

University of Dundee

Target 2035-update on the quest for a probe for every protein

Mueller, Susanne; Ackloo, Suzanne; Al Chawaf, Arij; Al-Lazikani, Bissan; Antolin, Albert; Baell, Jonathan B.

Published in:
RSC Medicinal Chemistry

DOI:
[10.1039/d1md00228g](https://doi.org/10.1039/d1md00228g)

Publication date:
2022

Licence:
CC BY

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Mueller, S., Ackloo, S., Al Chawaf, A., Al-Lazikani, B., Antolin, A., Baell, J. B., Beck, H., Beedie, S., Betz, U. A. K., Bezerra, G. A., Brennan, P. E., Brown, D., Brown, P. J., Bullock, A. N., Carter, A. J., Chaikuad, A., Chaineau, M., Ciulli, A., Collins, I., ... Arrowsmith, C. H. (2022). Target 2035-update on the quest for a probe for every protein. *RSC Medicinal Chemistry*, 13(1), 13-21. <https://doi.org/10.1039/d1md00228g>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

RSC Medicinal Chemistry

rsc.li/medchem

Towards Medicines
For All

TARGET
2035



ISSN 2632-8682

OPINION

Susanne Müller, Cheryl H. Arrowsmith *et al.*
Target 2035 - update on the quest for a probe for
every protein

Cite this: *RSC Med. Chem.*, 2022, 13, 13

Target 2035 – update on the quest for a probe for every protein

Susanne Müller, ^{id}*^{ab} Suzanne Ackloo,^c Arij Al Chawaf,^d Bissan Al-Lazikani,^{ef} Albert Antolin,^{eg} Jonathan B. Baell,^{id}^{hi} Hartmut Beck,^j Shaunna Beedie,^k Ulrich A. K. Betz,^l Gustavo Arruda Bezerra,^k Paul E. Brennan,^{id}^m David Brown,ⁿ Peter J. Brown,^c Alex N. Bullock,^k Adrian J. Carter,^o Apirat Chaikuad,^{ab} Mathilde Chaineau,^{id}^p Alessio Ciulli,^{id}^q Ian Collins,^g Jan Dreher,^j David Drewry,^{rs} Kristina Edfeldt,^t Aled M. Edwards,^c Ursula Egner,^u Stephen V. Frye,^{sv} Stephen M. Fuchs,^w Matthew D. Hall,^x Ingo V. Hartung,^{id}^y Alexander Hillisch,^j Stephen H. Hitchcock,^z Evert Homan,^{aa} Natarajan Kannan,^{ab} James R. Kiefer,^{ac} Stefan Knapp,^{ab} Milka Kostic,^{id}^{ad} Stefan Kubicek,^{ae} Andrew R. Leach,^{af} Sven Lindemann,^{id}^{ag} Brian D. Marsden,^{kah} Hisanori Matsui,^{ai} Jordan L. Meier,^{id}^{aj} Daniel Merk,^{aak} Maurice Michel,^{aa} Maxwell R. Morgan,^{id}^c Anke Mueller-Farnow,^u Dafydd R. Owen,^{al} Benjamin G. Perry,^{id}^{am} Saul H. Rosenberg,^{an} Kumar Singh Saikatendu,^{ao} Matthieu Schapira,^{cap} Cora Scholten,^{aq} Sujata Sharma,^{ar} Anton Simeonov,^x Michael Sundström,^{as} Giulio Superti-Furga,^{aeat} Matthew H. Todd,^{id}^{au} Claudia Tredup,^{ab} Masoud Vedadi,^{id}^{cap} Frank von Delft,^{kavawax} Timothy M. Willson,^s Georg E. Winter,^{ae} Paul Workman^{fg} and Cheryl H. Arrowsmith^{*cay}

Twenty years after the publication of the first draft of the human genome, our knowledge of the human proteome is still fragmented. The challenge of translating the wealth of new knowledge from genomics into new medicines is that proteins, and not genes, are the primary executors of biological function. Therefore, much of how biology works in health and disease must be understood through the lens of protein function. Accordingly, a subset of human proteins has been at the heart of research interests of scientists over the centuries, and we have accumulated varying degrees of knowledge about approximately 65% of the human proteome. Nevertheless, a large proportion of proteins in the human proteome (~35%) remains uncharacterized, and less than 5% of the human proteome has been successfully targeted for drug discovery. This highlights the profound disconnect between our abilities to obtain genetic information and subsequent development of effective medicines. Target 2035 is an international federation of biomedical scientists from the public and private sectors, which aims to address this gap by developing and applying new technologies to create by year 2035 chemogenomic libraries, chemical probes, and/or biological probes for the entire human proteome.

Received 6th July 2021,
Accepted 21st September 2021

DOI: 10.1039/d1md00228g

rsc.li/medchem

^a Institute of Pharmaceutical Chemistry, Goethe University Frankfurt, Frankfurt 60438, Germany

^b Structural Genomics Consortium, BMLS, Goethe University Frankfurt, Frankfurt 60438, Germany. E-mail: susanne.mueller-knapp@bmls.de

^c Structural Genomics Consortium, University of Toronto, Toronto, Ontario, M5G 1L7, Canada. E-mail: Cheryl.Arrowsmith@uhnresearch.ca

^d University of Toronto, Toronto, Ontario, M5G 1L7, Canada

^e Department of Data Science, The Institute of Cancer Research, London, SM2 5NG, UK

^f CRUK ICR/Imperial Convergence Science Centre, London, SM2 5NG, UK

^g Cancer Research UK Cancer Therapeutics Unit, The Institute of Cancer Research, London SM2 5NG, UK

^h Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

ⁱ School of Pharmaceutical Sciences, Nanjing Tech University, No. 30 South Puzhu Road, Nanjing 211816, People's Republic of China

^j Research and Development, Bayer AG, Pharmaceuticals, 42103 Wuppertal, Germany

^k Centre for Medicines Discovery, University of Oxford, Old Road Campus Research Building, Roosevelt Drive, Oxford OX3 7DQ, UK

