death activity was associated with the occurrence of PM pores in HeLa cells.

Helper NLRs form Ca²⁺-permeable channels

Ca²⁺ influx is a hallmark of programmed cell death in both the animal and plant kingdoms

N-terminal motif recently implicated in CNL function (34), these two domains share regularly spaced, negatively charged residues. Such residues are critical for ion selectivity and permeability in Ca^{2+} channels (35, 36). We targeted the negatively charged residues within the first 16 amino acids of NRG1.1 DV and the auto-active WT ADR1 (37) for structurally conservative but uncharged *cis* mutations (NRG1.1 DV: D3N, E14Q, or D3N/E14Q; ADR1: D6N, D11N, or D6N/D11N).

We assayed Ca^{2+} influx triggered by these alleles in planta using the intracellular Ca^{2+} reporter GCaMP3 in transgenic *N. benthamiana* (38). We confirmed that either NRG1.1 DV or ADR1 expression triggered $[\text{Ca}^{2+}]_i$ increases in planta 2 to 3 hours after estradiol treatment (a typical time required for estradiol-induced protein accumulation), which was accompanied by cell death 16 hours after $[\text{Ca}^{2+}]_i$ increase (Fig. 4, A and B). LaCl_3 and GdCl_3

abolished cell death (Fig. 4A) and reduced Ca^{2+} influx (Fig. 4B). In both plant (Fig. 4, C to F) and HeLa cells (Fig. 4, G to J), mutations in the RNL-conserved N-terminal motif, especially E14Q (in NRG1.1 DV) and D11N (in ADR1), significantly reduced the rate of Ca^{2+} influx (Fig. 4, C, E, and G to J, and fig. S13). These mutations also reduced cell death induction (Fig. 4, D and F, and figs. S14 and S15).

We analyzed the electrophysiological properties of NRG1.1 DV and NRG1.1 DV D3N E14Q in human embryonic kidney (HEK) 293 cells using the patch-clamp technique (28, 29). Having confirmed that NRG1.1 DV drives Ca^{2+} influx in transfected HEK293 cells similar to HeLa cells using Fura-2 imaging (fig. S17), we proceeded to record whole-cell voltage-clamp currents with voltage ramp (from +50 mV to -200 mV). We recorded robust currents at both positive and negative potentials only in cells expressing NRG1.1 DV with CsCl and

CaCl_2 in the pipette and bath solution, respectively (Fig. 5, A and C). Both inward and outward currents were abolished when Cs^+ and Ca^{2+} were substituted for the channel-impermeable cation TEA⁺ (Fig. 5, B and C), indicating that NRG1.1 DV formed nonselective cation channels mediating Ca^{2+} influx rather than Cl^- efflux. The NRG1.1 D3N E14Q double mutant eliminated Ca^{2+} currents (Fig. 5, A to C), consistent with D3 and/or E14 acting as a selectivity filter. Substitution of Mg^{2+} for Ca^{2+} in the bath resulted in a similar permeability for Mg^{2+} and Ca^{2+} (Fig. 5D), confirming that NRG1.1 is a nonselective, cation-permeable channel.

Discussion

Ca^{2+} signaling is known to regulate plant immunity (26, 39). Constitutive cytoplasmic Ca^{2+} influx induces constitutive defense activation and cell death (40, 41). We found that the NRG1.1-signaling domain structure resembles the pore-forming domain of the cation channel MLKL (Fig. 1), oligomerizes in puncta on the PM (Fig. 2), and is sufficient to drive cytoplasmic Ca^{2+} influx in plants and human cells (Figs. 3 to 5 and fig. S2). Our data are consistent with NRG1.1 acting as a nonselective, Ca^{2+} -permeable cation channel.

All TNL immune receptors tested so far required the redundant RNLs of the ADR1 and NRG1 subfamilies (5). Thus, we propose that TNL activation induces RNL-dependent Ca^{2+} influx to initiate cell death and, likely, immune responses. Supporting this view, we observed that TNL RPS4 activation is associated with Ca^{2+} influx (fig. S18). Cyclic nucleotide-gated channels may also contribute to sustain the RNL-initiated Ca^{2+} influx because they have been shown to be activated upon immune response (42). It is plausible that cell death is a product of Ca^{2+} -responsive factors that execute a cell death program rather than being a consequence of Ca^{2+} cytotoxicity. Cell death and defense activation can be uncoupled during effector-triggered immunity, suggesting that cell death has additional signaling requirements beyond elevated $[\text{Ca}^{2+}]$ (2, 43). Plants have MLKL homologs that are involved in immunity and can trigger cell death (44), possibly by regulating ion homeostasis [Mg^{2+} rather than Ca^{2+} (18, 45), by analogy with human MLKL].

CNLs likely also function like RNLs and trigger cytoplasmic Ca^{2+} influx. Recently, the CNL ZAR1 was shown to be a Ca^{2+} -permeable channel that requires a negative charge ring for cation influx (21). It is thus likely that ZAR1 E11, NRG1.1 E14, and ADR1 D11 all form negative charge rings that act as cation selective filters to achieve fast Ca^{2+} influx and initiation of cell death and defense signaling. The combined data from our work and (21) provide a mechanistic explanation for how the two major

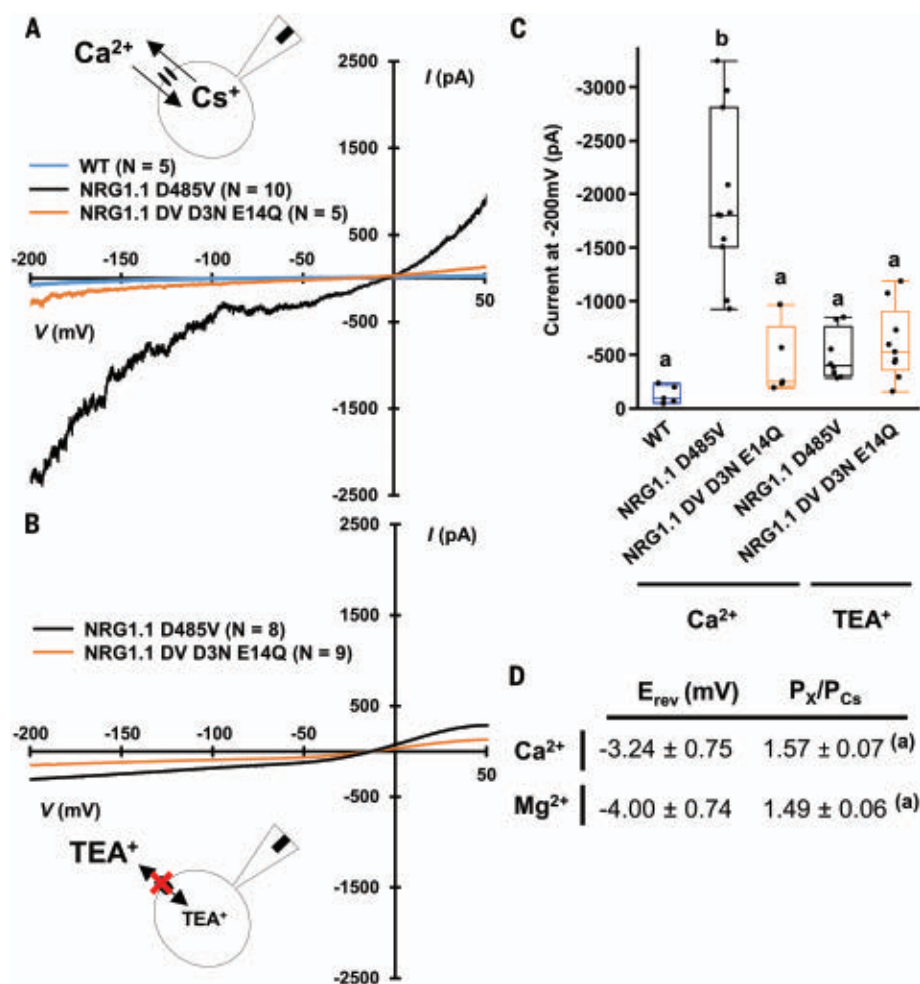


Fig. 5. NRG1.1 D485V forms nonselective cation channel. (A) Whole-cell current recorded in voltage ramp (+50 to -200 mV) in a CaCl_2 bath solution. (B) Whole-cell current recorded in voltage ramp (+50 to -200 mV) in a TEA-Cl bath solution. (C) Mean current recorded at -200 mV during experiments in (A) and (B). (D) Reversal potential (E_{rev}) and relative ion permeability (P_X/P_{Cs}) measured in voltage ramp experiments in CaCl_2 or MgCl_2 solutions in HEK293 cells expressing NRG1.1 D485V ($N = 10$ Ca^{2+} and $N = 8$ Mg^{2+}). Letters indicate statistical difference (ANOVA with post hoc Tukey, $P < 0.05$).

classes of plant intracellular innate immune receptors, the TNIs and CNs, control cell death triggered by pathogen recognition.

REFERENCES AND NOTES

1. J. D. Jones, R. E. Vance, J. L. Dangl, *Science* **354**, aaf6395 (2016).
2. B. Laflamme et al., *Science* **367**, 763–768 (2020).
3. J. Tamborski, K. V. Krasileva, *Annu. Rev. Plant Biol.* **71**, 355–378 (2020).
4. I. M. L. Saur, R. Panstruga, P. Schulze-Lefert, *Nat. Rev. Immunol.* **21**, 305–318 (2020).
5. S. C. Saile et al., *PLOS Biol.* **18**, e3000783 (2020).
6. L. M. Jubic, S. Saile, O. J. Furzer, F. El Kasmi, J. L. Dangl, *Curr. Opin. Plant Biol.* **50**, 82–94 (2019).
7. D. Lapin, D. D. Bhandari, J. E. Parker, *Annu. Rev. Phytopathol.* **58**, 253–276 (2020).
8. C. Van Ghelder et al., *Sci. Rep.* **9**, 11614 (2019).
9. D. T. N. Tran et al., *Curr. Biol.* **27**, 1148–1160 (2017).
10. Z. Wu et al., *New Phytol.* **222**, 938–953 (2019).
11. S. J. Williams et al., *Mol. Plant Microbe Interact.* **24**, 897–906 (2011).
12. W. I. L. Tameling et al., *Plant Physiol.* **140**, 1233–1245 (2006).
13. T. Qi et al., *Proc. Natl. Acad. Sci. U.S.A.* **115**, E10979–E10987 (2018).
14. D. Lapin et al., *Plant Cell* **31**, 2430–2455 (2019).
15. A. R. Benthams, R. Zdrzalek, J. C. De la Concepcion, M. J. Banfield, *Plant Cell Physiol.* **59**, 2398–2408 (2018).
16. A. Daskalov et al., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2720–2725 (2016).
17. J. Wang et al., *Science* **364**, eaav5870 (2019).
18. B. Xia et al., *Cell Res.* **26**, 517–528 (2016).
19. Z. Cai et al., *Nat. Cell Biol.* **16**, 55–65 (2014).
20. J. Wang et al., *Science* **364**, eaav5868 (2019).
21. G. Bi et al., *Cell* **184**, 3528–3541.e12 (2021).
22. C. Seuring et al., *PLOS Biol.* **10**, e1001451 (2012).
23. D. R. Herr et al., *Neuromolecular Med.* **22**, 293–303 (2020).
24. J. Ding et al., *Nature* **535**, 111–116 (2016).
25. M. Huysmans, S. Lema A. N. S. Coll, M. K. Nowack, *Curr. Opin. Plant Biol.* **35**, 37–44 (2017).
26. W. Moeder, V. Phan, K. Yoshioka, *Plant Sci.* **279**, 19–26 (2019).
27. M. Grant et al., *Plant J.* **23**, 441–450 (2000).
28. M. J. Caterina et al., *Nature* **389**, 816–824 (1997).
29. F. Yuan et al., *Nature* **514**, 367–371 (2014).
30. X. Gao et al., *PLOS Pathog.* **9**, e1003127 (2013).
31. J. Castro, E. C. Aromataris, G. Y. Rychkov, G. J. Barritt, *Biochem. J.* **418**, 553–566 (2009).
32. K. De Vriese, A. Costa, T. Beeckman, S. Vanneste, *Int. J. Mol. Sci.* **19**, 1506 (2018).
33. J. R. Inglefield, R. D. Schwartz-Bloom, *Methods Enzymol.* **307**, 469–481 (1999).
34. H. Adachi et al., *eLife* **8**, e49956 (2019).
35. L. Tang et al., *Nature* **505**, 56–61 (2014).
36. M. Fan et al., *Nature* **582**, 129–133 (2020).
37. O. X. Dong et al., *New Phytol.* **210**, 960–973 (2016).
38. T. A. DeFalco et al., *Plant Cell Physiol.* **58**, 1173–1184 (2017).
39. S. Stael et al., *Trends Plant Sci.* **20**, 3–11 (2015).
40. K. Yoshioka et al., *Plant Cell* **18**, 747–763 (2006).
41. C. Zhao et al., *New Phytol.* **230**, 1078–1094 (2021).
42. W. Tian et al., *Nature* **572**, 131–135 (2019).
43. W. Gassmann, *Mol. Plant Microbe Interact.* **18**, 1054–1060 (2005).
44. L. K. Mahdi et al., *Cell Host Microbe* **28**, 813–824.e6 (2020).
45. U. Ros et al., *Cell Rep.* **19**, 175–187 (2017).

ACKNOWLEDGMENTS

We thank B. Staskawicz (UCB), R. Martin (UCB), J. Schroeder (UCSD), and S. Grant (UNC) for reading the manuscript; members of the Dangl laboratory for discussions about the work; B. Staskawicz and T. Qi for *N. benthamiana nrg adrl* seeds; K. Yoshioka (Univ. of Toronto) for *N. benthamiana GcAMP3* seeds; and V. Madden and K. White for electron microscopy sample preparation. UNC Microscopy Services are supported in part by Cancer Center Core Support Grant P30 CA016086 to the UNC Lineberger Comprehensive Cancer Center. **Funding:** This work was supported by the National Science Foundation (grant IOS-1758400 to J.L.D. and grant IOS-1457257 to Z.-M.P.), the National Institutes of Health (grants GM137286 and GM135218 to M.R.R.) and the Howard Hughes Medical Institute (HHMI). J.L.D. is an HHMI investigator. N.-H.K. was partially supported by the Basic Science Research Program through the National Research Foundation of Korea Fellowship funded by the Ministry of Education (grant 2014R1A6A3A03058629). F.E.-K. was supported by the University of Tübingen, the Deutsche Forschungsgemeinschaft (SFB/CRC1101

project D09), and the Reinhard Frank Stiftung (Project “helperless plant”) to F.E.-K. and J.L.D. L.W. was supported by National Key Laboratory of Plant Molecular Genetics, the Institute of Plant Physiology and Ecology/Center for Excellence in Molecular Plant Sciences, and the Chinese Academy of Sciences Strategic Priority Research Program (type B; project number XDB27040214).

Author contributions: P.J., N.H.K., L.W., and J.L.D. conceived the project. P.J., N.H.K., F.W., F.E.-K., Y.C., W.G.W., O.J.F., A.D.L., S.S., K.K., Z.-M.P., and L.W. generated and analyzed data and generated figures. M.R.R., Z.-M.P., and J.L.D. provided funding and project management. P.J., N.H.K., Z.-M.P., L.W., and J.L.D. wrote the paper. All other authors edited and commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Data are available in the manuscript, the supplementary materials, or in the Protein Data Bank (PDB) files. NRG1.1 N-terminal domain atomic

coordinates have been deposited in the PDB with accession codes 7L7W (2K/E native) and 7L7V (7K(R)/E native).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/420/suppl/DC1
Materials and Methods
Tables S1 to S4
Figs. S1 to S18
References (46–66)
MDAR Reproducibility Checklist

28 January 2021; accepted 7 June 2021
Published online 17 June 2021
10.1126/science.abg7917

ASTROPARTICLE PHYSICS

Peta-electron volt gamma-ray emission from the Crab Nebula

The LHAASO Collaboration*†

The Crab Nebula is a bright source of gamma rays powered by the Crab Pulsar’s rotational energy through the formation and termination of a relativistic electron-positron wind. We report the detection of gamma rays from this source with energies from 5×10^{-4} to 1.1 peta-electron volts with a spectrum showing gradual steepening over three energy decades. The ultrahigh-energy photons imply the presence of a peta-electron volt electron accelerator (a pevatron) in the nebula, with an acceleration rate exceeding 15% of the theoretical limit. We constrain the pevatron’s size between 0.025 and 0.1 parsecs and the magnetic field to ≈ 110 microgauss. The production rate of peta-electron volt electrons, 2.5×10^{36} ergs per second, constitutes 0.5% of the pulsar spin-down luminosity, although we cannot exclude a contribution of peta-electron volt protons to the production of the highest-energy gamma rays.

The Crab Nebula is a remnant of the explosion of a massive star. The nebula formed on 7 July 1054, recorded in Chinese chronicles as a “guest star” (1). It is the brightest known pulsar wind nebula—an extended nonthermal structure powered by the ultrarelativistic electron-positron wind from the central neutron star (the Crab Pulsar). This makes the Crab Nebula one of the brightest gamma-ray (γ -ray) sources in the sky, which has been observed for decades at tera-electron volt energies (2). At the transition from giga-electron volt to tera-electron volt energies, the spectral energy distribution [(SED) $E^2 dN / dE$, where dN / dE is the differential flux of radiation at the photon energy E] reaches a maximum of ~ 100 GeV (3, 4). Previous observations reported a detection at ≈ 80 TeV (5) and measured the spectrum up to 300 TeV (6–8). The angular size of the γ -ray source at tera-electron volt energies has been reported as ≈ 50 arc sec (9). The Crab Nebula was among the sources de-

tected at energies above 100 TeV using the Large High Altitude Air Shower Observatory (LHAASO) (10).

LHAASO is a dual-purpose complex of particle detectors, designed for the study of cosmic rays (CRs) and γ rays in the sub-tera-electron volt to 1000-PeV energy range (11, 12). When a very-high-energy extraterrestrial particle enters Earth’s atmosphere, it initiates a cascade consisting of secondary hadrons, leptons, and photons known as an air shower. The LHAASO detectors record different components of these air showers, which are used to reconstruct the type, energy, and arrival direction of the primary particles.

LHAASO consists of three arrays (13). The largest is the square kilometer array (KM2A), comprising surface counters and subsurface muon detectors. KM2A is designed to detect CRs from 10 TeV to 100 PeV and distinguish γ -ray photons from CRs using the muon detector array. The muon content in an air shower event can be used to effectively reject hadronic showers, which are initiated by CRs. The surface array of KM2A is used to determine the energy and arrival direction of the primary particles. KM2A operates as an ultrahigh-energy [(UHE) $E_\gamma > 0.1$ PeV] γ -ray telescope (14) with energy resolution of $\leq 20\%$ and angular resolution of 0.25° . For γ rays above 50 TeV,

*Corresponding authors: Zhen Cao (caozh@ihep.ac.cn); S. Z. Chen (chensz@ihep.ac.cn); S. J. Lin (linsj6@mail.sysu.edu.cn); S. S. Zhang (zhangss@ihep.ac.cn); M. Zha (zhama@ihep.ac.cn); Cong Li (licong@ihep.ac.cn); L. Y. Wang (wangly@ihep.ac.cn); L. Q. Yin (yinq@ihep.ac.cn); F. Aharonian (felix.aharonian@mpi-hd.mpg.de); R. Y. Liu (ryliu@nju.edu.cn) †The LHAASO Collaboration authors and affiliations are listed in the supplementary materials.

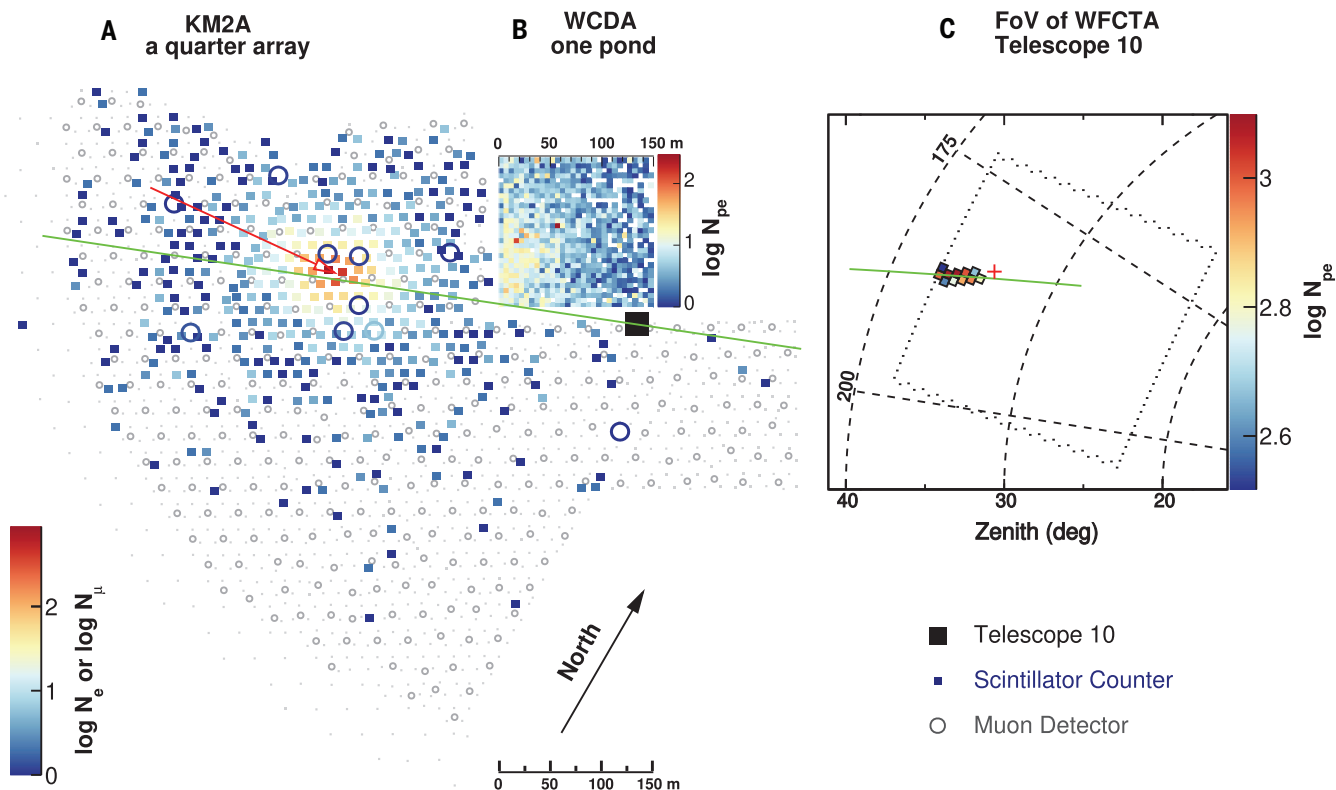


Fig. 1. The 0.88-PeV γ -ray event from the Crab recorded by the LHAASO detectors. (A) Squares indicate the scintillator counters of KM2A, colored according to the logarithm of the number of detected particles N_e (color bar). The open circles indicate the 10 muon detectors of KM2A triggered by the shower. The position of the core is indicated by the red arrow, which is orientated in the arrival direction of the primary photon. (B) Map of WCDA detector units. The logarithm of the number of photoelectrons recorded in each unit is indicated by the color. The scale is represented by the color bar. Telescope 10 of WFCTA (black square), which also detected the event, is located on the southern side of WCDA. (C) Telescope FOV

outlined by the dotted lines. The dashed arcs indicate zenith angles of 20°, 30°, and 40° from right to left, and dashed lines indicate azimuth angles of 175° and 200° counterclockwise. The shower image, composed of 11 pixels, started about 34° in zenith and stretched to the edge of the FOV at 38°. The color scale shows the logarithm of the number of photoelectrons (N_{pe}) in each pixel. The main axis of the image in the FOV of the telescope indicates the shower-telescope plane, which is consistent with the event direction [indicated by the red cross in (C)] reconstructed using KM2A. The green line in (A) is the intersection of the shower plane and the ground, which is consistent with the shower core.

the sensitivity of KM2A reaches the flux level of $10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ for a point source like the Crab Nebula in 1 year of observations (12).

The Water Cherenkov Detector Array (WCDA) consists of three ponds with a total area of 0.08 km^2 that are sensitive to γ rays down to 0.1 TeV with an angular resolution of $\delta\psi \approx 0.2^\circ$ (15). WCDA is designed to perform deep surveys of very-high-energy γ -ray sources, including the Crab Nebula for both pulsed and unpulsed signals. The sensitivity at energies $>2 \text{ TeV}$ reaches the flux level of $10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1}$ for a point source like the Crab Nebula in 1 year of observations (12). At energies above 0.1 PeV, WCDA also serves as an additional muon detector to enhance the determination between electromagnetic and hadronic showers detected by KM2A.

KM2A and WCDA are complemented by the Wide-Field-of-view Cherenkov Telescope Array (WFCTA) (16), designed to consist of 18 telescopes (of which 14 have been constructed) that detect Cherenkov radiation emitted by air showers induced by CRs, with energy rang-

ing from 0.1 to 1000 PeV. Cherenkov light in showers initiated by γ rays at energies above 0.1 PeV is recorded by WFCTA telescopes which cover a 32° by 112° region of the northern sky, through which the Crab Nebula passes every day.

LHAASO observations of the Crab Nebula

On 11 January 2020 at 17:59:18 coordinated universal time (UTC), a giant air shower was recorded by all three LHAASO detectors. The shower arrived from the direction of the Crab Nebula and was centered in the western part of KM2A (Fig. 1A). WCDA was also triggered, despite being 150 m away from the shower core. The event occurred after local midnight, when eight WFCTA telescopes were operational. The Cherenkov radiation from the air shower appeared in the field of view (FOV) of telescope no. 10 and triggered it (Fig. 1C).

We identify the event as a γ ray-induced shower, on the basis of 4996 particles (electrons, photons, muons, and hadrons) recorded

by 395 surface detectors and 15 muons recorded by 10 under-surface detectors of KM2A. The probability for this event to be a misidentified CR is estimated as 0.1% (13). Two independent estimates of the shower energy were derived from the KM2A and WFCTA data: $0.88 \pm 0.11 \text{ PeV}$ and $0.92^{+0.28}_{-0.20} \text{ PeV}$, respectively. The former value has previously been reported as the maximum energy of γ rays detected from the Crab Nebula by LHAASO (10).

Approximately 1 year later, on 4 January 2021 at 16:45:06 UTC, another shower was registered by KM2A at even higher energy— $1.12 \pm 0.09 \text{ PeV}$. This event occurred at zenith angle 12.9° , closer to vertical and therefore better measured by KM2A than the previous shower. Unfortunately, the primary photon arrived 1 hour before the Crab entered the FOV of the WFCTA telescopes. Although the number of detected secondary particles (5094) exceeded those in the previous event, the number of muons (14) was fewer, so we also identified this event as a γ ray-induced shower with a misidentification probability of 0.03%.

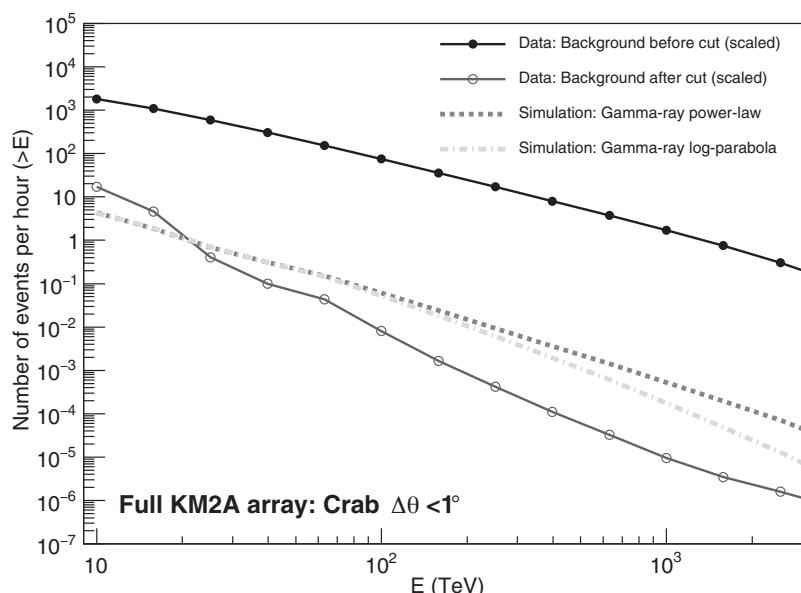


Fig. 2. The rates of detection of γ rays from the Crab and the CR background events above the shower energy E by the 1-km² array in a cone of 1° centered at the Crab direction. The rates correspond to the number of events per hour of observation of the source within the FOV of KM2A. E is the reconstructed shower energy. The dash-dotted and dashed lines represent the integrated detection rates of γ rays from the Crab, based on log-parabola and power-law models fitted to the measured fluxes (see Fig. 3), respectively. The filled circles show the rate of CR events accumulated by KM2A. The integrated rate can be approximately described as power law with index of -1.6 . The open circles represent the integrated rate of CR events after the muon-cut filter.

KM2A operated for 314 days at half of its design capacity and a further 87 days at three-quarters of its design capacity. In that time, a total of 89 UHE γ rays with energy exceeding 0.1 PeV were detected from the Crab. Figure 2 shows the integrated γ -ray detection rate, re-normalized to the nominal 1-km² array and 1-hour exposure time, assuming the Crab is within the FOV of KM2A (approximately 7.4 hours per day with zenith angle $<50^\circ$). Above 0.1 PeV, we find ~ 0.05 events per hour, equivalent to 135 events per year. Figure 2 also shows the detection rate of the CR-induced showers within a cone of 1° centered on the Crab, both before and after applying the CR shower rejection on the basis of the muon content in the showers. This muon-cut filter requires that the number of muons detected by KM2A in the shower must be less than $1/_{230}$ of the number of particles detected by the KM2A surface counters. The cut reduces the CR background by factors of 1000 and 500,000 at 50 TeV and 1 PeV, respectively. At energies above 0.1 PeV, the detection rate of γ rays from the Crab exceeds the CR-induced background by an order of magnitude. Because the point spread function (PSF) of KM2A is $\delta\psi \approx 0.25^\circ$, the CR background could be lower by an additional factor of $(1^\circ/\delta\psi)^2 \approx 16$.

Spectral energy distribution

The γ -ray fluxes in the energy range from 0.5 to 13 TeV were measured using the WCDA pond.

From September 2019 to October 2020, the total exposure was 343.5 transits of the Crab Nebula. The KM2A measurements in the observation period reported above cover the higher energy range from 10 TeV to 1.6 PeV. We combined the WCDA and KM2A data to determine the SED of the Crab Nebula, shown in Fig. 3A. The two independent measurements are consistent with a simple SED functional form, $dN/dE = (8.2 \pm 0.2) \times 10^{-14} (E/10 \text{ TeV})^{-\Gamma} \text{ cm}^{-2} \text{ s}^{-1} \text{ TeV}^{-1}$, where N is the number of γ rays, E is the γ -ray energy, and Γ is the energy-dependent spectral index. The two measurements are connected smoothly in the narrow overlapping region around 12.5 TeV. In this energy bin, the discrepancy between flux measured by KM2A and WCDA is 1.3σ . The overall χ^2/dof is 9.3/14, where dof is the number of degrees of freedom. No systematic deviation between the two segments of the SED is found. The functional form of the spectral index, $\Gamma = (2.90 \pm 0.01) + (0.19 \pm 0.02) \log_{10}(E/10 \text{ TeV})$, implies a gradual steepening of the spectrum characterized by the local index Γ , from ≈ 2.5 at 1 TeV to 3.7 at 1 PeV (Fig. 3B).

Figure 3A also shows previous measurements using atmospheric Cherenkov telescopes (5, 6, 17) and air shower arrays (7, 8, 18). These are consistent with the WCDA and KM2A data from sub-tera-electron volt to multi-tera-electron volt energies, including the flux at the highest energy previously reported, $E_\gamma \approx 300 \text{ TeV}$ (8).

Origin of the peta-electron volt photons and electrons

Photons from the Crab Nebula have been detected over 22 decades of energy, from megahertz radio to UHE γ rays, in both pulsed and unpulsed components. γ rays can be produced in three physically distinct sites—in the pulsar's magnetosphere, in the ultrarelativistic electron-positron (hereafter, simply electron) wind, and in the nebula. UHE γ rays are absorbed in the strong magnetic field of the pulsar, so a pulsed γ -ray component from the magnetosphere has been predicted only in the energy range of mega-electron volts to giga-electron volts. However, the detection of pulsed tera-electron volt γ rays from the Crab (19, 20) indicates that pulsed γ -ray emission occurs in the wind (21–23). Extension of this component to UHE γ rays is theoretically unexpected. We nevertheless searched for a pulsed component in the LHAASO data but did not detect one. We therefore assume that the entire flux detected by LHAASO consists of an unpulsed component that is produced in the nebula.

The broad-band nonthermal emission of the Crab Nebula is dominated by two mechanisms—synchrotron radiation and inverse Compton (IC) scattering of relativistic electrons interacting with the ambient magnetic and radiation fields (24, 25). In the standard paradigm, the acceleration of electrons is initiated by the termination of the wind, at a standing reverse shock located a distance $R \approx 0.1 \text{ pc}$ from the pulsar (26, 27). Although the details of the acceleration mechanism remain unknown, the detection of peta-electron volt photons allows us to estimate the accelerator size l , the magnetic field strength B , and the minimum acceleration rate η .

In the Crab Nebula, several radiation fields supply target photons for the IC scattering of electrons. However, at energies above 100 TeV, the 2.7-K cosmic microwave background radiation (CMBR) dominates the γ -ray production (25, 28). Because the CMBR is well quantified, the γ -ray data provide direct information about the parent electrons. γ -ray production above 100 TeV proceeds in the Klein-Nishina regime, where the energies of the upscattered photon E_γ and the parent electron E_e are linked through the simple relation $E_\gamma = 0.37(E_e/1 \text{ PeV})^{1.3} \text{ PeV}$, which, over two energy decades, from 30 TeV to 3 PeV, is accurate to within 10% (see fig. S7) or, equivalently

$$E_e \approx 2.15(E_\gamma/1 \text{ PeV})^{0.77} \text{ PeV} \quad (1)$$

Thus, for the 1.1-PeV photon, the energy of the parent electron is 2.3 PeV. Correspondingly, the mean energies of the synchrotron (ϵ_{syn}) and IC (E_γ) photons produced by the same electron in the ambient magnetic field B are related by

$$\epsilon_{\text{syn}} = 9.3(E_\gamma/1 \text{ PeV})^{1.5} (B/100 \mu\text{G}) \text{ MeV} \quad (2)$$

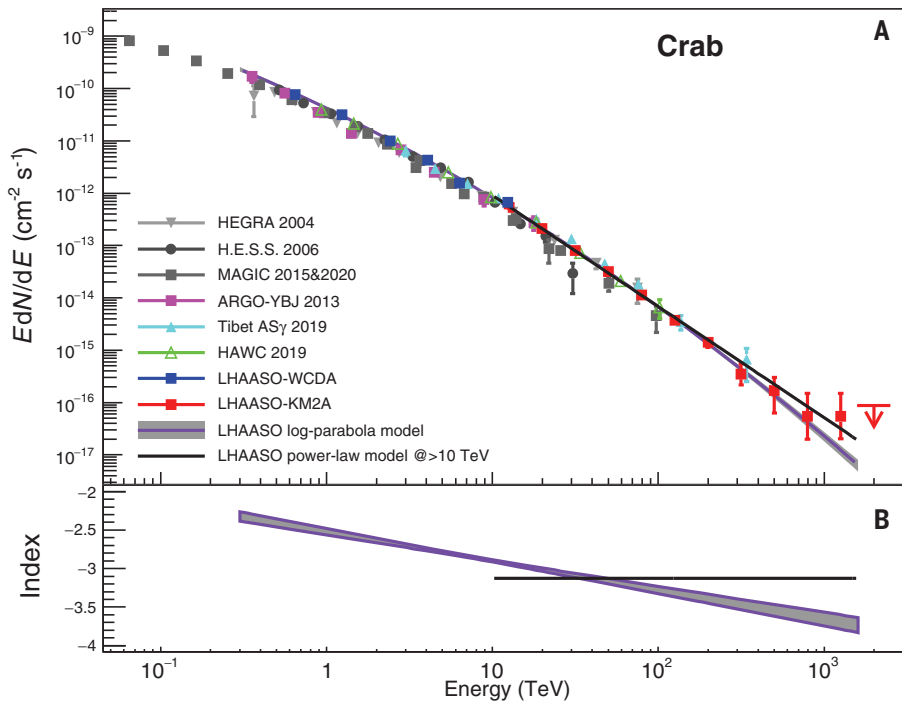


Fig. 3. γ -ray flux of the Crab measured by LHAASO and spectral fitting. (A) Tera-electron volt to peta-electron volt γ -ray fluxes of the Crab plotted as EdN/dE . The red squares and blue squares are the spectral points measured using KM2A and WCDA, respectively. The spectral points above 100 TeV were obtained in the signal-dominated regime, with 89 detected γ rays and two events expected from CR-induced (hadronic) air showers after the muon cuts. No events were detected in the 1.6- to 2.5-PeV bin, where an arrow indicates the flux upper limit at 90% confidence level. The purple line shows the fitting using a log-parabola model in the 0.3-TeV to 1.6-PeV interval (χ^2/dof :9.3/14). For comparison, the black line shows the fitting using a simpler power-law model in the 10-TeV to 1.6-PeV interval (χ^2/dof :5.4/9). Also plotted are previous observations of the Crab by other facilities: High Energy Gamma Ray Astronomy (5), High Energy Stereoscopic System (17), Major Atmospheric Gamma Imaging Cherenkov Telescope (4, 6), Astrophysical Radiation by Ground-based Observation at Yang Ba Jing (18), High Altitude Water Cherenkov Detector (7), and Tibet Air Shower array (8). **(B)** Energy-dependent local power-law index Γ derived by the log-parabola model fitting, as indicated by the purple band. For comparison, the black line shows the photon index 3.12 ± 0.03 derived from the simpler power-law model fitting. Error bars represent one standard deviation.