^l Merck KGaA, Darmstadt, Germany

^m Alzheimer's Research UK Oxford Drug Discovery Institute, Centre for Medicines Discovery, University of Oxford, Oxford, OX3 7FZ, UK

ⁿ Institut Recherches de Servier, 125 Chemin de Ronde, 78290 Croissy, France

^o Discovery Research, Boehringer Ingelheim, 55216 Ingelheim am Rhein, Germany
^p Early Drug Discovery Unit (EDDU), Montreal Neurological Institute-Hospital, McGill University, Montreal, QC, Canada

^q School of Life Sciences, Division of Biological Chemistry and Drug Discovery, University of Dundee, James Black Centre, Dundee, UK

^r Structural Genomics Consortium, UNC Eshelman School of Pharmacy, Chapel Hill, NC, USA

^s Lineberger Comprehensive Cancer Center, Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA



Target 2035: conceptual framework

The Target 2035 concept arose from discussions among scientists at the Structural Genomics Consortium (SGC) and like-minded colleagues in industry, government and academia^{2,5-7} who recognised the slow progress towards understanding and exploiting the plethora of human proteins ('dark proteome') despite their suspected or potential role in disease states.¹⁻⁴ This large untapped reservoir may provide new targets to address unmet medical needs. We recognised the lack of tools to study protein function as the main obstacle to progress in this area. Therefore, we launched Target 2035 as an ambitious open science initiative to discover and make available chemogenomic libraries,⁸⁻¹¹ chemical probes, and/or functional antibodies for nearly all human proteins by the year 2035.^{2,12}

Target 2035 provides a strategy for the global research community to work together to create the technologies and research tools required to interrogate the function and therapeutic potential of all human proteins. We are especially interested in the discovery of pharmacological modulators (chemical probes and chemogenomic compounds), given that availability of such compounds is known to have a transformative effect on the ability to study protein biology and target proteins for drug discovery purposes.² It is expected that the resulting research tools and knowledge will significantly enhance our ability to translate the tremendous recent advances in genomics into new medicines. Importantly, the Target 2035 model is founded on open science, collaboration, and data sharing between interested scientists from both the public and private sectors (Fig. 1A).

The SGC has assumed an initial leadership and organisational role and, in consultation with the broader community, formulated key short-term and long-term strategic priorities (Fig. 1B). The short-term priorities (phase I) focus on establishing a roadmap around which international groups of like-minded scientists can assemble and collaborate towards the following goals: (1) collecting, characterising, and distributing existing pharmacological modulators; (2) generating novel chemical probes for 'druggable' proteins; (3) developing a centralised infrastructure to facilitate data collection, curation, dissemination, and mining; and (4) creating centralised facilities to streamline ligand discovery for currently 'undruggable' targets. Long-term priorities would build on the achievements of phase I to transition into a more formalised federation and accelerate efforts towards creating solutions for the 'dark proteome' (Fig. 1B).

Here, we provide an update on the progress towards establishing phase I of Target 2035 as a robust collaborative network of scientists to develop and apply new technologies towards generating specific pharmacological modulators for all human proteins.

Outreach activities towards engaging the global research community

Target 2035 is currently an informal alliance of biomedical scientists, and its success will require much wider participation, collaboration, and innovation from the global research community.² Thus, a key challenge facing Target 2035 in this launch phase is increasing awareness of the

¹ Structural Genomics Consortium, Department of Medicine, Karolinska University Hospital and Karolinska Institutet, Stockholm, Sweden

² Nuvisan Innovation Campus Berlin GmbH, Müllerstraße 178, 13353, Berlin, Germany

³ Center for Integrative Chemical Biology and Drug Discovery, Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA

⁴ Institute for Protein Innovation, Boston, MA, USA

⁵ National Center for Advancing Translational Sciences, National Institutes of Health, Rockville, Maryland 20850, USA

⁶ Medicinal Chemistry, Global R&D, Merck Healthcare KGaA, Frankfurter Straße 250, 64293, Darmstadt, Germany

⁷ Takeda California, 9625 Towne Centre Drive, San Diego, California 92121, USA

⁸ Science for Life Laboratory, Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden

⁹ Institute of Bioinformatics and Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA, USA

¹⁰ Genentech, Inc., 1 DNA Way, South San Francisco, California 94080, USA

¹¹ Department of Cancer Biology and Chemical Biology Program, Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, MA 02215, USA

¹² CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria

¹³ European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire CB10 1SD, UK

¹⁴ Strategic Innovation, Global R&D, Merck Healthcare KGaA, Frankfurter Straße 250, 64293, Darmstadt, Germany

¹⁵ Kennedy Institute of Rheumatology, NDORMS, University of Oxford, UK

¹⁶ Neuroscience Drug Discovery Unit, Research, Takeda Pharmaceutical Company Limited, Fujisawa, Kanagawa, Japan

¹⁷ Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD, USA

¹⁸ LMU Munich, Department of Pharmacy, Chair of Pharmaceutical and Medicinal Chemistry, 81377 Munich, Germany

¹⁹ Discovery Network Group, Pfizer Medicine Design, Cambridge, MA 02139, USA

²⁰ Drugs for Neglected Diseases initiative, 15 Chemin Camille Vidart, Geneva 1202, Switzerland

²¹ AbbVie, North Chicago, Illinois, USA

²² Global Research Externalization, Takeda California, Inc., 9625 Towne Center Drive, San Diego, CA 92121, USA

²³ Department of Pharmacology & Toxicology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

²⁴ Research and Development, Bayer AG, Pharmaceuticals, 13353 Berlin, Germany

²⁵ Structural & Protein Sciences, Discovery Sciences, Janssen Research & Development, 1400 McKean Rd, Spring House, PA 19477, USA

²⁶ Division of Rheumatology, Department of Medicine Solna, Karolinska University Hospital and Karolinska Institutet, Stockholm, Sweden

²⁷ Center for Physiology and Pharmacology, Medical University of Vienna, Vienna, Austria