Simultaneous modeling (13) of the synchrotron and IC components constrains the magnetic field strength within a narrow interval, $B \approx 112^{+15}_{-13} \mu\text{G}$ (see Fig. 4 and discussion below). The upper limit is set by the requirement for the synchrotron radiation of the same electrons responsible for the production of 1-PeV photons to not overshoot the measured mega-electron volt flux (29). The lower limit is set by requiring the electron energy E_e and the accelerator's linear size, l , to meet the condition that the electron gyroradius $R_g = E_e/eB$ (where e is the charge of electron) cannot exceed l . Using Eq. 1, we find

$$(B/100 \mu\text{G})(l/1 \text{ pc}) \geq 0.023(E_\gamma/1 \text{ PeV})^{0.77} \quad (3)$$

Magnetohydrodynamic (MHD) models of the Crab Nebula (26, 27) postulate that electrons are accelerated at the termination of an electron-positron wind, then advected into the nebula through the MHD outflow. X-ray imaging of

the inner parts of the nebula (30) indicates that the acceleration site is located close to the termination shock, at $R \approx 0.1$ to 0.14 pc from the pulsar (28). The linear size of the accelerator must exceed the gyroradius, $l \geq R$; this imposes a lower limit on the magnetic field from Eq. 3, $B \geq 20 \mu\text{G}$. On the other hand, the standard one-zone model, which assumes that both the synchrotron and IC radiation components are produced in the same region by the same electron population, gives $B \approx 112 \mu\text{G}$ (Fig. 4). Then, from Eq. 3, we find that $l \geq 0.025 \text{ pc}$. These constraints are inconsistent with estimates of the characteristic size and magnetic field in the region(s) where the flares of the mega-electron volt or giga-electron volt γ -ray emissions ("Crab flares") (31) originate. Those variations of γ -ray fluxes on time scales of days are interpreted as fast acceleration and synchrotron cooling of peta-electron volt electrons in compact ($R \leq 0.01 \text{ pc}$) highly magnetized ($B \geq 1 \text{ mG}$) regions [see (32) for a review]. In the presence of such a large

magnetic field, the IC γ -ray component is suppressed; however, this does not exclude an indirect link between the peta-electron volt electrons responsible for the UHE γ -ray emission and the synchrotron mega-electron volt or giga-electron volt flares.

The detection of ~ 1 -PeV photons implies an acceleration rate that overcomes the synchrotron losses of the parent electrons up to peta-electron volt energies. The acceleration rate of electrons is $\dot{E} = e\mathcal{E}c = \eta eBc$, where η is the ratio of the projection of the electric field \mathcal{E} , averaged over the particle trajectory, to the magnetic field, $\eta = \mathcal{E}/B$. This parameter characterizes the acceleration efficiency and is always smaller than 1; in objects where $\eta \rightarrow 1$, the accelerator proceeds at the maximum rate allowed by classical electrodynamics and ideal MHD (33). The maximum energy of electrons is then determined by the balance between the acceleration and the energy loss rates: $E_{e,\text{max}} \approx 5.8\eta^{1/2}(B/100 \mu\text{G})^{-1/2} \text{ PeV}$. Using the relation between E_γ and E_e given by Eq. 1, we find

$$\eta = 0.14(B/100 \mu\text{G})(E_\gamma/1 \text{ PeV})^{1.54} \quad (4)$$

Thus, for the detected $E_\gamma = 1.1 \text{ PeV}$ and the magnetic field $B \approx 112 \mu\text{G}$ derived from the one-zone model (Fig. 4), we find that the acceleration proceeds at the efficiency $\eta \approx 0.16$. For comparison, at the diffusive shock acceleration in young supernova remnants (34), η is smaller by at least three orders of magnitude.

For the distance to the Crab Nebula, $d \approx 2 \text{ kpc}$ (35), the luminosity in peta-electron volt γ rays is estimated as $L_{\gamma,\text{PeV}} = 4\pi d^2 F_\gamma \approx 5 \times 10^{31} \text{ erg s}^{-1}$, where $F_\gamma \approx 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$ is the energy flux of 0.5- to 1.1-PeV γ rays (Fig. 4). In the Klein-Nishina limit, the IC cooling time of electrons in 2.7-K CMBR is $t_{\text{IC}} \approx 5 \times 10^{11}(E_e/1 \text{ PeV})^{0.7} \text{ s}$ (36) or, for the given E_γ , using Eq. 1 we have $t_{\text{IC}} \approx 8 \times 10^{11}(E_\gamma/1 \text{ PeV})^{0.54} \text{ s}$. Thus, the total energy held by the ≥ 1 -PeV electrons responsible for production of ≥ 0.5 -PeV γ rays is estimated as $W_{e,\text{PeV}} = L_{\gamma,\text{PeV}} t_{\text{IC}} \approx 4 \times 10^{43} \text{ erg}$. Because the overall cooling of peta-electron volt electrons is dominated by synchrotron losses ($t_{\text{syn}} \ll t_{\text{IC}}$), the injection rate of peta-electron volt electrons is $\dot{W}_{e,\text{PeV}} = W_{e,\text{PeV}}/t_{\text{syn}} \approx 2 \times 10^{36}(B/100 \mu\text{G})^2 \text{ erg s}^{-1}$. Thus, within the framework of the standard one-zone model with $B = 112 \mu\text{G}$, the acceleration power of peta-electron volt electrons constitutes an $\approx 0.5\%$ fraction of the pulsar's spin-down luminosity, $L_{\text{SD}} \approx 5 \times 10^{38} \text{ erg s}^{-1}$ (27).

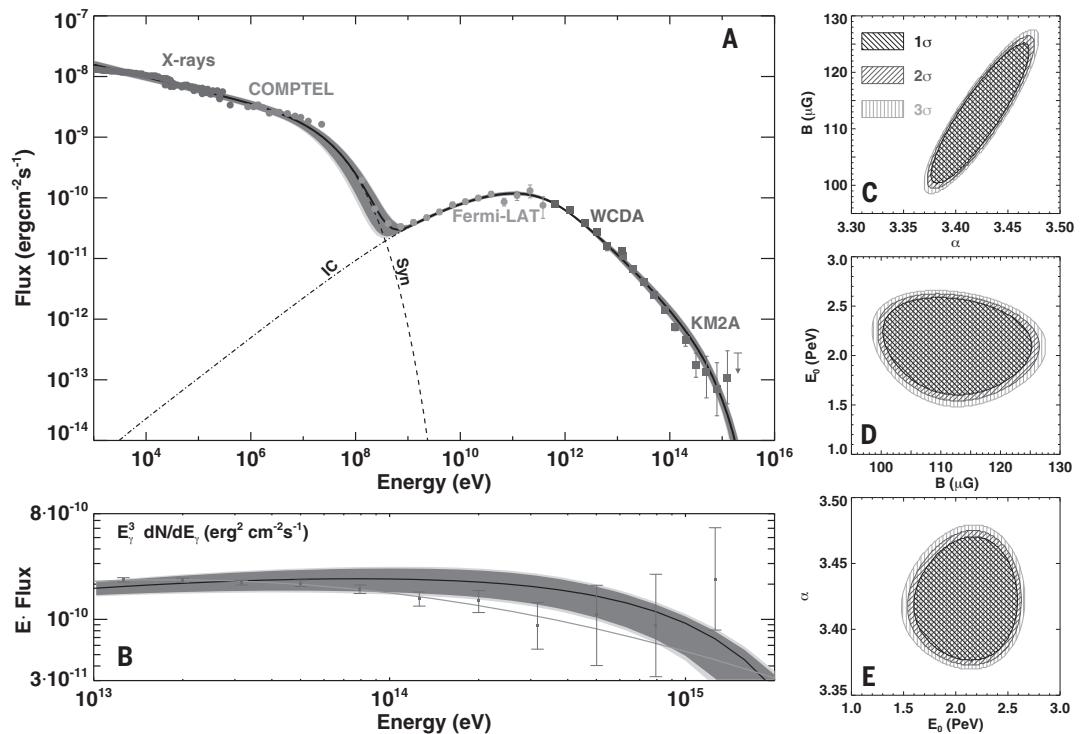
Multiwavelength modeling

We consider whether the detection of peta-electron volt photons agrees with predictions from the standard MHD paradigm of the Crab Nebula, which assumes that nonthermal emission from x-rays to multi-tera-electron volt γ rays is produced by electrons accelerated at the termination of the pulsar wind

Fig. 4. The spectral energy distribution of the Crab Nebula.

(A) The black curves represent the fluxes of the synchrotron (Syn) and IC components of radiation of an electron population calculated within the one-zone model. The electron spectrum above 1 TeV is assumed to be a power-law function terminated by a super-exponential cutoff, $E^{-\alpha} \exp[-(E/E_0)^2]$. The best-fitting model parameters are: $\alpha = 3.42 \pm 0.05$, $E_0 = 2.15^{+0.55}_{-0.65}$ PeV, and $B = 112^{+15}_{-13}$ μ G. The total energy in electrons above 1 TeV is $W_e = 7.7 \times 10^{47}$ erg. A break in the electron spectrum at $E_b = 0.76$ TeV is assumed to provide a consistency with the giga-electron volt γ -ray and low-frequency synchrotron data (see supplementary text and fig. S6). The dark gray and light gray shaded regions show the 1σ and 3σ uncertainty regions, respectively. The circles on the left

show the x-ray and the mega-electron volt emission of the Crab Nebula (29). The circles in the middle represent the Crab observations by Fermi–Large Area Telescope (LAT) in the nonflare state (3). The squares on the right represent WCDA and KM2A measurements reported in this work. (B) Zoom-in of the fluxes above 10 TeV, plotted as $E^3 dN/dE$. The thin curve presents the log-parabola spectral fitting shown in Fig. 3. (C to E) Two-dimensional projected parameter spaces of the free parameters α , B , and E_0 , with uncertainty regions indicated by the hatching shown in the legend in (C).



(25, 28, 37–39). We modeled the Crab’s multi-wavelength radiation within the idealized synchrotron-IC one-zone model, assuming a homogeneous spatial distribution of the magnetic field and electrons (Fig. 4). For $E_\gamma \geq 100$ TeV γ rays, the dominant target for IC scattering is the 2.7-K CMBR, with properties that are known more precisely than the targets at lower energies. For a steady-state electron energy distribution, above 1 TeV, we assumed a power-law function terminated by a super-exponential cutoff at the high-energy end. Figure 4 shows the SED model using three free parameters—the power-law slope $\alpha = 3.42$, cutoff energy $E_0 = 2.15$ PeV, and magnetic field $B = 112$ μ G, which reproduces the observations from the x-rays to multi-mega-electron volt γ rays with synchrotron radiation and the tera-electron volt to peta-electron volt γ rays with IC radiation. Below 1 TeV, the electron spectrum must undergo a break to avoid conflicting with the synchrotron radiation at optical to radio frequencies and to provide a smooth transition of the IC radiation from tera-electron volt to giga-electron volt energies (13).

The one-zone model in Fig. 4 is tightly constrained; the 3σ limits on the index α (3.37 to 3.47) are constrained by the uncertainty of the x-ray data. Because ≥ 10 -TeV electrons are cooled on short time scales, the index of the initial (acceleration) spectrum must be

close to 2.4. The magnetic field is determined by the flux ratio of the synchrotron and IC components, with 3σ range from 100 to 130 μ G. For these values of α and B , the cutoff energy E_0 is set by the synchrotron and IC spectra above 10 MeV and 100 TeV, respectively. The derived ranges of magnetic field and the power-law index of electrons are consistent with previous studies based on the synchrotron-IC one-zone model (28, 39) as well as MHD flow models (25, 38, 40) in which the magnetic field’s radial distribution is determined by the wind-magnetization parameter σ_B . The magnetic field $B \simeq 112$ G derived for the production region of multi-tera-electron volt to peta-electron volt γ rays is a factor of 2 to 3 times smaller than the average nebular magnetic field, consistent with the MHD flow model (27). The latter predicts reduced magnetic field at the termination shock for a broad range of the magnetization parameter σ_B between 0.001 and 0.01 (27).

Within the one-zone model, the IC γ -ray spectrum can be precisely calculated. Whereas the KM2A spectral points from 10 TeV to 1 PeV agree with the one-zone model within the statistical uncertainties, there are possible deviations from its predictions. Between 60 and 500 TeV, the two differ with a significance of 4σ , with the observational data having a steeper spectrum than the one-zone model predic-

tions. The possible excess around 1 PeV indicates an opposite tendency—a hardening of the spectrum. A hardening of the electron spectrum is difficult to accommodate theoretically with plausible assumptions. The problem of suppression of the one-zone spectrum at 1 PeV can be circumvented by introducing a second population of peta-electron volt electrons. This could also explain the inconsistency of the synchrotron part of the SED with the one-zone model (fig. S4) by decoupling the highest-energy synchrotron and IC components, assuming that the mega-electron volt synchrotron radiation is predominantly produced in compact, highly magnetized regions, whereas the peta-electron volt IC photons originate from regions with $B \leq 100$ μ G. A second electron component could extend the SED to a few peta-electron volts but not much further. From Eq. 4, it follows that, even for $\eta = 1$ and minimum allowed magnetic field, $B \geq 20$ μ G, then the maximum energy of photons cannot exceed 4 PeV. Any detection of γ rays well beyond 1 PeV would require a nonleptonic origin for the extra component of radiation—e.g., multi-peta-electron volt protons and atomic nuclei in the nebula.

Because of the limited energy budget available for acceleration of protons, hadronic interactions cannot be responsible for the overall broad-band γ -ray luminosity. However, the contribution of hadronic interactions

at peta-electron volt energies could be non-negligible. The γ -ray production efficiency of hadronic interactions—i.e., the fraction of the proton kinetic energy converted to γ rays—is determined by the ratio $\kappa = t_{\text{esc}}/t_{\text{pp}}$, where t_{esc} is the confinement time of protons inside the nebula, and $t_{\text{pp}} \approx 1.5 \times 10^8 (n/1 \text{ cm}^{-3})^{-1}$ years is the cooling time of protons through the production and decay of the secondary π^0 -mesons. For the average density of the nebular gas $n \approx 10 \text{ cm}^{-3}$, the radiation efficiency is low; even for the most effective confinement of 10-PeV protons, $t_{\text{esc}} \leq 250$ years (13), thus $\kappa \leq 5 \times 10^{-5}$. To explain the peta-electron volt γ -ray luminosity, $L_\gamma \approx 5 \times 10^{31} \text{ erg s}^{-1}$, the acceleration power of ~ 10 -PeV parent protons would need to be $\dot{W}_p = W_p/t_{\text{esc}} = \kappa^{-1} L_\gamma \approx 10^{36} \text{ erg s}^{-1}$ or, in the case of a broad E^{-2} -type proton spectrum, an order of magnitude larger. These estimates are supported by numerical calculations (fig. S5). They avoid exceeding the theoretical constraints on the proton fraction (41) only if there is effective proton confinement in the nebula.

REFERENCES AND NOTES

1. K. Lundmark, *Publ. Astron. Soc. Pac.* **33**, 225 (1921).
2. S. Funk, *Annu. Rev. Nucl. Part. Sci.* **65**, 245–277 (2015).
3. M. Arakawa, M. Hayashida, D. Khangulyan, Y. Uchiyama, *Astrophys. J.* **897**, 33 (2020).
4. J. Aleksić et al., *J. High Energy Astrophys.* **5–6**, 30–38 (2015).
5. F. Aharonian et al., *Astrophys. J.* **614**, 897 (2004).
6. MAGIC Collaboration, *Astron. Astrophys.* **635**, A158 (2020).
7. A. U. Abeysekara et al., *Astrophys. J.* **881**, L34 (2019).
8. M. Amenomori et al., *Phys. Rev. Lett.* **123**, 051101 (2019).
9. H.E.S.S. Collaboration, *Nat. Astron.* **4**, 167–173 (2020).
10. Z. Cao et al., *Nature* **594**, 33–36 (2021).
11. Z. Cao et al., *Chinese Phys. C* **34**, 1852 (2010).
12. H. He, LHAASO Collaboration, *Radiat. Detect. Technol. Methods* **2**, 7 (2018).
13. Materials and methods are available as supplementary materials.
14. LHAASO Collaboration, *Chinese Phys. C* **45**, 025002 (2021).
15. LHAASO Collaboration, Performance of LHAASO-WCDA and Observation of Crab Nebula as a Standard Candle. arXiv:2101.03508 [astro-ph.IM] (2021).
16. LHAASO Collaboration, Construction and On-site Performance of the LHAASO WFCTA Camera. arXiv:2012.14622 [physics.ins-det] (2020).
17. F. Aharonian et al., *Astron. Astrophys.* **457**, 899–915 (2006).
18. B. Bartoli et al., *Astrophys. J.* **779**, 27 (2013).
19. VERITAS Collaboration, *Science* **334**, 69–72 (2011).
20. S. Ansoldi et al., *Astron. Astrophys.* **585**, A133 (2016).
21. F. A. Aharonian, S. V. Bogovalov, D. Khangulyan, *Nature* **482**, 507–509 (2012).
22. I. Mochol, J. Petri, *Mon. Not. R. Astron. Soc.* **449**, L51–L55 (2015).
23. P. K. H. Yeung, *Astron. Astrophys.* **640**, A43 (2020).
24. O. C. de Jager, A. K. Harding, *Astrophys. J.* **396**, 161 (1992).
25. A. M. Atoyan, F. A. Aharonian, *Mon. Not. R. Astron. Soc.* **278**, 525–541 (1996).
26. M. J. Rees, J. E. Gunn, *Mon. Not. R. Astron. Soc.* **167**, 1–12 (1974).
27. C. F. Kennel, F. V. Coroniti, *Astrophys. J.* **283**, 694 (1984).
28. M. Meyer, D. Horns, H. S. Zechlin, *Astron. Astrophys.* **523**, A2 (2010).
29. L. Kuiper et al., *Astron. Astrophys.* **378**, 918–935 (2001).
30. M. C. Weisskopf et al., *Astrophys. J.* **536**, L81–L84 (2000).
31. M. Tavani et al., *Science* **331**, 736–739 (2011).
32. R. Bühler, R. Blandford, *Rep. Prog. Phys.* **77**, 066901 (2014).
33. F. A. Aharonian, A. A. Belyanin, E. V. Derishev, V. V. Kocharovskiy, V. V. Kocharovskiy, *Phys. Rev. D* **66**, 023005 (2002).
34. M. A. Malkov, L. O. Drury, *Rep. Prog. Phys.* **64**, 429–481 (2001).
35. D. L. Kaplan, S. Chatterjee, B. M. Gaensler, J. Anderson, *Astrophys. J.* **677**, L201–L215 (2008).
36. D. Khangulyan, F. A. Aharonian, S. R. Kelner, *Astrophys. J.* **783**, 100 (2014).
37. O. C. de Jager et al., *Astrophys. J.* **457**, 253 (1996).
38. D. Volpi, L. Del Zanna, E. Amato, N. Bucciantini, *Astron. Astrophys.* **485**, 337–349 (2008).
39. D. Khangulyan, M. Arakawa, F. Aharonian, *Mon. Not. R. Astron. Soc.* **491**, 3217–3224 (2020).
40. N. Bucciantini, J. Arons, E. Amato, *Mon. Not. R. Astron. Soc.* **410**, 381–398 (2011).
41. J. Arons, *Astrophys. J.* **589**, 871–892 (2003).

ACKNOWLEDGMENTS

We thank all staff members who work at the LHAASO site, located above 4400 m above sea level, year-round to maintain the detector and keep the electricity power supply and other components of the experiment operating smoothly. **Funding:** This work was supported in China by the National Key R&D program of China under grants 2018YFA0404201, 2018YFA0404202, 2018YFA0404203, and 2018YFA0404204; by NSFC under grant nos. 12022502, 11905227, U1831208, U1931112, 11635011, 11761141001, 11905240, 11675204, 11475190, U2031105, and U1831129; and in Thailand by grant RTA6280002 from Thailand Science Research and Innovation. Chengdu Management Committee of Tianfu New Area provided financial support for research with LHAASO data. **Author contributions:** Z. Cao and F. Aharonian led the drafting of the text about the experiment and interpretation, respectively. S. Z. Chen led the KM2A data analysis with the team members C. Li and L. Y. Wang and others. L. Y. Wang carried out event-by-event analysis and corresponding simulation. S. J. Lin and M. Zha led teams to conduct the spectrum analysis using WCDA data and corresponding cross checking. S. S. Zhang led the WFCTA team, including L. Q. Yin, to perform the combined analysis of all data from WFCTA, WCDA, and KM2A for commonly registered events. L. Q. Yin carried out the specific multicomponent analysis and corresponding simulation. F. Aharonian and R. Y. Liu conducted the theoretical interpretation. R. Y. Liu produced all corresponding calculations and all figures and edited the manuscript. Z. Cao is the spokesperson of the LHAASO Collaboration and the principal investigator of the LHAASO project in China and led the specific working group for this paper involving all corresponding authors. The Editorial Board of the LHAASO

Collaboration was led by S. M. Liu and D. della Volpe, who organized internal review before submission. B. D'Elterre Piazzoli participated in manuscript revision and corresponding discussion. All other authors participated in data analysis, including event reconstruction; simulation and event building with multicomponents; detector calibration; and operation and maintenance of all scintillator counters, muon detectors, water ponds, and Cherenkov telescopes. The detector arrays were still under construction while the data were collected for this paper. Many authors participated in the detector construction and deployment. **Competing interests:** There are no conflicts of interests to declare for any collaboration members. **Data and materials availability:** Data and software to reproduce these results are available on the LHAASO public data web page at <http://english.ihep.cas.cn/doc/4035.html>, including the event list, results of the background rate calculations, numerical values of the derived WCDA and KM2A spectra (Fig. 3) and SED (Fig. 4), and the code used to produce the SED and significance maps (fig. S1). The software we wrote for simulation of the detectors and background rate calculations is restricted by the terms of a grant from China's National Commission of Development and Reform, due to legal restrictions imposed during the construction phase of LHAASO. Readers who are willing to become associated members of the LHAASO Collaboration can request a copy under the condition that any resulting paper is authored by all collaboration members, including associated members. Interested readers should contact the corresponding authors for details.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/425/suppl/DC1
LHAASO Collaboration Author List
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 and S2
References (42–64)

12 January 2021; accepted 23 June 2021
Published online 8 July 2021
10.1126/science.abg5137

REPORTS

QUANTUM DEVICES

Coherent manipulation of an Andreev spin qubit

M. Hays^{1*}, V. Fatemi^{1*}, D. Bouman^{2,3}, J. Cerrillo^{4,5}, S. Diamond¹, K. Serniak^{1†}, T. Connolly¹, P. Krogstrup⁶, J. Nygård⁶, A. Levy Yeyati^{5,7}, A. Geresdi^{2,3,8}, M. H. Devoret^{1*}

Two promising architectures for solid-state quantum information processing are based on electron spins electrostatically confined in semiconductor quantum dots and the collective electrodynamic modes of superconducting circuits. Superconducting electrodynamic qubits involve macroscopic numbers of electrons and offer the advantage of larger coupling, whereas semiconductor spin qubits involve individual electrons trapped in microscopic volumes but are more difficult to link. We combined beneficial aspects of both platforms in the Andreev spin qubit: the spin degree of freedom of an electronic quasiparticle trapped in the supercurrent-carrying Andreev levels of a Josephson semiconductor nanowire. We performed coherent spin manipulation by combining single-shot circuit–quantum-electrodynamics readout and spin-flipping Raman transitions and found a spin-flip time $T_S = 17$ microseconds and a spin coherence time $T_{2E} = 52$ nanoseconds. These results herald a regime of supercurrent-mediated coherent spin-photon coupling at the single-quantum level.

A weak link between two superconductors hosts discrete, fermionic modes known as Andreev levels (1). These levels govern the physics of the weak link on the microscopic scale, ultimately giving rise to macroscopic phenomena such as the Josephson supercurrent. Superconducting electrodynamic qubits rely on the nonlinearity of the supercurrent in Josephson tunnel junctions, a manifestation of the ground-state properties of

millions of Andreev levels acting in concert (2). Although the vast majority of conduction electrons undergo collective nonlinear bosonic oscillations in the superconducting condensate, each Andreev level can be occupied by a fermionic electronic excitation known as a Bogoliubov quasiparticle.

It was proposed that quantum information could be stored in the spin of a quasiparticle trapped in a weak link possessing a spin-orbit

interaction (3). This Andreev spin qubit (ASQ) would carry a spin-dependent supercurrent, opening paths for spin manipulation and measurement unavailable to electrostatically confined spin qubits (4–7). In particular, this supercurrent could provide strong coupling between the spin and a superconducting microwave resonator (8–13). In the context of such supercurrent-based integration of Andreev levels with circuit quantum electrodynamics (cQED), pairs of quasiparticles have been coherently generated and recombined (14, 15). However, because Andreev levels in most weak links are paired into spin-degenerate doublets, quasiparticle spin manipulation has remained elusive.

Weak links composed of superconductor-proximitized semiconductor nanowires (“Josephson nanowires” for short) offer a platform to investigate spinful Andreev physics (16–18). The Andreev spectra of Josephson nanowires was recently observed, owing to the development of an atomically perfect superconductor-semiconductor interface. This in turn revealed a rich interplay between superconductivity, spin-orbit interaction, electromagnetic field effects, and device geometry. Properties of Andreev levels in superconductor-semiconductor nanowires have been exploited to realize gate-tunable weak links for superconducting qubits (19, 20), probe non-abelian Andreev levels known as Majorana zero modes (21), and relevant to this experiment, investigate spin-split doublets without a Zeeman field (18, 22).

We performed coherent manipulation of the spin of an individual quasiparticle excitation of a superconductor. A naturally occurring quasiparticle is stochastically trapped in the Andreev levels of a Josephson nanowire (14, 15), where it then resides predominantly in the two spin states of the lowest-energy Andreev doublet with roughly equal probability (22). First, we initialized this ASQ by post-selecting on a single-shot cQED spin measurement (22). We then achieved coherent control of the ASQ by driving Raman transitions in a Λ system formed by the two spin states and an auxiliary higher-energy Andreev level. We observed spin

lifetimes of up to $T_S = 17 \mu\text{s}$ and a spin coherence time $T_{2E} = 52 \text{ ns}$, which appears to be limited by a spinful bath.

Our realization of the ASQ hinges on the interplay between spin-orbit interaction in the semiconductor nanowire and the superconducting phase bias across the weak link (Fig. 1A) (3–6, 18, 22). In a conventional weak link, a trapped quasiparticle is restricted to spin-degenerate Andreev doublets, and therefore the spin cannot be coherently manipulated. In a Josephson nanowire, however, an intersubband spin-orbit interaction can cause spin to hybridize with translational degrees of freedom (this hybridized spin is sometimes known as pseudospin; we will continue to refer to it as “spin” for simplicity). Because of this interaction between spin and motion, the two spin states of an Andreev doublet carry equal and opposite supercurrent $\pm I_S/2$ at $\varphi = 0$, with I_S being doublet dependent (5). Then, because energy varies linearly with $|Id\varphi$, the doublet degeneracy can be lifted with a nonzero phase bias: Perturbatively near $\varphi = 0$, the spin splitting is $\epsilon_s = I_S\varphi\Phi_0/2\pi$.

Microwave quantum optics techniques are well suited to achieve quasiparticle spin manipulation, given such a phase-induced spin

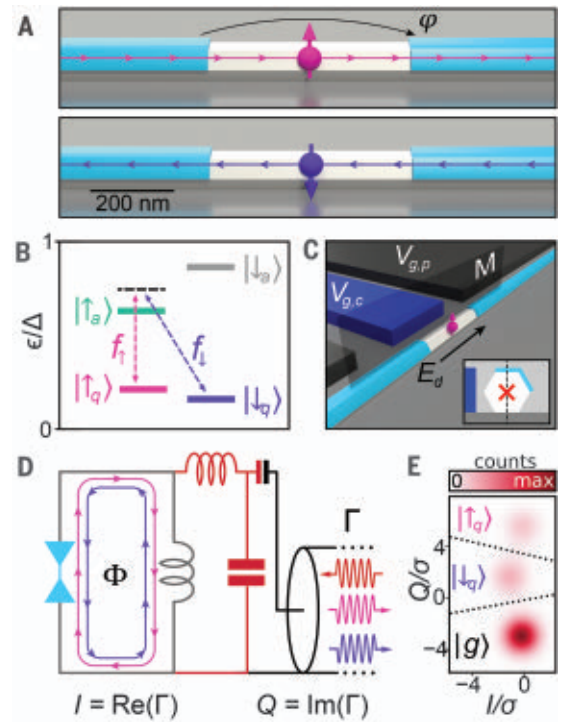
splitting. In this experiment, the two spin states $|\downarrow_q\rangle, |\uparrow_q\rangle$ of one Andreev doublet form the qubit basis (Fig. 1B), while a second, higher-energy doublet provides auxiliary states $|\downarrow_a\rangle, |\uparrow_a\rangle$ critical for both qubit control and measurement (22). The number of Andreev doublets is given roughly by the ratio between the weak link length (here, $\sim 500 \text{ nm}$) and the superconducting coherence length (typically $\sim 200 \text{ nm}$ for the presented device parameters) (5, 6, 18). To manipulate the ASQ, we used the qubit states $|\downarrow_q\rangle, |\uparrow_q\rangle$ in conjunction with $|\uparrow_a\rangle$ as a Λ system. We applied simultaneous microwave drives to both the $|\uparrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ transition (drive frequency f_1) and $|\downarrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ (drive frequency f_2). By equally detuning the two drives from their respective transitions, a Raman process is induced so that the $\{|\downarrow_q\rangle, |\uparrow_q\rangle\}$ manifold can be coherently manipulated while $|\uparrow_a\rangle$ remains minimally populated.

The success of the Raman process is contingent on driving both the spin-conserving transition $|\uparrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ and the spin-flipping transition $|\downarrow_q\rangle \leftrightarrow |\uparrow_a\rangle$. In principle, a radio frequency (rf) electric field can flip spin by coupling to the spatial character of the spin-orbit hybridized Andreev levels (23). However, as we outline here [(24), section 1], a broken spatial

Fig. 1. Principle of the ASQ.

(A) Illustration of a semiconductor nanowire (white) coated with epitaxial superconducting leads (light blue). The pair potential Δ of the leads confines the quasiparticle to the weak link. Because of spin-orbit interaction, if the quasiparticle is in the spin-up state (top), supercurrent flows to the right near zero phase bias $\varphi = 0$, whereas in the spin-down state (bottom), supercurrent flows to the left. Nonzero φ thus breaks spin degeneracy. (B) Energies of two Andreev doublets tuned to a Λ configuration. Two microwave drives (frequencies f_1, f_2) are equally detuned from $|\uparrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ and $|\downarrow_q\rangle \leftrightarrow |\uparrow_a\rangle$, inducing a Raman process between the qubit states $|\uparrow_q\rangle$ and $|\downarrow_q\rangle$ by way of a virtual level (black dashed line).

(C) Both drives induce an rf electric field E_d between the superconducting leads. But for a nanowire symmetric across the plane M (only nanowire + substrate), drive-induced spin-flips would be forbidden. (Inset) However, the mirror symmetry is broken by both



the partial aluminum shell as well as the cutter (blue, bias $V_{g,c}$) and plunger gates (black, bias $V_{g,p}$). (D) The Josephson nanowire (light blue) is embedded in a superconducting loop (gray), which enables phase bias through an external flux $\varphi \approx 2\pi\Phi/\Phi_0$ as well as inductive coupling to a superconducting microwave resonator (dark red). The resonator reflection coefficient $\Gamma = I + iQ$ is probed with a tone near its fundamental frequency $f_r = 9.188 \text{ GHz}$. (E) Histogram of repeated $1.9\text{-}\mu\text{s}$ measurements of Γ clustered into three distributions, corresponding to $|\uparrow_q\rangle, |\downarrow_q\rangle$, and $|g\rangle$ (standard deviation σ). The system state was assigned according to thresholds indicated with the black dotted lines.

¹Department of Applied Physics, Yale University, New Haven, CT 06520, USA. ²QuTech and Delft University of Technology, 2600 GA Delft, Netherlands. ³Kavli Institute of Nanoscience, Delft University of Technology, 2600 GA Delft, Netherlands. ⁴Área de Física Aplicada, Universidad Politécnica de Cartagena, E-30202 Cartagena, Spain. ⁵Departamento de Física Teórica de la Materia Condensada C-V, Universidad Autónoma de Madrid, E-28049 Madrid, Spain. ⁶Center for Quantum Devices, Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark. ⁷Condensed Matter Physics Center (IFIMAC) and Instituto Nicolás Cabrera, Universidad Autónoma de Madrid, E-28049 Madrid, Spain. ⁸Quantum Device Physics Laboratory, Department of Microtechnology and Nanoscience, Chalmers University of Technology, SE 41296 Gothenburg, Sweden.

*Corresponding author. Email: max.hays@yale.edu (M.H.); valla.fatemi@yale.edu (V.F.); michel.devoret@yale.edu (M.H.D.)

†Present address: MIT Lincoln Laboratory, 244 Wood Street, Lexington, MA 02420, USA.

symmetry of the Josephson nanowire system is also required. Our hexagonal nanowire was made of [001] wurtzite indium arsenide grown by use of molecular beam epitaxy. Such a nanowire lying alone on a substrate would possess a transverse mirror symmetry (Fig. 1C); this property would then be inherited by the levels of the nanowire so that one spin state of each doublet would be mirror-symmetric and the other spin state mirror-antisymmetric (mirror character coincides with spin component character). Because we applied the drive voltage along the weak link, the rf electric field did not break the mirror symmetry—it points along the nanowire—and therefore would not flip spin.

In our device, the mirror symmetry is broken by both the superconducting leads and

the electrostatic gates (Fig. 1C) as well as any symmetry-breaking disorder present in the nanowire. The superconducting leads consist of 10-nm-thick epitaxial aluminum, of which a 500-nm length was removed to form the weak link. The aluminum only covers two of six nanowire facets, breaking the mirror symmetry of the nanowire-substrate system. Because the gates are fabricated on one side of the nanowire, they also break the mirror symmetry. Both the cutter and plunger gates were used to tune the transparency of the weak link and were biased to $V_{gc} = -71.9$ mV and $V_{gp} = 4.0$ mV, respectively, unless otherwise noted [system tune up is provided in (24), section 2].

As mentioned above, the ASQ appears spontaneously when a quasiparticle is stochasti-

cally trapped in the Josephson nanowire. These nonequilibrium quasiparticles are ubiquitous in superconducting circuits and likely originate from background ionizing radiation and infrared photons (25, 26). By embedding the Josephson nanowire in a cQED architecture (Fig. 1D), the effect of spin-orbit interaction can be harnessed to determine whether a trapped quasiparticle is in the “spin down” state $|\downarrow_q\rangle$ or the “spin up” state $|\uparrow_q\rangle$ or whether the weak link is in the ground state $|g\rangle$, where no quasiparticles are present (Fig. 1E) (22). For the bias conditions presented in this work, a quasiparticle remained trapped in the weak link on average 22 ± 1 μ s and had a spin-flip lifetime $T_S = 17 \pm 1$ μ s [(24), section 6]. Below, we present data in terms of spin state occupation probabilities P_\uparrow, P_\downarrow computed on the basis of the thresholds displayed in Fig. 1E. Because the two spin states were occupied with roughly equal probability, under any coherent manipulation $|\downarrow_q\rangle \leftrightarrow |\uparrow_q\rangle$ the observed spin state populations would not change. Throughout this work, we overcame this problem by preparing the quasiparticle in $|\uparrow_q\rangle$ by means of an initial readout pulse and post-selection $|\downarrow_q\rangle$ post-selection is provided in (24)].

The first step in driving the Raman process (Fig. 1B) was to locate the two transitions that defined the Λ system: $|\uparrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ and $|\downarrow_q\rangle \leftrightarrow |\uparrow_a\rangle$. After breaking spin degeneracy with $\Phi = -0.010\Phi_0$ (Fig. 1D), we measured the spectrum (Fig. 2A) using two-tone spectroscopy, without spin initialization. The dip in P_\uparrow at 13.00 GHz corresponds to the drive coming into resonance with the $|\uparrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ transition, resulting in population transfer out of $|\uparrow_q\rangle$ and into $|\uparrow_a\rangle$. Similarly, the dip in P_\downarrow at 13.68 GHz corresponds to the $|\downarrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ transition. Taking the difference yields the spin splitting $\epsilon/\hbar = 680$ MHz.

After initializing the quasiparticle in $|\uparrow_q\rangle$, we applied two simultaneous Gaussian pulses with variable respective carrier frequencies f_\uparrow and f_\downarrow and then measured the final state (Fig. 2B) [a wider frequency range is available in (24), section 3]. Along the line $f_\downarrow = f_\uparrow + 610$ MHz, we observed an increased $|\downarrow_q\rangle$ population that we attribute to a Raman process. As expected for Raman transitions, the slope of the line is unity because a shift of one drive frequency must be compensated by an equal shift of the other. The discrepancy between the spin splitting $\epsilon/\hbar = 680$ MHz and the 610 MHz offset is due to an uncontrolled jump of the Andreev spectrum that occurred between the measurements shown in Fig. 2, A and B (24).

We then proceeded to coherent manipulation of the spin of an individual quasiparticle. Parking f_\uparrow, f_\downarrow on resonance with the Raman process [(24), section 4], we varied the amplitudes A_\uparrow, A_\downarrow of the two Gaussian pulses before determining the final quasiparticle state (Fig. 3A). The observed oscillations in the spin population difference are characteristic of a

Fig. 2. Raman transitions of a trapped quasiparticle. (A) Two-

tone spectroscopy of the $|\uparrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ and $|\downarrow_q\rangle \leftrightarrow |\uparrow_a\rangle$ transitions consisting of a saturation pulse (gray, 1 μ s long with variable carrier frequency f_d) followed by a readout pulse (dark red). (B) The quasiparticle was first prepared in $|\uparrow_q\rangle$ by means of an initial readout pulse and postselection. Simultaneous Gaussian pulses [235 ns full width at half maximum, 30 dB more power than used in (A)] with variable frequencies f_\uparrow, f_\downarrow were then applied, followed by a final readout pulse. Color shows probability of measuring $|\downarrow_q\rangle$, and the peak in the final $|\downarrow_q\rangle$ population lies along $f_\downarrow = f_\uparrow + 610$ MHz (black dashed line). (C) Full Γ histograms of the final readout pulse for the two subsets of measurements enclosed by the gray and black solid lines in (B). Data accrued in the region enclosed by the black line (right) show appreciable population transfer $|\uparrow_q\rangle \rightarrow |\downarrow_q\rangle$ compared with data enclosed by the gray line (left).

Fig. 3. Coherent Λ -Rabi oscillations of the ASQ. $f_\downarrow = 13.280$ GHz; $f_\uparrow =$

13.964 GHz. Color represents the probability difference between $|\uparrow_q\rangle$ and $|\downarrow_q\rangle$ after the drive.

(A) Varying the amplitudes A_\uparrow, A_\downarrow of the simultaneous drive pulses (94 ns full width at half maximum) resulted in oscillations between $|\uparrow_q\rangle$ and $|\downarrow_q\rangle$.

(B) Simulated dynamics of the quasiparticle under the action of the drive pulses. The reduced contrast observed in (A) is taken into account by using the measured readout fidelities.

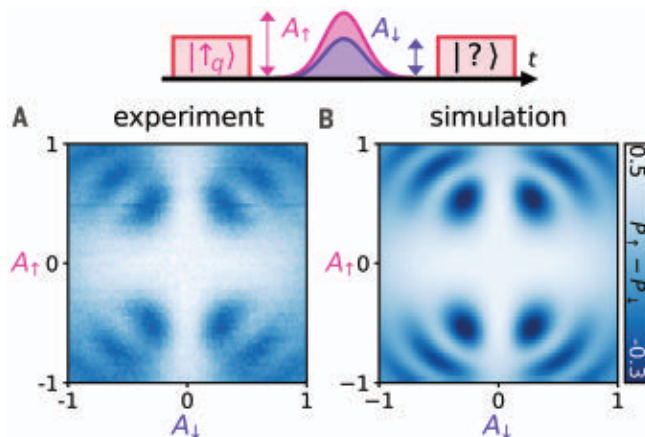
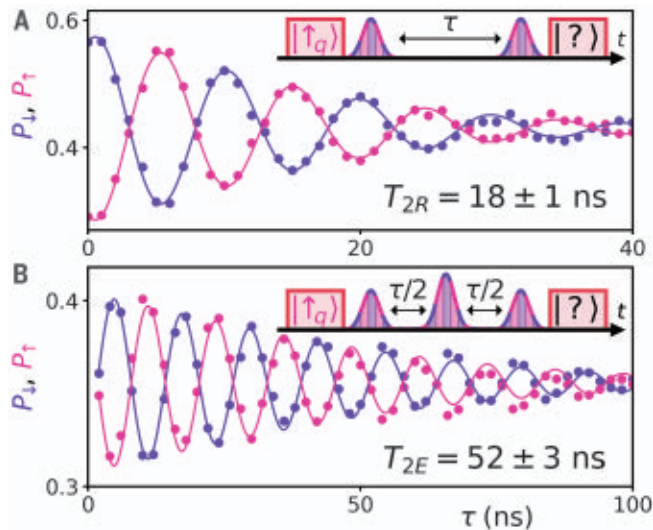


Fig. 4. Coherence decay of the ASQ.

$V_{g,c} = -59.1$ mV; $V_{g,p} = -33.3$ mV; $\Phi = -0.115\Phi_0$. (A) Ramsey and (B) Hahn-echo experiments reveal $T_{2R} = 18 \pm 1$ ns and $T_{2E} = 52 \pm 3$ ns, respectively (supplementary materials, materials and methods). Oscillations were introduced in both cases by adding a phase proportional to τ to the final Raman pulse. The smaller value of $P_{\uparrow} + P_{\downarrow}$ in (B) as compared with (A) is likely due to additional quasiparticle detrapping $|\downarrow_q\rangle, |\uparrow_q\rangle \rightarrow |g\rangle$ caused by the echo pulse.



coherent Raman process. Qualitatively, when either $A_{\uparrow} = 0$ or $A_{\downarrow} = 0$, there is no population transfer because both drives are required for the Raman process. As the amplitudes of both drives are increased (roughly along the diagonals $|A_{\uparrow}| \cong |A_{\downarrow}|$), the spin population difference undergoes coherent oscillations. As expected, the data are symmetric under $A_{\uparrow} \rightarrow -A_{\uparrow}$ and $A_{\downarrow} \rightarrow -A_{\downarrow}$. Quantitatively, we verified this interpretation by using a Lindblad master equation to numerically calculate the dynamics induced by the drive pulse (Fig. 3B) (24). We then fit the simulation to the measured data by varying the four interdoublet transition matrix elements, a slight detuning from the Raman resonance condition (fitted value 5.5 ± 0.1 MHz), and a phenomenological drive-induced detrapping rate (fitted value 10.8 ± 0.9 MHz at $|A_{\uparrow}| = |A_{\downarrow}| = 1$) to capture the measured increase of $|\downarrow_q\rangle, |\uparrow_q\rangle \rightarrow |g\rangle$ for larger drive powers.

We then characterized the coherence lifetime of an Andreev spin by performing Ramsey (Fig. 4A) and Hahn-echo experiments (Fig. 4B). The measured coherence times, $T_{2R} = 18 \pm 1$ ns and $T_{2E} = 52 \pm 3$ ns, are similar to what has been measured for spin-orbit qubits (23, 27). However, we found that coherence lifetimes of so-called “pair transitions” were systematically an order of magnitude longer [(24), section 5]. In a pair transition, two quasiparticles are excited out of the condensate into both levels of a doublet so that the transition frequency is agnostic to baths that couple to spin (14, 15, 18). It therefore appears that the much shorter ASQ coherence lifetime is limited by such a spin-coupled bath—perhaps hyperfine interactions with the spinful nuclei of indium and arsenic [although nuclear baths typically have a lower frequency spectrum than the measured ratio $\frac{T_{2E}}{T_{2R}} = 2.9$ would indicate (28)], phonon-induced fluctuations of spin-orbit coupling, or paramagnetic impurities on the nanowire surface (7).

At the heart of the ASQ lies a spin-dependent supercurrent, presenting a natural mechanism for coherent integration of spins and electrodynamic modes at the single-quantum level. This fusion of superconducting electrodynamic qubits and semiconductor spin qubits has the potential to inherit the benefits of both platforms while leaving behind certain shortcomings. In modern many-qubit superconducting processors, the qubit size is comparable with that of the readout resonators (29). The ASQ presents the possibility to shrink the qubit area by a factor of more than a thousand without sacrificing coupling strengths. The small spatial footprint of the ASQ also protects it from dielectric loss mechanisms prevalent in superconducting qubits, and experiments following the path carved by electrostatically confined spins (such as the dynamics in tesla-scale magnetic fields) may diagnose the decoherence mechanisms observed here (7). The effects of dephasing and quasiparticle de-trapping could be mitigated by the use of a resonant stimulated Raman adiabatic passage (STIRAP) protocol, which would reduce the required pulse lengths and amplitudes (30). Last, our single-spin manipulation techniques could be used in conjunction with a superconducting bus to achieve gates between spatially distant spin qubits, a long-standing goal in the spin qubit community (8–13).

REFERENCES AND NOTES

1. I. Kulić, *J. Exp. Theor. Phys.* **30**, 944 (1969).
2. M. H. Devoret, R. J. Schoelkopf, *Science* **339**, 1169–1174 (2013).
3. N. M. Chitchevskiy, Y. V. Nazarov, *Phys. Rev. Lett.* **90**, 226806 (2003).
4. C. Padurariu, Y. V. Nazarov, *Phys. Rev. B Condens. Matter Mater. Phys.* **81**, 144519 (2010).
5. A. A. Reynoso, G. Usaj, C. A. Balseiro, D. Feinberg, M. Avignon, *Phys. Rev. B Condens. Matter Mater. Phys.* **86**, 214519 (2012).
6. S. Park, A. L. Yeyati, *Phys. Rev. B* **96**, 125416 (2017).
7. R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, L. M. K. Vandersypen, *Rev. Mod. Phys.* **79**, 1217–1265 (2007).

8. K. D. Petersson et al., *Nature* **490**, 380–383 (2012).
9. N. Samkharadze et al., *Science* **359**, 1123–1127 (2018).
10. X. Mi et al., *Nature* **555**, 599–603 (2018).
11. T. Cubaynes et al., *npj Quantum Inf.* **5**, 47 (2019).
12. F. Borjans, X. G. Croot, X. Mi, M. J. Gullans, J. R. Petta, *Nature* **577**, 195–198 (2020).
13. A. Wallraff et al., *Nature* **431**, 162–167 (2004).
14. C. Janvier et al., *Science* **349**, 1199–1202 (2015).
15. M. Hays et al., *Phys. Rev. Lett.* **121**, 047001 (2018).
16. P. Krogstrup et al., *Nat. Mater.* **14**, 400–406 (2015).
17. D. J. van Woerkom et al., *Nat. Phys.* **13**, 876–881 (2017).
18. L. Tosi et al., *Phys. Rev. X* **9**, 011010 (2019).
19. T. W. Larsen et al., *Phys. Rev. Lett.* **115**, 127001 (2015).
20. G. de Lange et al., *Phys. Rev. Lett.* **115**, 127002 (2015).
21. V. Mourik et al., *Science* **336**, 1003–1007 (2012).
22. M. Hays et al., *Nat. Phys.* **16**, 1103–1107 (2020).
23. S. Nadj-Perge, S. M. Frolov, E. P. Bakkers, L. P. Kouwenhoven, *Nature* **468**, 1084–1087 (2010).
24. Materials and methods are available as supplementary materials.
25. L. J. Swenson et al., *Appl. Phys. Lett.* **96**, 263511 (2010).
26. M. Houzet, K. Serniak, G. Catelani, M. H. Devoret, L. I. Glazman, *Phys. Rev. Lett.* **123**, 107704 (2019).
27. J. W. van den Berg et al., *Phys. Rev. Lett.* **110**, 066806 (2013).
28. F. K. Malinowski et al., *Nat. Nanotechnol.* **12**, 16–20 (2017).
29. F. Arute et al., *Nature* **574**, 505–510 (2019).
30. J. Cerrillo, M. Hays, V. Fatemi, A. Levy Yeyati, *Phys. Rev. Res.* **3**, 022012 (2020).

ACKNOWLEDGMENTS

We thank G. de Lange for assistance with device design and thank N. Frattini and V. Sivak for providing us with a superconducting nonlinear asymmetric inductive element (SNAIL) parametric amplifier. We are grateful to M. Goffman, C. Metzger, H. Pothier, L. Tosi, and C. Urbina for sharing their experimental results and hypotheses. We acknowledge useful discussions with N. Frattini, L. Frunzio, L. Glazman, M. Houzet, P. Kurilovich, V. Kurilovich, and C. Marcus. **Funding:** This research was supported by the US Army Research Office (ARO) grant W911NF-18-1-0212. M.H. acknowledges partial support from the ARO (W911NF-18-1-0020). S.D. acknowledges partial support from the ARO (W911NF-16-1-0349). The view and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ARO or the US government. The US government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation herein. D.B. acknowledges support by the Netherlands Organisation for Scientific Research (NWO) and Microsoft Corporation Station Q. J.C. acknowledges the support from MICINN (Spain) (“Beatriz Galindo” Fellowship BEAGAL18/00081). J.N. acknowledges support from the Danish National Research Foundation. Some of the authors acknowledge the European Union’s Horizon 2020 research and innovation program for financial support; A.G. received funding from the European Research Council, grant 804988 (SIMS); and A.G., A.L.Y., J.C., and J.N. further acknowledge grant 828948 (AndQC) and QuANTERA project 127900(SuperTOP). A.L.Y. acknowledges support by Spanish MICINN through grants FIS2017-84860-R and through the “María de Maeztu” Programme for Units of Excellence in R&D (grant MDM-2014-0377). **Author contributions:** M.H., V.F., K.S., D.B., T.C., A.G., and M.H.D. designed the experimental setup. P.K. and J.N. developed the nanowire materials. D.B. and A.G. fabricated the device. M.H. and V.F. performed the measurements. V.F., M.H., J.C., and A.L.Y. developed the symmetry analysis and microscopic modeling. M.H., J.C., V.F., and A.L.Y. developed and performed the Raman simulations. M.H., V.F., K.S., S.D., and M.H.D. analyzed the data. M.H., V.F., and M.H.D. wrote the manuscript, with feedback from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or supplementary materials.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/430/suppl/DC1
Supplementary Text
Figs. S1 to S9
Table S1
References (31–37)

17 January 2021; accepted 27 May 2021
10.1126/science.abf0345

PLANETARY SCIENCE

Upper mantle structure of Mars from InSight seismic data

Amir Khan^{1,2*}, Savas Ceylan¹, Martin van Driel^{1,3}, Domenico Giardini¹, Philippe Lognonné⁴, Henri Samuel⁴, Nicholas C. Schmerr⁵, Simon C. Stähler¹, Andrea C. Duran¹, Quancheng Huang⁵, Doyeon Kim⁵, Adrien Broquet^{6,7}, Constantinos Charalambous⁸, John F. Clinton⁹, Paul M. Davis¹⁰, Mélanie Drilleau¹¹, Foivos Karakostas⁵, Vedran Lekic⁵, Scott M. McLennan¹², Ross R. Maguire⁵, Chloé Michaut^{13,14}, Mark P. Panning¹⁵, William T. Pike⁸, Baptiste Pinot¹¹, Matthieu Plasman⁴, John-Robert Scholz¹⁶, Rudolf Widmer-Schmidrig¹⁷, Tilman Spohn¹⁸, Suzanne E. Smrekar¹⁵, William B. Banerdt¹⁵

For 2 years, the InSight lander has been recording seismic data on Mars that are vital to constrain the structure and thermochemical state of the planet. We used observations of direct (*P* and *S*) and surface-reflected (*PP*, *PPP*, *SS*, and *SSS*) body-wave phases from eight low-frequency marsquakes to constrain the interior structure to a depth of 800 kilometers. We found a structure compatible with a low-velocity zone associated with a thermal lithosphere much thicker than on Earth that is possibly related to a weak *S*-wave shadow zone at teleseismic distances. By combining the seismic constraints with geodynamic models, we predict that, relative to the primitive mantle, the crust is more enriched in heat-producing elements by a factor of 13 to 20. This enrichment is greater than suggested by gamma-ray surface mapping and has a moderate-to-elevated surface heat flow.

The Interior Exploration Using Seismic Investigations, Geodesy and Heat Transport (InSight) mission (1) touched down on Mars in Elysium Planitia (2) at the end of 2018 and has been acquiring high-quality seismic data with the Seismic Experiment for Interior Structure (SEIS) instrument (3) since early 2019 (4, 5). SEIS's main instrument is a three-component very-broadband seismometer (6), and to date (sol 676), >1000 distinct seismic events have been identified by the Marsquake Service (MQS) (7). A primary goal of the InSight mission (1) is to image the interior structure of Mars from observations of seismic events and to use this to improve our understanding of its formation and evolution.

The compilation of travel-time tables of seismic body waves by Jeffreys and Bullen (8) was of fundamental importance for establishing the first radial average structure of Earth's interior. Earth's seismic velocity structure has been constrained by a plethora of seismic observations including body waves, surface waves, and normal modes (9–12). However, replicating this progress on the Moon has been more difficult. For the Moon, strong scattering in the shallow parts of the lunar crust has largely limited seismic observations to the main *P*- and *S*-wave arrivals (13) and core reflections (14, 15), from which averaged radial seismic

profiles of the crust, mantle, and core could be deduced (16). Relative to the four-station seismic array that operated on the Moon, travel-time inversions on Mars are complicated because accurate epicentral distance and origin-time measurements of marsquakes are difficult to obtain with only one seismic station. Only event distance can be estimated from arrival-time measurements of direct *P*- and *S*-waves (5). To obtain information on interior structure, we need observations of additional seismic phases such as planet-circling surface waves (17, 18), normal modes (19, 20), or reflected and refracted body waves that have interacted with the surface or internal boundaries. As we have yet to positively identify planet-circling waves or normal modes on Mars, we used observations of *P*- and *S*-wave differential travel-time measurements together with surface-reflected body-wave phases *PP*, *PPP*, *SS*, and *SSS* (fig. S1) to jointly invert for both epicentral distance and interior structure. This strategy allows us to construct a quantitative model of the seismic velocity structure of the martian mantle.

Marsquakes have been divided into two main categories based on their frequency content. The low-frequency (LF) events have energy dominantly below 1 Hz, and high-frequency (HF) events have energy dominantly above 1 Hz (5, 7). InSight has recorded far more HF

events, and these events often lack clearly identifiable *P*- and *S*-wave arrivals and polarizations. Here, we focused on eight high-quality LF events of the 43 recorded to date by InSight (21), labeled by mission sol of occurrence and sublabelled alphabetically for sols with more than one event: S0167b, S0173a, S0185a, S0189a, S0235b, S0325a, S0407a, and S0484b. Each of these events has a high signal-to-noise ratio with identifiable *P*- and *S*-waves, which allows for epicentral distance estimation (7). We determined the back azimuth of three events (S0173a, S0235b, and S0325a) from polarization and found that all were located in the Cerberus Fossae region (7). These eight events occurred in the distance range ~25° to 75° and had moment magnitudes M_W between 3.0 and 4.0 (7).

An example broadband-filtered seismogram from marsquake event S0235b is shown in Fig. 1A and includes our picks of the direct *P*- and *S*-wave arrivals. Marsquake waveforms are characterized by codas produced by scattering (4, 5), which, like their lunar seismogram counterparts, complicate the identification of seismic body-wave phases. Consequently, we do not identify seismic arrivals directly in the time series, but instead use narrowband-filtered time-domain envelopes (Fig. 1B) called filter banks (6). Figure 1B shows a zoom-in around the main *P*- and *S*-wave arrivals for event S0235b and our identification of surface-reflected body-wave phases *PP*, *PPP*, and *SS* that were made on the filter banks (bold black lines). We made our picks on the envelope peaks rather than the phase onsets because arrival onsets are more difficult to discern for the later-arriving seismic phases. Filter banks for the eight LF events considered here are shown in fig. S9.

To provide an independent verification of the filter bank-based picks, we used complementary methods for the identification of seismic phases. These include (i) polarization filtering and vespagrams, and (ii) waveform matching (6). Despite the independent nature of the three analysis methods, our picks (table S4) are consistent within the overlapping uncertainties as illustrated in fig. S9. To ensure that the arrivals identified in the envelopes were not due to noise, we compared the observed waveforms to the pulse shape of the direct *P*- and *S*-wave arrivals. As we expected for multiply reflected body waves, the waveforms we observed are consistent with shifts

¹Institute of Geophysics, ETH Zürich, Zürich, Switzerland. ²Physik-Institut, University of Zürich, Zürich, Switzerland. ³Mondaic AG, Zypressenstrasse 82, 8004 Zürich, Switzerland. ⁴Université de Paris, Institut de Physique du Globe de Paris, CNRS, Paris, France. ⁵Department of Geology, University of Maryland, College Park, MD, USA. ⁶Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721, USA. ⁷Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Nice, France. ⁸Department of Electrical and Electronic Engineering, Imperial College London, London, UK. ⁹Swiss Seismological Service, ETH Zürich, Zürich, Switzerland. ¹⁰Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, CA, USA. ¹¹Institut Supérieur de l'Aéronautique et de l'Espace SUPAERO, Toulouse, France. ¹²Department of Geosciences, Stony Brook University, Stony Brook, NY, USA. ¹³Institut Universitaire de France, Paris, France. ¹⁴Laboratoire de Géologie, Terre, Planètes, Environnement, Lyon, France. ¹⁵Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. ¹⁶Max Planck Institute for Solar System Research, Göttingen, Germany. ¹⁷Black Forest Observatory, Institute of Geodesy, University of Stuttgart, Stuttgart, Germany. ¹⁸International Space Science Institute, Bern, Switzerland.

*Corresponding author. Email: akhan@ethz.ch

in phase of 90° and 180° for PP/SS and PPP/SSS , respectively (6). On the basis of absence of surface waves and dispersion in the LF seismograms, we assumed that the events occurred below the crust-mantle interface (5, 22) at a depth of 50 km. We found that varying assigned event depth in the range 50 to 70 km produced little difference in inversion results. We note that for some events, possible depth phases are discernable (Fig. 1B) but are not considered in the inversion.

We used the travel-time picks based on the visual inspection of filter banks (table S4, bold numbers) as our preliminary differential travel-time dataset to simultaneously invert for epicentral distance (Δ) and elastic seismic wave velocity structure (18, 23). We modeled Mars as a spherically symmetric planet, although we acknowledge that it varies laterally in crustal thickness (24, 25) and possibly in seismic properties (20, 26). Global and regional synthetic seismograms that we computed for one- and three-dimensional (1D, 3D) Mars models (23, 26) yielded generally small travel-time differences between 1D and 3D models for the direct P - and S -wave arrivals (<3 s). Thus, picking uncertainty (5 to 10 s) should be larger than what we expect for 3D effects. This might not be the case for the multiply reflected waves. For instance, 3D effects could potentially be more pronounced for the surface-reflected phases because they travel in the more heterogeneous shallow structure beneath Elysium Planitia (1, 4).

We carried out two separate inversions to determine interior structure using a seismic and a geophysical parameterization (6). The seismic parameterization is based on P - and S -wave velocities in a layered planet model (23), whereas the geophysical parameterization (18, 27) relies on an unified description of phase equilibria, seismic properties, and thermochemical parameters. We assumed a homogeneous bulk composition derived from the martian meteorites (28) and computed P - and S -wave velocities as a function of temperature, composition, and pressure using Gibbs free energy minimization (29). On the basis of stagnant-lid models that include a crust enriched in heat-producing elements (30–33), we parameterized the martian geotherm using variable conductive crustal and lithospheric geotherms for which the underlying mantle can be assumed to be adiabatic. The exact nature of the crustal geotherm is less important here because we relied on a seismic parameterization of the crust (6). We computed mantle adiabats (isentropes) self-consistently from the entropy of the lithology at the pressure and temperature of the bottom of the thermal lithosphere (6). To solve the inverse problem (6), we used a stochastic sampling algorithm (34) that produces a suite of models that fit the differential travel times, within uncertainty, and enable estimates of model

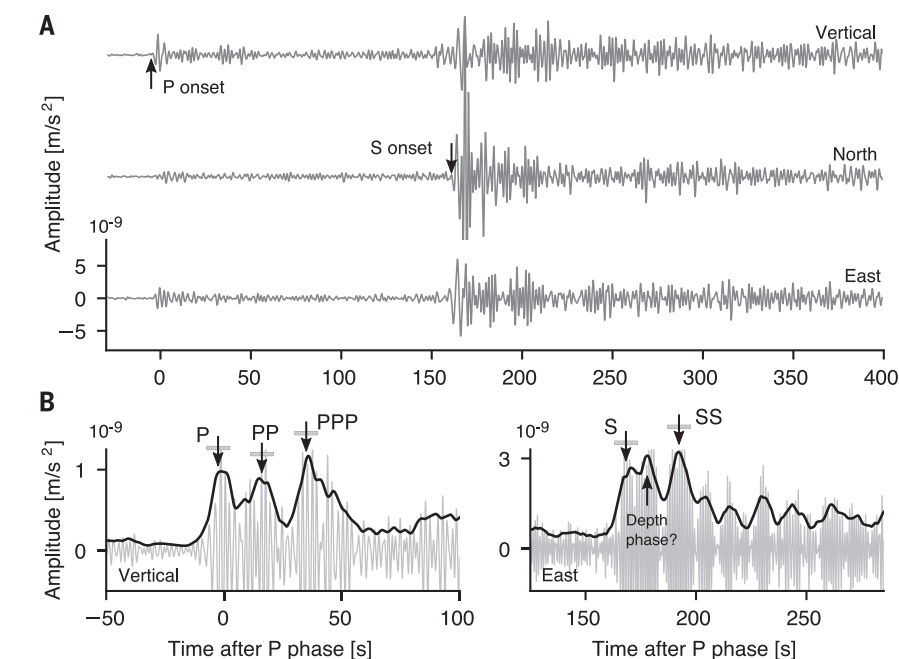


Fig. 1. Marsquake recorded by the InSight seismometer. (A) Three-component broadband-filtered (1.5 to 8 s) seismogram of event S0235b. The main direct P - and S -wave arrivals are indicated by the vertical arrows. (B) Close-up view of the P -wave (left) and S -wave (right) arrivals, showing the direct and surface-reflected body waves (PP , PPP , and SS). Because arrival onsets are less palpable for the surface-reflected phases, arrivals are picked on the peak rather than the onset (arrows). The arrival of a possible depth phase is also marked. Bold black lines indicate the time-domain envelopes that are used for picking phase arrivals. The P - and S -wave envelopes are filtered at 2 s and 1.4 s, respectively, and are 0.5 octave wide on either side. Horizontal gray bars indicate the measurement uncertainty on the picks.

parameter uncertainty. In total we sampled 10^5 models, of which $\sim 10^4$ were retained to ensure model independence.

We inverted for radial P - and S -wave velocity and geothermal profiles (Fig. 2, A and B), using both the geophysical parameterization (red and blue models) and the seismic parameterization (gray-shaded models). The seismic velocity profiles we obtained from both inversions are very similar, with slightly wider posterior model parameter ranges in the case of the seismic parameterization. This reflects the inclusion of mineral physics information in the geophysical parameterization. On the basis of ray path geometry and sensitivity (Fig. 2C), we found that the differential travel-time data constrained the structure to 800-km depth, with sensitivity below ~ 500 km provided primarily by the two events located at $\Delta > 60^\circ$ (S0185a and S0167b). Below 800-km depth, the distribution of inverted profiles starts to resemble the prior P - and S -wave velocity distributions (fig. S10, gray-shaded areas), signaling loss of resolution. Comparison of prior (fig. S13) and posterior (Fig. 2, A and B) distributions of geothermal and P - and S -wave velocity profiles based on the geophysical parameterization confirms that the differential travel-time data constrain upper mantle structure. For both parameterizations, S -wave velocity is

better constrained than P -wave velocity, which reflects the more abundant number of S -wave picks relative to P -wave picks. We found good agreement between the inverted epicentral distances for both parameterizations (fig. S11), with slightly wider Δ distributions for the seismic parameterization because of the broader velocity distributions. We focus on the implications of the geophysical parameterization because it provides a good fit to the observations (Fig. 2, D and E).

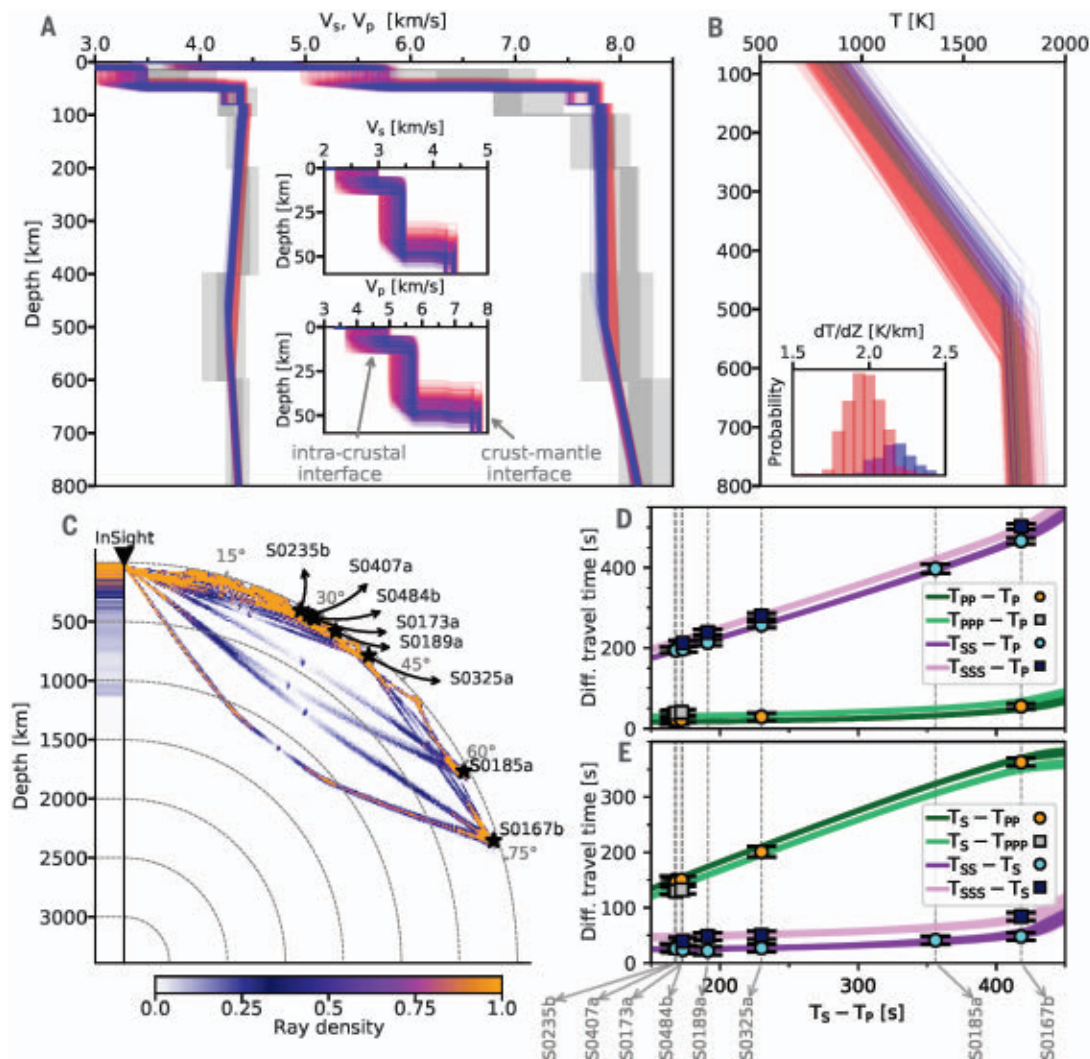
We grouped the geotherms, including lithospheric geothermal gradients (Fig. 2B, inset), and seismic wave velocity profiles into families according to lithospheric thickness (Z_{lit}): $Z_{lit} < 400$ km (fig. S13, green models), $Z_{lit} = 400$ to 500 km (blue models), and $Z_{lit} = 500$ to 600 km (red models). We found that only models that group in a Z_{lit} range of 400 to 600 km (blue and red models), corresponding to a thermal gradient between 1.6 and 2.5 K/km (mantle potential temperatures between 1600 and 1700 K), fit the data. These thermal gradients are consistent with those obtained on the basis of radar observations of the faint lithospheric flexure beneath the north polar cap (1.3 to 3.2 K/km), which constrains the thermal state of the present-day lithosphere (35).

Our crustal velocity models (Fig. 2A, insets) include an upper-crustal low-velocity layer with

Fig. 2. Summary of Mars' upper mantle structure.

(A and B) Inverted *S*- and *P*-wave velocity and geothermal profiles. Colored (red and blue) and gray-shaded models are obtained from the geophysical and seismic inversions, respectively. Insets show the distribution of sampled crustal *S*- and *P*-wave velocity structure and lithospheric geothermal gradients (dT/dz), respectively. Profiles are color-coded according to lithospheric thickness: 400 to 500 km (blue) and 500 to 600 km (red). The lithospheric thermal gradient is determined from the temperature at the crust-mantle interface and at the bottom of the lithosphere and from the difference in depth between the two points. For comparison, prior sampled models are shown in fig. S13. Gray-shaded contours in (A) indicate the 50%, 75%, and 90% credible intervals computed from the distribution of models inverted using a purely seismic parameterization (6). (C) Body-wave ray path geometry for the eight events (labeled S0167b, S0185a, etc.) considered in this study. Color bar denotes ray path density (i.e., number of rays passing through a given area)

based on inverted models shown in (A), which explains the diffuseness of ray paths and source locations. The column to the left of "InSight" shows radial sensitivity, computed as the integrated ray path density with epicentral distance. (D and E) Differential body-wave travel-time misfits for all sampled models shown in (A). Green and purple lines denote differential travel times computed using the inverted models; squares and circles indicate the observations including error bars. For the travel-time calculations performed here, we always pick the first arrival. See fig. S14 for a more detailed version of the differential travel-time misfit.



S- and *P*-wave velocities in the ranges 2 to 3 km/s and 3.5 to 5 km/s, respectively, separated from the lower crust by an intracrustal discontinuity around 5- to 10-km depth, which had been observed previously (4), and a larger discontinuity around 30- to 50-km depth that could be the crust-mantle interface. Lateral variations in crustal structure can result in differences between near-station structure and profiles inverted from our differential travel-time dataset, which averages structure across a wide geographic region, yet may not be entirely representative of the average structure of Mars. The crustal discontinuities beneath the InSight lander are discussed in more detail in (36).

At greater depth, the lithospheric structure is characterized by constant negative *S*- and neutral-to-positive *P*-wave velocity gradients, respectively, to depths of 400 to 600 km. The

S-wave velocity decrease, an inherent feature of the models and consistent with the data, is followed by an increase to 800-km depth. This behavior defines an *S*-wave low-velocity zone (LVZ) in the martian upper mantle. The LVZ results from the large thermal gradient across the lithosphere that arises when a relatively thick stagnant conductive thermal boundary layer sits on top of a convective mantle, as previously postulated (37). A possible manifestation of an LVZ is the presence of a weak *S*-wave shadow zone at epicentral distances of ~40° to 60°. There is no equivalent LVZ for compressional waves because *P*-wave velocity is less affected by temperature. The LVZ on Mars results from the decrease in seismic velocity associated with increasing temperature with depth dominating over the opposing increase in seismic velocity associated with the

increase in pressure with depth. Other mechanisms such as fluid and melt in the asthenosphere (38) and redox effects (39) are also thought to play a role in shaping Earth's LVZ (40), but we lack the resolution to address these issues.

To provide an observational constraint on the strength and extent of the LVZ, we considered the alignment of the seismic traces (fig. S9A) to qualitatively estimate the *S*-wave amplitude behavior with distance. From the alignment, we make the following observations: (i) For $\Delta < 40^\circ$, both *P*- and *S*-wave arrivals are identifiable (S0235b, S0407a, S0484b, S0173a, S0189a, and S0325a); (ii) for $40^\circ < \Delta < 59^\circ$, only a *P*-wave arrival is visible above the background noise on the vertical component, although there is a signal consistent with a low-amplitude *S*-wave (S0183a); and (iii) for $\Delta > 59^\circ$, *P*- and

S-wave arrivals are again distinctly visible (S0185a and S0167b). We interpreted the recovery in *S*-wave amplitude as the bottom of the LVZ because the increase in *S*-wave amplitude is commensurate with velocity switching to an increase with depth. This sequence provides tentative observational evidence for a weak *S*-wave shadow zone in the $\sim 40^\circ$ to 59° epicentral distance range.

We compared the InSight observations to reflectivity synthetics (6) for a range of models covering the inverted blue and red families (Fig. 3). The synthetic relative *S*- to *P*-wave amplitude ratios illustrate the effect of the LVZ and lead us to predict that the *S*- to *P*-wave amplitude ratio falls off with increasing epicentral distance, but stabilizes where the velocity gradient begins to increase with depth. The drop in synthetic relative amplitudes between $\sim 53^\circ$ and 59° is compatible with the low-amplitude *S*-wave event S0183a (not included in our travel-time inversions) and allows us to revise the event's distance and the models obtained from the travel-time inversion. The colored box (beige and brown) shows the allowed range ($\sim 40^\circ$ to 59°) based on alignment (fig. S9A); the brown portion corresponds to the location ($\sim 54^\circ$ to 59°) for which synthetically predicted amplitudes drop, as observed for S0183a. Comparison with InSight observations (Fig. 3, gray boxes) indicates that the blue model family predicts an amplitude behavior that is in line with the location of both the observed amplitude drop (S0183a) and amplitude increase (S0185a). The red model family would appear to put S0183a (drop) in slight contradiction with the inferred location of S0185a (increase).