²⁸ School of Pharmacy, University College London, London, WC1N 1AX, UK

²⁹ Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot OX11 0QX, UK

³⁰ Department of Biochemistry, University of Johannesburg, Auckland Park 2006, South Africa

³¹ Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot OX11 0FA, UK

³² Princess Margaret Cancer Centre, Toronto, Ontario, M5G 1L7, Canada



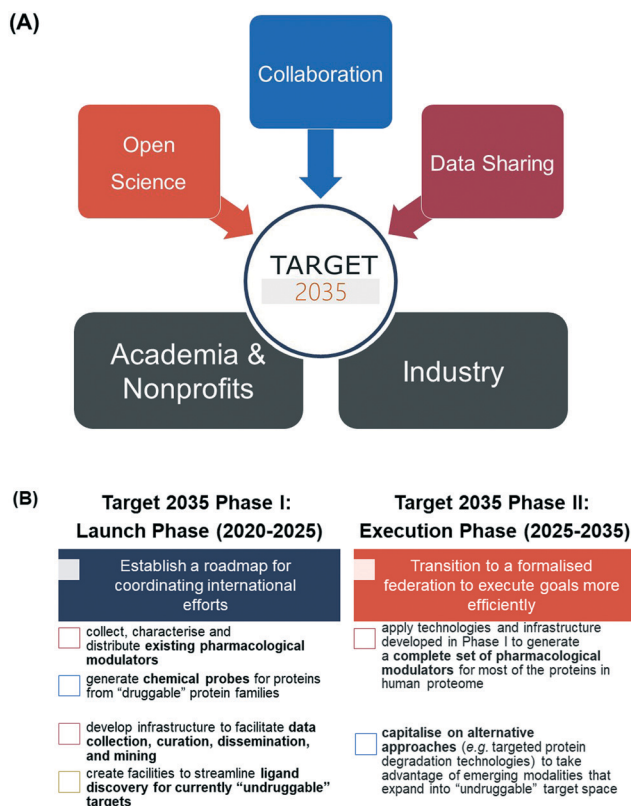


Fig. 1 Conceptual framework and strategic priorities for Target 2035. (A) Target 2035 integrates principles of open science, collaboration, and data sharing between scientists from academic, non-profit and industrial institutions. (B) Target 2035 will be executed in two phases (phase I and II), with clearly defined strategic priorities aligned with the ultimate goal to develop chemogenomic libraries, pharmacological and biochemical tool compounds (chemical probes), and/or functional antibodies for nearly all human proteins by the year 2035.

initiative in the biomedical research community, encouraging participation, and enabling knowledge exchange and collaborative opportunities. The website <https://www.target2035.net> is the main gateway currently being used to share information and communicate with the research community. Additionally, we have published an initial announcement of the initiative,² launched a series of webinars (see below) and leverage social media (#Target2035) as strategies to inform, engage and nucleate a Target 2035 community.

Webinar series

In order to engage the research community, Target 2035 hosts free (open) monthly webinars amenable to multiple global time zones. The webinars are also recorded and available online at <https://www.target2035.net/archived-recordings> and on YouTube; they are valuable resources for the scientific community including trainees in medicinal chemistry and chemical biology. The kick-off webinars in November 2020 featured key chemical biology and drug discovery opinion leaders who discussed the challenges and impact of Target

2035. Subsequent monthly webinars have focused on specific themes including targeted protein degradation and proximity-induced pharmacology, covalent ligand screening, computational and artificial intelligence (AI) methods for ligand discovery and chemoproteomic methods.¹³

Overall, feedback from the scientific community has been positive, with new collaborations arising from research highlighted in the webinars. Further, the Target 2035 community is active on social media and online tools (#Target2035, https://en.wikipedia.org/wiki/Target_2035), with much discussion of webinars, speakers, and recent publications. By facilitating the online conversation, we hope to gain an increase in awareness and involvement and continue to build the community.

New initiatives and research projects

Various relevant large-scale initiatives are already underway, several with the expressed objective of contributing to Target 2035. These consortia and groups are generally multidisciplinary, international and open access. Notably many groups – such as members of the Illuminating the Druggable Genome (IDG)⁴ project whose output is published in *Pharos*,¹⁴ the SGC and EUBOPEN (described below) – are taking a protein family approach, which is an efficient and practical way to tackle the human proteome.^{2,15} These efforts are often building on other enabling initiatives, where key reagents and protocols for an entire target class are being generated. The ReSOLUTE initiative, for example, is already well on its way to unlock solute carriers (SLCs) for chemical probe discovery by the community.¹⁶ The outputs of ReSOLUTE include robust assays for most SLCs in the genome, as well as the creation of other empowering tools, such as thousands of tailored cell lines, and making all data, protocols, and reagents publicly available through their portal or open distributors (re-solute.eu). The IDG project, supported by the NIH Common Fund,¹⁷ is developing chemical tools, assays, expression data, interaction maps, and knock-out mice for many dark members¹⁸ of the highly druggable GPCR, kinase, and ion channel protein families. All data, reagents, and tools generated by IDG are accessible through its portal.¹⁹ The IDG is also developing a suite of informatics tools that can be used to interrogate the biological function of these dark proteins and their potential role in human diseases.

Due to space restrictions, we will highlight here just a few of the additional initiatives that directly impact Target 2035 goals.

EUBOPEN

Having identified the need for high quality chemical probes for the research community, the EUBOPEN partnership of excellence (www.EUBOPEN.org) came together to 'Enable and Unlock biology in the OPEN' as part of the Innovative Medicines Initiative (IMI).²⁰ EUBOPEN aims to generate the largest freely available set of high-quality chemical



modulators for human proteins. These compounds will complement existing chemical probe resources such as the SGC's kinase, epigenetic and donated chemical probes collections.^{21–23} The EUBOPEN compound collections will include chemical probes for high impact target areas relating to human health, including solute carriers (SLCs) and ubiquitin ligases (E3s), where high-quality small-molecule binders have historically been lacking compared to other families *e.g.* protein kinases. It will also compile a chemogenomic library (CGL) of ~4–5000 compounds covering one third of the current druggable genome (Fig. 2). All compounds will be comprehensively characterised by defining their selectivity, potency, and cellular activity. The compounds will be annotated with activity data from a series of established and novel biochemical and cell-based assays, including many derived from primary patient cells. The development of new technologies to discover and characterise chemogenomics compounds and chemical probes is a core component of the project.