The weak *S*-wave shadow zone is based on an absent or low-amplitude *S*-wave from a single event (S0183a), which could potentially arise from the source radiation pattern. However, the relatively long (several minutes) coda that is observed throughout suggests that near-source scattering is substantial and that *S*-wave energy is emitted in all directions. Because the ray paths of the direct and coda *S*-waves are similar, the simultaneous absence of both is best explained by geometrical spreading such as that produced by a shadow zone. Attenuation may also result in additional complexity if the quality factors (*Q*) are strongly depth-dependent, such that the effective attenuation along the path is different for the surface-reflected phases. Our initial observations from the direct phases suggest an effective $Q \approx 200$ to 300 to distances of 45° (4, 5), and a similar value is expected for the surface-reflected phases. Alternatively, if we assume that attenuation increases with depth, the more deeply diving *S*-waves would simply weaken and would therefore be unable to replicate the amplitude effects of a LVZ, as evidenced by S0183a and S0185a.

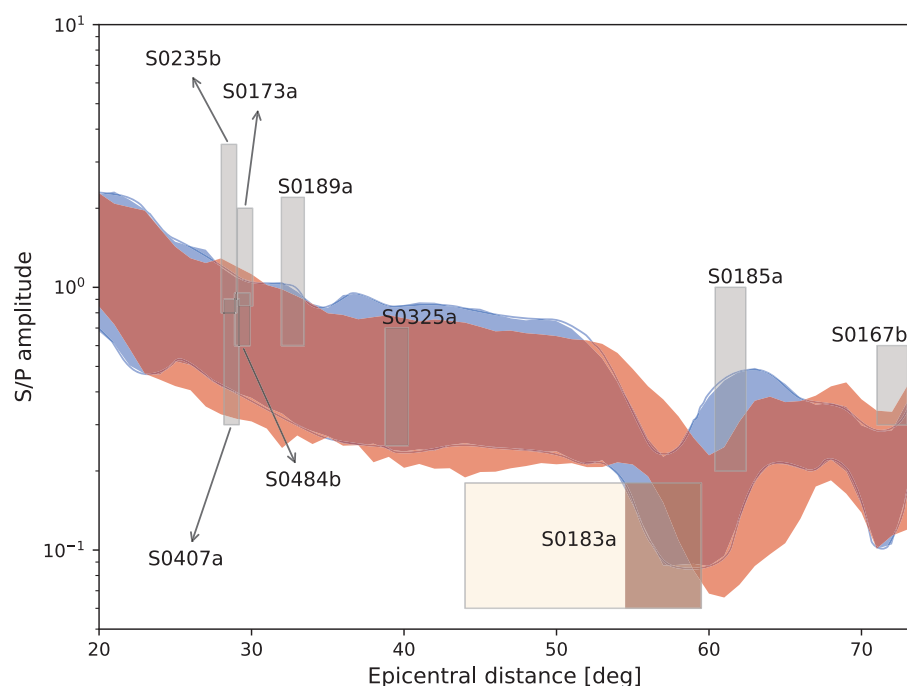


Fig. 3. Seismic amplitude behavior with distance. Blue and red bands show the predicted *S/P* amplitude ratio computed using a range of models covering the inverted blue and red model families (Fig. 2) that correspond to models with lithospheric thicknesses in the ranges 400 to 500 km and 500 to 600 km, respectively. The gray boxes indicate the observed *S/P* amplitude ratio for the events that have been determined from the time-domain envelopes (fig. S9A). The height of the gray boxes indicates the uncertainty on the observed amplitudes, whereas their location and width are based on the inverted epicentral distance distributions (fig. S11), except for S0183a. For S0183a, the widths of the beige and brown boxes show the ranges based on the alignment (fig. S9A) and the synthetically predicted amplitudes, respectively.

The presence of the LVZ has important implications for the thermal evolution of the martian mantle, the formation of the crust, and the planet's surface heat flow. To identify plausible geophysical parameters that pertain to the dynamical evolution of Mars, we computed present-day thermal profiles using a parameterized stagnant-lid mantle and core convection model that simulates 4.5 billion years of planetary evolution (32, 41–43). We explored all relevant geodynamic model parameters in computing the thermal evolution of Mars (6) and compared the resulting present-day lithospheric thermal profiles and crust and lithospheric thicknesses with the models constrained seismically (Fig. 2B).

We found that the seismic results are reproducible by parameterized convection models for a restricted range of geodynamic model parameters. In particular, the models point toward an initially relatively cold state (mantle potential temperature in the range 1630 to 1720 K) and a moderately sluggish mantle (with a reference viscosity $\eta_0 = 10^{20.2}$ to $10^{21.8}$ Pa·s). The bulk heat-producing element (HPE) content of the primitive mantle could be comparable to estimates based on the martian meteorites (28, 44) or 25 to 50% more enriched (45). Moreover, to match the seismic results requires

a crust that is more enriched in HPEs than the primitive mantle by a factor of 13 to 20, leading to an average surface heat flow of 14 to 29 mW/m²—a value that is generally higher than previously estimated (27, 31–33, 46–48). These enrichment levels also call into question models of crustal composition (6, 49, 50). Models that estimate crustal HPE from orbital gamma-ray near-surface mapping (51) predict HPE enrichments relative to the primitive mantle by a factor of no more than 12 (50). Accordingly, deeper crust may mostly consist of petrologically more evolved lithologies, such as those observed at Gale crater and in martian brecciated meteorites (52) and in TES Surface Type 2 (53).

We related mantle thermal structure to surface heat flow using radial models with an average crustal thickness, neglecting regional variations that could vary across the surface of Mars, perhaps by a factor of 2 or more (33, 47). Fortunately, the area sampled by the seismic data, between Cerberus Fossae and the InSight landing site, has been predicted to be relatively homogeneous in terms of surface heat flow and close to the average planetary value (33). Thus, our approach appears justified given the large uncertainty in current surface heat flow predictions.

As of sol 676, the SEIS instrument has operated at optimal noise conditions during the

martian northern spring and summer between sols 100 and 520 (1, 5). Since then, the average wind speed has increased to a level that would make all but the three largest events (S0173a, S0235b, and S0325a) unobservable. Seismic monitoring conditions have improved again since sol 780 (February 2021), and we expect ~10 additional *P-S* travel-time observations during the extended mission. Meanwhile, our preliminary radial velocity model for the upper mantle of Mars will help to guide and inform searches for other seismic arrivals, particularly core-related phases [the core is considered in more detail in (54)]; possibly aid in aligning currently unlocated LF events to complement our dataset; and constrain a variety of geophysical and geochemical models as well as models of planetary evolution.

REFERENCES AND NOTES

- W. B. Banerdt *et al.*, *Nat. Geosci.* **13**, 183–189 (2020).
- M. Golombek *et al.*, *Nat. Commun.* **11**, 1014 (2020).
- P. Lognonné *et al.*, *Space Sci. Rev.* **215**, 12 (2019).
- P. Lognonné *et al.*, *Nat. Geosci.* **13**, 213–220 (2020).
- D. Giardini *et al.*, *Nat. Geosci.* **13**, 205–212 (2020).
- See supplementary materials.
- J. F. Clinton *et al.*, *Phys. Earth Planet. Inter.* **310**, 106595 (2021).
- H. Jeffreys, K. E. Bullen, *Seismological Tables* (British Association for the Advancement of Science, 1940).
- A. M. Dziewonski, D. L. Anderson, *Phys. Earth Planet. Inter.* **25**, 297–356 (1981).
- B. L. N. Kennett, E. R. Engdahl, *Geophys. J. Int.* **105**, 429–465 (1991).
- A. Schaeffer, S. Lebedev, *Geophys. J. Int.* **194**, 417–449 (2013).
- J. Ritsema, A. Deuss, H. J. van Heijst, J. H. Woodhouse, *Geophys. J. Int.* **184**, 1223–1236 (2011).
- C. Nunn *et al.*, *Space Sci. Rev.* **216**, 89 (2020).
- R. C. Weber, P.-Y. Lin, E. J. Garnero, Q. Williams, P. Lognonné, *Science* **331**, 309–312 (2011).
- R. F. Garcia, J. Gagnepain-Beyneix, S. Chevrot, P. Lognonné, *Phys. Earth Planet. Inter.* **188**, 96–113 (2011).
- R. F. Garcia *et al.*, *Space Sci. Rev.* **215**, 50 (2019).
- M. P. Panning *et al.*, *Icarus* **248**, 230–242 (2015).
- A. Khan *et al.*, *Phys. Earth Planet. Inter.* **258**, 28–42 (2016).
- P. Lognonné *et al.*, *Planet. Space Sci.* **44**, 1237–1249 (1996).
- F. Bissig *et al.*, *Space Sci. Rev.* **214**, 114 (2018).
- InSight Marsquake Service, Mars Seismic Catalogue, InSight Mission; V4 2020-07-01. ETHZ, IPGP, JPL, ICL, ISAE-Supaero, MPS, University of Bristol Dataset (2020).
- N. Brinkman *et al.*, *J. Geophys. Res. Planets* **126**, e2020JE006546 (2021).
- M. Drilleau *et al.*, *Earth Space Sci.* **7**, e2020EA001118 (2020).
- M. A. Wiczeorek, M. T. Zuber, *J. Geophys. Res. Planets* **109**, E01009 (2004).
- G. A. Neumann *et al.*, *J. Geophys. Res. Planets* **109**, E08002 (2004).
- E. Bozdağ *et al.*, *Space Sci. Rev.* **211**, 571–594 (2017).
- A. Khan *et al.*, *J. Geophys. Res. Planets* **123**, 575–611 (2018).
- G. J. Taylor, *Geochemistry* **73**, 401–420 (2013).
- J. A. D. Connolly, *Geochem. Geophys. Geosyst.* **10**, Q10014 (2009).
- D. Breuer, T. Spohn, *J. Geophys. Res. Planets* **108**, 5072 (2003).
- M. Thiriet, C. Michaut, D. Breuer, A.-C. Plesa, *J. Geophys. Res. Planets* **123**, 823–848 (2018).
- H. Samuel, P. Lognonné, M. Panning, V. Lainey, *Nature* **569**, 523–527 (2019).
- A.-C. Plesa *et al.*, *J. Geophys. Res. Planets* **121**, 2386–2403 (2016).
- K. Mosegaard, A. Tarantola, *J. Geophys. Res. Solid Earth* **100**, 12431–12447 (1995).
- A. Broquet, M. A. Wiczeorek, W. Fa, *Geophys. Res. Lett.* **47**, e2019GL086746 (2020).
- B. Knapmeyer-Endrun *et al.*, *Science* **373**, 438–443 (2021).
- Y. Zheng, F. Nimmo, T. Lay, *Phys. Earth Planet. Inter.* **240**, 132–141 (2015).
- S.-i. Karato, *Phys. Earth Planet. Inter.* **228**, 300–306 (2014).
- C. J. Cline II, U. H. Faul, E. C. David, A. J. Berry, I. Jackson, *Nature* **555**, 355–358 (2018).
- H. Kawakatsu, H. Utada, *Annu. Rev. Earth Planet. Sci.* **45**, 139–167 (2017).
- T. Spohn, *Icarus* **90**, 222–236 (1991).
- D. J. Stevenson, T. Spohn, G. Schubert, *Icarus* **54**, 466–489 (1983).
- A. Morschhauser, M. Grott, D. Breuer, *Icarus* **212**, 541–558 (2011).
- H. Wänke, G. Dreibus, *Philos. Trans. R. Soc. London Ser. A* **349**, 285–293 (1994).
- T. Yoshizaki, W. F. McDonough, *Geochim. Cosmochim. Acta* **273**, 137–162 (2020).
- J. Ruiz, V. López, J. M. Dohm, *Icarus* **207**, 631–637 (2010).
- L. M. Parro, A. Jiménez-Díaz, F. Mansilla, J. Ruiz, *Sci. Rep.* **7**, 45629 (2017).
- S. E. Smrekar *et al.*, *Space Sci. Rev.* **215**, 3 (2019).
- H. Y. McSweeney Jr., G. J. Taylor, M. B. Wyatt, *Science* **324**, 736–739 (2009).
- S. R. Taylor, S. M. McLennan, *Planetary Crusts: Their Composition, Origin and Evolution* (Cambridge Univ. Press, 2008).
- G. J. Taylor *et al.*, *J. Geophys. Res. Planets* **111**, E03S10 (2006).
- V. Sautter *et al.*, *Lithos* **254–255**, 36–52 (2016).
- H. Y. McSweeney Jr., T. L. Grove, M. B. Wyatt, *J. Geophys. Res. Planets* **108**, 5135 (2003).
- S. C. Stähler *et al.*, *Science* **373**, 443–448 (2021).
- InSight Mars SEIS Data Service, SEIS raw data, InSight Mission. IPGP, JPL, CNES, ETHZ, ICL, MPS, ISAE-Supaero, LPG, MFSC (2019).
- InSight Mars SEIS Data Service, InSight SEIS Data Bundle. PDS Geosciences (GEO) Node (2019).
- InSight Marsquake Service, Mars Seismic Catalogue, InSight Mission; V4 2020-10-01 (2020).
- (V.L.): CNES (for the SEIS analysis) and an ANR grant (MAGIS, ANR-19-CE31-0008-08) (B.P., C.M., H.S., M.D., M.P., and P.L.); the UK Space Agency (W.T.P. and C.C.); NASA grant 80NSSC18K1622 (S.M.M.); and a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s922. **Author contributions:** S.C., D.G., A.K., P.L., N.C.S., J.F.C., S.C.S., M.v.D., and Q.H. analyzed the seismic data and made arrival-time picks; A.K. and A.C.D. performed the inversions; A.K., S.C., D.G., M.v.D., N.C.S., H.S., C.M., and S.E.S. participated and contributed to the interpretation of the results; and A.K. wrote the main paper with contributions from H.S., N.C.S., S.C., M.v.D., P.L., D.G., S.M.M., C.M., T.S., V.L., S.E.S., A.B., and R.R.M. Contributions to the supplement: section 1, S.C., D.G., A.K., N.C.S., Q.H., P.L., D.K., C.C., J.F.C., A.C.D., P.M.D., M.D., F.K., V.L., R.R.M., W.T.P., M.P., J.-R.S., and R.W.-S.; section 2, A.K. and A.C.D.; section 3, A.C.D., A.K., N.C.S., S.C., Q.H., and D.K.; sections 4 and 5, H.S., S.M.M., and A.B. The InSight mission is managed by W.B.B., S.E.S., and M.P.P. The SEIS instrument development was led by P.L., D.G., W.T.P., and W.B.B. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The InSight seismic event catalog (comprising all events, including phase picks, until July 2020) and waveform data are available from the IRIS-DMC, NASA-PDS, SEIS-InSight data portal and IPGP data center (55–57). Two representative interior structure models are available in digital format from DOI: 10.18715/IPGP.2021kpn925g.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/434/suppl/DC1
Materials and Methods
Figs. S1 to S18
Tables S1 to S9
References (58–91)

21 October 2020; accepted 14 May 2021
10.1126/science.abf2966

PLANETARY SCIENCE

Thickness and structure of the martian crust from InSight seismic data

Brigitte Knapmeyer-Endrun^{1*}, Mark P. Panning², Felix Bissig³, Rakshit Joshi⁴, Amir Khan^{3,5}, Doyeon Kim⁶, Vedran Lekić⁶, Benoît Tauzin^{7,8}, Saikiran Tharimena^{2,†}, Matthieu Plasman⁹, Nicolas Compaire¹⁰, Raphael F. Garcia¹⁰, Ludovic Margerin¹¹, Martin Schimmel¹², Éléonore Stutzmann⁹, Nicholas Schmerr⁶, Ebru Bozdağ¹³, Ana-Catalina Plesa¹⁴, Mark A. Wiczeorek¹⁵, Adrien Broquet^{16,15}, Daniele Antonangeli¹⁷, Scott M. McLennan¹⁸, Henri Samuel⁹, Chloé Michaut^{19,20}, Lu Pan²¹, Suzanne E. Smrekar², Catherine L. Johnson^{22,23}, Nienke Brinkman³, Anna Mittelholz³, Attilio Rivoldini²⁴, Paul M. Davis²⁵, Philippe Lognonné^{9,20}, Baptiste Pino¹⁰, John-Robert Scholz⁴, Simon Stähler³, Martin Knapmeyer¹⁴, Martin van Driel³, Domenico Giardini³, W. Bruce Banerdt²

A planet's crust bears witness to the history of planetary formation and evolution, but for Mars, no absolute measurement of crustal thickness has been available. Here, we determine the structure of the crust beneath the InSight landing site on Mars using both marsquake recordings and the ambient wavefield. By analyzing seismic phases that are reflected and converted at subsurface interfaces, we find that the observations are consistent with models with at least two and possibly three interfaces. If the second interface is the boundary of the crust, the thickness is 20 ± 5 kilometers, whereas if the third interface is the boundary, the thickness is 39 ± 8 kilometers. Global maps of gravity and topography allow extrapolation of this point measurement to the whole planet, showing that the average thickness of the martian crust lies between 24 and 72 kilometers. Independent bulk composition and geodynamic constraints show that the thicker model is consistent with the abundances of crustal heat-producing elements observed for the shallow surface, whereas the thinner model requires greater concentration at depth.

Planetary crusts form as a result of mantle differentiation and subsequent magmatic processes, including the partial melting of mantle reservoirs that may continue to the present day (1). For Mars, the cratering record shows that much of its

crust formed early in the planet's history and was accompanied by substantial volcanism (2, 3). During both the initial crystallization of a putative magma ocean as well as later-stage partial melting, incompatible components, including heat-producing elements and volatiles,

concentrated in the melt and were largely sequestered into the crust. The thickness of the crust of Mars thus provides fundamental constraints on how the planet differentiated, how incompatible elements were partitioned among the major silicate reservoirs, and how the planet evolved thermally and magmatically over geologic time (4–6).

Previous estimates of the crustal thickness of Mars and its spatial variations were made by modeling the relationship between gravity and topography. By assuming Airy isostasy and using a restrictive range of crustal densities of 2700 to 3100 kg m⁻³, the average crustal thickness of the planet was reported to be 57 ± 24 km (7). More recent analyses, however, have used elemental abundances of the surface (8) along with major element chemistry of martian meteorites to argue that the crust could be considerably denser, with values close to ~3300 kg m⁻³. If these higher densities were representative of the underlying crust, the gravity data would allow average crustal thicknesses up to 110 km (9). By contrast, bulk crustal densities lower than those previously assumed (~2600 kg m⁻³) have been inferred from gravity analyses and would allow a thinner average crustal thickness (10). Low densities were confirmed locally for the near-surface sediments in Gale crater (11) as well as the pyroclastic deposit of the Medusa Fossae Formation (12). Low bulk crustal densities could result from either substantial porosity or the presence of buried silica- and feldspar-rich rocks (13). Silica-rich magmatic rocks are potentially consistent with ancient evolved lithologies identified in martian meteorite breccias (14).

We used data from the Seismic Experiment for Interior Structure (SEIS) on NASA's Interior Exploration using Seismic Investigations, Geodesy and Heat Transport (InSight) mission (15) to provide an absolute measurement of Mars' crustal thickness and layering. Our assessment of the crustal structure at the landing site is based on a combination of methods using both converted

and reflected seismic phases to resolve trade-offs between the depth of a layer and its seismic velocity (16). By calculating receiver functions (17, 18), we extracted *P*-to-*S* conversions from the *P*-wave coda of three seismic events with the clearest *P*-wave onsets and polarizations. In addition, we applied seismic interferometric techniques by calculating autocorrelations of both ambient noise and event coda using the vertical component. Under the assumption of a diffuse wavefield, as expected in the case of noise from homogeneously distributed, uncorrelated sources as well as in the coda of high-frequency events, the correlations can be interpreted as zero-offset vertical reflection responses (19). By focusing on the reflected wavefield, the autocorrelations provide independent and complementary information to the receiver-function conversion-based methods that make use of the transmitted wavefield (20).

In a previous study (18), we already considered *P*-to-*S* receiver functions for two of the same events but only inverted for the properties of the interface at the base of the shallowest layer (interpreted there as a transition from fractured to unfractured basalt within the crust), causing the first converted arrival at 2.4 s. After including an additional event and applying extensive reanalysis to the data (16), the *P*-to-*S* receiver functions for nine different processing methods (16) show three consistent positive arrivals within the first 8 s but no clear and consistent negative arrivals or later phases (Fig. 1A). Because all three events are located at epicentral distances between 25° and 59° (21, 22), no strong move-out of either direct arrivals or multiple reflections is expected, which impedes the unambiguous identification of multiples. The third positive arrival at 7.2 to 7.5 s could either be simply a *PpPs* multiple of the first arrival at 2.4 s (ray path 3 in Fig. 2B) or contain additional energy from a direct conversion from a third, deeper discontinuity (ray path 3 in Fig. 2D). We applied two inversion approaches to the *P*-to-*S* re-

ceiver functions (16), and both can match the three clear peaks with either two (Fig. 2, A and B) or three interfaces (Fig. 2, C and D). In both inversion approaches, our models showed robust and consistent depths of the

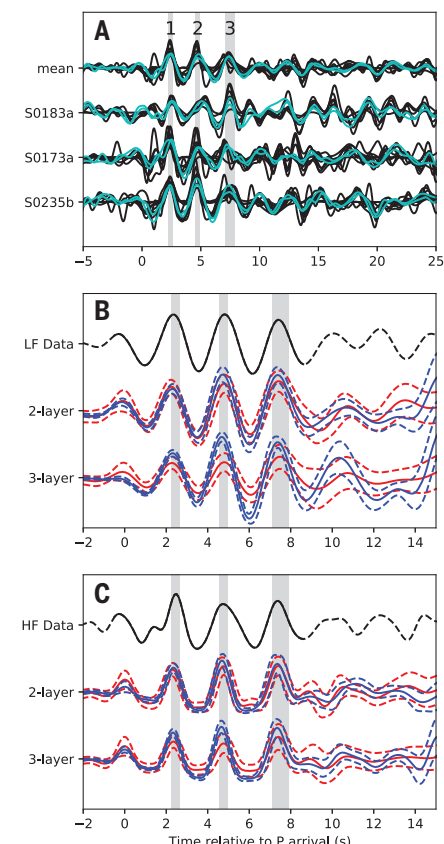


Fig. 1. Measured and modeled converted phases that constrain the crustal structure at the InSight landing site. (A) *P*-to-*S* receiver functions for the three events considered (S0183a, S0173a, and S0235b) and the summed trace. Different traces for each event correspond to different processing methods as described in the supplementary materials. Gray shading highlights the three clear positive phases within the first 8 s. Numbered labels correspond to predicted ray paths shown in Fig. 2, B and D. The two datasets used for model inversions shown in Fig. 2 are highlighted in cyan. (B) Comparison between the low-frequency (LF) representative receiver function sum trace and synthetic summed *P*-to-*S* receiver functions for the two- and three-layer models. Data are shown in black on top, with solid portion of the line representing the time window used in the inversion. Solid and dashed red lines show the synthetics computed by the range of models produced by inversion method A (16), whereas solid and dashed blue lines show the mean receiver functions with standard deviations based on the 5000 best-fitting receiver functions derived from inversion method B (16). Gray-shaded regions are the same as in (A). (C) Same as (B) but for the high-frequency (HF) receiver functions.

¹Bensberg Observatory, University of Cologne, Vinzenz-Pallotti-Str. 26, 51429 Bergisch Gladbach, Germany. ²Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., M/S 183-301, Pasadena, CA 91109, USA. ³Institute of Geophysics, ETH Zurich, Sonneggstr. 5, 8092 Zurich, Switzerland. ⁴Max Planck Institute for Solar System Research, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany. ⁵Physik-Institut, University of Zurich, Zurich, Switzerland. ⁶Department of Geology, University of Maryland, College Park, 8000 Regents Dr., College Park, MD, 20782-4211, USA. ⁷Université de Lyon, Université Claude Bernard Lyon 1, Ecole Normale Supérieure de Lyon, Université Jean Monet, CNRS, Laboratoire de Géologie de Lyon, Terre, Planètes, Environnement, F-69622 Villeurbanne, France. ⁸Research School of Earth Sciences, The Australian National University, Canberra, ACT 0200, Australia. ⁹Université de Paris, Institut de Physique du Globe de Paris, CNRS, 1 rue Jussieu, F-75005 Paris, France. ¹⁰Institut Supérieur de l'Aéronautique et de l'Espace SUPAERO, 10 Avenue Edouard Belin, 31400 Toulouse, France. ¹¹Institut de Recherche en Astrophysique et Planétologie, Université Toulouse III Paul Sabatier, CNRS, CNES, 14 Av. E. Belin, 31400 Toulouse, France. ¹²Geosciences Barcelona-CSIC, Barcelona, Spain. ¹³Department of Geophysics, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401, USA. ¹⁴Institute of Planetary Research, German Aerospace Center (DLR), 12489 Berlin, Germany. ¹⁵Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, 06304 Nice, France. ¹⁶Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721, USA. ¹⁷Sorbonne Université, Muséum National d'Histoire Naturelle, UMR CNRS 7590, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, IMPMC, 75005 Paris, France. ¹⁸Department of Geosciences, Stony Brook University, Stony Brook, NY 11794-2100, USA. ¹⁹Université de Lyon, Ecole Normale Supérieure de Lyon, Université Claude Bernard Lyon 1, Université Jean Monet, CNRS, Laboratoire de Géologie de Lyon, Terre, Planètes, Environnement, F-69007 Lyon, France. ²⁰Institut Universitaire de France, Paris, France. ²¹Center for Star and Planet Formation, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark. ²²Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. ²³Planetary Science Institute, Tucson, 1700 East Fort Lowell, Suite 106, Tucson, AZ 85719-2395, USA. ²⁴Royal Observatory of Belgium, Brussels, Belgium. ²⁵Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, CA 90095, USA.

*Corresponding author. Email: bknappmey@uni-koeln.de †Present address: University of Vienna, Althanstrasse 14, 1090 Vienna, Austria.

two shallowest interfaces. The first layer with a thickness of 6 to 11 km and an S -wave velocity between 1.2 and 2.1 km s⁻¹ is consistent with the previous results for the shallow crust (18), whereas a second interface is found at a depth of 15 to 25 km independent of the model parameterization. The third interface, the existence of which is supported but not abso-

lutely required by the data, showed greater variability in depth between different inversion choices and generally required a smaller velocity contrast at the base of this layer than for the shallower second interface (figs. S18 and S19). Based on the ensemble of models from the two inversion approaches, our results are consistent with either a local crustal thick-

ness at the InSight landing site of 15 to 25 km, when the base of layer 2 is the Moho (thin-crust models), or 27 to 47 km, when the base of layer 3 is the Moho (thick-crust models; Fig. 2 and figs. S18 and S19). S -to- P receiver functions can also be calculated for two events (S0173a and S0235b; figs. S4, S6, and S7), and both show a signal consistent with conversion at the first interface, whereas S0235b also shows possible arrivals consistent with deeper conversions (16). Further support for the P -to- S receiver function-derived models is provided by waveform fits in inversions for source mechanisms (16), where a strong interface at a depth around 24 km is required to match S precursors.

Vertical component autocorrelations based on different datasets and processing algorithms (16, 23) show consistent energy maxima in the 5- to 6-s, 10- to 11-s, and 20- to 21.5-s time ranges (Fig. 3). Comparison with predicted arrival times from representative models produced by the receiver-function inversion shows that these energy maxima can be explained by P -wave reflections in those models interacting with the first two interfaces, without any clear observations requiring the third interface. Previously published autocorrelations (24) contain an arrival near 10 s that is consistent with our results and which can be explained as a P -wave reflection from the bottom of the second layer at a depth around 22 km. A second arrival reported by (24) near 20 s, which is also present in many of the autocorrelation functions calculated here, is consistent with a multiple reflection from that layer (Fig. 3). These arrivals were interpreted by Deng and Levander (24) as P and S reflections, respectively, from a crust-mantle discontinuity at a depth of 35 km. However, we do not expect a strong S reflection in a vertical autocorrelation because vertically propagating S waves are horizontally polarized. Interpreting the second arrival as a multiple P reflection instead is consistent with our receiver function-derived results and more likely to be observed in a vertical component autocorrelation. The previously published crustal thickness estimate of 35 km based on autocorrelations (24) is consistent with the possible range of the thick-crust models, but the specific arrivals identified in that study are more consistent with a reflection and multiple from the shallower second interface at a depth around 20 km.

We inverted for the thickness of the crust at a global scale using the seismically estimated thickness at the InSight landing site and the observed gravity field as constraints (16). Our models consider the gravity of hydrostatic relief along density interfaces beneath the lithosphere, surface relief, variations in thickness of a constant density crust, and the low-density polar cap deposits (25). We used several different interior prelanding models (26) that specify the density profile of the mantle and

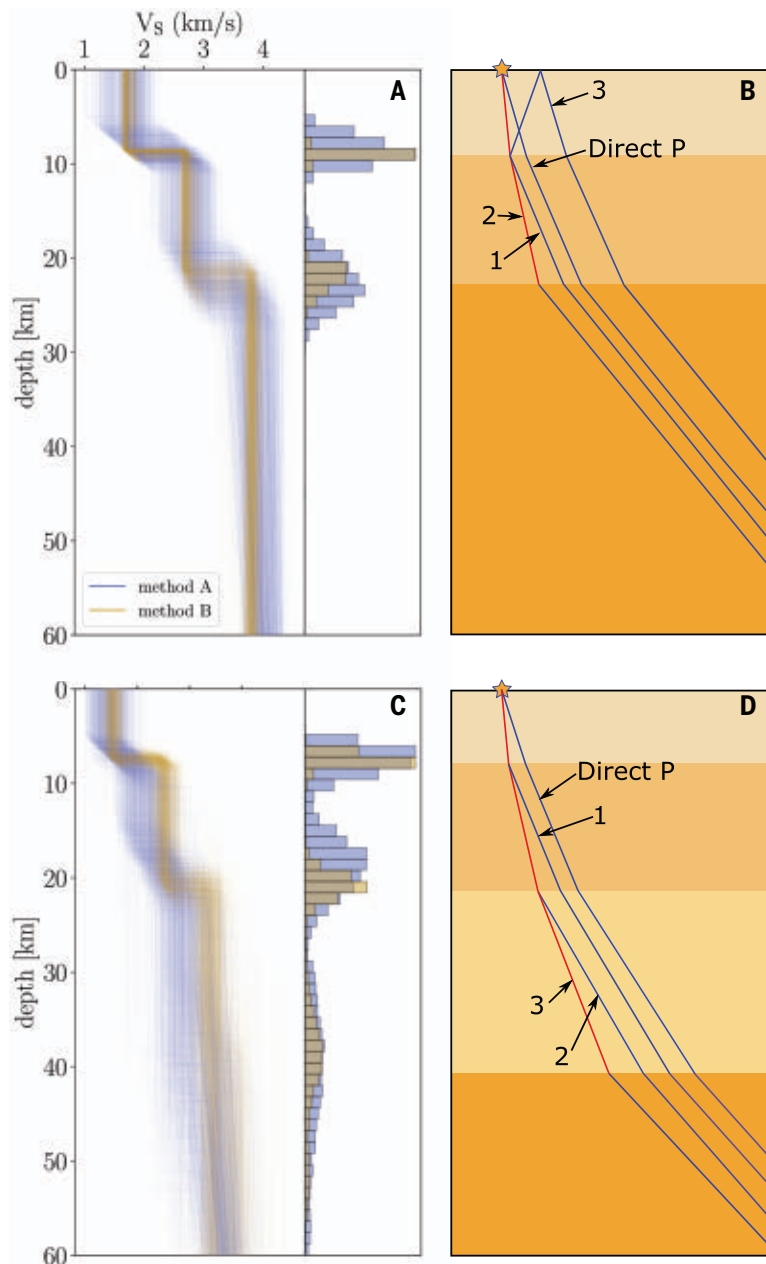


Fig. 2. Synopsis of the crustal structure of Mars at the InSight landing site from receiver-function analyses. (A) Inversion results for all three events using inversion method A (blue lines) and method B (brown lines) using a two-layer parameterization. (B) Cartoon showing the ray paths of the main direct and converted phases present in the data. Blue lines show P -phase paths, whereas red lines show conversions to S phases at the interfaces below the lander. Direct conversions and one P multiple are shown, and numbered labels correspond to arrivals identified in Fig. 1A. (C and D) Same as (A) and (B) except for the assumption of a three-layer model and exclusion of the multiple arrival.

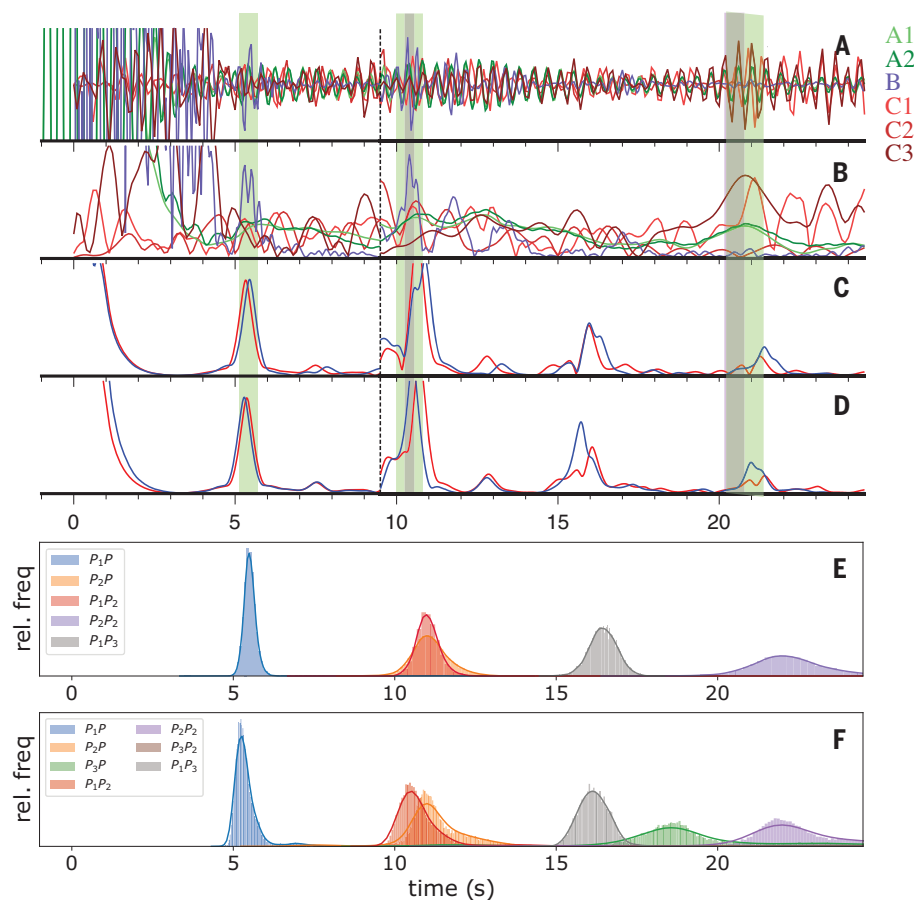


Fig. 3. Autocorrelation functions for different datasets, components, and processing methods. (A) Overlaid traces are from the three analysis methods discussed in the supplementary materials. The dashed bar at 9.5 s corresponds to a change in normalization in order to see smaller amplitude arrivals later in the trace. Green bars highlight areas where all methods are nearly in-phase and show potential arrivals, whereas purple bars highlight arrivals indicated from an independent study (24).

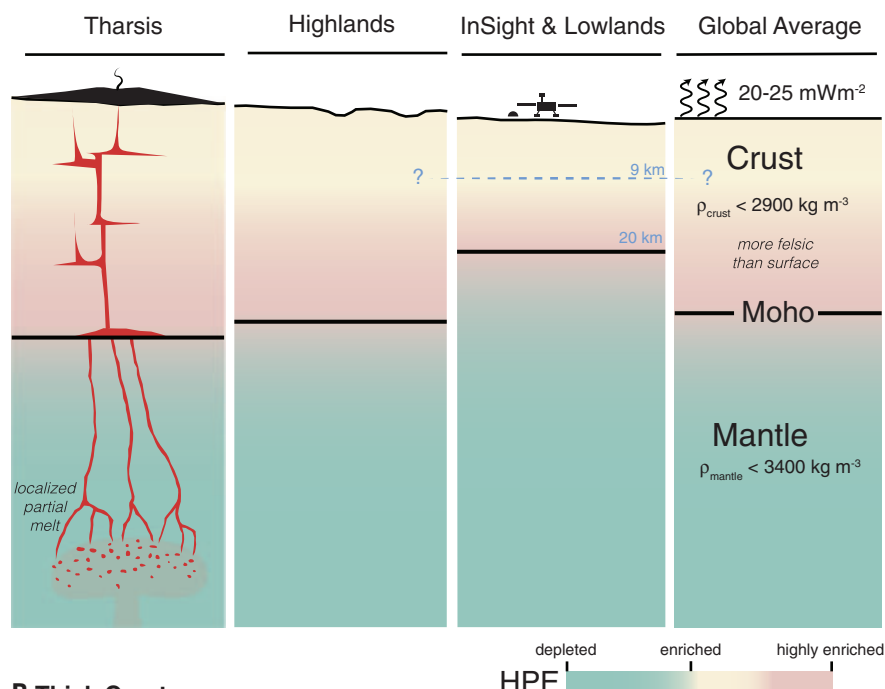
(B) Envelopes of the autocorrelation functions displayed in (A). (C) Envelopes of synthetic zero-offset Green's functions for a representative model from the family of two-layer models in Fig. 2A for method A in blue and method B in red. (D) Same as (C) but for the three-layer models from Fig. 2C. (E) Histograms of predicted arrivals from the family of two-layer models as shown in Fig. 2A. The first subscript of the arrival in the legend refers to the interface of reflection, and the second subscript (if present) represents a second or third bounce between the free surface and that interface. rel. freq., relative frequency. (F) Same as (E) but for the three-layer models in Fig. 2C.

core, and, for each, we constructed crustal thickness models for all permissible crustal densities. For a given seismic thickness, the mean thickness of the crust depends almost exclusively on the density contrast across the crust-mantle interface (fig. S22). To ensure that the thickness of the crust is positive within the major impact basins, each reference model has a maximum permissible crustal density. If the thin-crust seismic model is used as a constraint, the global mean crustal thickness is predicted to lie between 24 and 38 km and the maximum permissible density of the crust is 2850 kg m^{-3} (Fig. 4 and figs. S22 and S23). For the thick-crust seismic model, the average crustal thickness lies between 39 and 72 km and the maximum permissible crustal density is 3100 kg m^{-3} (Fig. 4 and figs. S22 and S23). For both seismic constraints, the crustal density is substantially less than would be expected based on the composition of surface materials (9), which is close to 3300 kg m^{-3} . The lower bulk densities are signatures of highly altered layers and can be accounted for by the presence of more than 5% porosity in the crust on average, the presence of fluids or low-density cements filling fractures and pore space, the existence of abundant petrologically evolved felsic rocks beneath the surface layer, or a combination thereof.

The seismic observations argue for a relatively thin crust, or at least thinner than some earlier predictions (9), providing constraints on crustal heat production and the degree of planetary silicate differentiation (Fig. 4). Because the present-day crustal thickness is the outcome of the planet's differentiation history (27, 28), geodynamic and geologic modeling can place constraints on the composition of the crust and of the mantle, and on the cooling rate of the planet (16). Our results indicate that average crustal thickness models that are consistent with the thick-crust seismic model are compatible with currently accepted bulk (29, 30) and crustal (8, 31) heat-producing element contents and the occurrence of present-day melting only in an ascending plume below the thickened crust of the Tharsis province (fig. S27). Such a scenario implies a crust that is about 13 times more enriched in heat-producing elements than the primitive mantle (fig. S24), consistent with 55 to 70% of the martian heat-producing elements being sequestered into the crust. By contrast, the thin-crust seismic model requires a crust that is about 21 times more enriched than a relatively cold primitive mantle (fig. S25). This is more than two times larger than estimates from gamma-ray spectroscopy data that constrain the surface layer of the crust (table S6) and

would point toward an enrichment in heat-producing elements beneath the surface layer (16). Furthermore, this would call for an efficient process of incompatible element extraction from the mantle, possibly by upward segregation during the solidification of a magma ocean or by a secondary differentiation mechanism, as for the continental crust of Earth. In both crustal models, assuming a Wänke and Dreibus (29) bulk composition, the present-day heat flux is predicted to lie between 20 and 25 mW m^{-2} (Fig. 4). The depth to the crust-mantle boundary, as well as layering in the crust, can further constrain crustal magnetization amplitudes, depending on whether the magnetization is carried in upper or lower crustal layers, or both (16). We can also investigate whether crustal thickness and density models are consistent with moment-of-inertia measurements and constraints on the properties of Mars core from the k_2 tidal Love number (16). Generally, these constraints are easier to match for most mantle composition models with the thick-crust seismic models, although some models also allow for the thin-crust model. Overall, when considering geodynamic, geochemical, and geodetic constraints, the thin-crust models place tighter constraints on the density and enrichment of heat-producing elements within the crust,

A Thin Crust



B Thick Crust

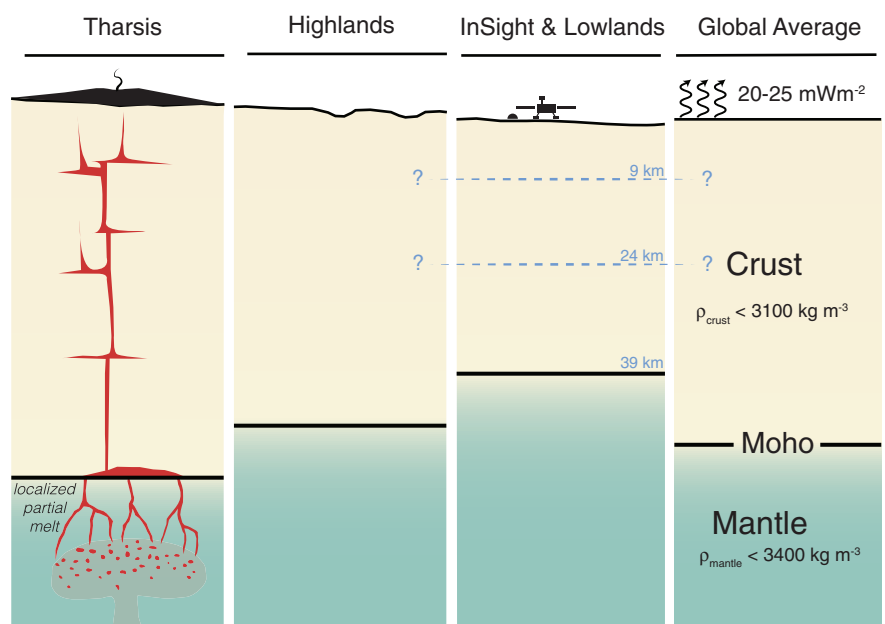


Fig. 4. Schematic interpretations of the geochemical and geodynamical implications for the thin- and thick-crust models. (A and B) To match geodynamic constraints, an enrichment of heat-producing elements (HPEs), shown in color, and lower density than observed from the surface are required in the thin-crust model (A), whereas the thick-crust model (B) is consistent with surface observations. ρ , density.

as well as on the mantle composition, than the thick-crust models, but neither of the two can be excluded.

REFERENCES AND NOTES

1. L. T. Elkins-Tanton, *Annu. Rev. Earth Planet. Sci.* **40**, 113–139 (2012).
2. M. Grott *et al.*, *Space Sci. Rev.* **174**, 49–111 (2013).
3. F. Nimmo, K. Tanaka, *Annu. Rev. Earth Planet. Sci.* **33**, 133–161 (2005).
4. D. Breuer, W. B. Moore, in *Physics of Terrestrial Planets*, vol. 10 of *Treatise on Geophysics*, G. Schubert, Ed. (Elsevier, ed. 2, 2015), pp. 255–305.
5. S. A. Hauck II, R. J. Phillips, *J. Geophys. Res.* **107**, 5052 (2002).
6. A.-C. Plesa *et al.*, *Geophys. Res. Lett.* **45**, 12198–12209 (2018).
7. M. A. Wieczorek, M. T. Zuber, *J. Geophys. Res.* **109**, E01009 (2004).
8. B. C. Hahn, S. M. McLennan, E. C. Klein, *Geophys. Res. Lett.* **38**, L14203 (2011).
9. D. Baratoux *et al.*, *J. Geophys. Res.* **119**, 1707–1727 (2014).
10. S. Goossens *et al.*, *Geophys. Res. Lett.* **44**, 7686–7694 (2017).
11. K. W. Lewis *et al.*, *Science* **363**, 535–537 (2019).
12. L. Ojha, K. Lewis, *J. Geophys. Res.* **123**, 1368–1379 (2018).
13. V. Sautter *et al.*, *Nat. Geosci.* **8**, 605–609 (2015).
14. M. Humayun *et al.*, *Nature* **503**, 513–516 (2013).
15. P. Lognonné *et al.*, *Space Sci. Rev.* **215**, 12 (2019).
16. Methods and additional materials are available as supplementary materials.
17. C. A. Langston, *J. Geophys. Res.* **84**, 4749–4762 (1979).
18. P. Lognonné *et al.*, *Nat. Geosci.* **13**, 213–220 (2020).
19. K. Wapenaar, D. Draganov, R. Snieder, X. Campman, A. Verdel, *Geophysics* **75**, 75A195–75A209 (2010).
20. D. Kim, V. Lekić, *Geophys. Res. Lett.* **46**, 13722–13729 (2019).
21. D. Giardini *et al.*, *Nat. Geosci.* **13**, 205–212 (2020).
22. A. Khan *et al.*, *Science* **373**, 434–438 (2021).
23. N. Compaire *et al.*, *J. Geophys. Res.* **126**, e2020JE006498 (2021).
24. S. Deng, A. Levander, *Geophys. Res. Lett.* **47**, e2020GL089630 (2020).
25. M. A. Wieczorek, M. Beuthe, A. Rivoldini, T. Van Hoolst, *J. Geophys. Res. Planets* **124**, 1410–1432 (2019).
26. S. E. Smrekar *et al.*, *Space Sci. Rev.* **215**, 3 (2019).
27. H. Samuel, P. Lognonné, M. Panning, V. Lainey, *Nature* **569**, 523–527 (2019).
28. M. Thiriet, C. Michaut, D. Breuer, A. C. Plesa, *J. Geophys. Res. Planets* **123**, 823–848 (2018).
29. H. Wänke, G. Dreibus, *Philos. Trans. R. Soc. London Ser. A* **349**, 285–293 (1994).
30. G. J. Taylor, *Geochemistry* **73**, 401–420 (2013).
31. S. R. Taylor, S. M. McLennan, *Planetary Crusts: Their Composition, Origin and Evolution* (Cambridge Univ. Press, 2009).
32. InSight Mars SEIS Data Service, SEIS raw data, InSight Mission, IPGP, JPL, CNES, ETHZ, ICL, MPS, ISAE-Supaero, LPG, MFSC, SEIS (2019); https://doi.org/10.18715/SEIS.INSIGHT.XB_2016.
33. B. Knapmeyer-Endrun *et al.*, Crustal models of Mars at the InSight landing site from InSight seismic data (2021); <https://doi.org/10.18715/IPGP.2021.kqwc4yaf>.

ACKNOWLEDGMENTS

This is Insight Contribution Number (ICN) 187. We acknowledge NASA, CNES, their partner agencies and institutions (UKSA, SSO, DLR, JPL, IPGP-CNRS, ETHZ, IC, and MPS-MPG), and the flight operations team at JPL, SISMO, MSDS, IRIS-DMC, and PDS for providing SEED SEIS data. **Funding:** M.P.P., S.T., E.B., S.E.S., and W.B.B. were supported by the NASA InSight mission and funds from the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. F.B. was supported by research grant ETH-05 17-1. A.K., D.G., M.v.D., and S.S. acknowledge funding by the Swiss National Science Foundation and the Swiss State Secretariat for Education, Research and Innovation, and support from ETHZ through the ETH+ funding scheme (ETH-02 19-1). V.L. and D.K. acknowledge funding from a Packard Foundation Fellowship to V.L. B.T. is supported by the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement 793824. French co-authors acknowledge the support of CNES and ANR (MAGIS, ANR-19-CE31-0008-08). N.S. was supported by NASA grant 80NSSC18K1628. E.B. was funded through NASA Participating Scientist Program grant 80NSSC18K1680. A.-C.P. gratefully acknowledges the financial support and endorsement from the DLR Management Board Young Research Group Leader Program and the Executive Board Member for Space Research and Technology. Geodynamical models used in this work were performed on the supercomputer ForHLR funded by the Ministry of Science, Research and the Arts Baden-Württemberg and by the Federal Ministry of Education and Research. S.M.M. was funded through NASA InSight Participating Scientist Program award no. 80NSSC18K1622. C.M. acknowledges the support of the Institut Universitaire de France (IUF). C.L.J. and A.M. acknowledge support from the InSight Mission, the Canadian Space Agency, and ETH Zurich (ETH fellowship 19-2 FEL-34). N.B. is supported by research grant ETH-06 17-02. The work of A.R. was financially supported by the Belgian PRODEX program managed by the European Space Agency in collaboration with the Belgian Federal Science Policy Office. **Author contributions:** B.K.-E., M.P.P., F.B., R.J., A.K., D.K., V.L., B.T., S.T., M.P., N.C., R.F.G., L.M., M.S., É.S., N.S., A.-C.P., M.W., D.A., S.M.M., H.S., C.M., L.P., S.E.S., C.L.J., P.L., M.K., D.G., and W.B.B. contributed to the conceptualization of this study. The applied methodologies were developed by B.K.-E., F.B., R.J., A.K., D.K., V.L., B.T., S.T., M.P., N.C., R.F.G., L.M., M.S., É.S., A.-C.P., M.W., H.S., C.M., L.P., P.M.D., P.L., B.P., and J.-R.S. B.K.-E., F.B., R.J., A.K., D.K., N.C., M.S., É.S., A.-C.P., M.W., H.S., C.M., P.M.D., P.L., B.P., and J.-R.S. developed, implemented, and tested the software that was used. B.K.-E., M.P.P., N.S., N.B., and M.v.D. provided validation. Formal analyses were conducted by B.K.-E., R.B., R.J., A.K., D.K., V.L., B.T., S.T., M.P., N.C., M.S., É.S., A.-C.P., M.W., A.B., H.S., C.M., C.L.J., N.B., A.M., A.R., P.M.D., P.L., B.P.,

J.-R.S., and S.S. B.K.-E., F.B., R.J., A.K., D.K., V.L., S.T., M.P., N.C., M.S., É.S., A.-C.P., M.W., S.M.M., H.S., C.M., C.L.J., N.B., A.M., A.R., P.M.D., P.L., B.P., J.-R.S., and S.S. provided investigations. Data curation activities were done by P.L., S.S., M.v.D., and D.G. B.K.-E., M.P.P., F.B., R.J., A.K., D.K., V.L., B.T., S.T., N.C., R.F.G., L.M., M.S., É.S., A.-C.P., M.W., S.M.M., H.S., C.M., C.L.J., N.B., A.M., A.R., and P.M.D. contributed to the writing of the original draft. Review and editing were performed by B.K.-E., M.P.P., N.S., E.B., A.B., D.A., S.E.S., C.L.J., and M.K. B.K.-E., M.P.P., F.B., R.J., A.K., D.K., V.L., B.T.,

N.C., M.S., É.S., N.S., A.-C.P., M.W., H.S., C.M., A.M., A.R., and P.M.D. worked on visualizing the results. B.K.-E., M.P.P., A.K., R.F.G., L.M., N.S., C.L.J., P.L., and D.G. supervised junior scientists involved in the project. The InSight mission is managed by W.B.B., M.P.P., and S.E.S. The SEIS instrument development was led by P.L., D.G., and W.B.B. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All InSight SEIS data (32) used in this paper are available from the IGP Data Center, IRIS-DMC, and NASA PDS. Crustal models derived in this study (Fig. 2) are available from MSDS (33).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/438/suppl/DC1
Materials and Methods
Figs. S1 to S29
Tables S1 to S6
References (34–118)

26 November 2020; accepted 21 May 2021
10.1126/science.abf8966

PLANETARY SCIENCE

Seismic detection of the martian core

Simon C. Stähler^{1*}, Amir Khan^{1,2}, W. Bruce Banerdt³, Philippe Lognonné⁴, Domenico Giardini¹, Savas Ceylan¹, Mélanie Drilleau⁵, A. Cecilia Duran¹, Raphaël F. Garcia⁵, Quancheng Huang⁶, Doyeon Kim⁶, Vedran Lekic⁶, Henri Samuel⁴, Martin Schimmel⁷, Nicholas Schmerl⁶, David Sollberger¹, Éléonore Stutzmann⁴, Zongbo Xu⁴, Daniele Antonangeli⁸, Constantinos Charalambous⁹, Paul M. Davis¹⁰, Jessica C. E. Irving¹¹, Taichi Kawamura⁴, Martin Knapmeyer¹², Ross Maguire⁶, Angela G. Marusiak³, Mark P. Panning³, Clément Perrin¹³, Ana-Catalina Plesa¹², Attilio Rivoldini¹⁴, Cédric Schmelzbach¹, Géraldine Zenhäusern¹, Éric Beucler¹³, John Clinton¹⁵, Nikolaj Dahmen¹, Martin van Driel¹, Tamara Gudkova¹⁶, Anna Horleston¹¹, W. Thomas Pike⁹, Matthieu Plasman⁴, Suzanne E. Smrekar³

Clues to a planet's geologic history are contained in its interior structure, particularly its core. We detected reflections of seismic waves from the core-mantle boundary of Mars using InSight seismic data and inverted these together with geodetic data to constrain the radius of the liquid metal core to 1830 ± 40 kilometers. The large core implies a martian mantle mineralogically similar to the terrestrial upper mantle and transition zone but differing from Earth by not having a bridgmanite-dominated lower mantle. We inferred a mean core density of 5.7 to 6.3 grams per cubic centimeter, which requires a substantial complement of light elements dissolved in the iron-nickel core. The seismic core shadow as seen from InSight's location covers half the surface of Mars, including the majority of potentially active regions—e.g., Tharsis—possibly limiting the number of detectable marsquakes.

The core of a planet plays a prominent role because it governs many of the fundamental processes—from dynamo action and magnetic-field generation to mantle convection—that affect the surface through volcanic and tectonic activity and may influence the early climate through magnetic shielding of the atmosphere. The size of the martian core is of particular interest because of its notable impact on the planet's evolution, which differs from that

of Earth primarily as a result of Mars' smaller size and, therefore, accelerated differentiation and core formation and cooling-off that resulted in a rigid-shell, one-plate planet (1, 2).

A small core with a radius between 1300 and 1600 km would be deficient in light elements and would accommodate a martian lower mantle similar to that of Earth, which is dominated by bridgmanite-structure silicates (3). A large core with a radius range between 1800 and 1900 km would instead be enriched in light elements and exclude the presence of a phase transition equivalent to the 660-km discontinuity that marks the onset of Earth's lower mantle (4). It would exert a markedly different dynamic control over the martian mantle (5–8), with implications for an early martian dynamo (1) that could explain the observed highly magnetized crust in the southern highlands (9). Direct constraints on the core and deep interior of Mars, however, are scarce and limited to global geophysical measurements, including mass, moment of inertia, and tidal response (10, 11), in addition to geochemical data based on achondritic basaltic meteorites that originated from the surface of Mars (12–14). Collectively, these observations suggest a liquid core with a radius in the range of ~1700 to 1900 km (4, 8, 11, 15, 16). However, without additional observations, the range of potential martian core sizes en-

compasses a large enough range to allow for either the presence or the absence of a lower mantle.

The estimate of core size can be improved with the direct detection of core-interacting seismic phases, which constrain the core size of Earth and the Moon. Earth's core had been predicted in the late 18th century—on the basis of the recognition that the density of near-surface rocks is substantially lower than the average density of Earth (17)—but was not confirmed until global observations of earthquakes became possible (18). Subsequent seismic measurements led to unexpected discoveries in Earth's deep interior, such as the inner core (19) and deep mantle layering (20, 21). Similarly, the Apollo lunar seismic data (22) were used to establish the existence of a lunar core (23, 24). The detection of seismic waves reflected from the core stands not only to refine the insights gained from studying the chemistry of the martian rocks, whose siderophile element depletion and isotopic signature point to a core-forming event early in Mars' history (25, 26), but also to considerably improve our understanding of the deep interior of Mars.

After a full martian year, the Interior Exploration using Seismic Investigations, Geodesy, and Heat Transport (InSight) mission (27) and its seismometer SEIS (28) have recorded a multitude of seismic events, which have been located and classified by the Marsquake Service (29–31). Of these, the low-frequency events with main energy below 1 Hz and waves traveling through the mantle can be used to characterize the crustal and mantle structure of Mars (32, 33).

To investigate the core of Mars, we analyzed seismic data from 11 low-frequency marsquakes. Six of them were found in a suitable distance range [27° to 38.5° (30)], with high enough signal-to-noise ratio (SNR) to identify potential core-reflected *S* waves (*ScS* waves) (Table 1). Because *S* waves cannot propagate in a fluid medium, the core-mantle boundary (CMB) acts as a polarization filter, reflecting horizontally polarized *S* waves (*SH* waves) back into the mantle, whereas vertically polarized *S* waves (*SV*) lose some energy as a result of *S*-to-*P* conversion. *S* waves reflected from the CMB are therefore expected to be predominantly horizontally polarized, with an azimuth orthogonal to the source-receiver direction.

¹Institute of Geophysics, ETH Zürich, Zürich, Switzerland.

²Physik-Institut, University of Zürich, Zürich, Switzerland.

³Jet Propulsion Laboratory, California Institute of Technology,

Pasadena, CA, USA. ⁴Université de Paris, Institut de

physique du globe de Paris, CNRS, Paris, France. ⁵Institut

Supérieur de l'Aéronautique et de l'Espace SUPAERO,

Toulouse, France. ⁶Department of Geology, University of

Maryland, College Park, MD, USA. ⁷Geosciences Barcelona –

CSIC, Barcelona, Spain. ⁸Sorbonne Université, Muséum

National d'Histoire Naturelle, UMR CNRS 7590, Institut de

Minéralogie, de Physique des Matériaux et de Cosmochimie,

IMPMC, Paris, France. ⁹Department of Electrical and

Electronic Engineering, Imperial College, London, UK.

¹⁰Department of Earth, Planetary, and Space Sciences,

University of California Los Angeles, Los Angeles, CA, USA.

¹¹School of Earth Sciences, University of Bristol, Bristol, UK.

¹²DLR Institute of Planetary Research, Berlin, Germany.

¹³Laboratoire de Planétologie et Géodynamique (LPG), UMR

CNRS 6112, Université de Nantes, Université d'Angers,

France. ¹⁴Royal Observatory of Belgium, Brussels, Belgium.

¹⁵Swiss Seismological Service (SED), ETH Zürich, Zürich,

Switzerland. ¹⁶Schmidt Institute of Physics of the Earth RAS,

Moscow, Russia.

*Corresponding author. Email: simon.staehler@erdw.ethz.ch

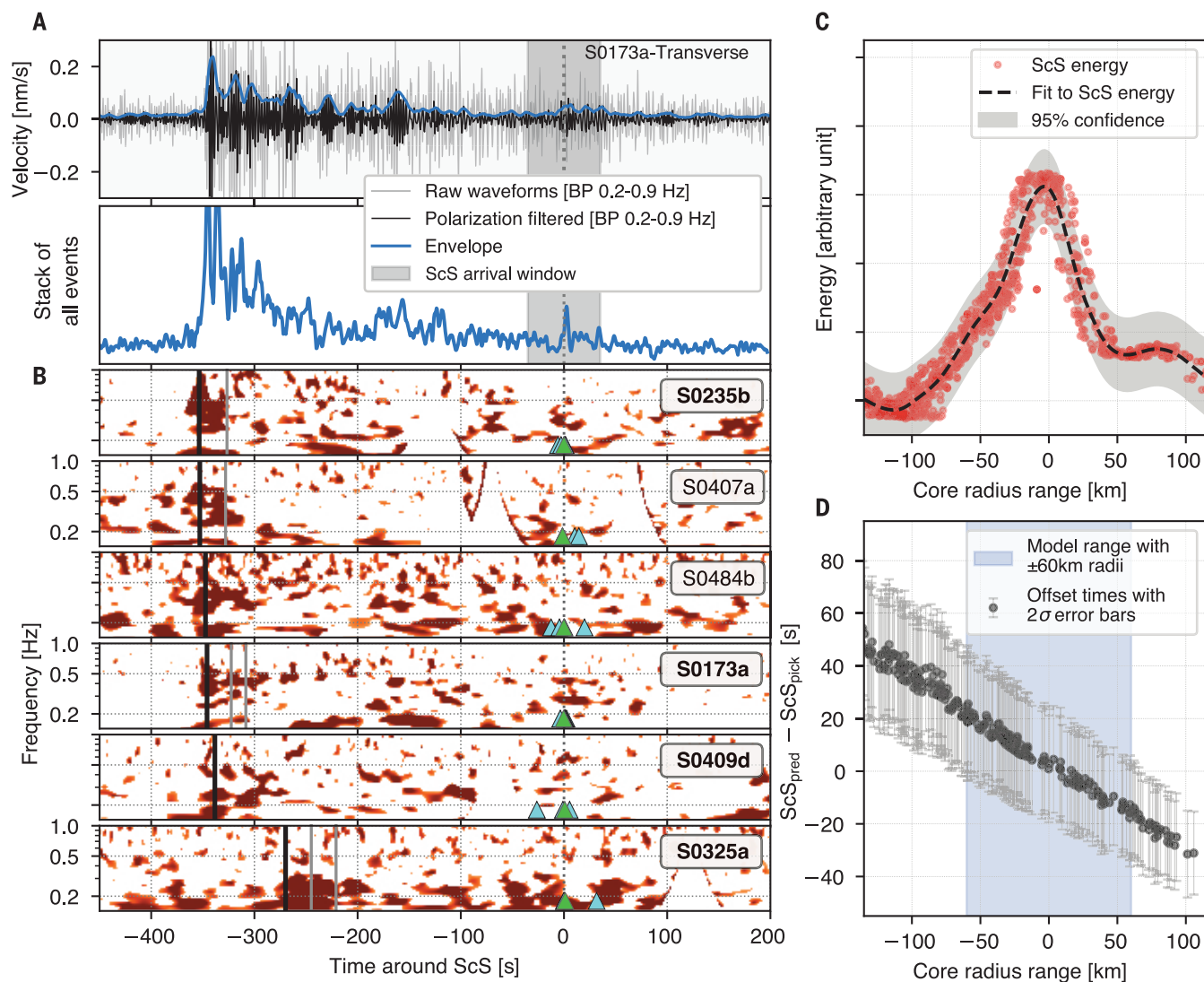


Fig. 1. Summary of data processing and identification of core-reflected S waves. (A) Seismogram and envelope of event S0173a before and after polarization filtering (top) and envelope stack for all six events (bottom). The individual event envelopes are shown in fig. S1-4 (38). (B) Polarization-filtered spectrograms for each of the six events considered here. Cyan triangles mark proposed ScS picks based on the processing methods described in the main text, and green triangles mark the final set of picks summarized in Table 1. Light gray lines mark SS and SSS picks from (32).

All events show energy around the predicted arrival time of ScS, using a model from (32) with a core radius of 1830 km, in agreement with the ScS observation for S0173a shown in (A). Bold event labels (e.g., S0235b) indicate events with strong ScS energy. (C) Stacked energy in a 10-s time window around ScS as predicted for 5000 models from (32) with core radii centered around 1830. (D) Residual travel times of the models presented in (C) compared with the picks in (B), show that the ScS picks are able to constrain the core radius to within ± 60 km (80).

The events for which the direction could be determined from *P*- and *S*-wave polarization (S0173a and S0235b) appear to originate in the Cerberus Fossae graben system (29, 34), with backazimuths of 70° to 90° . The focal mechanisms inferred for these events confirmed normal faulting (35), resulting in relatively strong radiation of *SH* waves toward the core.

All low-frequency marsquakes show an *S*-wave coda dominated by receiver-side scattering (Fig. 1A) (31). Within this coda, the marsquake with the highest SNR observed so far (S0173a) shows a peak around 350 s after the main

S-wave arrival. We used polarization filtering (36) to enhance signals that are linearly and horizontally polarized, as expected for ScS waves (Fig. 1A). After filtering, we observed matching signals for five other events (S0235b, S0407a, S0409d, S0484a, and S0325a) at a similar distance (Fig. 1B). To identify energy pulses that are consistent with ScS, we performed a slant stack (37) for the events, using predicted travel times in 5000 mantle models compatible with surface-reflected seismic body waves (32). The stack of energy in a 10-s time window (filtered between 0.3 and 0.8 Hz) around the predicted ScS arrival time

shows the presence of a signal (Fig. 1C), at the time when a reflection from an ~ 1600 -km depth interface is expected. This result is confirmed by a model-agnostic vespagram stack showing a low-slowness, steep incidence-angle arrival compatible with a deep reflector (38). To refine the arrival time estimate, the time window identified in the stacking was further examined using two independent methods: (i) manual picking in polarization-filtered filter banks to discriminate seismic phases and potential artifacts from instrument or wind and (ii) coda correlation using the *S* wave-form as matching filter, assuming that the

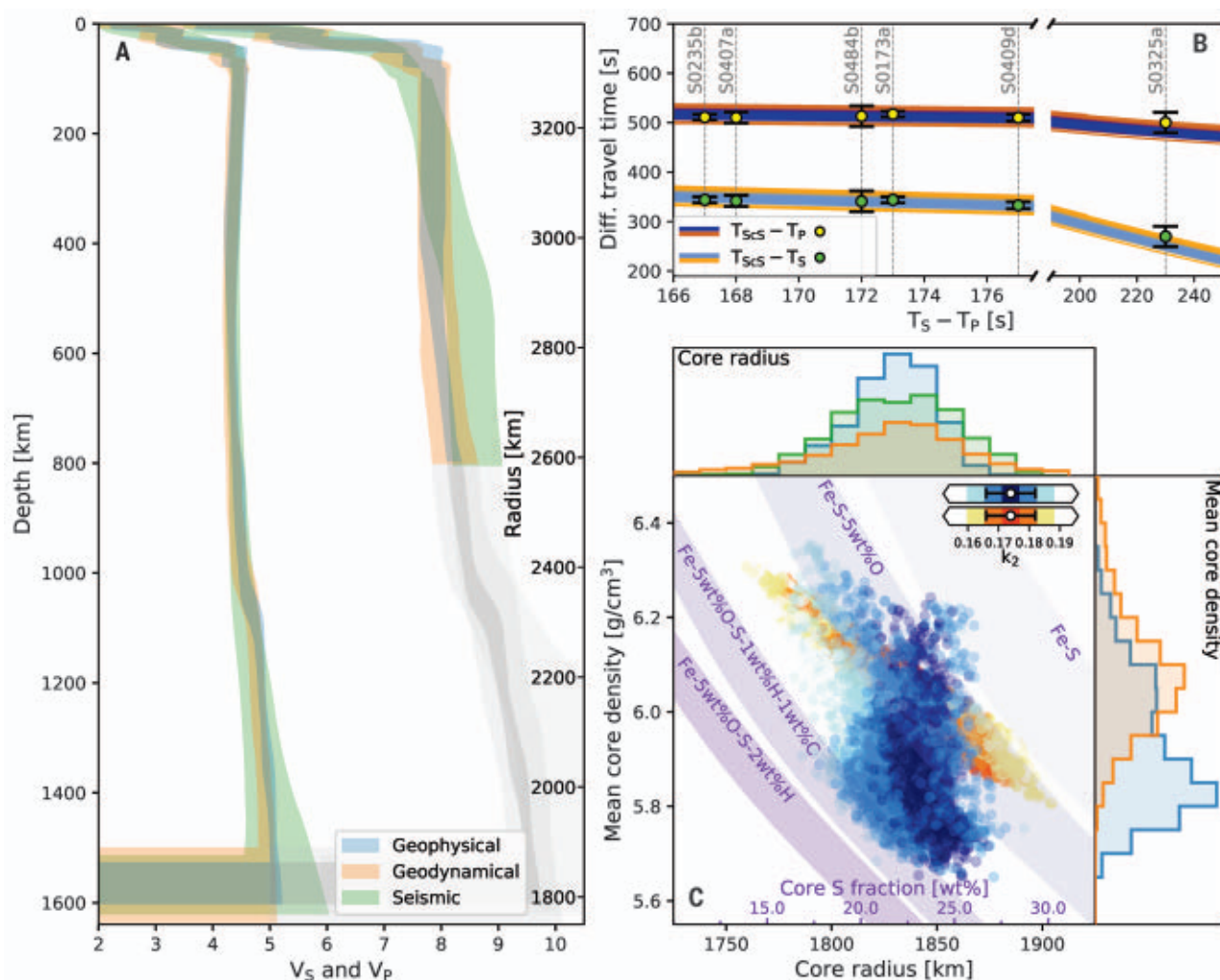


Fig. 2. Mars' mantle and core structure. (A) Inverted seismic wave velocity profiles (95% credible intervals) based on the three inversion methods defined here (see legend). The pale gray shaded area below 800-km depth for the *P*-wave velocity profiles indicates that no direct information is available for this region. (B) Differential body wave travel time misfits for all sampled models obtained from the geophysical (blue shades) and geodynamical (orange shades) inversions shown in (A). Yellow and green circles indicate the observations including error bars. A detailed version of the misfit plot is shown in fig. S9-1 (38). (C) Sampled core properties. The middle plot shows

the mean core density versus core radius for the geophysical (blue) and geodynamical (orange) methods, whereas their marginal distributions are shown as histograms to the right and on the top. The seismic method only constrains core radius. The blue and orange models are color-coded according to their fit to the tidal response in the form of the observed degree-2 Love number k_2 (11), defined by the white circles and horizontal error bars. Relying on Fe-S models, the purple bands indicate the variation of core sulfur (S) content (purple axis) with mean core density for four different iron (Fe)-light element (S, O, H, and C) assemblages (65–68).

S-wave train has high similarity to *ScS* (38). This resulted in a set of picks for each event (Fig. 1D). The results from the two methods were compared with the event slant-stacks to remove misidentified signals and were consolidated into a single set of picks (Table 1).

Whereas the SNR of the *ScS* phases is around 2 or less, the signal is significant compared with the energy of the *S*-wave coda, wind noise, and known instrument artifacts (38, 39). From travel-time tables computed in seismic velocity models compatible with observations so far (8, 32, 40), no other seismic phase is predicted to arrive this late in the seismogram with similar move-out and polar-

ization. Love waves traveling along the surface would have similar polarization but arrive much earlier and show a large move-out and dispersion. Multiply-reflected body waves should also arrive earlier. We therefore interpreted this signal as an *S*-wave reflection from the CMB that, when using seismic velocity models from (32), corresponds to a core with a radius in the range 1770 to 1890 km. We also searched for other phases, such as core-reflected *P* waves (*PcP*), but were unable to find consistent arrivals. This was not unexpected because of the lower *P*-wave reflection coefficient at the CMB. A shear wave that is converted into a *P* wave at the CMB (*ScP*) is predicted to arrive 290 s after *P* for S0325a,

and we observed a weak phase by correlation analysis. This arrival is only 2 s after the identified *SSS* arrival reported in (32), so the peak, although consistent, is not used for inversion here.

The spectral character and the travel time of the deep-diving direct *S* wave for the most distant event (S0167b) located to date (30, 32) were consistent with and comparable to those of the *ScS* phase with similar travel times (fig. S6-1) (38). This supports the identification of the latter as a core reflection that has traversed the entire attenuating mantle. The attenuation-corrected power of the *ScS* phases is below that of the corresponding *S*-wave arrivals

by 10 to 20 dB. This value is slightly below the ray-theoretical prediction from geometrical spreading (fig. S6-2) (38), assuming full reflection of *SH* waves at the CMB, possibly because of scattering at other interfaces and three-dimensional structure. For four of the six events we investigated, (32) reports a second arrival within 20 s after the main *S* wave. We considered these secondary arrivals to be the depth phase (*sS*), resulting in marsquakes that occur in the depth range 20 to 35 km.

We inverted differential *ScS* travel times (Table 1) with respect to *P* together with the travel times of *S*, *PP*, *SS*, *PPP*, and *SSS* from (32) and geodetic data in the form of the degree-2 Love number (*I*1), mean density (*I*6), and mean moment of inertia (*I*2) for mantle *P*- and *S*-wave velocity and density, epicentral distance for all considered events, and radius and mean density of the core. We conducted three separate inversions that relied on a pure seismic parameterization and two mineralogically constrained inversions: a geodynamic (41, 42) and a geophysical parameterization (4). The seismic parameterization considers a layered model of Mars described by *P*- and *S*-wave velocity gradients, respectively, and does not use the geodetic data. The geodynamic parameterization depends on quantities that influence the thermochemical evolution of the planet and accounts for 4.5 billion years of planetary evolution. The geophysical parameterization relies on a unified description of phase equilibria, seismic properties, and thermochemical parameters. The parameterizations (38) reflect, in going from seismic over geodynamic to geophysical parameterization, a decrease in the number of degrees of freedom, as the two latter parameterizations depend increasingly on mineral physics information and therefore better-resolved parameters. Because the mean density of the core depends on that of the mantle and therefore on the bulk mantle composition, we considered six different model martian compositions (12, 14, 43–47) as part of the geophysical inversion. To solve the inverse problem, we employed a stochastic algorithm (48) that samples models that fit the differential body wave travel times within uncertainties and are consistent with prior information (38).