EUBOPEN partners include scientists from academia and industry developing openly available research reagents and exploring new protein target classes. Of key importance, the EUBOPEN project has a strong commitment to open data and open science, which will accelerate the dissemination of knowledge and project outputs, including chemogenomic library sets, chemical probes, assay protocols and associated research. To boost the project, EUBOPEN partners interact extensively with other large consortia (*e.g.* ReSOLUTE,¹⁶ Illuminating the Druggable Genome,²⁴ Open Targets²⁵) thus minimising the duplication of effort and maximizing global coordination. Finally, sustainability of the project's resources will be ensured through partnerships with chemical vendors and cheminformatics/database providers, which will ensure the long-term availability of the new chemical tool sets, all associated profiling data and established protocols. By working openly, and engaging the wider research community, this project will deliver pharmacological tool compounds for previously understudied targets, contributing to the Target 2035 campaign, and form the foundation for future efforts to generate chemical modulators for the entire druggable genome.

The SGC laboratory at the University of Toronto has launched a pilot project, *Target 2035: WDR proteins as a technology test-bed for illuminating the dark proteome*. WD40 repeat (WDR) domains are protein–protein interaction domains comprising one of the largest and most disease-associated protein families in the human proteome.^{26,27} This open science, industry-partnered programme is comparing different experimental strategies for progressing WDR targets from gene to chemical modulator, including AI-guided methods to accelerate discovery of new chemical tools for this under-explored protein family.

Open chemistry networks projects

The creation of such a large number of chemical modulators will require a significant amount of synthetic medicinal and organic chemistry to design, make and test compounds that are not commercially available. The SGC's new Open Chemistry Networks initiative aims to create opportunities for the community to drive the development of these novel probes through a highly distributed, open chemistry network in which chemical resources (compounds, ideas, laboratory space and time) are contributed on a patent-free, open access basis in return for biological evaluation of those compounds by the Target 2035 network, as well as exposure to the educational opportunities afforded by Target 2035. All crowdsourced chemistry inputs and Target 2035-generated assay data will be openly available to the research community without restrictions, under a novel set of open terms, analogous to a Creative Commons Attribution 4.0 International license for chemistry, including for use by participants towards co-authored publications and grant proposals. Based on prior experience with such distributed chemistry,²⁸ different cohorts of scientists are likely to be interested in different types of projects. The synthetic routes will range from simpler (for example, the synthesis of close analogues of a hit compound by straightforward coupling reactions) to more demanding modifications (for example, significant changes to a compound's core scaffold to optimise specific properties). The former might be of interest to undergraduate student cohorts while the latter may be a

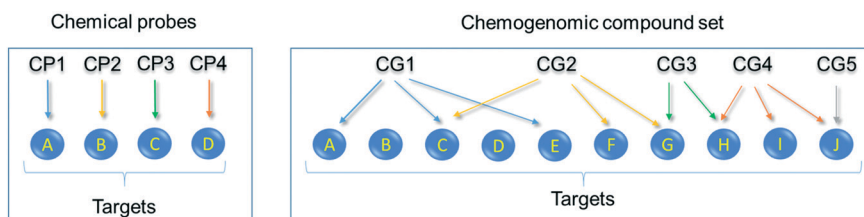


Fig. 2 Relationship between chemical compounds and their targets for chemical probes *versus* chemogenomic compound sets. Chemical probes (CPs) are small molecules that are potent, selective and cell-active. Ideally, a chemical probe negative control is identified and used in parallel in phenotypic assays. A chemical probe ideally binds to a single target but can also be useful if it targets just a small number (~2–3) of very closely sequence-related proteins. A chemogenomics library (CGL) comprises chemogenomic (CG) compounds of different chemotypes for each target; the chemotypes have complementary selectivity profiles, and preferably distinct modes of action (such as orthosteric or allosteric modulation). In the schematic, a CGL was screened against targets A–J yielding phenotypic data for 8 targets. These data are interpreted in the context of the magnitude of the phenotypic response, the mode of action (colour of arrow) and the CG chemical structures and accompanying characterisation.



more time-intensive project suitable for a MSc or PhD thesis. Each project will be led by a science team that could involve a variety of stakeholders/contributors such as: 1) the discoverer of the hit; 2) a pharmaceutical industry mentor, tasked with ensuring the project is kept on track and key data are collected; and 3) a champion acting as the project manager in areas such as managing meetings, the project 'To Do list' and community outreach. The online infrastructure needed (e.g. online lab notebook,²⁹ chemical registration system, project coordination website) will be easy to use and freely accessible to all participants. To ensure a clear description of project status and to ensure a connection between project activity and paper authorship, a novel form of 'living paper' will operate for each project: a series of preprints, where each 'version' is posted upon reaching a project milestone. Progress towards implementing this system has already been made and the first project has been posted.³⁰ Others will come online as Target 2035 accelerates.

EU-OPENSREEN³¹

EU-OPENSREEN³¹ is an open screening platform that offers access to a broad range of technologies and tools for the systematic screening of chemical inhibitors. Scientists from academia and industry can implement their screening projects at EU-OPENSREEN's partner laboratories using the European Chemical Biology Library (ECBL), a diversity compound collection of 140 000 commercially available as well as academic compounds, crowdsourced through chemical biology networks. All generated data and tool compounds are made available in an open access database for use by the wider life science as well as drug discovery communities.

FAIRplus data sharing

As outlined above, Target 2035 aims can only be reached with mass collaboration, including openness and data sharing. Given the vast number of different types of biological data sets produced during the chemical probe development process, and differences in data management between institutions, data outputs can be highly variable. The FAIRplus project aims to change and sustain the data management culture by providing tools to ensure that data produced follows community standards and ontologies, along with well-defined data dissemination pathways. As an example, EubOPEN's mandate is to share its project outputs in a sustainable manner. Therefore, the project is committed to ensuring that all outputs are compliant with FAIR principles:³² data is findable, accessible, interoperable, and reusable. To support this effort, EubOPEN works directly with the FAIRplus IMI consortium who are providing advice and support to ensure the outputs can be used by the research community. The collaboration is part of EubOPEN's drive to develop sustainable infrastructures and feeds the FAIRplus vision of making all life science data FAIR.

The Chemical Probes Portal

The Chemical Probes Portal (Portal; <https://www.chemicalprobes.org>) is a long-standing resource for the biomedical research community that provides expert-curated information on the quality of chemical probes, and enables users to identify the best probes for a biological experiment or study.⁵ The Portal was very recently re-launched to provide the community with an enhanced and extended user-friendly interface and integrate chemical compound information from multiple relevant databases to enable user decision-making in terms of the right chemical modulator for a given target and a given biological context. Although there are several complementary online resources available that automatically score the quality of probes using computational approaches,^{33–38} such as Probe Miner,³⁹ thus far the Portal is the only expert-curated resource.³³ The Portal currently features more than 600 compounds covering 300 targets and draws on more than 170 chemical biology experts to provide three reviews for each probe, commenting on their characterisation data and properties and providing recommendations on their usefulness, flaws, and strengths. While the majority of probes so far are inhibitors or antagonists, the Portal provides guidelines and information pages about new types of chemical modulators such as PROTACs and molecular glue degraders, and recommendations and guidelines about evaluating their performance.⁴⁰ Importantly, the Portal also contains a list of historical but still prevalent compounds that typically are non-selective or not sufficiently potent compared to other available chemical probes in order to discourage their use when interrogating the biological activity of specific targets. The new user interface will facilitate the submission of probes and reviews, streamlining the process to facilitate an increase in target coverage.

Considerations for the open sharing of chemical matter

The discovery of new chemical matter is often associated with IP protection, based on the expectation that IP protection is required for further translation, application, or exploitation of the discovery. However, IP protection can also be a barrier to collaboration and open sharing of information or materials, and therefore requires new approaches, creativity and novel ways of thinking. A number of innovative approaches have been chosen by different stakeholders. Broadly, they fall into two categories. The first, is for a single 'closed' discovery team (such as a single lab or company) to establish their IP position for a chemical probe (to be patented or not patented), and then share the chemical probe with the community for unrestricted research use. For example, as part of the DCP program, pharmaceutical companies provide the SGC with a license to make data and compounds available to the public, which is also extended to selected trusted commercial vendors in order to guarantee a sustainable supply of the chemical probe matter as well as quality and structural identity of the distributed compounds. In a similar strategy, Boehringer Ingelheim makes some of their well-characterized compounds



available *for free via* their *opnMe* program,⁴¹ where researchers can order molecules directly *via* the website for unencumbered research use. The second, alternative approach is that taken by the SGC for the discovery of novel chemical probes. In this scenario multi-institutional groups collaborating on a given chemical probe project agree, *in advance*, that the output of their research will be an unencumbered public domain tool. This enables a more open and rapid exchange of information and materials, and does not preclude the future development of therapeutic products based on the initial chemical probe. Such products may emerge through public and philanthropic investment to advance the initial probe into later-stage development^{42,43} or through a more traditional proprietary drug development paradigm, for example by advancing patented, optimized derivatives of the initial probe structure⁴⁴ or novel chemical matter developed through further studies enabled by the probe.

Future outlook

Since the launch of Target 2035 in 2019, we have witnessed the success of the open science model in the global fight against the COVID-19 pandemic. The swift sharing of SARS-CoV2 viral sequences⁴⁵ and other data has led to extraordinarily rapid development of new vaccines⁴⁶ and small molecule antiviral trials.^{47–50} A recent example is the fast development of anti-viral small molecules against the SARS-CoV-2 protease, as part of an initiative called the COVID Moonshot.⁵¹ The rapid release of the SARS-CoV-2 Mpro protease crystal structure⁵² and subsequent crystallography-based fragment screening campaign at the Diamond Light Source XChem facility set the starting point for this drug discovery endeavour.⁵³ In collaboration with laboratories based in several continents, some of these compounds have been characterised and progressed into anti-viral probes that are shortly entering clinical trials. Similarly, global open collaborative initiatives in drug discovery addressing different SARV-CoV-2 targets have also significantly advanced the development new anti-virals.^{50,54} The events of the past year and a half have demonstrated how scientists in academia and industry can work together around the world at a pace and scale to address important biomedical problems with open science driving the early stages of drug/vaccine research.

Target 2035 aspires to replicate this success but on a human-proteome scale, which will require building a large robust scientific community that works together using open science principles to achieve its ambitious goals and to avoid redundancies and wasting of scarce resources. We have already made significant progress in developing the underlying infrastructure, *e.g.*, with respect to data handling and mining.

Future plans include expansion of the Target 2035 webpage to become a gateway for information sharing and communication among interested scientists. This will include a collaboration marketplace where researchers will be able to share information on their research projects and identify potential collaborators with complementary expertise,

infrastructure, and capabilities. We invite those interested to engage with the current Target 2035 team and consider getting involved. Novel technologies, such as machine learning and high throughput chemistry, proteome-wide analyses, and collaborative research approaches will need to be developed to accelerate the rate of ligand discovery and characterisation. The challenges of finding a high quality tool for any protein of interest by 2035 are immense. However, if we work together and share the results freely, we will succeed in reaching this goal as seen from the success of the Human Genome Project, also a highly collaborative, global effort. We are optimistic that major breakthroughs are around the corner and that we will enable high quality biological data linking specific protein targets with phenotypic outcomes and new therapeutic hypotheses for tomorrow's medicines.

Conflicts of interest

A. A. A., B. A.-L., I. C., and P. W. are employees of the Institute of Cancer Research (ICR) which operates a Rewards to Inventors scheme whereby employees of the ICR may receive financial benefit following commercial licensing of their research. P. W. is a consultant/scientific advisory board member for Nextech Invest Ltd, Storm Therapeutics, Astex Pharmaceuticals, Black Diamond Therapeutics, CV6 and Vividion Therapeutics, and holds stock in Chroma Therapeutics, NextInvest, and Storm Therapeutics. P. W. is also a non-executive director of Storm Therapeutics and the Royal Marsden NHS Trust; a board member and executive director of the non-profit Chemical Probes Portal; and a former employee of AstraZeneca. P. W. has received research funding from Vernalis, Astex Therapeutics, Merck KGaA, BACIT/Sixth Element Capital/CRT Pioneer Fund. B. A.-L. is/was a consultant/scientific advisory board member for GSK, Open Targets, Astex Pharmaceuticals, Nuvectis Pharma and Astellas Pharma, and is an ex-employee of Inpharmatica Ltd. A. A. A., B. A.-L., and P. W. have been instrumental in the creation/development of canSAR and Probe Miner. B. A.-L. was instrumental in the creation of ChEMBL and is a director of the non-profit Chemical Probes Portal. I. C. is/was a consultant to Epidarex LLP, AdoRx Therapeutics, and Enterprise Therapeutics, and is a director of the non-profit Chemical Probes Portal. I. C. has received research funding from Astex, Merck KGaA, Janssen Biopharma, Monte Rosa Therapeutics, and Sixth Element Capital/CRT Pioneer Fund. I. C. holds stock in Monte Rosa Therapeutics AG and is a former employee of Merck Sharp & Dohme. M. K. is a paid consultant for Life Science Editors (LSE). A. C. receives or has received sponsored research support from Almirall, Amphista therapeutics, Boehringer Ingelheim, Eisai, Nurix therapeutics, and Ono Pharmaceuticals. A. C. is a scientific founder, shareholder, and consultant of Amphista therapeutics. A. R. L. has consulted for Astex Therapeutics and has received research funding from Novo Nordisk. S. Ku. is a co-founder and shareholder of Proxygen GmbH and Solgate GmbH. G. E. W. is a founder and shareholder of Proxygen and Solgate



Therapeutics. He is on the Research Review Committee of Almirall and coordinates a research collaboration between CeMM and Pfizer. S. F. reports equity ownership and membership on the Meryx Board of Directors and consulting or SAB relationships with Artios, Astex, Cullgen, Design Therapeutics, Flare, Mitokinin, Pathios, ReViral and eFFector. This communication reflects the views of the authors and neither IMI nor the European Union, EFPIA or any Associated Partners are liable for any use that may be made of the information contained herein.

Acknowledgements

The SGC is a registered charity (number 1097737) that receives funds from Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genentech, Genome Canada through Ontario Genomics Institute [OGI-196], EU/EFPIA/OICR/McGill/KTH/Diamond Innovative Medicines Initiative 2 Joint Undertaking (EubOPEN grant number 875510), Janssen, Merck KGaA (aka EMD in Canada and US), Pfizer, Takeda. A. A. A. is primarily supported by a Wellcome Trust Sir Henry Wellcome Postdoctoral Fellowship (204735/Z/16/Z); the People Programme (Marie Curie Actions) of the seventh Framework Program of the European Union (FP7/2007–2013) under REA grant agreement no. 600388 (TECNIOspring program); and the Agency for Business Competitiveness of the Government of Catalonia, ACCIO. B. A.-L., I. C., and P. W. are funded by a Cancer Research UK (CRUK) program grant to the CRUK Cancer Therapeutics Unit (C309/A11566). B. A.-L. and P. W. are funded by a Wellcome Trust biomedical resource and technology grant to sustain and develop the Chemical Probes Portal (212969/Z/18/Z) and a CRUK Drug Discovery Committee strategic award to sustain and develop canSAR (C35696/A23187). P. W. is a CRUK Life Fellow and is funded by the Chordoma Foundation and Mark Foundation. P. W. and B. A.-L. are members of the CRUK ICR/Imperial Convergence Science Centre (A26234). Institute of Cancer Research (ICR) authors acknowledge infrastructure support from CRUK for the ICR CRUK Centre and NHS funding to the NIHR Biomedical Research Centre at the ICR and Royal Marsden NHS Foundation Trust. M. K. is partly supported by the National Cancer Institute, USA (NCI) T32 CA236754. A. C. laboratory receives funding from the Innovative Medicines Initiative 2 (IMI2) Joint Undertaking under grant agreement No 875510 (EubOPEN project). A. R. L. acknowledges support from the Member States of the European Molecular Biology Laboratory (EMBL).

References

- 1 T. I. Oprea, C. G. Bologa, S. Brunak, A. Campbell, G. N. Gan, A. Gaulton, S. M. Gomez, R. Guha, A. Hersey, J. Holmes, A. Jadhav, L. J. Jensen, G. L. Johnson, A. Karlson, A. R. Leach, A. Ma'ayan, A. Malovannaya, S. Mani, S. L. Mathias, M. T. McManus, T. F. Meehan, C. von Mering, D. Muthas, D. T. Nguyen, J. P. Overington, G. Papadatos, J. Qin, C. Reich, B. L. Roth, S. C. Schurer, A. Simeonov, L. A. Sklar, N. Southall, S. Tomita, I. Tudose, O. Ursu, D. Vidovic, A. Waller, D. Westergaard, J. J. Yang and G. Zahoranzky-Kohalmi, *Nat. Rev. Drug Discovery*, 2018, **17**, 377.
- 2 A. J. Carter, O. Kraemer, M. Zwick, A. Mueller-Fahrnow, C. H. Arrowsmith and A. M. Edwards, *Drug Discovery Today*, 2019, **24**, 2111–2115.
- 3 M. E. Berginski, N. Moret, C. Liu, D. Goldfarb, P. K. Sorger and S. M. Gomez, *Nucleic Acids Res.*, 2021, **49**, D529–D535.
- 4 T. Sheils, S. L. Mathias, V. B. Siramshetty, G. Bocci, C. G. Bologa, J. J. Yang, A. Waller, N. Southall, D. T. Nguyen and T. I. Oprea, *Curr. Protoc. Bioinf.*, 2020, **69**, e92.
- 5 C. H. Arrowsmith, J. E. Audia, C. Austin, J. Baell, J. Bennett, J. Blagg, C. Bountra, P. E. Brennan, P. J. Brown, M. E. Bunnage, C. Buser-Doepner, R. M. Campbell, A. J. Carter, P. Cohen, R. A. Copeland, B. Cravatt, J. L. Dahlin, D. Dhanak, A. M. Edwards, M. Frederiksen, S. V. Frye, N. Gray, C. E. Grimshaw, D. Hepworth, T. Howe, K. V. Huber, J. Jin, S. Knapp, J. D. Kotz, R. G. Kruger, D. Lowe, M. M. Mader, B. Marsden, A. Mueller-Fahrnow, S. Muller, R. C. O'Hagan, J. P. Overington, D. R. Owen, S. H. Rosenberg, B. Roth, R. Ross, M. Schapira, S. L. Schreiber, B. Shoichet, M. Sundstrom, G. Superti-Furga, J. Taunton, L. Toledo-Sherman, C. Walpole, M. A. Walters, T. M. Willson, P. Workman, R. N. Young and W. J. Zuercher, *Nat. Chem. Biol.*, 2015, **11**, 536–541.
- 6 R. M. Garbaccio and E. R. Parmee, *Cell Chem. Biol.*, 2016, **23**, 10–17.
- 7 A. Mullard, *Nat. Rev. Drug Discovery*, 2019, **18**, 733–736.
- 8 D. H. Drewry, C. I. Wells, D. M. Andrews, R. Angell, H. Al-Ali, A. D. Axtman, S. J. Capuzzi, J. M. Elkins, P. Ettmayer, M. Frederiksen, O. Gileadi, N. Gray, A. Hooper, S. Knapp, S. Laufer, U. Luecking, M. Michaelides, S. Muller, E. Muratov, R. A. Denny, K. S. Saikatendu, D. K. Treiber, W. J. Zuercher and T. M. Willson, *PLoS One*, 2017, **12**, e0181585.
- 9 J. M. Elkins, V. Fedele, M. Szklarz, K. R. Abdul Azeez, E. Salah, J. Mikolajczyk, S. Romanov, N. Sepetov, X. P. Huang, B. L. Roth, A. Al Haj Zen, D. Fourches, E. Muratov, A. Tropsha, J. Morris, B. A. Teicher, M. Kunkel, E. Polley, K. E. Lackey, F. L. Atkinson, J. P. Overington, P. Bamborough, S. Muller, D. J. Price, T. M. Willson, D. H. Drewry, S. Knapp and W. J. Zuercher, *Nat. Biotechnol.*, 2016, **34**, 95–103.
- 10 R. B. A. Quinlan and P. E. Brennan, *RSC Chem. Biol.*, 2021, **2**(3), 759–795.
- 11 L. H. Jones and M. E. Bunnage, *Nat. Rev. Drug Discovery*, 2017, **16**, 285–296.
- 12 K. Colwill, G. Renewable Protein Binder Working and S. Graslund, *Nat. Methods*, 2011, **8**, 551–558.
- 13 A. Mullard, *Nat. Rev. Drug Discovery*, 2021, **20**, 247–250.
- 14 <https://pharos.nih.gov/>.
- 15 K. D. Barnash, L. I. James and S. V. Frye, *Nat. Chem. Biol.*, 2017, **13**, 1053–1056.
- 16 <https://re-solute.eu/>.
- 17 <https://commonfund.nih.gov/idg>.
- 18 <https://druggablegenome.net/IDGProteinList>.
- 19 <https://druggablegenome.net/IDGResourceTables>.



- 20 <https://www.imi.europa.eu/apply-funding/closed-calls/imi2-call-17#:~:text=IMI2%20%2D%20Call%2017%20was%20launched,posed%20by%20human%20medicinal%20products>.
- 21 <https://www.thesgc.org/chemical-probes>.
- 22 <https://www.sgc-ffm.uni-frankfurt.de/>.
- 23 S. Muller, S. Ackloo, C. H. Arrowsmith, M. Bauser, J. L. Baryza, J. Blagg, J. Bottcher, C. Bountra, P. J. Brown, M. E. Bunnage, A. J. Carter, D. Damerell, V. Dotsch, D. H. Drewry, A. M. Edwards, J. Edwards, J. M. Elkins, C. Fischer, S. V. Frye, A. Gollner, C. E. Grimshaw, A. P. IJzerman, T. Hanke, I. V. Hartung, S. Hitchcock, T. Howe, T. V. Hughes, S. Laufer, V. M. Li, S. Liras, B. D. Marsden, H. Matsui, J. Mathias, R. C. O'Hagan, D. R. Owen, V. Pande, D. Rauh, S. H. Rosenberg, B. L. Roth, N. S. Schneider, C. Scholten, K. Singh Saikatendu, A. Simeonov, M. Takizawa, C. Tse, P. R. Thompson, D. K. Treiber, A. Y. Viana, C. I. Wells, T. M. Willson, W. J. Zuercher, S. Knapp and A. Mueller-Fahrnow, *eLife*, 2018, 7, e34311.
- 24 <https://drugablegenome.net/>.
- 25 <https://www.opentargets.org/>.
- 26 M. Schapira, M. Tyers, M. Torrent and C. H. Arrowsmith, *Nat. Rev. Drug Discovery*, 2017, 16, 773–786.
- 27 Y. Wang, X. J. Hu, X. D. Zou, X. H. Wu, Z. Q. Ye and Y. D. Wu, *Nucleic Acids Res.*, 2015, 43, D339–344.
- 28 A. E. Williamson, P. M. Ylloja, M. N. Robertson, Y. Antonova-Koch, V. Avery, J. B. Baell, H. Batchu, S. Batra, J. N. Burrows, S. Bhattacharyya, F. Calderon, S. A. Charman, J. Clark, B. Crespo, M. Dean, S. L. Debbert, M. Delves, A. S. Dennis, F. Deroose, S. Duffy, S. Fletcher, G. Giaever, I. Hallyburton, F. J. Gamo, M. Gebbia, R. K. Guy, Z. Hungerford, K. Kirk, M. J. Lafuente-Monasterio, A. Lee, S. Meister, C. Nislow, J. P. Overington, G. Papadatos, L. Patiny, J. Pham, S. A. Ralph, A. Ruecker, E. Ryan, C. Southan, K. Srivastava, C. Swain, M. J. Tarnowski, P. Thomson, P. Turner, I. M. Wallace, T. N. Wells, K. White, L. White, P. Willis, E. A. Winzeler, S. Wittlin and M. H. Todd, *ACS Cent. Sci.*, 2016, 2, 687–701.
- 29 <https://openlabnotebooks.org/>.
- 30 <https://github.com/StructuralGenomicsConsortium/CNP1-PKIS-5>.
- 31 www.eu-openscreen.eu.
- 32 M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. T. Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S. A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao and B. Mons, *Sci. Data*, 2016, 3, 160018.
- 33 A. A. Antolin, P. Workman and B. Al-Lazikani, *Future Med. Chem.*, 2021, 13, 731–747.
- 34 <https://www.probes-drugs.org>.
- 35 <https://labsyspharm.shinyapps.io/smallmoleculesuite>.
- 36 www.bindingdb.org.
- 37 www.guidetopharmacology.org.
- 38 www.ebi.ac.uk/chembl.
- 39 A. A. Antolin, J. E. Tym, A. Komianou, I. Collins, P. Workman and B. Al-Lazikani, *Cell Chem. Biol.*, 2018, 25, 194–205e195.
- 40 R. Chopra, A. Sadok and I. Collins, *Drug Discovery Today: Technol.*, 2019, 31, 5–13.
- 41 <https://openme.com/>.
- 42 D. Ensan, D. Smil, C. A. Zepeda-Velazquez, D. Panagopoulos, J. F. Wong, E. P. Williams, R. Adamson, A. N. Bullock, T. Kiyota, A. Aman, O. G. Roberts, A. M. Edwards, J. A. O'Meara, M. B. Isaac and R. Al-Awar, *J. Med. Chem.*, 2020, 63, 4978–4996.
- 43 M. R. Morgan, O. G. Roberts and A. M. Edwards, *Wellcome Open Res.*, 2018, 3, 154.
- 44 R. E. Gold and M. Morgan, *Wellcome Open Res.*, 2019, 3, 154.
- 45 R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W. J. Liu, D. Wang, W. Xu, E. C. Holmes, G. F. Gao, G. Wu, W. Chen, W. Shi and W. Tan, *Lancet*, 2020, 395, 565–574.
- 46 P. K. Rai, Z. Mueed, A. Chowdhury, R. Deval, D. Kumar, M. A. Kamal, Y. S. Negi, S. Pareek and N. K. Poddar, *Curr. Pharm. Biotechnol.*, 2021, DOI: 10.2174/1389201022666210509022313.
- 47 R. L. Hoffman, R. S. Kania, M. A. Brothers, J. F. Davies, R. A. Ferre, K. S. Gajiwala, M. He, R. J. Hogan, K. Kozminski, L. Y. Li, J. W. Lockner, J. Lou, M. T. Marra, L. J. Mitchell, Jr., B. W. Murray, J. A. Nieman, S. Noell, S. P. Planken, T. Rowe, K. Ryan, G. J. Smith, 3rd, J. E. Solowiej, C. M. Steppan and B. Taggart, *J. Med. Chem.*, 2020, 63, 12725–12747.
- 48 <https://clinicaltrials.gov/ct2/show/NCT04535167?term=covid-19+pfizer+intravenous&draw=2&rank=1>.
- 49 J. Qiao, Y. S. Li, R. Zeng, F. L. Liu, R. H. Luo, C. Huang, Y. F. Wang, J. Zhang, B. Quan, C. Shen, X. Mao, X. Liu, W. Sun, W. Yang, X. Ni, K. Wang, L. Xu, Z. L. Duan, Q. C. Zou, H. L. Zhang, W. Qu, Y. H. Long, M. H. Li, R. C. Yang, X. Liu, J. You, Y. Zhou, R. Yao, W. P. Li, J. M. Liu, P. Chen, Y. Liu, G. F. Lin, X. Yang, J. Zou, L. Li, Y. Hu, G. W. Lu, W. M. Li, Y. Q. Wei, Y. T. Zheng, J. Lei and S. Yang, *Science*, 2021, 371, 1374–1378.
- 50 C. B. Apaydin, G. Cinar and G. Cihan-Ustundag, *Curr. Drug Targets*, 2021, DOI: 10.2174/1389450122666210215112150.
- 51 J. Chodera, A. A. Lee, N. London and F. von Delft, *Nat. Chem.*, 2020, 12, 581.
- 52 Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao and H. Yang, *Nature*, 2020, 582, 289–293.
- 53 A. Douangamath, D. Fearon, P. Gehertz, T. Krojer, P. Lukacik, C. D. Owen, E. Resnick, C. Strain-Damerell, A.



Aimon, P. Abranyi-Balogh, J. Brandao-Neto, A. Carbery, G. Davison, A. Dias, T. D. Downes, L. Dunnett, M. Fairhead, J. D. Firth, S. P. Jones, A. Keeley, G. M. Keseru, H. F. Klein, M. P. Martin, M. E. M. Noble, P. O'Brien, A. Powell, R. N. Reddi, R. Skyner, M. Snee, M. J. Waring, C.

Wild, N. London, F. von Delft and M. A. Walsh, *Nat. Commun.*, 2020, **11**, 5047.

54 A. Khalili Yazdi, F. Li, K. Devkota, S. Perveen, P. Ghiabi, T. Hajian, A. Bolotokova and M. Vedadi, *SLAS Discovery*, 2021, 24725552211008863, DOI: 10.1177/24725552211008863.