We plotted the results from the joint inversion of the differential body wave travel times and the geophysical data (Fig. 2). The *S*-wave velocity profiles (Fig. 2A) we obtained from the three parameterizations were found to be in good agreement. More scatter exists in the *P*-wave velocity profiles, which reflects fewer *P*-wave observations, and structure is only constrained to 800-km depth (Fig. 2A). All parameterizations provide a good fit to the *ScS-P* travel time observations (Fig. 2B). Above 800-km depth, the velocity profiles are similar to those obtained by the upper mantle

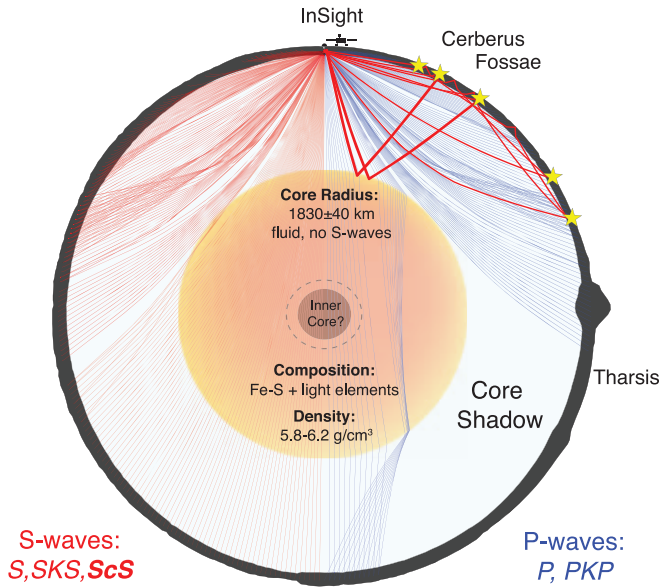


Fig. 3. Schematic diagram of Mars' interior structure. The cross section depicts the core-induced shadow zone for seismic waves. The surface topography is a cut through the MOLA map (81) on a great circle arc from InSight through Olympus Mons. The *S*-wave shadow zone is minimal and probably filled by diffracted *S* waves (*S*dif), whereas the *P*-wave shadow zone is large and contains specifically the Tharsis region. The existence of an inner core cannot be determined by current data, and the seismic ray paths shown assume no inner core. Topography and InSight lander are exaggerated in scale.

inversion (32), and below, the *S*-wave velocity profiles show a distinct increase around 1050-km depth, equivalent to the 410-km seismic discontinuity in Earth's mantle that marks the onset of the mantle transition zone, where the dominant upper mantle mineral olivine transforms to wadsleyite. The CMB occurs between 1520- and 1600-km depth, corresponding to CMB pressures of 18 to 19 GPa and temperatures in the range of ~1900 to 2000 K. These conditions are unfavorable for the stabilization of bridgmanite and imply that the lower mantle of Mars is mineralogically comparable to Earth's mantle transition zone. This means that a relatively dense and thermally insulating lower mantle is absent in Mars, which favors the development of an early thermally driven dynamo as a means of explaining crustal magnetism because of elevated core heat flux (1, 49, 50).

In agreement with geodetic observations that require a liquid core (10), the observation of *ScS* with substantial relative amplitudes compared with direct *S* waves rules out a solid outer core because reflection coefficients would be too small at a solid-solid interface [fig. S6-2 (38)]. The separate inversions converge on the same mean radius but show more spread in mean core density (Fig. 2C), which reflects the trade-off with mantle density through bulk mantle composition (indicated by the blue circles in Fig. 1C). On the basis of the distributions, we estimate core radius to be 1830 ± 40 km, at the upper end of premission estimates (4, 8, 15, 16) that were based on an earlier and slightly lower degree-2 Love number [0.169 ± 0.006 (57)], and mean core density in the range 5.7 to 6.3 g/cm³. We also conducted separate inversions using the geodynamic method to consider the influence

Table 1. Consolidated differential travel times of *S* and *ScS* for the events used in this study. Magnitudes, *M_w*, are from the Marsquake Service catalog, version 6 (30), as defined by Böse et al. (82). Depth estimates are based on the identification of the depth phase *sS* (see the main text). The events are labeled by mission Sol of occurrence and sublabelled alphabetically for Sols with more than 1 event.

Event	tS-tP (s)	tScS-tP (s)	sigma(tScS)	Depth (km)	M _w
S0235b	167	511	3	24 ± 5	3.5
S0407a	168	510	10	25 ± 5	3.0
S0484b	172	513	20	33 ± 5	2.9
S0173a	173	512	3	24 ± 5	3.6
S0409d	177	510	5	25 ± 5	3.1
S0325a	230	500	20	30 ± 5	3.7

of individual datasets on the retrieved core properties (38). These inversions showed that the mean core radius changed from 1836 km (seismic data only) to 1815 km (geodetic data only), whereas the mean core density remained unchanged. To test the influence of the source depth, we fixed it to 50 km for all events—consistent with (29, 32)—and found that this would change the core radius to 1820 ± 40 km, within the above range.

Compositional constraints on the core typically derive from geochemical models coupled with metal-silicate partitioning and mass balance arguments (52–55) but depend on the assumed compositions of the building blocks (56, 57). Although sulfur is commonly considered the main light element (14, 44, 45, 55) because of its abundance in the mantle as determined from the martian meteorites (58) and its siderophile nature at the P-T- fO_2 conditions of the formation of Mars' core (59), additional light elements, including C, O, Si, N, and H, are all potentially viable candidates, as in Earth's core (60–63). Ni is also expected to be a core constituent based on meteorite compositions (64) and should make up 5 to 6 wt % (4). The purple-shaded areas in Fig. 2C indicate how the mean core density varies with S content in the Fe-S, Fe-S-O, Fe-S-O-H, and Fe-S-O-H-C systems, based on thermodynamic solution models constructed using experimental data (65–68). For a core composed of Fe and S, sulfur contents surpass 25 wt %, which is above the value of the sulfur-richest meteorites (EH chondrites) and in excess of what is deduced from geo- and cosmochemical models [<13 to 19 wt % (52, 55)]. To bring the S content in line with the cosmochemical constraints, additional light elements (e.g., C, O, and H) in the core are needed (fig. S11–1) (38). For geochemically defensible amounts of S, O, and H, the mean core density has to be >6 g/cm³, which encompasses the upper end of our predictions (Fig. 2C). Because our core density estimate is tied to reasonable assumptions about the composition and temperature of the martian mantle, a higher mean core density is possible and requires less light elements as a result of a lower bulk mantle FeO content (46) or higher temperatures as seen in some geodynamic models (69). The influence of a lower bulk mantle FeO content on mean core density is reflected by the cloud of blue circles with a mean core density >6.1 g/cm³ (Fig. 2C) that were obtained on the basis of the bulk mantle composition of (46). Thus, a lower bulk mantle FeO content seems to provide a better match than the canonical martian compositions with FeO contents exceeding 17 wt % (12, 14, 44, 45), which had also been found to be at odds with geophysical constraints (4, 57). As a preliminary observation, our results can be construed as pointing to an

Fe-Ni core that is composed of, in order of abundance, S (10 to 15 wt %), O (<5 wt %), and H and C (<1 wt % level) (38). Although such O, H, and C contents represent upper limits, they serve to emphasize the need for light elements in Mars' core.

A chemical composition close to the eutectic (~ 15 wt % S in the Fe-FeS binary at the CMB conditions of Mars) likely prevented crystallization of a bottom-up inner core because of the temperatures (~ 1200 K) required to drop below the liquidus (70). This is compatible with the absence of a current martian geodynamo (71). Strong crustal remnant magnetization in the southern hemisphere of Mars (72) and observations of further magnetized units suggest a dynamo that was active between 4.5 and 3.7 billion years ago (9). The dynamo would have been thermally driven in the first few hundred million years (49, 50) and possibly followed by a compositionally driven dynamo that may resuscitate through FeO exsolution (73) or inner-core crystallization (71, 74, 75). This, however, depends critically on the light-element content and thermal state of the core.

As a consequence of the large core, the seismic core shadow on Mars (Fig. 3) commences at closer epicentral distances (94° to 98°) than on Earth (100° (76)). This makes up half of the planet, including 33% of total extensional faults, specifically 75% of those in terrain younger than 600 million years (77). For marsquakes in Tharsis, the region presumed to host most of the recent tectonic activity (78), direct *P* and *S* waves are therefore unobservable, and this may result in an underestimation of the global seismic activity of Mars as seen from InSight's location (27).

The differentiation of Mars into a primordial crust, mantle, and core is likely the result of early magma ocean crystallization and solidification that could potentially result in compositional stratification of the mantle (69, 79). There is, however, no direct evidence for this on the basis of current observations. Continued analysis using marsquakes observed during the extended mission will be required to delve into the question of deep mantle layering. To determine the compressional wave speeds of the core itself and further refine light-element content, an unequivocal observation of an event beyond the core shadow, with clear core-crossing phases like *PKP* or *SKS*, will be necessary. This would allow us to employ velocity-density systematics (61) as a means of gaining further insight into the total light-element content of the core. In the interim, the new martian seismic data and models presented here provide a wealth of new insights into the interior structure of Mars, which contain the clues needed to unravel the planetary building blocks (57); the physical and chemical conditions during assembly (52); and the chronology of crust, mantle, and core formation (25).

REFERENCES AND NOTES

- D. J. Stevenson, *Nature* **412**, 214–219 (2001).
- M. T. Zuber et al., *Science* **287**, 1788–1793 (2000).
- C. M. Bertka, Y. Fei, *Earth Planet. Sci. Lett.* **157**, 79–88 (1998).
- A. Khan et al., *J. Geophys. Res. Planets* **123**, 575–611 (2018).
- D. Breuer, D. A. Yuen, T. Spohn, *Earth Planet. Sci. Lett.* **148**, 457–469 (1997).
- P. van Thienen, N. J. Vlaar, A. P. van den Berg, *Phys. Earth Planet. Inter.* **142**, 61–74 (2004).
- T. Ruedas, P. J. Tackley, S. C. Solomon, *Phys. Earth Planet. Inter.* **216**, 32–58 (2013).
- S. E. Smrekar et al., *Space Sci. Rev.* **215**, 3 (2019).
- A. Mittelholz, C. L. Johnson, J. M. Feinberg, B. Langlais, R. J. Phillips, *Sci. Adv.* **6**, eaba0513 (2020).
- C. F. Yoder, A. S. Konopliv, D. N. Yuan, E. M. Standish, W. M. Folkner, *Science* **300**, 299–303 (2003).
- A. S. Konopliv et al., *Geophys. Res. Lett.* **47**, e2020GL090568 (2020).
- G. Dreibus, H. Wänke, in *Proceedings of the 27th International Geological Congress*, vol. 11 (1984), pp. 1–20.
- H. Y. McSweeney Jr., K. Keil, *Geochim. Cosmochim. Acta* **64**, 2155–2166 (2000).
- G. J. Taylor, *Geochemistry* **73**, 401–420 (2013).
- A. Bagheri, A. Khan, D. Al-Attar, O. Crawford, D. Giardini, *J. Geophys. Res. Planets* **124**, 2703–2727 (2019).
- A. Rivoldini, T. Van Hoolst, O. Verhoeven, A. Mocquet, V. Dehant, *Icarus* **213**, 451–472 (2011).
- C. Hutton, *Phil. Trans. R. Soc.* **68**, 689–788 (1778).
- R. D. Oldham, *Q. J. Geol. Soc.* **62**, 456–475 (1906).
- I. Lehmann, *Seismol. Internat. Strasbourg Public. Bur. Centr. Sci.* **14**, 87–115 (1936).
- K. E. Bullen, *Phys. Chem. Earth* **1**, 68–93 (1956).
- C. G. Dahm, *Eos Trans. AGU* **15**, 80–83 (1934).
- G. Latham et al., *Science* **174**, 687–692 (1971).
- R. F. Garcia, J. Gagnepain-Beyneix, S. Chevrot, P. Lognonné, *Phys. Earth Planet. Inter.* **188**, 96–113 (2011).
- R. C. Weber, P.-Y. Lin, E. J. Garnero, Q. Williams, P. Lognonné, *Science* **331**, 309–312 (2011).
- N. Dauphas, A. Pourmand, *Nature* **473**, 489–492 (2011).
- K. Mezger, V. Debaille, T. Kleine, *Space Sci. Rev.* **174**, 27–48 (2013).
- W. B. Banerdt et al., *Nat. Geosci.* **13**, 183–189 (2020).
- P. Lognonné et al., *Space Sci. Rev.* **215**, 12–12 (2019).
- D. Giardini et al., *Nat. Geosci.* **13**, 205–212 (2020).
- InSight Marsquake Service, Mars Seismic Catalogue, InSight Mission: V6 2021-04-01 (ETHZ, IPGP, JPL, ICL, MPS, Univ. Bristol, 2021); <https://doi.org/10.12686/a11>.
- P. Lognonné et al., *Nat. Geosci.* **13**, 213–220 (2020).
- A. Khan et al., *Science* **373**, 434–438 (2021).
- B. Knapmeyer-Endrun et al., *Science* **373**, 438–443 (2021).
- J. Taylor, N. A. Teanby, J. Wookey, *J. Geophys. Res. Planets* **118**, 2570–2581 (2013).
- N. Brinkman et al., *J. Geophys. Res. Planets* **126**, e2020JE006546 (2021).
- M. Schimmel, J. Gallart, *Bull. Seismol. Soc. Am.* **94**, 1016–1035 (2004).
- P. S. Schultz, J. F. Claerbout, *Geophysics* **43**, 691–714 (1978).
- Materials and methods are available as supplementary materials.
- J.-R. Scholz et al., *Earth Space Sci.* **7**, 1–31 (2020).
- A.-C. Plesa et al., *J. Geophys. Res. Planets* **126**, e2020JE00675514 (2021).
- M. Drilleau, H. Samuel, A. Rivoldini, M. Panning, P. Lognonné, *Geophys. J. Int.* **226**, 1615–1644 (2021).
- H. Samuel, P. Lognonné, M. Panning, V. Lainey, *Nature* **569**, 523–527 (2019).
- A. Khan, J. A. D. Connolly, *J. Geophys. Res.* **113**, E07003 (2008).
- K. Ladders, B. Fegley Jr., *Icarus* **126**, 373–394 (1997).
- C. Sanloup, A. Jambon, P. Gillet, *Phys. Earth Planet. Inter.* **112**, 43–54 (1999).
- T. Yoshizaki, W. F. McDonough, *Geochim. Cosmochim. Acta* **273**, 137–162 (2020).
- C. B. Agee, D. S. Draper, *Earth Planet. Sci. Lett.* **224**, 415–429 (2004).
- K. Mosegaard, A. Tarantola, *J. Geophys. Res.* **100**, 12431–12447 (1995).
- D. Breuer, T. Spohn, *J. Geophys. Res.* **108**, 5072 (2003).
- J.-P. Williams, F. Nimmo, *Geology* **32**, 97–100 (2004).
- A. S. Konopliv, R. S. Park, W. M. Folkner, *Icarus* **274**, 253–260 (2016).
- M. C. Brennan, R. A. Fischer, J. C. E. Irving, *Earth Planet. Sci. Lett.* **530**, 115923 (2020).
- K. Righter, N. L. Chabot, *Meteorit. Planet. Sci.* **46**, 157–176 (2011).

54. D. C. Rubie *et al.*, *Icarus* **248**, 89–108 (2015).
55. E. S. Steenstra, W. van Westrenen, *Icarus* **315**, 69–78 (2018).
56. C. Fitoussi, B. Bourdon, X. Wang, *Earth Planet. Sci. Lett.* **434**, 151–160 (2016).
57. C. Liebske, A. Khan, *Icarus* **322**, 121–134 (2019).
58. G. Dreibus, H. Wänke, *Meteoritics* **20**, 367–381 (1985).
59. N. Rai, W. van Westrenen, *J. Geophys. Res. Planets* **118**, 1195–1203 (2013).
60. J. Badro, J. P. Brodholt, H. Piet, J. Siebert, F. J. Ryerson, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12310–12314 (2015).
61. F. Birch, *J. Geophys. Res.* **69**, 4377–4388 (1964).
62. J.-P. Poirier, *Phys. Earth Planet. Inter.* **85**, 319–337 (1994).
63. V. N. Zharkov, *Sol. Syst. Res.* **30**, 456 (1996).
64. A. E. Ringwood, *Geochim. J.* **11**, 111–135 (1977).
65. J. V. Badding, H. K. Mao, R. J. Hemley, in *High-Pressure Research: Application to Earth and Planetary Sciences* (American Geophysical Union, 1992), pp. 363–371.
66. T. Komabayashi, *J. Geophys. Res. Solid Earth* **119**, 4164–4177 (2014).
67. G. Morard *et al.*, *Am. Mineral.* **103**, 1770–1779 (2018).
68. H. Terasaki *et al.*, *J. Geophys. Res. Planets* **124**, 2272–2293 (2019).
69. L. T. Elkins-Tanton, E. M. Parmentier, P. C. Hess, *Meteorit. Planet. Sci.* **38**, 1753–1771 (2003).
70. Y. Mori *et al.*, *Earth Planet. Sci. Lett.* **464**, 135–141 (2017).
71. A. J. Stewart, M. W. Schmidt, W. van Westrenen, C. Liebske, *Science* **316**, 1323–1325 (2007).
72. M. H. Acuña *et al.*, *Science* **284**, 790–793 (1999).
73. K. Tsuno, E. Ohtani, H. Terasaki, *Phys. Earth Planet. Inter.* **160**, 75–85 (2007).
74. D. Breuer, W. B. Moore, in *Treatise on Geophysics*, G. Schubert, Ed. (Elsevier, ed. 2, 2015), pp. 255–305.
75. D. J. Hemingway, P. E. Driscoll, *J. Geophys. Res. Planets* **126**, e2020JE006663 (2021).
76. B. L. N. Kennett, E. R. Engdahl, *Geophys. J. Int.* **105**, 429–465 (1991).
77. M. Knapmeyer *et al.*, *J. Geophys. Res.* **111**, E11006 (2006).
78. R. C. Anderson *et al.*, *J. Geophys. Res.* **106**, 20563–20585 (2001).
79. H. Samuel *et al.*, *J. Geophys. Res. Planets* **126**, e2020JE006613 (2021).
80. A. G. Marusiak, N. C. Schmerr, M. E. Banks, I. J. Daubar, *Icarus* **335**, 113396 (2020).
81. D. E. Smith *et al.*, *J. Geophys. Res.* **106**, 23689–23722 (2001).
82. M. Böse *et al.*, *Bull. Seismol. Soc. Am.*, 10.1785/0120210045 (2021).
83. J. D. Hunter, *Comput. Sci. Eng.* **9**, 90–95 (2007).
84. L. Krischer *et al.*, *Comput. Sci. Discov.* **8**, 014003 (2015).
85. C. R. Harris *et al.*, *Nature* **585**, 357–362 (2020).
86. P. Virtanen *et al.*, *Nat. Methods* **17**, 261–272 (2020).
87. InSight Mars SEIS Data Service, SEIS raw data, InSight Mission (IPGP, JPL, CNES, ETHZ, ICL, MPS, ISAE-Supero, LPG, MFSC, 2019); <https://doi.org/10.18715/SEIS.INSIGHT.XB> 2016.
88. S. C. Stähler, M. Drilleau, A. C. Duran, A. Khan, H. Samuel, Interior Models of Mars from inversion of seismic body waves (2021); <https://doi.org/10.18715/IPGP.2021.kpmqmrz8>.

ACKNOWLEDGMENTS

This is InSight contribution 200. We acknowledge NASA, CNES, and partner agencies and institutions (UKSA, SSO, ESA-PRODEX, DLR, JPL, IPGP-CNRS, ETHZ, IC, and MPS-MPG) for the development of SEIS. Numerical simulations were supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s922 as well as HPC resources of CINES under the allocation A0090407341, made by GENCI. We thank B. Dintrans, director of CINES, for his efficient handling of our request for computational time. Figures were created using matplotlib (83), seismic data processing was done in ObsPy (84), and numerical evaluation was done in NumPy and SciPy (85, 86). **Funding:** S.C.S., A.K., D.G., J.C., A.C.D., G.Z., and N.D. acknowledge support from ETHZ through the ETH+ funding scheme (ETH+2 19-1: “Planet MARS”). S.C.S. acknowledges funding from ETH research grant ETH-10 17-3. W.B.B., A.G.M., M.P.P., and S.E.S. were supported by the NASA InSight mission and funds from the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). D.A. has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement 724690). The French teams acknowledge support from CNES as well as Agence Nationale de la Recherche (ANR-14-CE36-0012-02 and ANR-19-CE31-0008-08). A.R. was financially supported by the Belgian PRODEX program managed by the European Space Agency in collaboration with the Belgian Federal Science Policy Office. M.S. wishes to thank SANIMS (RTI2018-095594-B-I00).

M.v.D. received support from the ERC under the European Union’s Horizon 2020 program (grant no. 714069). D.S. and C.S. acknowledge funding from ETH research grant ETH-06 17-02. J.C.E.I. acknowledges support from NASA grant 80NSSC18K1633. N.S., D.K., Q.H., R.M., V.L., and A.G.M. acknowledge NASA grant 80NSSC18K1628 for support. V.L. acknowledges support from the Packard Foundation. W.T.P. and C.C. received funding from the UK Space Agency, grant ST/S001239/1. A.H. was funded by the UK Space Agency (grant ST/R002096/1). A.-C.P. acknowledges the financial support and endorsement from the DLR Management Board Young Research Group Leader Program and the Executive Board Member for Space Research and Technology. **Author contributions:** S.C.S., D.G., S.C., R.F.G., Q.H., D.K., V.L., M.S., N.S., D.S., É.S., and G.Z. analyzed the seismic data and made ScS arrival time picks. S.C.S., P.L., D.G., Z.X., C.C., and W.T.P. performed the statistical analysis of the observed signals. S.C.S., Q.H., N.S., R.M., and A.G.M. identified the arrivals as ScS waves based on interior models from A.K., H.S., and A.R. A.K., M.D., A.C.D., and H.S. performed the inversions. S.C.S., A.K., P.L., D.G., D.A., J.C.E.I., M.K., C.P., A.-C.P., A.R., T.G., and S.E.S. participated and contributed to the interpretation of the results. Review of the continuous data and detection of marsquakes was done by S.C.S., S.C., G.Z., C.C., N.D., J.C., M.v.D., T.K., M.P., and A.H. with operational support by É.B., C.P., and P.M.D. S.C.S. and A.K. wrote the central part of the paper with contributions from H.S., N.S., D.A., J.C.E.I., A.G.M., A.-C.P., A.R., J.C., and M.v.D. J.C.E.I., R.M., M.K., and V.L. reviewed the contributions to the supplementary materials. The InSight mission is managed by W.B.B., M.P.P., and S.E.S. The SEIS

instrument development was led by P.L., D.G., W.T.P., and W.B.B. Supplementary section 1 was written by M.S., D.S., and É.S. with contributions from S.C.S., C.S., and Z.X. Supplementary section 2 was written by D.K. and V.L. with contributions from J.C.E.I. and N.S. Supplementary section 3 was written by M.S. and É.S. Supplementary section 4 was written by R.F.G. with contributions from M.D. Supplementary section 5 was written by Q.H. with contributions from N.S. Supplementary section 6 was written by S.C.S. with contributions from the authors of the other supplements. Supplementary section 7 was written by Z.X. and C.C. with contributions from P.L. and W.T.P. Supplementary section 8 was written by A.K., M.D., A.C.D., and H.S. Supplementary section 9 was written by M.D. Supplementary section 10 was written by A.C.D., A.K., and M.D. Supplementary section 11 was written by D.A. and A.R. with contributions from A.K. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** We thank the operators of JPL SISMOC, MSDS, IRIS-DMC, and PDS for providing SEED SEIS data (87). Three hundred interior models derived in this study are available from MSDS (88).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/443/suppl/DC1
Materials and Methods
Figs. S1-1 to S11-1
Tables S0 to S8-3
References (89–125)

1 April 2021; accepted 14 June 2021
10.1126/science.abi7730

PHOTOPHYSICS

Inhibited nonradiative decay at all exciton densities in monolayer semiconductors

Hyungjin Kim^{1,2,†}, Shiekh Zia Uddin^{1,2,†}, Naoki Higashitarumizu^{1,2}, Eran Rabani^{2,3,4}, Ali Javey^{1,2,*}

Most optoelectronic devices operate at high photocarrier densities, where all semiconductors suffer from enhanced nonradiative recombination. Nonradiative processes proportionately reduce photoluminescence (PL) quantum yield (QY), a performance metric that directly dictates the maximum device efficiency. Although transition metal dichalcogenide (TMDC) monolayers exhibit near-unity PL QY at low exciton densities, nonradiative exciton-exciton annihilation (EEA) enhanced by van-Hove singularity (VHS) rapidly degrades their PL QY at high exciton densities and limits their utility in practical applications. Here, by applying small mechanical strain (less than 1%), we circumvented VHS resonance and markedly suppressed EEA in monolayer TMDCs, resulting in near-unity PL QY at all exciton densities despite the presence of a high native defect density. Our findings can enable light-emitting devices that retain high efficiency at all brightness levels.

Exciton recombination in transition metal dichalcogenide (TMDC) monolayers depends on the exciton generation rate (G), background carrier concentration, and electronic band structure. The combined effect of the first two factors has been studied thoroughly (*1*), and the photocarrier G and background carrier concentration were tuned by varying the incident pump power and the gate voltage (V_g), respectively, in a capacitor structure. Given pronounced Coulomb inter-

actions, background carriers turn photogenerated excitons into charged trions (*2*) that mostly recombine nonradiatively (*1, 3*). In the absence of background carriers, at low exciton densities, neutral excitons in intrinsic monolayers can recombine completely radiatively even in the presence of defects (*1, 4*). However, at high exciton densities, recombination of neutral excitons is dominated by exciton-exciton annihilation (EEA), in which an exciton nonradiatively recombines while colliding with another exciton (*4–6*). All excitonic materials found in nature (*7*) exhibit EEA, which is the primary cause of the efficiency roll-off observed in all organic (*8*) and some inorganic (*9, 10*) light-emitting devices. The EEA effect has similarities to Auger recombination, which is commonly observed in conventional free-carrier systems and is also a prominent cause

¹Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, USA. ²Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. ³Department of Chemistry, University of California, Berkeley, CA 94720, USA. ⁴The Raymond and Beverly Sackler Center of Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 69978, Israel.

*Corresponding author. Email: ajavey@berkeley.edu

†These authors contributed equally to this work.

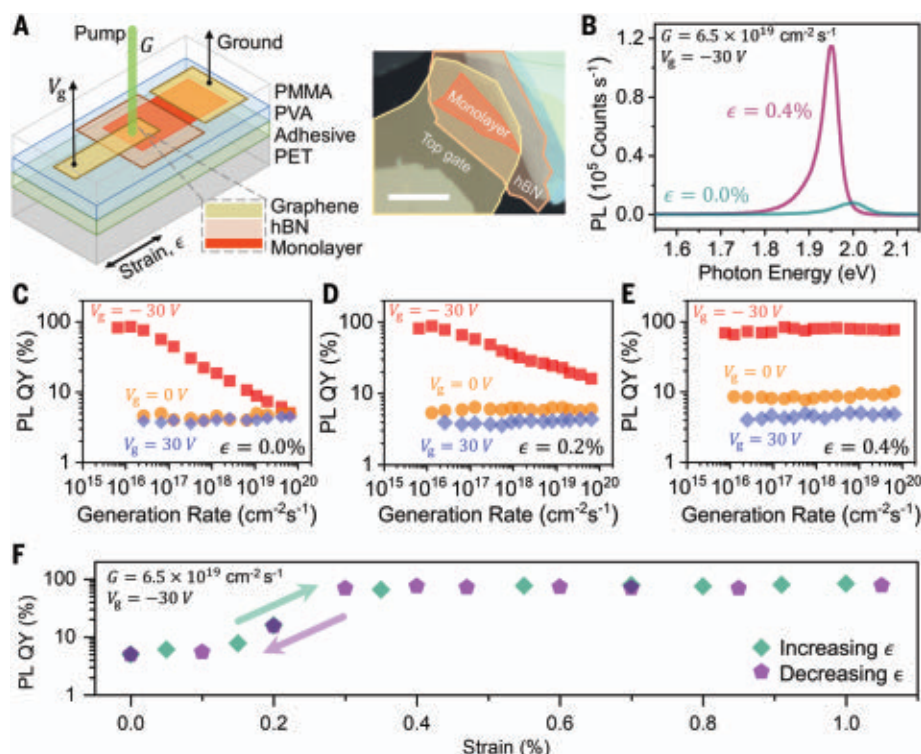


Fig. 1. Near-unity PL QY in WS₂ at all generation rates (Gs). (A) Schematic and optical micrograph of the device structure. Scale bar, 20 μm . A two-terminal capacitor structure with graphene as source and gate and hexagonal boron nitride as insulator was fabricated on a flexible polymer substrate. (B) Comparison of PL spectra of unstrained and 0.4% strained monolayer WS₂ at a high G of $6.5 \times 10^{19} \text{ cm}^{-2} \text{ s}^{-1}$ and a gate voltage of $V_g = -30 \text{ V}$. (C to E) PL QY of monolayer WS₂ as a function of gate voltage, G , and strain. (F) PL QY approaching unity with the application of strain at a high G of $6.5 \times 10^{19} \text{ cm}^{-2} \text{ s}^{-1}$.

of the efficiency roll-off observed in light-emitting diodes and solar cells (*11, 12*). Through the conservation of momentum and energy in the participating quasiparticles, EEA depends on the aforementioned third factor: detailed band structure (*6, 13, 14*).

In this work, we modulated these three factors simultaneously. Along with V_g and pump power, we altered the electronic band structure by applying a uniaxial strain (ϵ). With appropriate strain, all neutral excitons recombined radiatively even at high concentrations in monolayers of WS₂, WSe₂, and MoS₂, resulting in near-unity quantum yield (QY) at all measured G s. The density of states of electrons in a two-dimensional (2D) periodic crystal are topologically constrained to exhibit logarithmic van Hove singularities (VHSs) arising from saddle points in the energy dispersion (*15*). When the energy of a transition is near the VHS, weak interactions are often intensified by the enhanced density of states (DOS) (*16*). On the basis of the energy and momentum conservation, we found that as-exfoliated TMDC monolayers exhibited enhanced EEA because the final energy of this process coincided with inherent VHS. Strain drove the final energy away from the VHS resonance and substantially re-

duced EEA in both sulfur- and selenide-based TMDCs. We could uniformly suppress all non-radiative recombination in a centimeter-scale chemical vapor deposition (CVD)-grown WS₂ monolayer at all G s.

Monolayer field effect transistor devices were fabricated on plastic flexible substrate, enabling simultaneous modulation of carrier concentration and strain while performing quantitative photoluminescence (PL) QY measurements (device fabrication details are available in the supplementary materials). Figure 1A shows the schematic and optical micrograph of such a device using WS₂ monolayers. Mechanically exfoliated monolayers were transferred to a polyvinyl alcohol layer attached to polyethylene terephthalate (PET)-handling substrate, a combination selected for its superior strain transfer efficiency (*17*). Exfoliated hexagonal boron nitride (70 to 100 nm in thickness) and graphene (2 to 5 nm in thickness) were transferred sequentially on the monolayer WS₂ to serve as a gate insulator and gate electrode, respectively. The WS₂ monolayer was electrically grounded and V_g was applied to the top graphene. Bending the PET substrate with positive curvature applied tensile strain to the monolayer in the direction of bending (*18*). Applied strain was

calculated from geometrical considerations. G is the number of excitons created or the number of photons absorbed per unit area per unit time, which can be tuned by changing the laser intensity.

As-exfoliated monolayer WS₂ is known to be electron rich. Application of a negative V_g (-30 V) electrostatically compensated for that unintentional electron doping and ensured that the recombination process was dominated by neutral excitons (*1*). At this biasing condition, we compared PL spectra of an unstrained and strained WS₂ monolayer at a high G of $6.5 \times 10^{19} \text{ cm}^{-2} \text{ s}^{-1}$ (Fig. 1B). We observed both a redshift of PL spectra and ~ 15 times enhancement in PL intensity with the application of $\epsilon = 0.4\%$ tensile strain at this high G . Note that G is independent of strain because absorption at pump wavelength does not change with strain (fig. S1).

Calibrated PL measurements at room temperature were performed to quantitatively extract the QY as a function of ϵ , V_g , and G (Fig. 1, C to F, and fig. S2), and Fig. 1C shows PL QY as a function of V_g and G when no strain is applied. At zero and positive gate voltages, the background electron concentration of monolayer WS₂ was large and led to the formation of negative trions that predominantly recombined nonradiatively, thus yielding a low PL QY. At negative V_g , the background electrons were removed, and recombination of neutral excitons dominated. At low G , neutral excitons recombined completely radiatively despite native defect density.

At high G , PL QY rolled off because of EEA. These same results were observed previously for other substrates such as polymethyl methacrylate and SiO₂ and are quantitatively the same as the results here (*1*). With the application of 0.2% tensile strain, the PL QY drop-off at high exciton G was notably reduced for negative V_g (Fig. 1D). For a tensile strain of 0.4%, no PL QY drop-off at high G was observed (Fig. 1E). Thus, all nonradiative recombination processes in as-exfoliated monolayer WS₂ were suppressed by applying tensile strain and V_g .

Suppression of EEA is further elucidated in Fig. 1F, where we show PL QY as a function of strain at $V_g = -30 \text{ V}$ and at a high G of $6.5 \times 10^{19} \text{ cm}^{-2} \text{ s}^{-1}$. For no applied strain, PL QY was low at this G , but for a threshold strain of 0.3%, PL QY increased sharply and asymptotically approached unity. This PL QY enhancement by strain was also reversible and repeatable, because PL QY traces for increasing and decreasing strain fully overlapped (fig. S3). Such reversibility and repeatability indicate an absence of slippage, so applied strain is equal to the actual strain transferred to the monolayer. High PL QY at all pump powers persisted even for a tensile strain of $\epsilon = 1.0\%$ (Fig. 1F and fig. S4). In addition to electrostatic counterdoping, high PL QY was achieved by applying tensile

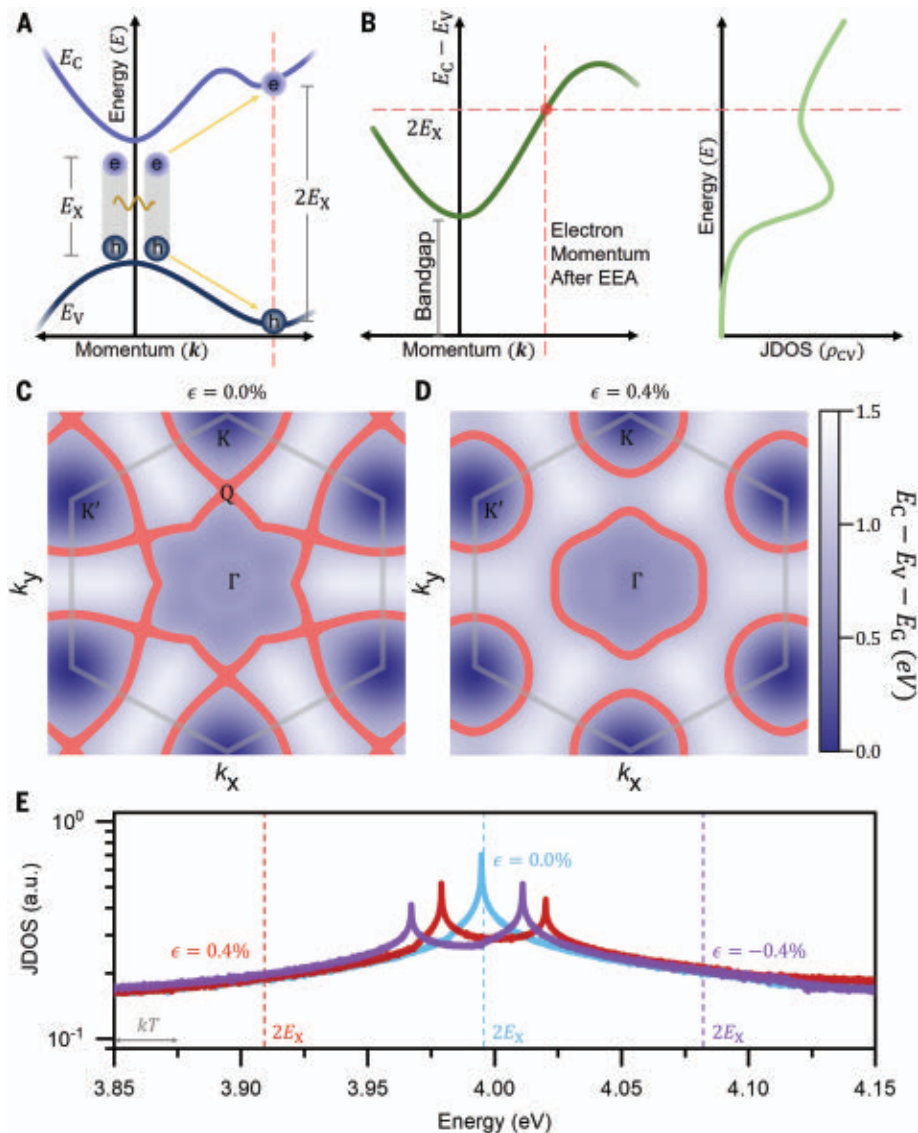


Fig. 2. EEA suppression by strain. (A) Schematic describing the microscopic mechanism of EEA, which occurs when one exciton ionizes another exciton by nonradiatively transferring its energy. (B) Momentum and energy conservation dictate that the electron from the ionized exciton ends up with a wavevector in which the energy difference between the conduction and valence bands is equal to twice the exciton transition energy. The JDOS at that energy determines the rate of EEA. (C and D) Energy difference between the conduction and valence bands for unstrained and 0.4% strained monolayer WS₂, respectively. Gray hexagon denotes the first Brillouin zone (each side is 19.946 nm⁻¹), and red areas are the possible final wavevectors. Possible final wavevectors in unstrained monolayer WS₂ include the saddle points (“Q”) where there are VHSs but strained samples do not. (E) Calculated JDOS for monolayer WS₂. JDOS $\rho_{cv}(E)$ at $E = E_C - E_V = 2E_X$ determines EEA rate. Dashed lines denote values of $2E_X$ at the corresponding strain.

strain in monolayer WS₂ chemically counter-doped by Nafion, a known hole dopant (fig. S5).

The PL QY for neutral excitons ($V_g = -30$ V) can be written as the ratio of radiative recombination rate of to the total recombination rate as follows:

$$QY = \frac{R_r}{R_r + R_{nr}} \quad (1)$$

where R_r and R_{nr} are the exciton radiative and nonradiative recombination rates, respectively.

Nonradiative recombination for neutral excitons is predominantly attained through the EEA process (1). Therefore, $R_{nr} = R_{EEA}$, where R_{EEA} is the nonradiative EEA rate. Even with the application of strain, the semiconductor remained strongly excitonic (6), so strain did not change the PL QY versus G response at low G , and R_r was independent of strain (Fig. 1, D to F). However, because strain changed the QY at high pump powers, R_{EEA} must depend strongly on the strain. Using the experimen-

tally measured PL QYs at a high G of 6.5×10^{19} cm⁻² s⁻¹, we found that

$$\frac{R_{EEA}(\epsilon = 0.4\%)}{R_{EEA}(\epsilon = 0.0\%)} \approx 7 \times 10^{-3} \quad (2)$$

which would indicate a roughly two orders of magnitude decrease in EEA rate at the highest G . Although the EEA rate was not rigorously zero, it was decreased by the application of strain such that $R_r \gg R_{EEA}$, and the radiative relaxation dominated the total recombination.

The photophysics of these monolayer semiconductors as a function of V_g and G arises from the quasiparticle interaction and can be captured by a simple kinetic model described elsewhere (1). The exciton-to-trion ratio can be tuned with V_g , activating their respective recombination routes, whereas high densities of neutral excitons activate the EEA process. However, the mechanism by which strain suppresses EEA at high pump necessitated a closer look at the EEA process itself. EEA occurs when one exciton ionizes another exciton by non-radiatively transferring its energy (Fig. 2A) (6). The initial state consists of two excitons, with center-of-mass momenta \mathbf{K}_1 and \mathbf{K}_2 and energies E_1 and E_2 , respectively. The exciton energy and momentum are related by (13) as follows:

$$E_1 = E_G - E_B + \frac{\hbar^2 |\mathbf{K}_1|^2}{2M_X} \quad (3)$$

$$E_2 = E_G - E_B + \frac{\hbar^2 |\mathbf{K}_2|^2}{2M_X} \quad (4)$$

In the above relations, E_G is the fundamental bandgap, E_B is the exciton binding energy, and M_X is the exciton mass. The final state consists of a high-energy electron and hole, with crystal momenta \mathbf{k}_e and \mathbf{k}_h and energies E_e and E_h , respectively.

Irrespective of the details of the interaction potential, two quantities are conserved in the EEA process: total momentum and energy (13, 19, 20). The condition for conservation of crystal momentum yields (13) is

$$\mathbf{k}_e + \mathbf{k}_h = \mathbf{K}_1 + \mathbf{K}_2 \quad (5)$$

Because \mathbf{K}_1 and \mathbf{K}_2 are determined by the thermal motion of excitons, they are negligible compared with \mathbf{k}_e and \mathbf{k}_h , implying that the electron and the hole momentum in the final state should be almost opposite to each other

$$\mathbf{k}_e \approx -\mathbf{k}_h \quad (6)$$

Therefore, momentum conservation dictates that the electron and hole from the ionized exciton have opposing crystal wavevectors (on

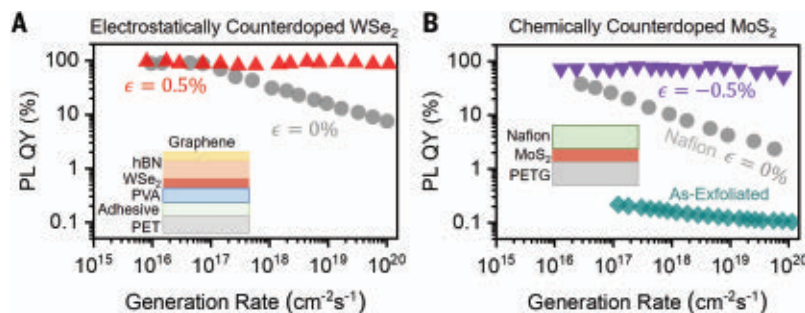


Fig. 3. General nature of EEA suppression. Near-unity PL QY at all generation rates (Gs) in electrostatically counterdoped WSe₂ (A) and chemically counterdoped MoS₂ (B) by tensile and compressive strain, respectively, demonstrates the universality of the conditions that suppress nonradiative recombination.

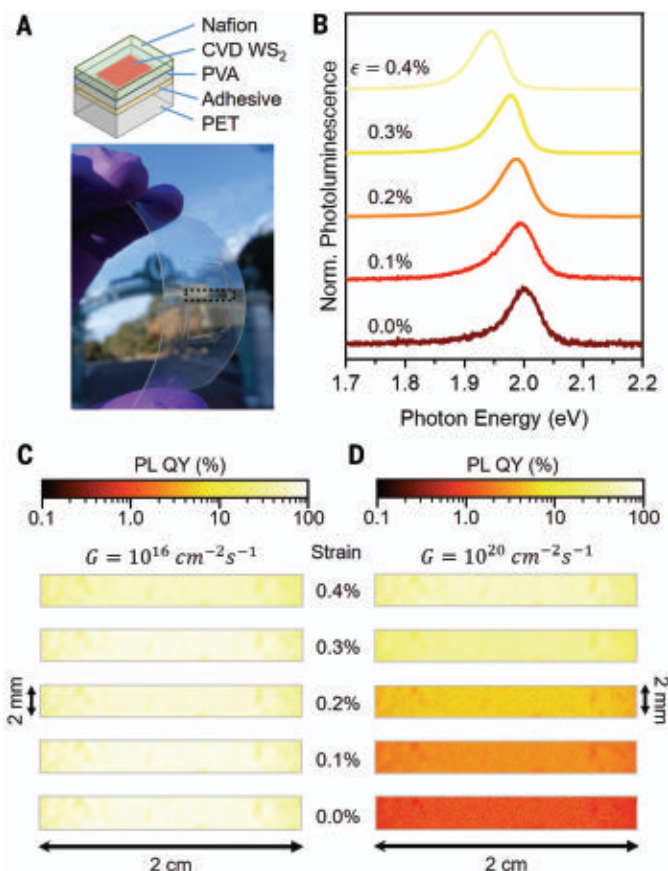


Fig. 4. High PL QY on large-area CVD-grown WS₂. (A) Schematic (top) and photograph (bottom) of the flexible substrate used to strain CVD grown, centimeter-scale WS₂. (B) Normalized PL spectra of a typical spot at different strains. (C and D) Spatial mapping of PL QY of a 2-mm × 2-cm area of WS₂ grown at a low G of $10^{16} \text{ cm}^{-2} \text{ s}^{-1}$ and a high G of $10^{20} \text{ cm}^{-2} \text{ s}^{-1}$, respectively, shows that high PL QY is uniformly achieved by strain at the highest G while maintaining high PL QY at low pump.

top of each other in the energy dispersion, shown with the red dashed line in Fig. 2A). The condition for conservation of crystal momentum yields (13) is

$$E_e + E_h = E_1 + E_2 \quad (7)$$

Because the exciton center of mass wavevector is negligible, $E_1 \approx E_2 \approx E_G - E_B = E_X$, where E_X is

the exciton transition energy. Therefore, energy conservation stipulates that the energy difference between the electron and hole must be $2E_X$

$$E_e + E_h \approx 2E_X \quad (8)$$

If we denote the conduction and valence band of the semiconductor as E_C and E_V , then com-

bined momentum and energy conservation can be written as

$$\begin{aligned} E_e + E_h &= E_C(\mathbf{k}_e) - E_V(-\mathbf{k}_h) \\ &= E_C(\mathbf{k}_e) - E_V(\mathbf{k}_e) = 2E_X \end{aligned} \quad (9)$$

Therefore, any wavevector for which the energy difference between the conduction and valence band is equal to twice the exciton transition energy can be the final wavevector of the electron from the ionized exciton (Fig. 2B). Note that the exciton transition energy is different from the fundamental bandgap because of enhanced electron-hole interaction in TMDC monolayers. By Fermi's golden rule, the joint DOS (JDOS) at twice the exciton transition energy determines the strength of the EEA process (13, 20).

Typically, the effective mass approximation is invoked to determine the energy of the ionized electron. However, the ionized electron and hole have high energy, so this is no longer a valid approximation. We first calculated the band structure of monolayer TMDC with an 11-band tight-binding theory based on Wannier transformation of ab initio density functional theory calculations [details of the band structure are in (21) and (22)]. We then calculated the energy difference between the conduction and valence band ($E_C - E_V$) for unstrained and strained monolayer WS₂, respectively. The first Brillouin zone is indicated by a gray hexagon in Fig. 2, C and D; red areas are the final states of the electron from the ionized exciton where the conservation laws are satisfied.

Possible final wavevectors in unstrained monolayer WS₂ include the saddle points (indicated by “Q” in Fig. 2C), but in 0.4% strained WS₂, they do not include the saddle points (Fig. 2D). Saddle points in the band structure of a 2D semiconductor create VHS and result in a logarithmically diverging JDOS (15, 23, 24). The JDOS for monolayer WS₂ $\rho_{CV}(E)$ at $E = E_C - E_V = 2E_X$ determines the strength of EEA. Figure 2E shows the JDOS of 0.4% strained and unstrained WS₂. In unstrained samples, there was a VHS at twice the exciton transition energy, resulting in an expedited EEA. Strain shifted the exciton transition energy E_X such that $2E_X$ did not overlap VHS resonance and reduced EEA.

We also show the JDOS of a 0.4% compressively strained WS₂ monolayer in Fig. 2E. We note that compressive strain also drove the EEA process off VHS resonance by changing the exciton transition energy, for both uniaxial strain applied in any direction and biaxial strain (figs. S6 to S8). The choice of tensile versus compressive strain to suppress all nonradiative recombination should ensure that the system remains direct bandgap with the application of strain to avoid nonradiative recombination through momentum-dark indirect excitons

(25). Compressive strain can make some TMDC monolayers such as WS₂ indirect.

Strain has also been used to reduce traditional Auger recombination in conventional 3D semiconductors, but the mechanism is different (26). In 3D semiconductors, bandgap renormalization and effective mass equalization by applied strain can lead to one order of magnitude reduction of the conventional Auger rate (27). However, in the case of 2D TMDCs, EEA is inhibited by shifting the exciton transition energy E_X such that $2E_X$ did not overlap VHS resonance. Because of the existence of saddle points, logarithmic VHSs always characteristically appeared in 2D semiconductors and were not found in 3D, leading to a much stronger response to strain in monolayer TMDCs (15).

These principles applied equally to other TMDC semiconductors. Like WS₂, exfoliated monolayer WSe₂ also exhibited near-unity PL QY at all G s when it was made intrinsic by electrostatic counterdoping and tensile strain was being applied (Fig. 3A and fig S9). As-exfoliated monolayer MoS₂ has a PL QY in the range of 0.1 to 1.0%, which markedly increased at low pump after chemical counterdoping by Nafion (Fig. 3B). Unlike WS₂ and WSe₂, monolayer MoS₂ became an indirect-gap material when tensile strain was applied and remained a direct-gap material when compressive strain was applied (17).

We found that, rather than creating compressive strain, downward bending of flexible substrate with negative curvature resulted in buckling of monolayer TMDC. Instead, we used the thermal coefficient of expansion mismatch between the glycol-modified PET substrate and MoS₂ to apply compressive strain. A chemically counterdoped, 0.5% compressively strained monolayer MoS₂ also exhibited roll-off free PL QY at all G s.

The optoelectronic quality of large-area 2D TMDCs must be improved for their use as next-generation devices. The principles of suppressing all nonradiative recombination can also be applied to achieve high PL QY on centimeter-scale WS₂ monolayer grown by CVD. We first transferred a large-area CVD-grown WS₂ (28) onto a flexible substrate and spin-coated it with Nafion (Fig. 4A). The normalized PL spectra redshifted with applied tensile strain (Fig. 4B). Nafion counterdoping led to a strain-independent PL QY of $70 \pm 10\%$ (mean \pm SD) at low pump powers, as can be seen in a spatial map of the PL QY of a 2-mm \times 2-cm area taken at a G of 10^{10} cm⁻² s⁻¹ (Fig. 4C). At a high G of 10^{20} cm⁻² s⁻¹, if no strain was applied, a low PL QY of $1.6 \pm 0.8\%$ was observed uniformly, consistent with as-exfoliated monolayer WS₂ (Fig. 4D). However, at 0.4% applied tensile strain, PL QY of the same area reaches $59 \pm 10\%$, indicating an ~ 38 times uniform enhancement of PL QY for a large-scale grown sample at a high G of 10^{20} cm⁻² s⁻¹.

This demonstration shows that ideal optoelectronic quality can be achieved in large-area-grown monolayers by relatively simple and scalable application of counterdoping and strain.

REFERENCES AND NOTES

1. D.-H. Lien *et al.*, *Science* **364**, 468–471 (2019).
2. J. S. Ross *et al.*, *Nat. Commun.* **4**, 1474 (2013).
3. F. Withers *et al.*, *Nat. Mater.* **14**, 301–306 (2015).
4. M. Amani *et al.*, *Science* **350**, 1065–1068 (2015).
5. Y. Yu *et al.*, *Phys. Rev. B* **93**, 201111 (2016).
6. F. Wang, Y. Wu, M. S. Hybertsen, T. F. Heinz, *Phys. Rev. B Condens. Matter Mater. Phys.* **73**, 245424 (2006).
7. S. Z. Uddin, E. Rabani, A. Javey, *Nano Lett.* **21**, 424–429 (2021).
8. C. Murawski, K. Leo, M. C. Gather, *Adv. Mater.* **25**, 6801–6827 (2013).
9. W. Zou *et al.*, *Nat. Commun.* **9**, 608 (2018).
10. W. K. Bae *et al.*, *Nat. Commun.* **4**, 2661 (2013).
11. J. Iveland, L. Martinelli, J. Peretti, J. S. Speck, C. Weisbuch, *Phys. Rev. Lett.* **110**, 177406 (2013).
12. M. A. Green, *IEEE Trans. Electron Dev.* **31**, 671–678 (1984).
13. G. M. Kavoulakis, G. Baym, *Phys. Rev. B Condens. Matter* **54**, 16625–16636 (1996).
14. O. B. Aslan, M. Deng, T. F. Heinz, *Phys. Rev. B* **98**, 115308 (2018).
15. L. Van Hove, *Phys. Rev.* **89**, 1189–1193 (1953).
16. J. E. Hirsch, D. J. Scalapino, *Phys. Rev. Lett.* **56**, 2732–2735 (1986).
17. Z. Li *et al.*, *Nat. Commun.* **11**, 1151 (2020).
18. S. B. Desai *et al.*, *Nano Lett.* **14**, 4592–4597 (2014).
19. P. Landsberg, *Recombination in Semiconductors* (Cambridge Univ. Press, 2003).
20. V. N. Abakumov, V. I. Perel, I. N. Yassievich, *Nonradiative Recombination in Semiconductors* (Elsevier, 1991).
21. S. Fang, S. Carr, M. A. Cazalilla, E. Kaxiras, *Phys. Rev. B* **98**, 075106 (2018).
22. S. Fang *et al.*, *Phys. Rev. B Condens. Matter Mater. Phys.* **92**, 205108 (2015).
23. F. Bassani, G. P. Parravicini, R. A. Ballinger, J. L. Birman, *Phys. Today* **29**, 58–59 (1976).
24. M. Cardona, P. Y. Yu, *Fundamentals of Semiconductors* (Springer, Heidelberg, ed. 3, 2005).

25. J. Madéo *et al.*, *Science* **370**, 1199–1204 (2020).
26. M. Silver, E. P. O'Reilly, A. R. Adams, *IEEE J. Quantum Electron.* **33**, 1557–1566 (1997).
27. A. Haug, *J. Phys. C Solid State Phys.* **16**, 4159–4172 (1983).
28. J. Cho *et al.*, *Adv. Funct. Mater.* **30**, 1907941 (2020).

ACKNOWLEDGMENTS

We thank J. Cho for providing large-area CVD-grown WS₂ samples. **Funding:** This work was funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under contract no. DE-AC02-05-CH11231 (EMAT program KC1201). E.R. acknowledges support from the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under contract no. DE-AC02-05-CH11231 within the Physical Chemistry of Inorganic Nanostructures Program (KC3103). H.K. acknowledges support from a Samsung Scholarship. N.H. acknowledges support from Postdoctoral Fellowships for Research Abroad of the Japan Society for the Promotion of Science. **Author contributions:** S.Z.U., H.K., and A.J. conceived the idea for the project and designed the experiments. H.K. and S.Z.U. performed optical measurements. H.K., S.Z.U., and N.H. fabricated devices. H.K., S.Z.U., and A.J. analyzed the data. S.Z.U. and E.R. performed analytical modeling. S.Z.U., H.K., N.H., and A.J. wrote the manuscript. All authors discussed the results and commented on the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in this study are present in the main text or the supplementary materials. Materials that support the findings of this study are available from the corresponding author upon reasonable and well-intentioned request.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/448/suppl/DC1
Materials and Methods
Figs. S1 to S10
References (29, 30)
MDAR Reproducibility Checklist

8 April 2021; accepted 21 June 2021
10.1126/science.abi9193

SPECTROSCOPY

Atomically resolved single-molecule triplet quenching

Jinbo Peng^{1*}, Sophia Sokolov¹, Daniel Hernangómez-Pérez², Ferdinand Evers³, Leo Gross⁴, John M. Lupton¹, Jascha Repp^{1*}

The nonequilibrium triplet state of molecules plays an important role in photocatalysis, organic photovoltaics, and photodynamic therapy. We report the direct measurement of the triplet lifetime of an individual pentacene molecule on an insulating surface with atomic resolution by introducing an electronic pump-probe method in atomic force microscopy. Strong quenching of the triplet lifetime is observed if oxygen molecules are coadsorbed in close proximity. By means of single-molecule manipulation techniques, different arrangements with oxygen molecules were created and characterized with atomic precision, allowing for the direct correlation of molecular arrangements with the lifetime of the quenched triplet. Such electrical addressing of long-lived triplets of single molecules, combined with atomic-scale manipulation, offers previously unexplored routes to control and study local spin-spin interactions.

Many macroscopic phenomena in chemistry, biology, and materials science can only be understood on the basis of the underlying microscopic nature of the interacting molecules. Ensemble averaging inherent to macroscopic measurements severely obscures interactions on the atomistic level, which in turn are those most readily grasped by theory. This limitation particularly holds true if two different compounds can interact in a large variety of spatial constellations. The optical detection of

single molecules, developed more than 30 years ago, marked a leap forward in overcoming molecular inhomogeneities but is inherently

¹Institute for Experimental and Applied Physics and Regensburg Center for Ultrafast Nanoscopy (RUN), University of Regensburg, 93040 Regensburg, Germany.

²Department of Molecular Chemistry and Materials Science, Weizmann Institute of Science, Rehovot 7610001, Israel.

³Institute for Theoretical Physics, University of Regensburg, 93040 Regensburg, Germany. ⁴IBM Research–Zürich, 8803 Rüschlikon, Switzerland.

*Corresponding author. Email: jinbo.peng@ur.de (J.P.); jascha.repp@ur.de (J.R.)

limited to bright molecular excited states and does not readily entail atomic spatial resolution (1, 2). Atomistic information directly resolved in space has recently become available as a result of progress in scanning probe microscopy (3), providing a detailed picture of chemical structure (4–6), bond order (7), molecular orbitals (8, 9), hydration-promoted diffusion of ions (10), and many other single-molecule features. Beyond steady-state spectroscopy, all-electronic pump-probe techniques have opened the door to accessing dynamical properties on submolecular scales, as recently demonstrated by measurements of spin lifetimes of individual atoms (11). Only recently, however, has access to out-of-equilibrium electronic states of molecules been gained in scanning probe spectroscopy (12, 13), enabled by electronic decoupling from a conductive support.

Here we demonstrate how to probe the lifetime of the out-of-equilibrium dark state of a single pentacene molecule—the triplet state—by means of electronic pump-probe atomic force microscopy (AFM). Combining this approach with real-space atomic resolution, we revealed atomistic details in the excited-state quenching by coadsorbed oxygen molecules. Different atomic arrangements of pentacene with oxygen in the immediate vicinity were created and imaged, allowing a direct correlation between molecular arrangement and the rate of excitation-energy transfer. The precise placing of oxygen molecules with respect to the pentacene was crucial, providing unprecedented atomistic insight into the phenomenon of triplet quenching. Such quenching is of significance in photochemistry because it leads to the formation of reactive oxygen species, which limit the durability of materials in organic electronics and are crucial in applications as wide-ranging as super-resolution microscopy, photocatalysis, and photodynamic therapy. Despite many decades of extensive research into the mechanisms of photosensitized production of singlet oxygen, there is still no broad consensus on the underlying microscopic interaction pathways between excited states of molecules and oxygen (14).

The experimental setup is depicted in Fig. 1A. Pentacene was adsorbed on an insulating NaCl layer that was thick enough to prevent electron exchange between the molecule and the gold substrate beneath the NaCl layer. The conductive tip of the atomic force microscope allowed deterministic tunneling of individual electrons into and out of the molecule. The gate voltage, V_g , applied between the gold substrate and the metal tip determined the alignment of the chemical potential of the tip with respect to the molecular states. This alignment, in turn, controlled transitions between the different electronic states of the molecule, including out-of-equilibrium ones such as the triplet excited state T_1 (15), by single-electron tunneling.

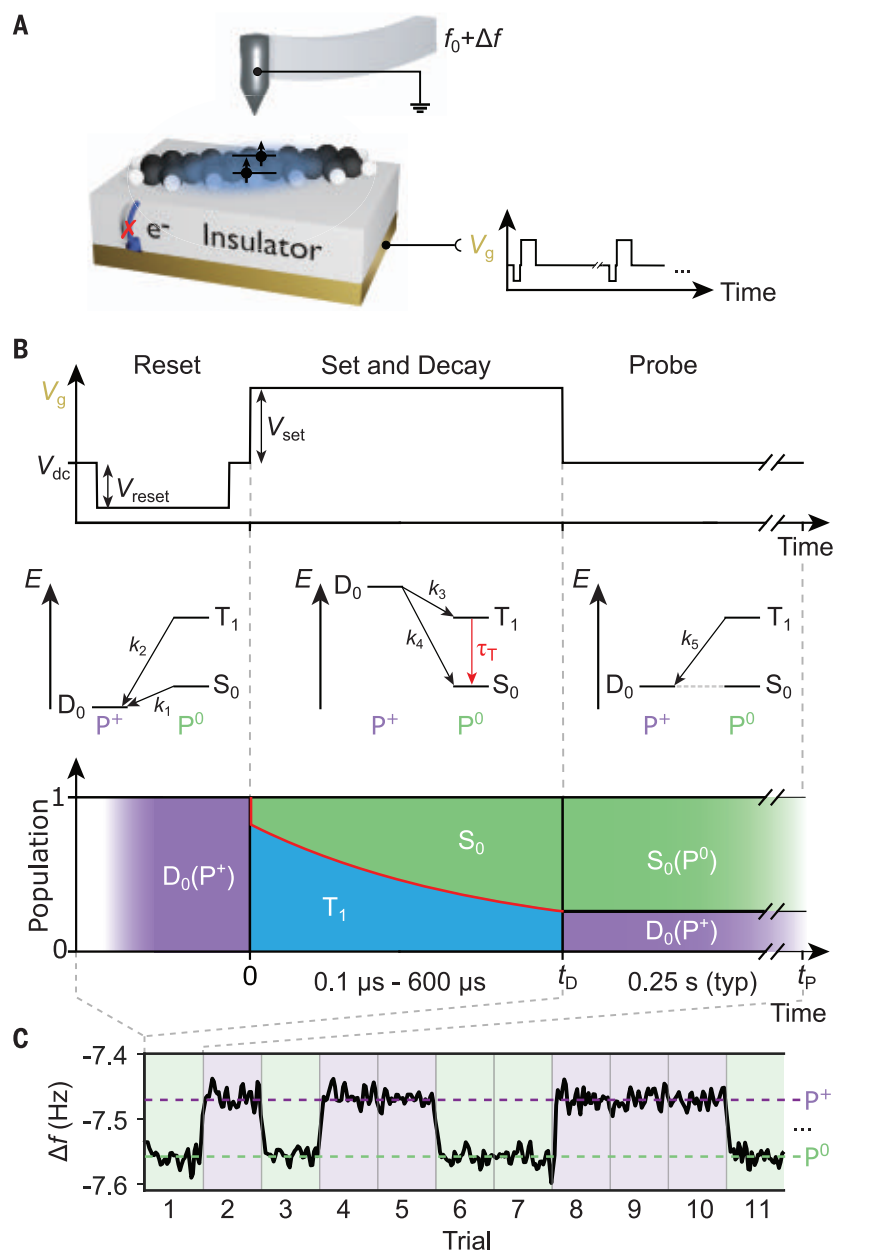
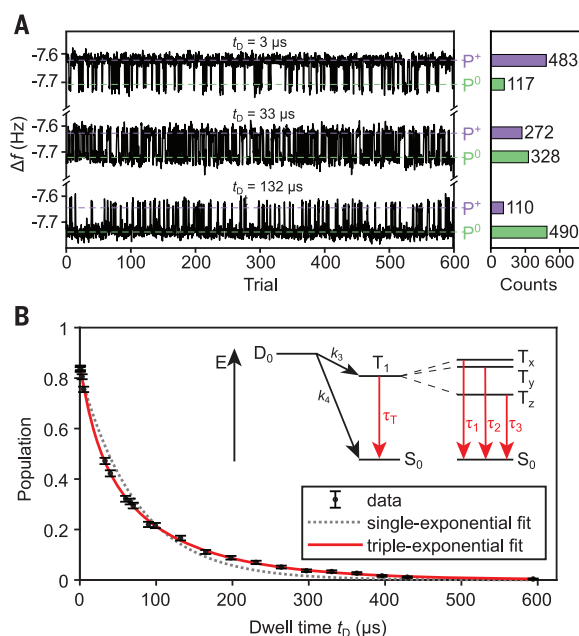


Fig. 1. Experimental procedure for measuring single-molecule triplet lifetimes by AFM.

(A) Sketch of the experimental setup. Charge transfer to the gold substrate was inhibited by a thick insulating NaCl layer. A time-dependent gate voltage, V_g , was applied to the gold substrate, driving single-electron exchange between the tip and the molecule. (B) (Top) Sequence of V_g controlling the three phases of a single measurement and the corresponding mutual energetic alignment of the cationic (P^+) doublet ground state D_0 , the neutral (P^0) singlet ground state S_0 , and the neutral triplet excited state T_1 (center). The charge-transfer rates k_{1-5} (black arrows) were chosen to be much faster than the triplet decay (red arrow). (Bottom) Example of the evolution of the state populations. $V_{dc} \approx -1.45$ V was chosen to render S_0 and D_0 degenerate. $V_{reset} = -1.38$ V and $V_{set} = +2.3$ V were added to V_{dc} . During the reset pulse, D_0 was rapidly occupied by a charge transferring from S_0 or T_1 . During the set-and-decay interval of duration t_D , this charged doublet population was swiftly transferred into a neutral state, either S_0 or T_1 . The latter decayed to S_0 with a characteristic lifetime of τ_T . During the probe interval, any remaining population in T_1 was quickly transferred back to D_0 . Owing to the reorganization energy, S_0 and D_0 do not interconvert. Thus, the populations of S_0 or T_1 were mapped to different charge states and detected. typ, typically. (C) Two discrete values of the shift in cantilever frequency Δf corresponded to the two charge states P^+ and P^0 and signified the two different outcomes of individual measurement trials.

Fig. 2. Extraction of the triplet lifetime from a single pentacene molecule. (A) Three Δf time traces, each representing a sequence of 600 trials for different dwell times t_D . The corresponding counts N_{P^+} and N_{P^0} of the two possible outcomes of charge transfer are shown on the right. The triplet population $P(t_D)$ was determined from $N_{P^+} / (N_{P^+} + N_{P^0})$. (B) Plot of $P(t_D)$ and extraction of individual lifetimes τ_1 , τ_2 , and τ_3 of the three triplet states T_x , T_y , and T_z , respectively. Each data point corresponds to 1800 individual trials, and the error bars are derived from the standard deviation of the binomial distribution $N_{P^+} / (N_{P^+} + N_{P^0})$ (see supplementary text). The triplet is zero-field split into three states (inset, splitting on the order of micro-electron volts, not drawn to scale). The measurements follow a triple-exponential decay. In contrast to optical experiments, all three triplet states were initially populated with equal probability by single-electron injection from the AFM tip.



Owing to the single-electron sensitivity of AFM, different charge states of the molecule could reliably be detected (16).

The approach advanced here to measure molecular triplet lifetimes by means of AFM was direct: By virtue of two subsequent tunneling events, we prepared the molecule in its first excited triplet state, T_1 , at a time $t = 0$. After a certain dwell time, t_D , we probed the fraction of the population remaining in the T_1 state and decaying into the singlet ground state, S_0 . The time t_D was chosen to span time scales from a small fraction to a large multiple of the excited state lifetime of the target molecule and has no relation to the cantilever oscillation frequency. To discriminate between S_0 and T_1 states by means of AFM, we transferred the population of T_1 to the cationic state by appropriate electronic level alignment with the gate voltage. In this way, the population fractions at time t_D were mapped onto different metastable charge states, which are associated with an appreciable shift Δf of the resonance frequency of the AFM cantilever (16).

The measurement scheme is depicted in greater detail in Fig. 1, B and C, and proceeds as follows. Before probing the triplet-state population, we ensured positive charging of pentacene by application of a reset pulse. This strong voltage pulse prompted one electron to tunnel out of the highest occupied molecular orbital (HOMO) of the pentacene molecule, transferring it into the cationic (P^+) doublet ground state, D_0 . A second voltage pulse of variable duration t_D , referred to as the set-and-decay pulse, shifted this cationic state D_0 energetically above the T_1 and S_0

levels of neutral pentacene (P^0). This shift prompted the molecule to accept a tunneling electron from the tip. The charge either refilled the HOMO of the molecule, yielding the S_0 ground state, or occupied the lowest unoccupied molecular orbital (LUMO), resulting in the T_1 excited state. The latter tunneling process occurred at a considerably higher energy [the T_1 state lies 0.86 eV in energy above S_0 (17)] and therefore experienced a lower remaining tunneling barrier and was thus more likely to occur (15). The rates of these two transitions could be controlled by the tip-to-molecule distance and were chosen such that the tunneling process arose within the first ~ 100 ns of the set-and-decay pulse (see fig. S1 for details). This time was short in comparison to the triplet lifetime of the molecule. After the tunneling process, the molecule was in its neutral charge state P^0 , either in the T_1 or S_0 state. The two states were found to be populated in a ratio of $\sim 4:1$ because of the aforementioned difference in tunneling rates, in agreement with the simulations shown in fig. S2. During the remainder of the set-and-decay pulse, the population of T_1 decayed into S_0 following the characteristic triplet lifetime τ_T . After the dwell time t_D , the voltage was set to a value V_{dc} , for which the D_0 and S_0 states were almost degenerate. Further details on how to choose V_{dc} to ensure this degeneracy are provided in fig. S3. Because of the reorganization energy associated with charging—considerably increased owing to the presence of NaCl (13)—no tunneling transition occurred between the S_0 and D_0 ground states (9, 12, 13, 15). However, the excited state T_1

was able to quickly decay into D_0 by the tunneling of an electron from the LUMO into the AFM tip. After the end of the set-and-decay pulse, the population of S_0 remained neutral, whereas the T_1 population converted to the charged D_0 state. Both ground states D_0 and S_0 were stable on macroscopic time scales under this biasing condition and were readily distinguished by virtue of the force acting on the tip and the resulting frequency shift Δf (16) (Fig. 1C and fig. S4). This measurement cycle was performed on the time scale of a fraction of a second.

In a single-molecule experiment, the molecule is only in one state at a given time. The measurement cycle therefore represented an individual Bernoulli trial with two possible outcomes, corresponding to S_0 or T_1 at t_D . The population $P(t_D)$ of different states for one particular dwell time t_D was extracted by repeating the experiment multiple times. Examples of three such sequences are shown in Fig. 2A. Repeating these sequences for different values of t_D yielded $P(t_D)$ (Fig. 2B and fig. S5), reflecting the decay of the T_1 population with a characteristic lifetime τ_T .

There are three triplet states, T_x , T_y , and T_z , as sketched in the inset of Fig. 2B, which are nondegenerate because of the zero-field splitting due to dipolar and spin-orbit interactions (18, 19). Although this energy splitting is tiny—on the order of several micro-electron volts—the triplet states have substantially different lifetimes, τ_1 , τ_2 , and τ_3 (18). Indeed, the decay of $P(t_D)$ with time t_D derived here was accurately described by a triple-exponential decay, with values of $\tau_1 = 11 \pm 5 \mu s$, $\tau_2 = 75 \pm 12 \mu s$, and $\tau_3 = 135 \pm 11 \mu s$ extracted for 23 datasets on several different individual molecules (see table S1 for further details). Note that in optical experiments, the initial population of the three triplet states is not uniform (18), such that the different rates can only be extracted indirectly. Here, however, the initial population was expected to be equal with respect to T_x , T_y , and T_z because the electron injection leading to triplet formation occurred with a random spin orientation. These lifetimes were comparable to those obtained by optical means for individual pentacene molecules in a host matrix (1, 20), suggesting that neither the presence of the AFM tip nor the interaction with the NaCl substrate considerably altered the triplet state. To scrutinize this conclusion further, we varied the tip-to-molecule distance and observed no appreciable change in the triplet lifetimes (fig. S6A). Additionally, no variation of lifetimes was observed when we varied the lateral tip position above the molecule either (fig. S6B).

We now turn to the possibility of tracking triplet-state quenching with atomic spatial resolution. It is well known that molecular oxygen, which itself has a triplet ground state,

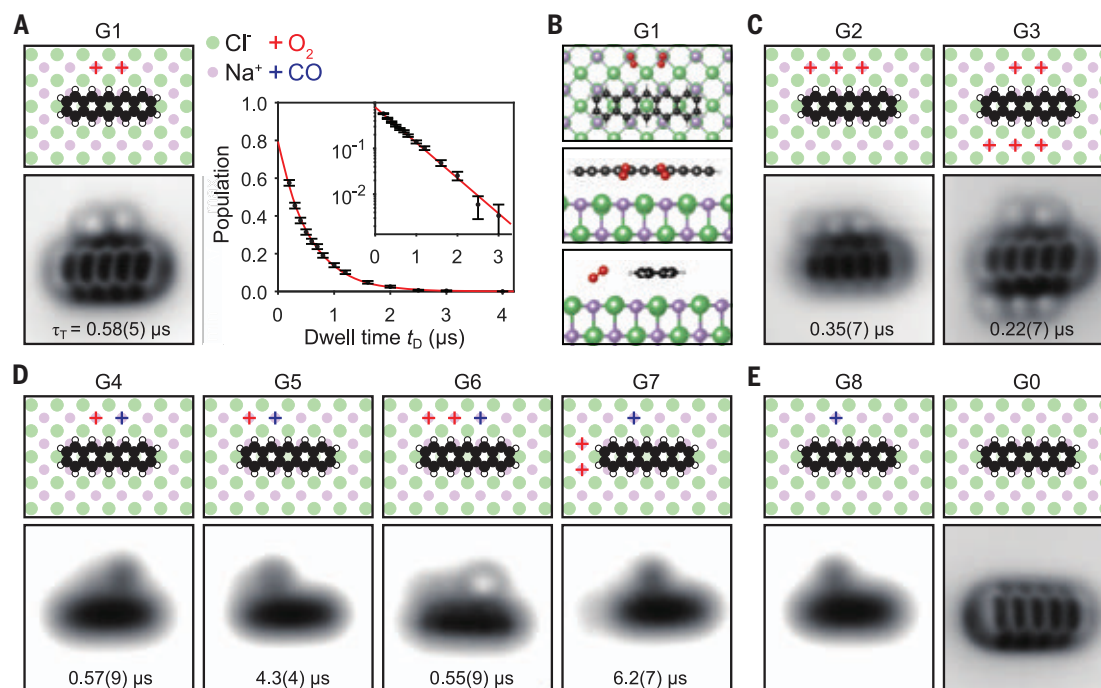


Fig. 3. Quenching of the pentacene triplet state by coadsorbed oxygen molecules for different well-defined geometries G0 to G8. The molecular structure on the NaCl lattice is shown with the corresponding constant-height AFM images. **(A)** Two O_2 molecules adsorbed close to the center of the pentacene molecule. The triplet decay is single exponential with $\tau_T = 0.58(5) \mu s$. Error bars are derived from the standard deviation of the binomial distribution $N_{P^+} / (N_{P^+} + N_{P^0})$ (see supplementary text). **(B)** On the basis of the experimental assignment of the atomic adsorption site, DFT calculations allowed for further specification of the atomic geometry. **(C)** Increasing the number of adsorbed O_2 molecules to three (G2) or five (G3) lowers the triplet lifetime even further. **(D)** For a given number of O_2 molecules, the relative position of the O_2

adsorbates strongly influenced the triplet-lifetime quenching, as shown for one (G4 versus G5) and two (G6 versus G7) O_2 molecules. CO molecules were used to stabilize oxygen arrangements that are unstable otherwise. These geometries had to be imaged at larger tip-to-sample distances, resulting in a reduced AFM contrast. The contrast reversal for CO and O_2 in the images resulted from the different tip-to-sample distances used, so that either attractive or repulsive interactions dominated. **(E)** The presence of a CO molecule did not affect the intrinsic triplet lifetime of the pentacene molecule (G8), as deduced from a comparison to measurements of an isolated pentacene molecule after removal of O_2 by atomic manipulation (G0). Size of the AFM images: 2.3 nm by 2.3 nm (G0 to G5, and G8), 2.2 nm by 2.2 nm (G6), and 2.5 nm by 2.5 nm (G7).

strongly quenches triplet excited states of other molecules (14). Indeed, after coadsorption of oxygen next to the pentacene molecule, we observed a drastic reduction of triplet lifetime by almost two orders of magnitude (Fig. 3A). In contrast to the case of isolated pentacene, the decay of the triplet population now followed a single exponential, which could be rationalized by the triplet-state lifetime being controlled solely by the energy-transfer rate to O_2 . Although the atomistic arrangement of triplet quenching had previously remained elusive experimentally, in this study, we resolved different geometries (G1 to G7) of the arrangement and measured the triplet lifetimes, as summarized in Fig. 3 and fig. S7. The adsorption site of O_2 is determined to be on top of a Na^+ ion (fig. S8). The experimental identification of the exact adsorption site was complemented by density functional theory (DFT) simulations, providing the minute details of the adsorption constellation (see Fig. 3B and fig. S9). The simulations confirmed that the adsorption site lay on top of a Na^+ ion and revealed a pronounced tilt (45° to 60°) of the O_2 molecule toward the pentacene. Further,

the simulations confirmed that the O_2 molecule was in a triplet ground state when adsorbed on NaCl (see fig. S10). This combination of atomically resolved experiments with atomistic simulations provided an unprecedented characterization of the process of triplet quenching for the different atomic geometries, demonstrating the full strength of our method. Our experiments showed that these strongly quenched lifetimes that were shorter than a microsecond were characteristic for the different local arrangements of molecules. Determination of these lifetimes was robust against variations of the tip height (fig. S11). Notably, the triplet lifetime varied strongly not only for an increasing number of oxygen molecules (Fig. 3C) but also for the same number of surrounding oxygen molecules if the O_2 molecules were positioned at different sites, as summarized in Fig. 3D.

Upon removal of the O_2 molecules in the direct vicinity of the pentacene molecule by atomic manipulation techniques (Fig. 3E, geometry G0), the same average of decay rates, $\langle k \rangle = \frac{1}{3} \sum_i \tau_i^{-1}$, as measured for pristine pentacene was recovered (table S1). The three

individual triplet lifetimes, τ_1 , τ_2 , and τ_3 , differed slightly, which was rationalized by the dipolar magnetic fields arising from the paramagnetic O_2 molecules adsorbed on the tip in the vicinity of the AFM junction, thus mixing the pristine triplet states T_x , T_y , and T_z to a new set of three eigenstates. Such an effect does not correspond to triplet quenching. Further details are discussed in the supplementary text section of the supplementary materials.

Finally, in a control experiment, we tested the effect of a single CO molecule adsorbed adjacent to the pentacene molecule (Fig. 3E). The triplet lifetimes of this geometry (G8) were identical to those of the isolated pentacene (G0) (fig. S12), confirming that the strong quenching was indeed characteristic of the presence of O_2 and that the pentacene triplet state was unaffected by CO.

Our experiments show how the interaction between individual molecules, and the impact of such interactions on molecular nonequilibrium properties, can be traced with atomic resolution directly in space. In view of the wealth of literature on triplet-based single-molecule electron-spin resonance (ESR), the

AFM-based method introduced here may pave the way for adding atomic-scale resolution to single-molecule ESR (21–23), including direct real-space mapping of the molecular hyperfine tensor (23) on insulators. Electronic and nuclear spins in molecules have long been considered building blocks for quantum-information science (24–27), particularly in molecular magnets of high electronic spin (27, 28). A key requisite for designing a molecular-level qubit is addressability on the single-molecule level, preferably electrical (29). Molecular triplet states in π -conjugated hydrocarbons are expected to exhibit long coherence times with a weak temperature dependence (30), suggesting their potential exploitation even at room temperature.

REFERENCES AND NOTES

1. J. Köhler *et al.*, *Nature* **363**, 242–244 (1993).
2. J. Wrachtrup, C. von Borczyskowski, J. Bernard, M. Orrit, R. Brown, *Nature* **363**, 244–245 (1993).
3. L. Gross, *Nat. Chem.* **3**, 273–278 (2011).
4. L. Gross, F. Mohn, N. Moll, P. Liljeroth, G. Meyer, *Science* **325**, 1110–1114 (2009).
5. D. G. de Oteyza *et al.*, *Science* **340**, 1434–1437 (2013).
6. J. Hulva *et al.*, *Science* **371**, 375–379 (2021).
7. L. Gross *et al.*, *Science* **337**, 1326–1329 (2012).
8. J. Repp, G. Meyer, S. M. Stojković, A. Gourdon, C. Joachim, *Phys. Rev. Lett.* **94**, 026803 (2005).
9. L. L. Patera, F. Queck, P. Scheuerer, J. Repp, *Nature* **566**, 245–248 (2019).
10. J. Peng *et al.*, *Nature* **557**, 701–705 (2018).
11. S. Loh, M. Etzkorn, C. P. Lutz, D. M. Eigler, A. J. Heinrich, *Science* **329**, 1628–1630 (2010).
12. W. Steurer, S. Fatayer, L. Gross, G. Meyer, *Nat. Commun.* **6**, 8353 (2015).
13. S. Fatayer *et al.*, *Nat. Nanotechnol.* **13**, 376–380 (2018).
14. C. Schweitzer, R. Schmidt, *Chem. Rev.* **103**, 1685–1757 (2003).
15. S. Fatayer *et al.*, *Phys. Rev. Lett.* **126**, 176801 (2021).
16. L. Gross *et al.*, *Science* **324**, 1428–1431 (2009).
17. J. Burgos, M. Pope, C. E. Swenberg, R. R. Alfano, *Phys. Status Solidi* **83**, 249–256 (1977).
18. J. Köhler, *Phys. Rep.* **310**, 261–339 (1999).
19. S. Sinnecker, F. Neese, *J. Phys. Chem. A* **110**, 12267–12275 (2006).
20. A. J. Van Strien, J. Schmidt, *Chem. Phys. Lett.* **70**, 513–517 (1980).
21. S. Baumann *et al.*, *Science* **350**, 417–420 (2015).
22. T. Choi *et al.*, *Nat. Nanotechnol.* **12**, 420–424 (2017).
23. P. Willke *et al.*, *Science* **362**, 336–339 (2018).
24. A. Ardavan *et al.*, *Phys. Rev. Lett.* **98**, 057201 (2007).
25. S. Bertaina *et al.*, *Nature* **453**, 203–206 (2008).
26. V. Filidou *et al.*, *Nat. Phys.* **8**, 596–600 (2012).
27. A. Gaita-Ariño, F. Luis, S. Hill, E. Coronado, *Nat. Chem.* **11**, 301–309 (2019).
28. S. Mishra *et al.*, *Nat. Nanotechnol.* **15**, 22–28 (2020).
29. N. Roch, S. Florens, V. Bouchiat, W. Wernsdorfer, F. Balestro, *Nature* **453**, 633–637 (2008).
30. W. J. Baker, D. R. McCamey, K. J. van Schooten, J. M. Lupton, C. Boehme, *Phys. Rev. B* **84**, 165205 (2011).
31. J. Peng *et al.*, Data archive of “Atomically resolved single-molecule triplet quenching,” University of Regensburg Publication Server (2021); <https://doi.org/10.5283/epub.45605>.

ACKNOWLEDGMENTS

We thank J. Wilhelm, L. Patera, and S. Fatayer for fruitful discussions and C. Rohrer for technical support. **Funding:** Funding from the Alexander von Humboldt Foundation and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; RE2669/6-2; RE2669/7-1; 314695032/CRC 1277 projects B01 and B03) is gratefully acknowledged. **Author contributions:** J.P., L.G., J.M.L., and J.R. conceived of the experiments, and J.P. and S.S. carried them out. J.P., S.S., and J.R. analyzed the experimental results. D.H.-P. and F.E. set up, performed, and analyzed the DFT calculations. J.P. and J.R. wrote the manuscript. All authors discussed the results and their interpretation and revised the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All

data are available from the University of Regensburg Publication Server (31).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/452/suppl/DC1
Materials and Methods

Supplementary Text
Figs. S1 to S12
Table S1
References (32–43)

20 February 2021; accepted 25 May 2021
10.1126/science.abh1155

ANIMAL CULTURE

Innovation and geographic spread of a complex foraging culture in an urban parrot

Barbara C. Klump^{1*}, John M. Martin², Sonja Wild^{1,3}, Jana K. Hörsch^{1,4}, Richard E. Major⁵, Lucy M. Aplin^{1,3*}

The emergence, spread, and establishment of innovations within cultures can promote adaptive responses to anthropogenic change. We describe a putative case of the development of a cultural adaptation to urban environments: opening of household waste bins by wild sulphur-crested cockatoos. A spatial network analysis of community science reports revealed the geographic spread of bin opening from three suburbs to 44 in Sydney, Australia, by means of social learning. Analysis of 160 direct observations revealed individual styles and site-specific differences. We describe a full pathway from the spread of innovation to emergence of geographic variation, evidencing foraging cultures in parrots and indicating the existence of cultural complexity in parrots. Bin opening is directly linked to human-provided opportunities, highlighting the potential for culture to facilitate behavioral responses to anthropogenic change.

Identifying persistent cultures in animal species—that is, population-specific behaviors acquired via social learning from knowledgeable individuals (1)—remains challenging. The behavior in question may be rare or hard to observe, and it is difficult to exclude genetic and ecological variables as alternative explanations for population-level differences (1, 2). Urban habitats provide opportunities to overcome these challenges, as community science allows for a wide spread of observations (3). Additionally, many urban resources are standardized (e.g., waste disposal) (4). Urban environments may also promote innovation [an associatively learned solution to a problem (5)] by providing novel resources and foraging opportunities (4, 6). Innovations by wild animals living in cities are well documented, and comparative analyses suggest that innovativeness may enable behavioral flexibility to anthropogenic changes (7). Observations of multiple individuals engaging in a novel behavior are often assumed to result from the spread of innovation (5, 8),

although direct evidence of this is lacking (9). However, if innovations spread through populations via social learning, they can lead to emergent adaptive cultures (8, 10)—socially learned behaviors that are beneficial for urban living and that differ between sites.

The cultural intelligence hypothesis predicts that animal culture has coevolved with larger brain size, high sociality, and slower life history (11). Supporting this, extensive cultures across multiple domains have been identified in social species with relatively complex cognition [e.g., primates, cetaceans, and corvids (2, 10, 12)]. Large parrots share many life history traits with these groups and also exhibit vocal cultures (13). However, there is a paucity of evidence for social learning or culture in the foraging behavior of large parrots. Sulphur-crested cockatoos, *Cacatua galerita* (henceforth: cockatoos), are a large-brained, long-lived, and highly social parrot native to eastern Australia. They are increasingly common in cities (3) and occur throughout our study area. In recent years there have been isolated reports of cockatoos opening the lids of household waste bins to scavenge food (henceforth: bin opening; Fig. 1). This foraging innovation exploits a widely available resource, but little else is known about it. We used a large-scale community science survey to map bin opening across the Sydney and Wollongong regions combined with direct observations at multiple sites to identify possible individual or site-level differences.

To map occurrence, we collected data via an online survey over two consecutive years (2018

¹Cognitive and Cultural Ecology Research Group, Max Planck Institute of Animal Behavior, Am Obstberg 1, 78315 Radolfzell am Bodensee, Germany. ²Taronga Institute of Science and Learning, Taronga Conservation Society Australia, Bradleys Head Rd, Mosman, NSW 2088, Australia. ³Centre for the Advanced Study of Collective Behaviour, University of Konstanz, Universitätsstrasse 10, 78464 Konstanz, Germany. ⁴Department of Biology, University of Konstanz, Universitätsstrasse 10, 78464 Konstanz, Germany. ⁵Australian Museum Research Institute, Australian Museum, 1 William Street, Sydney, NSW 2010, Australia.
*Corresponding author. Email: bkump@ab.mpg.de (B.C.K.); laplin@ab.mpg.de (L.M.A.)



Fig. 1. A paint-marked sulphur-crested cockatoo opening a household bin.

and 2019). We asked whether residents had observed cockatoos opening lids of household bins and, if so, when and where (supplementary text). We collected 1396 reports by 1322 participants across 478 suburbs, of which 338 reports from 44 suburbs described bin opening. Multiple cockatoos were present in 93.3% of the observed cases, highlighting the ample opportunities for birds to observe bin opening. In 88.8% of cases, birds opened general waste bins (identifiable by red lids).

To investigate spatial-temporal dynamics, we assigned the data to seven discrete time periods on the basis of survey reports (fig. S1; see supplementary text for sensitivity analysis on discretization of continuous data). We then created a spatial network of geographic distance (in meters) between suburbs, informed by knowledge of cockatoo dispersal distances, and ran a spatial network-based

diffusion analysis (14). This analysis assumes that if cumulative uptake in suburbs is due to social learning, new occurrences are more likely in suburbs closer to and within dispersal distance to where the behavior is established (14). Three suburbs in which bin opening was observed prior to 2018 were included as likely sites of origin. We included three suburb-level variables: (i) number of survey participants to control for reporting rate, (ii) suburb size (square kilometers) as an imperfect proxy for cockatoo population size, and (iii) number of dwellings as a proxy for resource availability (number of bins). Network models with social transmission received overwhelming support (summed Akaike weight, indicating relative support $\sum w_i > 0.99$). In the best model, bin opening was estimated to have spread by social learning to 93.9% (95% confidence interval 86.5 to 97.3%) of suburbs where it occurred. The models also revealed a secondary innovation of bin opening in the far north of Sydney (Narraweena) in late 2018, as shown by Fig. 2. From this, the behavior subsequently spread by social learning. The estimated social learning rate was negatively influenced by the number of dwellings ($\sum w_i = 0.87$) and positively by the number of survey participants ($\sum w_i > 0.99$), but no other suburb-level variable influenced learning rates (all $\sum w_i < 0.5$; table S1).

To evaluate the complexity of bin-opening behavior, we temporarily color marked 486 cockatoos across three hotspot sites identified from the online survey (Fig. 3, B and D, and fig. S2). Movie analysis of 160 successful opening sequences revealed that bin opening is a complex, multistep foraging behavior, consisting of five stages with several options at each stage (Fig. 3A and tables S2 and S3).

Several components of the opening sequences, such as walking direction (possibly related to

laterality) and whether or not the lid was pried open (possibly related to experience as prying is not necessary for successful bin opening), were strongly associated with individual cockatoos. Other components—e.g., those associated with the opening stage—differed more between sites (Fig. 3C). Overall, sequences by the same individual were more similar than sequences by different individuals (partial Mantel test, accounting for geographic distance: $r = 0.16$, $P < 0.001$). Controlling for individual identity, dissimilarity between sequences increased with increasing geographic distance between observations (partial Mantel test: $r = 0.21$, $P = 0.002$; Fig. 4 and Movie S1), suggesting the formation of local subcultures of the behavior [for a similar example in chimpanzees, see (15)].

To assess the importance of individual and social characteristics on bin-opening ability, we marked >90% of the birds at one hotspot site (Stanwell Park, Fig. 3B) and collected data on sex, age, weight, dominance rank, and social associations (16). Out of 114 identifiable individuals, 9 cockatoos were successful ($n = 112$ complete sequences) and 27 attempted but were unsuccessful ($n = 94$ sequences) in opening bins. Both adults and juveniles opened bins [cumulative link model (CLM) testing age bias: $\chi^2 = 0.86$, $P = 0.35$, $n = 84$], suggesting social transmission between age classes. Of cockatoos with known sex, most successful bin-openers (89%) and most birds that attempted to open bins (84%) were male (table S4). Individuals higher in the male dominance hierarchy were also more likely to attempt and succeed at bin opening (CLM: $\chi^2 = 7.68$, $P = 0.006$, $n = 37$; fig. S3). Finally, there was a nonsignificant tendency for heavier males to attempt bin opening (generalized linear model: $\chi^2 = 3.12$, $P = 0.08$, $n = 20$), although a small sample size for males with

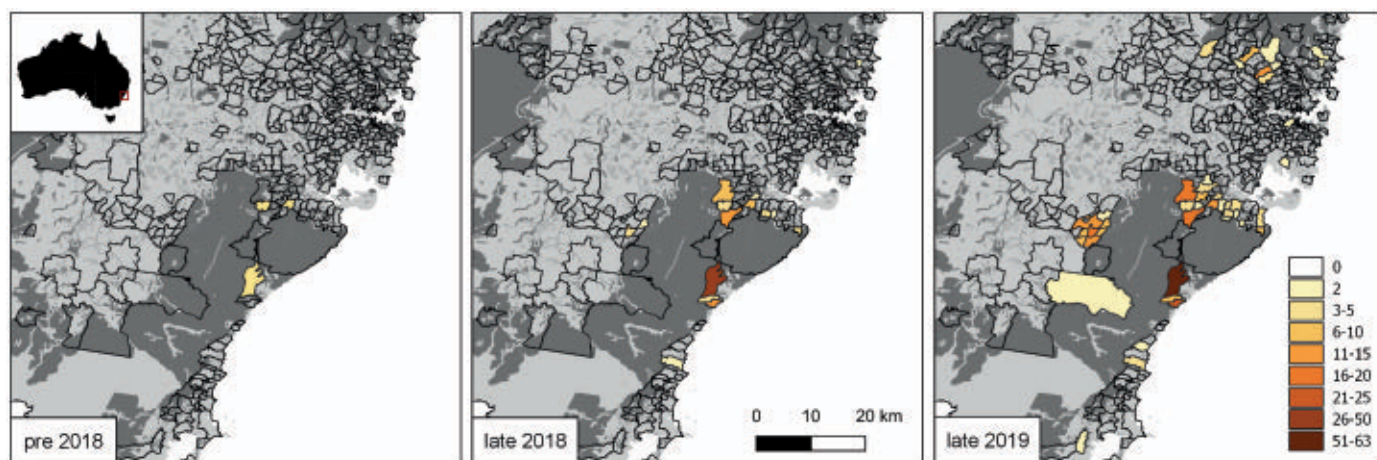


Fig. 2. Spread of bin opening across the Sydney and Wollongong regions. Reported in only three suburbs before 2018, bin-opening behavior had spread to 44 suburbs by late 2019. Suburbs outlined with black returned only negative reports, whereas suburbs with at least two positive reports for the respective time period are colored (cumulative over time). Forested areas (>9.6% of the area covered by trees 10 to 15 m high) are shown in dark gray. For all time periods, see fig. S1.

known weight precluded full analysis of potential interactions between rank and weight. Males tend to be heavier and more dominant than females (17), and males might have a strength advantage if bin opening is physically

demanding. Alternatively, high-ranking males might have preferential access to resources (bins); we frequently observed cockatoos displacing conspecifics off of bins. Our results contrast other studies suggesting that younger

and lower-ranking individuals have a higher probability of social learning [e.g., (18)] but demonstrate how resource distribution can influence social dynamics of social learning and culture (19). In the only other bird known

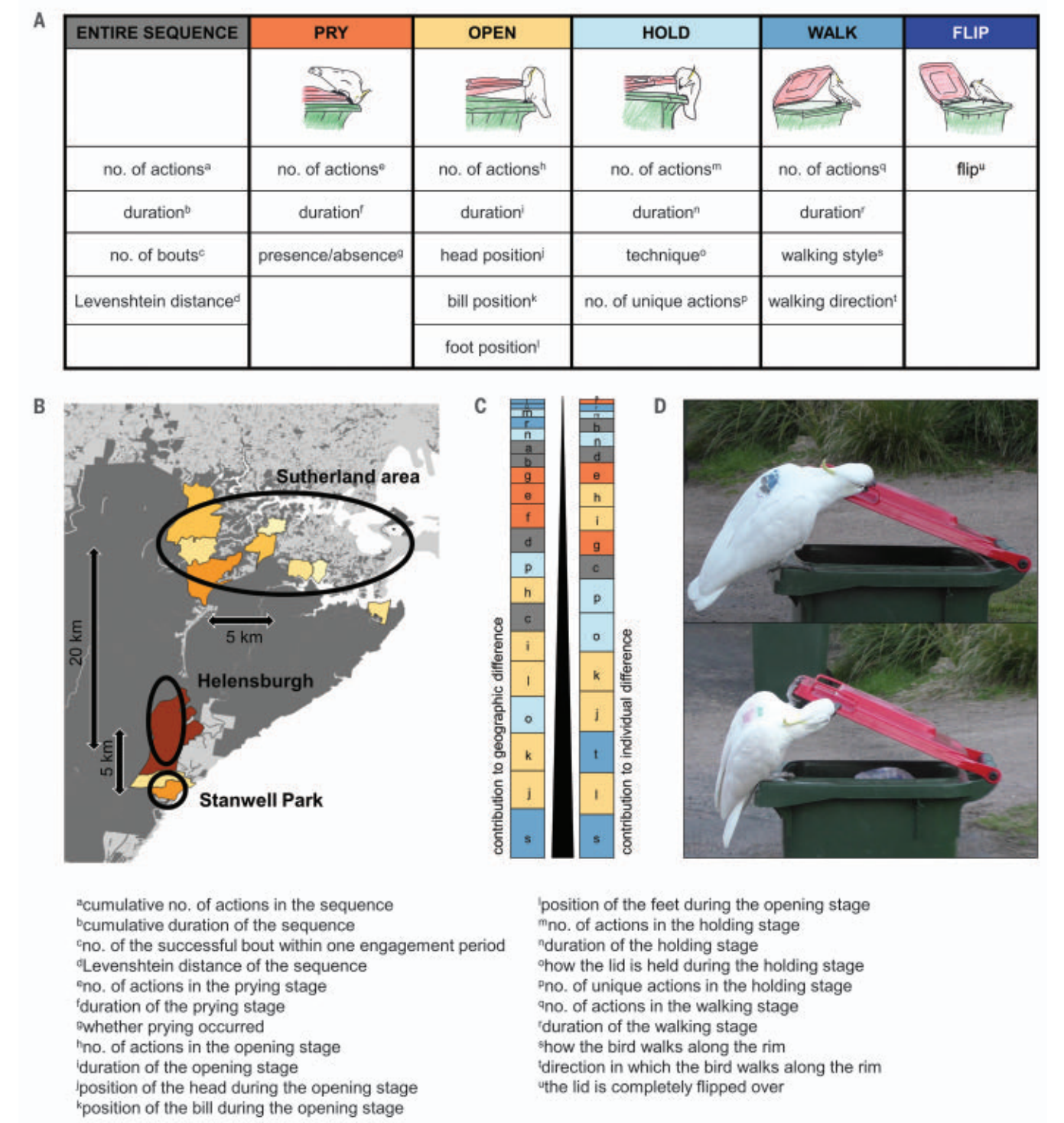


Fig. 3. Individual and geographic variation in bin-opening behavior. (A) The five stages of bin opening. Columns list the variables scored from the movies for each stage (for details, see tables S2 and S3). (B) Study areas for direct observations. (C) Relative contribution of each behavioral variable to geographic (left) and individual differences (right); most to least important, bottom to top, cumulative to 100%. Colors and letters are matched to those in (A). (D) Marked individuals holding the lid with the bill at the handle (top) and with bill and foot at the rim (bottom).

to open bins, the New Zealand kea (*Nestor notabilis*), it was also only males that opened lids (20). Further comparisons are difficult, however, as kea bin opening was only observed at one location and with low success rates.

Finally, social network analysis revealed a tendency for males [permutating within bin-opening status, linear model (LM): $t_{78} = 2.54$, $P = 0.018$] and attempting or successful bin-

openers to have a higher eigenvector centrality in the association network (permutating within sex, LM: $t_{78} = 1.07$, $P = 0.001$, fig. S4). As expected, the likelihood of higher similarity in bin-opening status was most strongly affected by sex [multiple regression quadratic assignment procedure (MRQAP): $est = 0.13$, $P < 0.001$, $n = 89$]. However, when males were considered separately, there was

a significant tendency for bin-openers to associate more (controlling for rank, MRQAP: $est = 0.97$, $P = 0.010$, $n = 34$). Although directionality cannot be ascertained from these data (supplementary text), these patterns are consistent with those expected from social learning, with associated and more socially central individuals having better access to social information and therefore a greater likelihood of learning [for a similar finding in primates, see (21)].

Our study combines the two classic approaches to the study of animal culture; namely, mapping the spread of an innovation over space and time (6, 10) and observing geographic variation in behavior (2). Our results show that the spread of innovation can not only result in establishment of culture, but can also further lead to emergent geographically distinct subcultures. Our study area is characterized by suburbs and forests. Data from color-marked birds suggest that forests may serve as barriers, as little movement through such areas was observed. This in turn might trap behaviors in suburban areas (within which they can spread). Founder effects and drift may then give rise to local behavioral variants at relatively small geographic scales; cultural evolutionary theory would predict that complex multistep traits, as observed here, could potentially further enhance this effect. This behavior has emerged in direct response to land-use change. It illustrates how, in the Anthropocene, animal culture can allow urban populations to access novel resources, potentially facilitating local adaptation (22, 23).

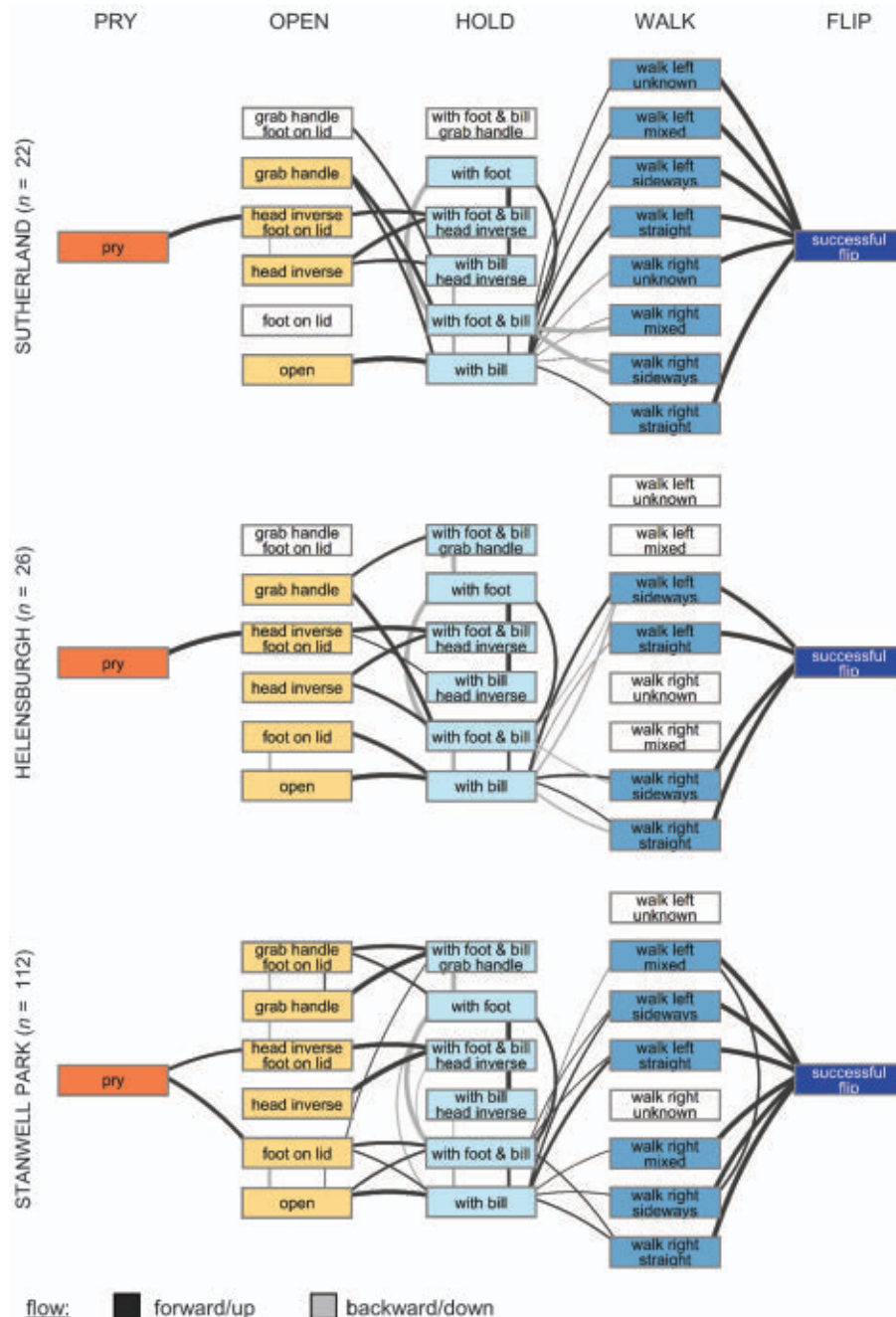


Fig. 4. Behavioral sequences of successful openings across sites. Bin-opening stages are depicted from left to right. Connections are colored by direction: advance to the next stage or upward within a stage (black), regress to the previous stage or downward within a stage (gray). The thickness of a connection represents the percent occurrence of this transition (for details, see tables S2 and S3).

REFERENCES AND NOTES

1. K. N. Laland, V. M. Janik, *Trends Ecol. Evol.* **21**, 542–547 (2006).
2. A. Whiten *et al.*, *Nature* **399**, 682–685 (1999).
3. L. M. Aplin, R. E. Major, A. Davis, J. M. Martin, *J. Anim. Ecol.* **90**, 222–232 (2021).
4. A. S. Griffin, K. Netto, C. Peneaux, *Curr. Opin. Behav. Sci.* **16**, 15–22 (2017).
5. S. Reader, K. N. Laland, in *Animal Innovation* (Oxford Univ. Press, 2003), pp. 3–38.
6. J. Fisher, R. A. Hinde, *Br. Birds* **42**, 347–357 (1949).
7. D. Sol, R. P. Duncan, T. M. Blackburn, P. Cassey, L. Lefebvre, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5460–5465 (2005).
8. L. M. Aplin, *Anim. Behav.* **147**, 179–187 (2019).
9. L. Lefebvre, *Behav. Processes* **34**, 43–53 (1995).
10. J. Allen, M. Weinrich, W. Hoppitt, L. Rendell, *Science* **340**, 485–488 (2013).
11. C. P. van Schaik, J. M. Burkart, *Philos. Trans. R. Soc. London B Biol. Sci.* **366**, 1008–1016 (2011).
12. G. R. Hunt, R. D. Gray, *Proc. Biol. Sci.* **70**, 867–874 (2003).
13. T. F. Wright, C. R. Dahlin, *Emu* **118**, 50–66 (2018).
14. M. J. Hasenjager, E. Leadbeater, W. Hoppitt, *J. Anim. Ecol.* **90**, 8–26 (2021).
15. C. Boesch *et al.*, *Nat. Hum. Behav.* **4**, 1–7 (2020).
16. D. R. Farine, H. Whitehead, *J. Anim. Ecol.* **84**, 1144–1163 (2015).
17. S. Noske, thesis, University of New England, Armidale (1980); <https://hdl.handle.net/1959.11/19443>.
18. L. M. Aplin, B. C. Sheldon, J. Morand-Ferron, *Anim. Behav.* **85**, 1225–1232 (2013).
19. M. Rodriguez-Santiago *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 18566–18573 (2020).

20. G. K. Gajdon, N. Fijn, L. Huber, . *Anim. Cogn.* **9**, 173–181 (2006).
21. N. Claidière, E. J. Messer, W. Hoppitt, A. Whiten, *Curr. Biol.* **23**, 1251–1255 (2013).
22. A. Whiten, *Interface Focus* **7**, 20160142 (2017).
23. T. Gruber *et al.*, *Palgrave Commun.* **5**, 64 (2019).
24. B. C. Klump *et al.*, Innovation and geographic spread of a complex foraging culture in an urban parrot, Edmond (2021); <https://dx.doi.org/10.17617/3.5t>.

ACKNOWLEDGMENTS

We are grateful for support from A. van Humbeeck and all field assistants (data collection), J. Penndorf (data analysis: dominance hierarchy), D. Farine (data analysis: behavioral sequences), G. Alarcón-Nieto (data preparation), and T. Jäger (movie scoring). We thank all survey participants, the CCE lab for discussion, G. Ruxton for comments on an earlier draft, and

anonymous reviewers for constructive feedback. **Funding:** This study was supported by a Max Planck Society Group Leader Fellowship to L.M.A. and a National Geographic Early Career Grant to B.C.K. L.M.A. and S.W. were supported by the Centre for the Advanced Study of Collective Behaviour funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2117 – 422037984). **Author contributions:** B.C.K., J.M.M., R.E.M., and L.M.A. designed the survey; B.C.K., J.K.H., R.E.M., and J.M.M. conducted fieldwork; B.C.K. scored movies and analyzed survey and movie data; sequence analysis was prepared by B.C.K. and conducted by L.M.A.; NBDA analysis was prepared by B.C.K. and L.M.A. and conducted by S.W.; J.K.H. rescored movies for interobserver reliability. B.C.K., S.W., and L.M.A. prepared figures, and B.C.K. wrote the paper, which was revised by all authors. L.M.A. and B.C.K. secured funding; all authors discussed the results and approved the final version. **Competing interests:** The authors

declare no competing interests. **Data and materials availability:** Data available at Edmond (24).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6553/456/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 to S5
References (25–50)
MDAR Reproducibility Checklist
Movie S1
Data S1 to S3

1 October 2020; accepted 9 June 2021
10.1126/science.abe7808



myIDP:
A career plan customized for you, by you.

For your career in science, 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.

 Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers In partnership with:



Australian National University

Research School of Biology

Employment opportunity

Classification: Academic Level B (Lecturer), or C (Senior Lecturer) or D (Associate Professor)

Salary package: \$99,809-\$152,124 per annum plus 17% Superannuation

Term: Continuing

The opportunity

We are seeking to appoint up to three new continuing faculty members at Academic Levels B, C or D:

- To address goals for gender equity, two positions are identified positions for applicants who identify as women.
- Position one-photosynthetic biology,
- Position two-host-microbe biology or membrane biology.
- The third position is open to all applicants and will be appointed in any area consistent with RSB research and teaching priorities.
- Applicants for all three positions should have received their PhDs within the last 15 years, allowing for career interruptions.

The school

The ANU Research School of Biology conducts research, research training (MPhil and PhD), Masters and undergraduate education in a wide range of biological and biomedical sciences, within three separate Division of Biomedical Science and Biochemistry (BSB); Division of Ecology and Evolution (E&E); Division of Plant Sciences (PS).

Further information

<https://jobs.anu.edu.au/cw/en/job/540890/group-leader>
or contact:
Professor Craig Moritz
Director, Research School of Biology
E: craig.moritz@anu.edu.au



Who's the top employer for 2020?

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

Science 2020 TOP EMPLOYER

By Anna Moyer

For those we've lost

Every day on my way to lab, I pass a bulletin board tacked with 4-by-6 photos of a petite brunette with twinkling eyes and a broad smile. In one, she's celebrating her successful qualifying exam. In another, she's *en pointe*, arms arched gracefully above her head. It was in this spot just over 2 years ago that I last spoke to Maria, who was standing in the hallway eating her lunch. "Why don't you go sit down in the break room to eat?" I laughed as I passed by. "If I sit down, I'll fall asleep," she answered. I walked on, concerned but telling myself she must have been tired from working long hours, as she always did.

A few weeks later, Maria passed away suddenly from a mysterious autoimmune disease. It was as though the life went out of our department. Maria had been a near-constant, vivid presence, and a bright light of welcome to everyone who entered the building. The hallways felt empty without the echo of her laughter. Her death was especially shocking because I, like many in the department, didn't even know she'd been sick.

Maria was by all accounts the model student: She had received a prestigious fellowship, was about to publish two high-profile research articles, and was deeply involved in serving the community. It was only after her death that I learned Maria had spent the last weeks of her life searching for answers to unexplained symptoms.

As I weighed her public accomplishments against her private hospitalizations and struggle to find a diagnosis, I felt like a hypocrite. For me, hiding the effects of my connective tissue disorder—my chronic pain, crushing fatigue, and skeletal deformities—was an art form. Summer of 2013, the classes taken during my recovery from major surgery to hide my summer "off." Summer of 2016, another major surgery and the pain of sitting through a short course too soon after having my rib cage pieced together like a jigsaw puzzle. Fall of 2016, weeks of 7 a.m. physical therapy scheduled to avoid missing classes. Summer of 2017, sleeping through my flight to a conference. Perhaps worse than any of the physical symptoms was my sense of alienation from my classmates, who seemed to know more about backpacking across the country and pulling all-nighters in the lab than appealing denied health insurance claims or juggling doctor's appointments. I wondered whether there was a place for me in academic science.

Maria's sudden death was a wake-up call. I began to reach



"Maria's sudden death was a wake-up call. ... I was only alone to the extent that I stayed silent."

I've now talked to countless students with a range of conditions and experiences. Ironically, the one feeling we all share is a deep sense of isolation. It can seem a privilege to be able to pass as nondisabled and avoid the potential consequences of disclosure, but it is also a curse when we choose to suffer in silence.

In 4 more years, the graduate students in our department will know Maria only as a name on a plaque. But the community can honor Maria's legacy by taking steps to ease the burdens of those with chronic illnesses and disabilities. Include us in diversity initiatives. Simplify access to health care and insurance. Promote a climate that prioritizes health above productivity. Foster a sense of community where we feel safe to disclose our conditions. Show us that we belong in science and don't let us fall through the cracks. ■

Anna Moyer is a graduate student at the Johns Hopkins University School of Medicine. Send your story to SciCareerEditor@aaas.org.

READY TO PUT THE SPOTLIGHT ON YOUR RESEARCH?

Submit your research:
cts.ScienceMag.org

ScienceImmunology



Twitter: @SciImmunology



Facebook: @ScienceImmunology



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

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE