

# 1 Fragment Binding to the Nsp3 Macrodome of SARS-CoV-2 Identified Through 2 Crystallographic Screening and Computational Docking

3  
4 Marion Schuller<sup>1†</sup>, Galen J. Correy<sup>2†</sup>, Stefan Gahbauer<sup>3†</sup>, Daren Fearon<sup>4†</sup>, Taiasean Wu<sup>5,6</sup>, Roberto Efraín  
5 Díaz<sup>2,7</sup>, Iris D. Young<sup>2,8</sup>, Luan Carvalho Martins<sup>9</sup>, Dominique H. Smith<sup>10</sup>, Ursula Schulze-Gahmen<sup>8</sup>, Tristan  
6 W. Owens<sup>8</sup>, Ishan Deshpande<sup>8</sup>, Gregory E. Merz<sup>8</sup>, Aye C. Thwin<sup>8</sup>, Justin T. Biel<sup>8</sup>, Jessica K. Peters<sup>8</sup>, Michelle  
7 Moritz<sup>8</sup>, Nadia Herrera<sup>8</sup>, Huong T. Kratochvil<sup>8</sup>, QCRG Structural Biology Consortium<sup>8‡</sup>, Anthony Aimon<sup>4</sup>,  
8 James M. Bennett<sup>11</sup>, Jose Brandao Neto<sup>4</sup>, Aina E. Cohen<sup>12</sup>, Alexandre Dias<sup>4</sup>, Alice Douangamath<sup>4</sup>, Louise  
9 Dunnett<sup>4</sup>, Oleg Fedorov<sup>11</sup>, Matteo P. Ferla<sup>13</sup>, Martin Fuchs<sup>14</sup>, Tyler J. Gorrie-Stone<sup>4</sup>, James M. Holton<sup>15,16,12</sup>,  
10 Michael G. Johnson<sup>17</sup>, Tobias Krojer<sup>11,18</sup>, George Meigs<sup>15,16</sup>, Ailsa J. Powell<sup>4</sup>, Johannes Gregor Matthias  
11 Rack<sup>1</sup>, Victor L Rangel<sup>11,18,19</sup>, Silvia Russi<sup>12</sup>, Rachael E. Skyner<sup>4</sup>, Clyde A. Smith<sup>12</sup>, Alexei S. Soares<sup>20</sup>,  
12 Jennifer L. Wierman<sup>12</sup>, Kang Zhu<sup>1</sup>, Natalia Jura<sup>21</sup>, Alan Ashworth<sup>10</sup>, John Irwin<sup>3</sup>, Michael C. Thompson<sup>22</sup>,  
13 Jason E. Gestwicki<sup>3,5</sup>, Frank von Delft<sup>4,11,18,23\*</sup>, Brian K. Shoichet<sup>3\*</sup>, James S. Fraser<sup>2\*</sup>, Ivan Ahel<sup>1\*</sup>

## 14 Affiliations

15  
16 <sup>1</sup> Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK  
17 <sup>2</sup> Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, CA, USA  
18 <sup>3</sup> Department of Pharmaceutical Chemistry, University of California San Francisco San Francisco, CA, USA  
19 <sup>4</sup> Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0DE, United Kingdom  
20 <sup>5</sup> Institute for Neurodegenerative Disease, University of California San Francisco, CA, USA  
21 <sup>6</sup> Chemistry and Chemical Biology Graduate Program, University of California San Francisco, CA, USA  
22 <sup>7</sup> Tetrad Graduate Program, University of California San Francisco, CA, USA  
23 <sup>8</sup> Quantitative Biosciences Institute (QBI) Coronavirus Research Group Structural Biology Consortium,  
24 University of California San Francisco, CA, USA  
25 <sup>9</sup> Biochemistry Department, Institute for Biological Sciences, Federal University of Minas Gerais. Belo  
26 Horizonte, Brazil  
27 <sup>10</sup> Helen Diller Family Comprehensive Cancer, University of California San Francisco, CA, USA  
28 <sup>11</sup> Centre for Medicines Discovery, University of Oxford, South Parks Road, Headington, OX3 7DQ, UK  
29 <sup>12</sup> Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Center, Menlo Park, CA 94025,  
30 USA  
31 <sup>13</sup> Wellcome Centre for Human Genetics, University of Oxford, Old Road Campus, Oxford OX3 7BN, UK  
32 <sup>14</sup> National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY, USA  
33 <sup>15</sup> Department of Biochemistry and Biophysics, University of California San Francisco, CA, USA  
34 <sup>16</sup> Department of Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory,  
35 Berkeley, CA, USA  
36 <sup>17</sup> ChemPartner Corporation, South San Francisco, CA, USA  
37 <sup>18</sup> Structural Genomics Consortium, University of Oxford, Old Road Campus, Roosevelt Drive, Headington  
38 OX3 7DQ, UK  
39 <sup>19</sup> School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, São Paulo, Brazil  
40 <sup>20</sup> Photon Sciences, Brookhaven National Laboratory, Upton, NY, USA  
41 <sup>21</sup> Department of Cellular and Molecular Pharmacology, University of California San Francisco, CA, USA  
42 <sup>22</sup> Department of Chemistry and Chemical Biology, University of California Merced, CA, USA  
43 <sup>23</sup> Department of Biochemistry, University of Johannesburg, Auckland Park, 2006, South Africa

44  
45 † These authors contributed equally

46 ‡ Full listing at the end of the manuscript

47 \* To whom correspondence should be addressed; Email: [ivan.ahel@path.ox.ac.uk](mailto:ivan.ahel@path.ox.ac.uk), [jfraser@fraserlab.com](mailto:jfraser@fraserlab.com),  
48 [bshoichet@gmail.com](mailto:bshoichet@gmail.com), [frank.von-delft@diamond.ac.uk](mailto:frank.von-delft@diamond.ac.uk)

49  
50

1 **Key Words:** COVID-19, SARS-CoV-2, coronavirus, ADP-ribose, macrodomain, non-structural protein 3  
2 (Nsp3), fragment-based drug discovery (FBDD), docking, virtual screening

### 3 ABSTRACT

4 The SARS-CoV-2 macrodomain (Mac1) within the non-structural protein 3 (Nsp3) counteracts host-mediated  
5 antiviral ADP-ribosylation signalling. This enzyme is a promising antiviral target because catalytic mutations  
6 render viruses non-pathogenic. Here, we report a massive crystallographic screening and computational  
7 docking effort, identifying new chemical matter primarily targeting the active site of the macrodomain.  
8 Crystallographic screening of diverse fragment libraries resulted in 214 unique macrodomain-binding  
9 fragments, out of 2,683 screened. An additional 60 molecules were selected from docking over 20 million  
10 fragments, of which 20 were crystallographically confirmed. X-ray data collection to ultra-high resolution and  
11 at physiological temperature enabled assessment of the conformational heterogeneity around the active site.  
12 Several crystallographic and docking fragment hits were validated for solution binding using three biophysical  
13 techniques (DSF, HTRF, ITC). Overall, the 234 fragment structures presented explore a wide range of  
14 chemotypes and provide starting points for development of potent SARS-CoV-2 macrodomain inhibitors.  
15

### 16 INTRODUCTION

17 Macrod domains are conserved protein domains found in all kingdoms of life and certain viruses (1). Viral  
18 macrodomains have been found to recognize and remove host-derived ADP-ribosylation, a post-translational  
19 modification of host and pathogen proteins (2–4). The innate immune response involves signalling by ADP-  
20 ribosylation, which contributes to the suppression of viral replication (4–9). Upon viral infection, ADP-  
21 ribosylation is catalyzed by an interferon-induced subset of mammalian ADP-ribosyltransferases (ARTs),  
22 collectively termed ‘antiviral poly(ADP-ribosyl) polymerases’ (PARPs) (4, 10). These enzymes transfer the  
23 ADP-ribose (‘ADPr’) moiety of NAD<sup>+</sup> onto target proteins (4, 10). For example, during coronavirus infection,  
24 PARP14 stimulates interleukin 4 (IL-4)-dependent transcription, which leads to the production of pro-  
25 inflammatory, antiviral cytokines (11). Viral macrodomains, which are found primarily in corona-, alpha-, rubi-  
26 and herpes-viruses, can counteract this host defense mechanism via their (ADP-ribosyl)hydrolase activity,  
27 contributing to the host-viral arms race for control of cell signalling (12).  
28

29 Coronaviruses (CoVs) are important pathogens of livestock and humans. Three strains out of seven known  
30 to infect humans have caused major outbreaks within the last two decades: the severe acute respiratory  
31 syndrome (SARS) coronavirus, causing the SARS epidemic from 2002-2004, the Middle East respiratory  
32 syndrome (MERS) coronavirus, causing outbreaks in 2012, 2015 and 2018, and SARS-CoV-2, causing the  
33 current COVID-19 pandemic (13, 14). The coronaviral conserved macrodomain (called ‘Mac1’ here; also  
34 known as ‘S2-MacroD’ or ‘X domain’) is encoded as part of the non-structural protein 3 (Nsp3), a 200 kDa  
35 multi-domain protein (15, 16). While cell culture experiments suggest that SARS Mac1 is dispensable for  
36 viral replication in some cell lines (6, 7, 17, 18), animal studies have shown that its hydrolytic activity promotes  
37 immune evasion and that it is essential for viral replication and pathogenicity in the host (8, 9). The critical  
38 role of macrodomains is further supported by experiments using catalytic null mutations of the murine  
39 hepatitis virus (MHV), which render that virus essentially non-pathogenic (6–8, 19). Collectively, these  
40 findings lend strong support to the idea that SARS-CoV-2 Mac1 is a promising drug target for disrupting the  
41 viral life-cycle.  
42

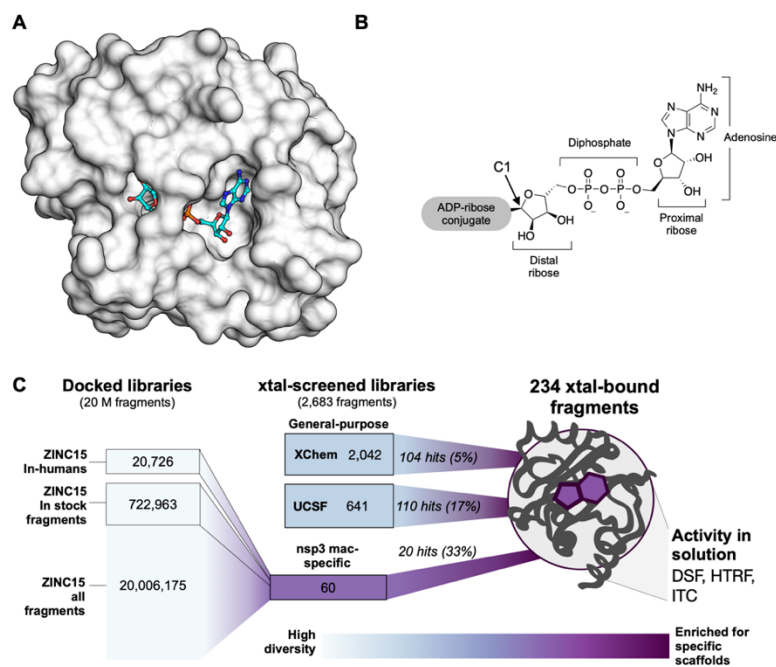
43 A barrier for macrodomain inhibitor discovery has been the lack of well-behaved inhibitors for this domain.  
44 Making matters worse, there are few biochemical assays suitable for screening for such inhibitors. Thus far,  
45 PDD00017273, an inhibitor of the poly(ADP-ribose)glycohydrolase (PARG), a macrodomain-type (ADP-

1 ribosyl)hydrolase, remains the only well-characterized inhibitor with convincing on-target pharmacology and  
2 selectivity (20). The initial hit was discovered by a homogeneous time-resolved fluorescence (HTRF)-based  
3 assay which measures PARG activity, and renders the assay unsuitable for macrodomains that lack this  
4 activity (21). A selective allosteric inhibitor targeting PARP14 was identified in an AlphaScreen-based high-  
5 throughput screen (HTS) (22). While this inhibitor showed on-target activity in cells, its unique allosteric  
6 binding site is difficult to translate to other macrodomains. While potential Mac1 inhibitors have emerged with  
7 the advent of SARS-CoV-2 (23–27), their binding mechanisms and efficacy remain unclear. The lack of a  
8 biochemical assay specific for Mac1 has hindered their development. Furthermore, structures of the new  
9 inhibitors bound to Mac1 have not yet been reported, making optimization of initial hits, however promising,  
10 difficult.

11  
12 To address the lack of chemical matter against Mac1, we turned to fragment-based ligand discovery using  
13 crystallography as a primary readout (**Fig. 1**). Fragment screens can efficiently address a large and relatively  
14 unbiased chemical space (28). Despite having generally weak overall binding affinity, fragments often have  
15 high ligand efficiency, and can provide templates for further chemical elaboration into lead-like molecules  
16 (29, 30). Crystallography can be used as a primary screening method for fragment discovery (31), and recent  
17 automation and processing software at synchrotron radiation sources has made this routinely possible at  
18 facilities like the XChem platform at Diamond Light Source (32–37). As part of Diamond’s contribution toward  
19 efforts to combat COVID-19, fragment screening expertise and infrastructure was made immediately  
20 available to any users working on SARS-CoV-2 targets (38). Similarly, synchrotron access for essential  
21 COVID-19-related research was also made available at the US Department of Energy light sources.

22  
23 Because crystallographic fragment screens can generate hits that bind anywhere on the protein surface, we  
24 wanted to supplement those screens with molecular docking intentionally targeting the active site. Docking  
25 has the additional benefit of exploring a much larger chemical space than an empirical fragment library. While  
26 an empirical library of ~1000-to-2000 fragments can represent a chemical space as large as, or larger, than  
27 that of a classic HTS library of several million compounds, exploration of chemotypes, including those that  
28 are well-suited to a particular target subsite, will inevitably be limited (39). Conversely, docking a much larger  
29 virtual library allows finer grained sampling around many chemotypes. A potential drawback of molecular  
30 docking is doubt about its ability to predict weakly-binding fragment geometries with high fidelity. While  
31 docking has identified potent ligands from libraries of lead-like molecules (250 to 350 amu) (40–46), such  
32 molecules offer more functional group handles for protein matching than do most fragments (150 to 250  
33 amu), and docking is thought to struggle with the smaller, less complex, and geometrically more promiscuous  
34 fragments (47, 48). Thus, the pragmatism of this approach has been uncertain (49, 50).

35  
36 Here, we present a combination of experimental crystallographic-based and computational docking-based  
37 fragment screens performed against Nsp3 Mac1 of SARS-CoV-2 (**Fig. 1**). Using X-ray crystallography, we  
38 screened fragment libraries of 2,683 compounds, yielding 214 unique fragment-bound Mac1 structures at  
39 atomic resolution. Docking of more than 20 million compounds prioritized 60 molecules for structure  
40 determination, yielding the structures of 20 additional compounds bound to Mac1. Additional X-ray data  
41 collection to ultra-high resolution and at physiological temperature illuminated the conformational  
42 heterogeneity in the Mac1 active site. We were able to confirm the binding of several fragments with  
43 differential scanning fluorimetry (DSF), isothermal titration calorimetry (ITC), and an ADPr-peptide  
44 displacement assay (HTRF), validating the activity of these molecules and providing a foundation for their  
45 optimization. The new fragments explore a wide range of chemotypes that interact with the catalytic site of  
46 the macrodomain. Together, these results create a roadmap for inhibitor development against Mac1, which  
47 may help to combat the pathogenicity of SARS-CoV-2.



## Figure 1. Overview of the fragment discovery approach for SARS-CoV-2 Nsp3 Mac1 presented in this study.

**A)** Surface representation of Nsp3 Mac1 with ADP-ribose bound (cyan) in a deep and open binding cleft. **B)** Nsp3 Mac1 possesses ADP-ribosylhydrolase activity which removes ADP-ribosylation modifications attached to host and pathogen targets. ADP-ribose is conjugated through C1 of the distal ribose. **C)** Summary of the fragment discovery campaign presented in this work. Three fragment libraries were screened by crystallography: two general-purpose (XChem and UCSF), and a third bespoke library of 60 compounds, curated for Mac1 by molecular docking of over 20M fragments. Crystallographic studies identified 214 unique fragments binding to Mac1, while the molecular docking effort yielded in 20 crystallographically confirmed hits. Several crystallographic and docking fragments were validated by ITC, DSF, and an HTRF-based ADPr-peptide displacement assay.

## 1 RESULTS

### 2 Two crystal forms of Nsp3 Mac1 reveal differences in active site accessibility

3 We sought a crystal system that enabled consistent ligand soaking for fragment screening and for testing  
4 docking predictions. Six Mac1 crystal forms have previously been reported (**Supplementary data set 1**).  
5 Initially, we designed a construct based on PDB entry 6VXS (25). This construct has been reported to  
6 crystallize in P1, C2 and P2<sub>1</sub> with either 1 or 2 molecules in the asymmetric unit (ASU) (**Supplementary data**  
7 **set 1**). This construct crystallized reproducibly in C2 with microseeding and diffracted to a maximum  
8 resolution of 0.77 Å (**Supplementary data set 1, Fig. S1, Fig. S2A**). This high resolution data yielded  
9 electron density maps at true atomic resolution with abundant alternative conformations (**Fig. S1**). The  
10 electron density maps also revealed features that are rarely observed in macromolecular crystallography,  
11 such as explicit hydrogen atoms, and covalent bond density (**Fig. S1**). Although the active site appears  
12 accessible (**Fig. S3**), efforts to soak ADP-ribose into the crystals were unsuccessful. Additionally, soaking  
13 revealed that this crystal form suffers from inconsistent DMSO tolerance (**Fig. S2**), which is problematic for  
14 fragment soaking. In attempts to overcome this problem, we experimented with lysine methylation (51), which  
15 unfortunately increases the occlusion of the active site, and dehydration, which increased DMSO tolerance  
16 at the cost of non-isomorphism (**Fig. S2**).

17

18 In parallel, we designed a new Mac1 construct that crystallized in P4<sub>3</sub> with two molecules in the ASU  
19 (**Supplementary data set 1**). This construct crystallized reproducibly with microseeding and diffracted to a  
20 maximum resolution of 0.85 Å (**Supplementary data set 1**). The sequence differences between the two  
21 constructs were slight (**Supplementary data set 1**), yet resulted in substantially different crystal packing  
22 (**Fig. S3**). Although the active site of protomer B was obstructed, the active site of protomer A was accessible  
23 (**Fig. S3**), and we were able to soak ADP-ribose into the crystals (**Fig. S4B,C**). This new structure also  
24 revealed a notable difference compared to previously reported Mac1-ADPr structures: the  $\alpha$ -anomer of the  
25 terminal ribose was observed instead of the  $\beta$ -anomer (**Fig. S4B-E**). Despite this, alignment of ADP-ribose



1 is excellent between all Mac1-ADPr structures (**Fig. S4A**). The DMSO tolerance of the P4<sub>3</sub> crystals was  
2 excellent (**Fig. S2**). Accordingly, most of our fragment soaking work proceeded with this construct.

3 Identifying new ligands for Nsp3 Mac1 using crystallographic fragment screening and docking

#### 4 *Characterization of experimental and virtual screening libraries*

5 Crystal soaking screens at the XChem facility were performed with the P4<sub>3</sub> crystals and a collection of  
6 fragment libraries (e.g. Diamond, SGC and iNEXT (DSI)-poised Library including 768 molecules (34) and the  
7 EU Open screen containing 969 molecules (52)) accumulating to 2,126 molecules (**Fig. S5, Supplementary**  
8 **data set 1**). Crystals were screened at the Diamond Light Source. At UCSF, a fragment library composed of  
9 Enamine's Essential Fragment library with 320 compounds, augmented by an additional 91 molecules from  
10 an in-house library (UCSF\_91), was screened against both the P4<sub>3</sub> and C2 crystal forms at the Advanced  
11 Light Source (ALS), the Stanford Synchrotron Radiation Lightsource (SSRL) and the National Synchrotron  
12 Light Source-II (NLS-II). On average, molecules across the X-Chem and UCSF collections had molecular  
13 weights of 192 ± 47 amu, cLogP -1.8 - 3.8, 13 ± 3 heavy atoms, and on average 2 rotatable bonds (**Fig. S5**).  
14

15 Two fragment libraries were computationally docked against the structure of the macrodomain (PDB 6W02):  
16 a library of 722,963 fragments "in-stock" at commercial vendors, and the entire ZINC15 fragment library of  
17 20,006,175 mainly make-on-demand fragments that have not been previously synthesized, but can readily  
18 be made, available predominantly from Enamine and Wuxi (25). Molecules from the ZINC15 fragment library  
19 had molecular weights ≤ 250 amu, cLogP ≤ 3.5, with an average of 4 rotatable bonds, and typically 4-19  
20 heavy atoms (**Fig. S5**). In addition, an "in-human" library of 20,726 drugs also comprising investigational new  
21 drugs, and metabolites beyond the fragment chemical space, that have been tested in humans were included  
22 into the docking screen, with a view to potential repurposing opportunities. All three sets can be downloaded  
23 from ZINC15 (<http://zinc15.docking.org>) (53).  
24

25 We investigated the fragment libraries for their diversity and their representation of chemotypes likely to bind  
26 at the adenine recognition site of the macrodomain (**Fig. S5**). Bemis-Murcko (BM) scaffold analysis revealed  
27 179 unique scaffolds in the UCSF libraries, and 809 distinct scaffolds in the XChem fragment libraries. The  
28 in-stock fragment docking library contained 69,244 distinct scaffolds, while 803,333 scaffolds were present  
29 in the entire ZINC15 20M fragment collection. Taken together, the experimentally screened libraries  
30 contained roughly two compounds per BM scaffold, while the docking libraries contained approximately 10  
31 fragments per scaffold, consistent with the expected higher granularity of the docking libraries afforded by  
32 their much larger size.  
33

34 Since adenine-containing compounds are the only structurally characterized binders of Mac1 and fragment  
35 libraries are intended to cover a wide chemotype space, we assessed the prevalence of pyrimidines in the  
36 libraries. We observed pyrimidines in 12 of the 411 fragments in the UCSF libraries and 72 of the 2,126  
37 XChem fragments (3.39% of the physically-screened fragments (**Fig. S5**). Pyrimidines were found in 41,531  
38 of the 722,963 (5.74%) in-stock fragments and in 890,199 molecules of the 20,006,175 compound fragment  
39 library (4.44%). While the percentages of molecules carrying the pyrimidine substructure were similar  
40 between the physical and docked fragments, the absolute numbers in the latter sets were far higher. Aside  
41 from bearing a pyrimidine substructure, they were otherwise diverse: among the 890,199 pyrimidine-  
42 containing docking fragments, 60,919 distinct BM scaffolds were identified. Adenine was present in 5,457  
43 fragments (582 different scaffolds). Furthermore, as ADP-ribose is negatively charged, anionic compounds  
44 were considered to exhibit favorable properties to bind to the macrodomain by targeting the diphosphate  
45 region. Fortuitously, a substantial fraction (35%) of the UCSF fragment libraries is anionic (**Fig. S5**).

## 1 *Hit rates and Mac1 interaction sites of fragments*

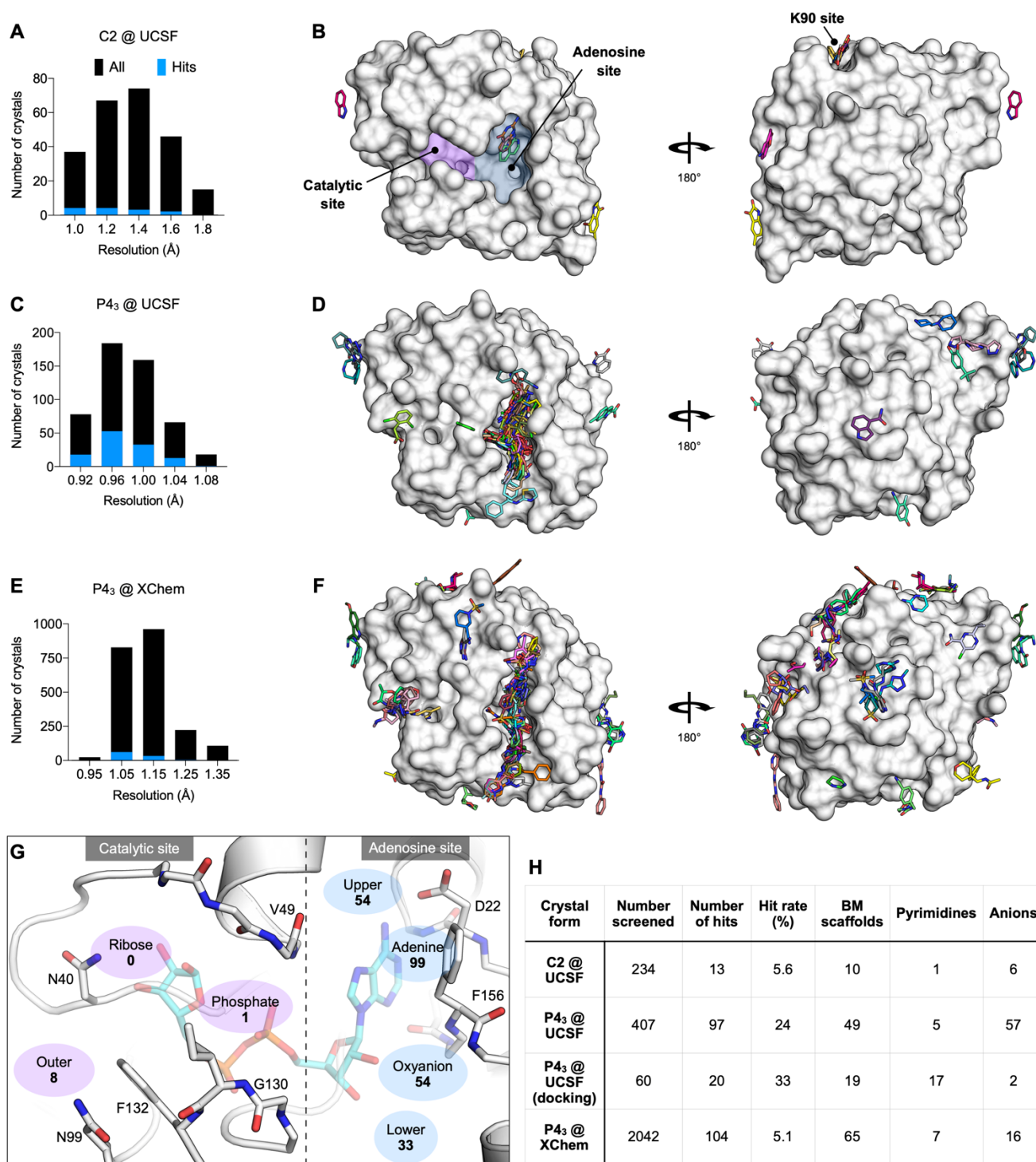
2 Across both crystal forms and facilities, we collected diffraction data for Mac1 crystals soaked with 2,850  
3 fragments (**Fig. 2**). The diffraction characteristics of the P4<sub>3</sub> crystals were excellent: the average resolution  
4 was 1.1 Å, and 98% of crystals diffracted beyond 1.35 Å (**Fig. 2C,E**). Although diffraction data was collected  
5 for 368 fragments soaked into the C2 crystals at UCSF, data pathologies meant that only 234 datasets could  
6 be analysed. The datasets collected from C2 crystals had a mean resolution of 1.4 Å and ranged from 1.0 to  
7 1.8 Å (**Fig. 2A**). In total, we identified 234 unique fragments binding to Mac1 using the PanDDA method  
8 (**Fig. 2, Supplementary data set 2**) (54). Of these, 221 were identified using P4<sub>3</sub> crystals (hit rate of 8.8%)  
9 and 13 using C2 crystals (hit rate of 5.6%). 80% of the fragments were identified in the Mac1 active site, near  
10 to or overlapping with the regions occupied by the nucleotide (the adenosine site) or the phospho-ribose (the  
11 catalytic site) (**Fig. 2G**). Additional fragments were scattered across the surface of the enzyme, with an  
12 enrichment at a distal macrodomain-conserved pocket near lysine 90 (the 'K90 site', 14 fragments) and with  
13 many others stabilized by crystal contacts (**Fig. 2B,D,F, Fig. S6**). Coordinates, structure factors, and  
14 PanDDA electron density maps for all the fragments have been deposited in the Protein Data Bank (PDB)  
15 and are available through the Fragalysis webtool (<https://fragalysis.diamond.ac.uk>).

16  
17 The unusually high hit rate for the adenosine site in the P4<sub>3</sub> form with the Enamine Essential library (21%)  
18 was in contrast to the relatively low hit rate with this library with the C2 form (1.3%). Of the five pairs of  
19 fragments identified in both crystal forms, two pairs were identified in the adenine subsite in both crystal  
20 forms, two in the adenine site in P4<sub>3</sub> crystals but in the K90 site in C2 crystals, and the remaining pair bound  
21 to a surface site in the P4<sub>3</sub> crystals and in the K90 site in the C2 crystals (**Supplementary data set 1**).  
22 Additional paired high quality datasets were available for 54 fragments that were bound within the P4<sub>3</sub>  
23 crystals, but all showed no density for fragments in the C2 crystals (**Supplementary data set 1**). It is possible  
24 that competition for binding with the N-terminal residues may have contributed to the relatively low hit rate  
25 for the C2 form (**Fig. S3**).

## 26 *Docking hits mimic the adenine recognition pattern*

27 Docking the entire (20 million) ZINC15 fragment library, after careful calibration of docking parameters using  
28 different control calculations (see Methods) (53, 55), was completed in just under 5 hours of elapsed time on  
29 a 500 core cluster. The 20,006,175 fragments were sampled in over 4.4 trillion complexes. Top-ranked  
30 molecules were inspected for their ability to form hydrogen bonds similar to adenine (e.g. with the side chain  
31 of Asp22 and with the backbones of Ile23 and Phe156) while molecules with internal molecular strain or  
32 unsatisfied hydrogen bond donors were deprioritized. Ultimately, we selected 54 fragments from the entire  
33 ZINC15 fragment library screen, 9 of which were immediately available for purchase from Enamine and 33  
34 of 45 make-on-demand molecules were successfully synthesized *de novo*. Furthermore, 8 fragments were  
35 purchased from the ZINC15 in-stock fragment library screen, and an additional 10 compounds were sourced  
36 based on the 'in-human' library docking (**Supplementary Table 1**).

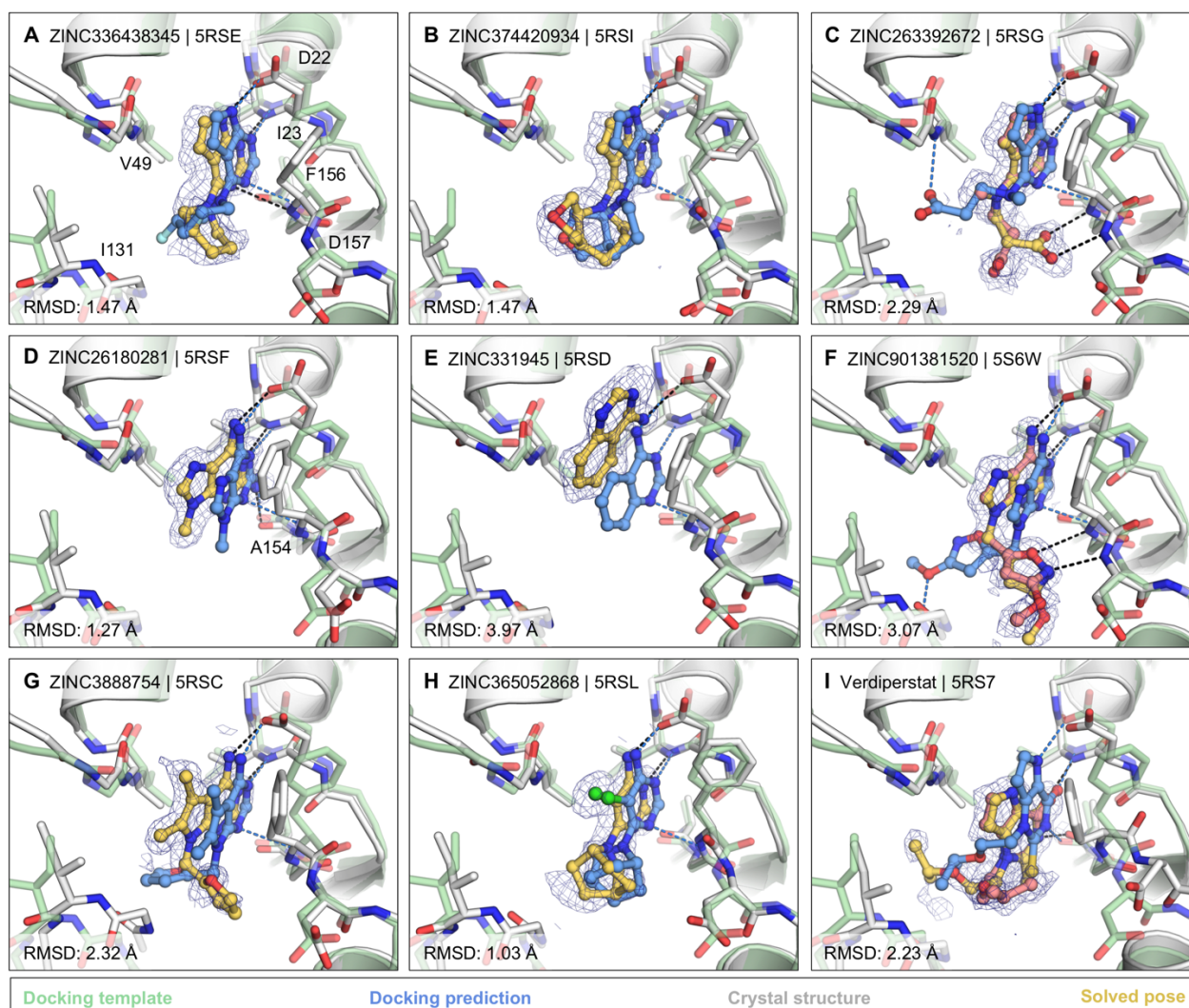
37  
38 Of the 60 molecules tested for complex formation by crystal soaking, 20 were observed with unambiguous  
39 electron density in complex with Mac1 (**Supplementary data set 1**). Here too, the crystals diffracted to  
40 exceptionally high resolution, between 0.94 and 1.01 Å. The predicted docking poses typically superposed  
41 well on the observed crystallographic results (Hungarian method root mean square deviations (56) ranging  
42 from 1-to-5 Å) and 19 out of the 20 docking hits bound to the adenosine sub-pocket of the Mac1, as targeted  
43 by docking (**Fig. 3 and Fig. S7**).



1 **Figure 2. Crystallographic screening identified 234 fragments bound to Mac1.** A,C,E) Histograms  
 2 showing the resolution of the crystallographic fragment screening data. The resolution of datasets where  
 3 fragments were identified are shown with blue bars. B,D,F) Surface representation of Mac1 with fragments  
 4 shown as sticks. G) The Mac1 active site can be divided based on the interactions made with ADP-ribose.  
 5 The 'catalytic' site recognizes the distal ribose and phosphate portion of the ADP-ribose, and harbours the  
 6 catalytic residue Asn40 (12). The 'adenosine' site recognizes adenine and the proximal ribose. The number  
 7 of fragments binding in each site is indicated. H) Summary of the fragments screened by X-ray  
 8 crystallography, including the number of Bemis-Murcko (BM) scaffolds and anionic fragments identified as  
 9 hits in each screen.



1 The most commonly observed scaffold among the docking hits was 7H-pyrrolo(2,3-d)pyrimidine occupying  
2 the adenine-binding pocket for ADP-ribose (**Fig. 3A-C**). This ring system is typically hydrogen bonded with  
3 Asp22, Ile23 and Phe156. Fragments with this scaffold revealed high fidelity between the docking results  
4 and the high resolution structures (RMSD 1.5 - 2.3 Å). Different substituents can be attached to this  
5 headgroup e.g. piperidine, adding a hydrophobic segment to the scaffold (e.g. ZINC336438345 (PDB  
6 5RSE)), occupying most of the adenosine binding pocket as shown in **Fig. 3A,B** and **Fig. S7A,B**. In addition  
7 to hydrophobic variations, ZINC263392672 (PDB 5RSG) attaches an anionic substituent to the  
8 pyrrolopyrimidine scaffold, offering additional hydrogen bonds within the binding pocket (**Fig. 3C**).  
9 Interestingly, while docking predicted the carboxylic acid of compound ZINC263392672 to insert into the  
10 phosphate binding tunnel, forming a hydrogen bond to Val49, the crystal structure instead revealed hydrogen  
11 bonds to the backbone amines of Phe156 and Asp157 which we defined as the 'oxyanion site' within the  
12 adenosine sub-pocket. Interactions with this backbone-defined oxyanion site were also observed for many  
13 other hits from both the docking and the crystallographically screened libraries (e.g. **Fig. 3F, Fig. S7E**).



14 **Figure 3. Docking hits confirmed by high-resolution crystal structures.** The protein structure (PDB  
15 6W02) (25), prepared for virtual screens is shown in green, predicted binding poses are shown in blue, the  
16 crystal protein structures are shown in grey, the solved fragment poses are shown in yellow, with alternative  
17 conformations shown in light pink. PanDDA event maps are shown as a blue mesh. Protein-ligand hydrogen  
18 bonds predicted by docking or observed in crystal structures are colored light blue or black, respectively.  
19 Hungarian RMSD values are presented between docked and crystallographically determined ligand poses  
20 (binding poses for additional docking hits are shown in **Fig. S7**).



1 For a set of smaller, mainly adenine-like docking hits, modeled to only occupy the adenine subpocket of the  
2 targeted adenosine binding site (**Fig. 3D,E, Fig. S7C,D**), the comparison between docked and experimental  
3 poses revealed deviations between 1.3 and 4 Å. Making these somewhat larger deviations harder to interpret  
4 was that for several fragments the crystallographically observed pose, e.g. ZINC331945 (RMSD 3.97 Å, **Fig.**  
5 **3E**) and ZINC763250 (RMSD 3.78 Å, **Fig. S7D**), is partially stabilized by interactions with the symmetry mate  
6 (*vide infra*).

7  
8 Another group of docking hits was selected as they mimicked the adenosine scaffold more closely (see **Fig.**  
9 **3F,G** and **Fig. S7I-L**). For these, the ultra-high resolution of the crystal structures was crucial, revealing that  
10 for four of these, initially the wrong purine isomer had been inadvertently synthesized, with alkyl derivatives  
11 from the N3 rather than the intended N9 nitrogen corresponding to the alkylation of adenine in ADP-ribose  
12 (**Fig. S7I-L**). Characterization of the original compound samples by HPLC/MS and NMR confirmed that the  
13 delivered compounds were >95% pure, mis-assigned positional isomers. For ZINC901381520 (**Fig. 3F**), both  
14 N3 (PDB 5RSK) and N9 (PDB 5S6W) isomers were synthesized in different batches and confirmed to bind  
15 to the targeted adenosine binding pocket forming equal hydrogen-bond interactions with the protein (**Fig.**  
16 **S7I**). ZINC3888754 (PDB 5RSC) (**Fig. 3G**) contains an adenine-like heterocycle extended by methyl-groups  
17 at the C7 and C8 positions, revealing opportunities for expanding purine scaffolds beyond the adenine  
18 subsite to achieve Mac1 selectivity over other adenine-binding proteins.

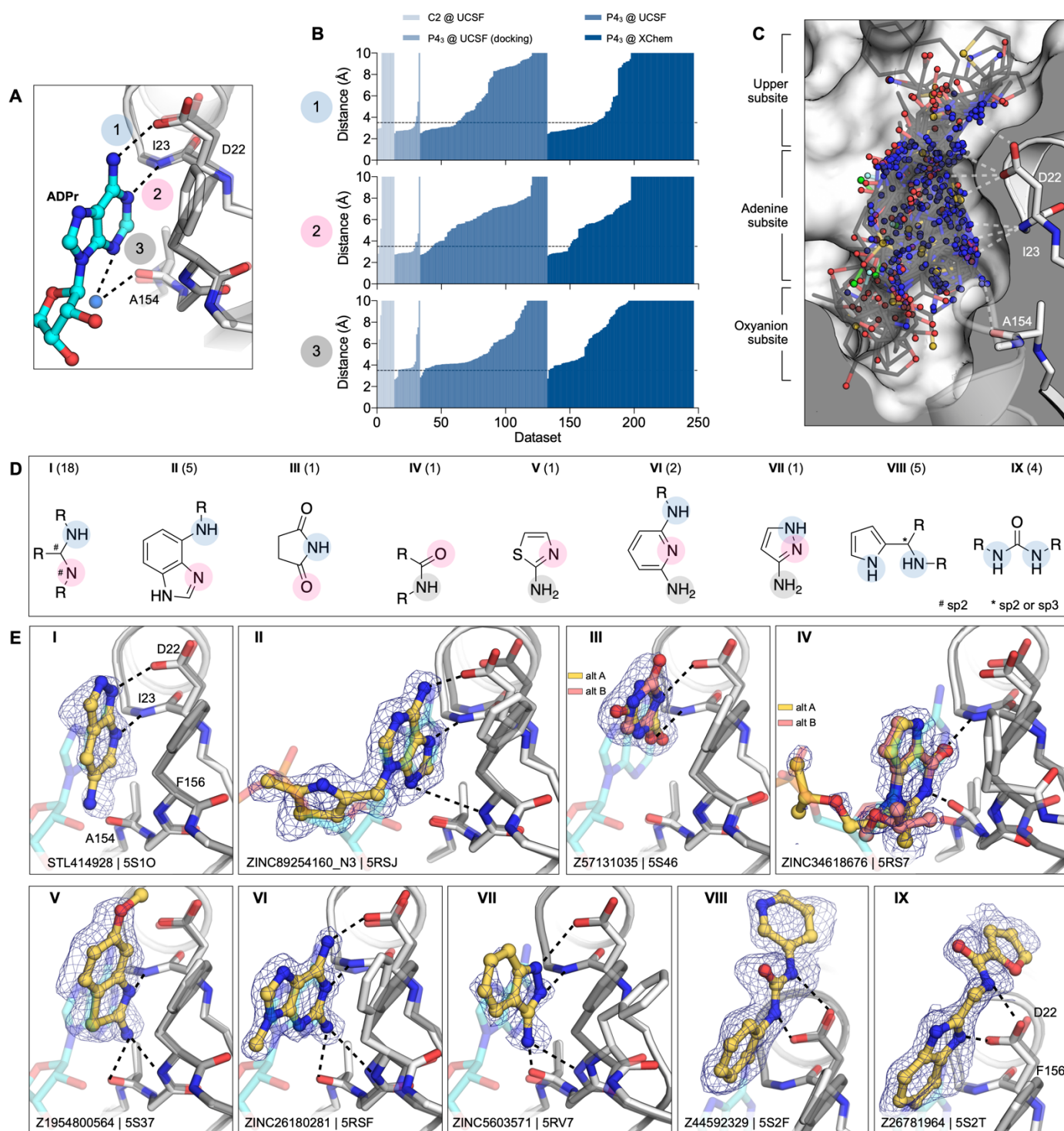
19  
20 In addition to hydrogen-bonding with residues involved in the adenine recognition of ADP-ribose, several  
21 docking hits hydrogen bond to the backbone carbonyl group of Ala154 (see **Fig. 3D,I** and **Fig. S7G**), thereby  
22 revealing an intriguing accessory polar contact within this sub-pocket. While most residues surrounding the  
23 adenosine-binding subpocket adopted similar conformations in the fragment-bound crystal structures as in  
24 the ADP-ribose-bound template protein structure used for docking (PDB 6W02) (25), Asp22 and Phe156 had  
25 increased dynamics in the electron density. In most fragment-bound crystal structures, Phe156 rotated by  
26 approximately 90° degrees, enabling improved  $\pi$ - $\pi$  stacking (face-to-face) against the aromatic moieties in  
27 the bound fragments, as shown in **Fig. 3C-G**. However, the docking template orientation of Phe156 was  
28 retained for other pyrimidine-containing fragment-bound crystal structures (**Fig. 3B,H**).

29  
30 Overall, two characteristics stand out from the docking screen: first, despite some important differences,  
31 there was high fidelity between the docking-predicted poses and those observed, to ultra-high resolution, by  
32 crystallography. The docking hits explored the adenine subsite to which they were targeted. Second, these  
33 hits did so with relatively dense variations around several chemotypes, something afforded by the granularity  
34 of a >20 million fragment library. This density can be explored further, for example, 9,170 fragments (888  
35 unique BM scaffolds) in the ZINC15 fragment library contained 7H-pyrrolo(2,3-d)pyrimidines, the functional  
36 group repeatedly observed in crystallographically confirmed docking hits (**Fig. 3A-C**).

37 Analysis of key interactions between Mac1 and fragments from the crystallographic screen

### 38 *Fragments binding to the adenine subsite*

39 While docking was successful in targeting the adenine binding pocket, crystallographic fragment screening  
40 has the advantages of being binding site agnostic and the potential to identify novel chemotypes at that site.  
41 Crystallographic screening identified 96 adenine-site binding fragments that form subsets of the three  
42 hydrogen bonds found between Mac1 and ADP-ribose (**Fig. 4A-C**). Fragments that formed at least two  
43 hydrogen bonds to the adenine subsite were separated into nine classes based on the number, nature and  
44 connectivity of atoms involved in such hydrogen bonding (**Fig. 4D**). The most common class consisted of a  
45 1,3-hydrogen bond donor/acceptor motif (**Fig. 4D,E,I**). This is reminiscent of the hinge binding motifs



1 **Figure 4. Fragments binding to the adenine subsite.** **A)** Stick representation showing the interaction of  
2 the adenosine moiety of ADP-ribose with Mac1. The key interactions are shown as dashed lines. **B)** Plot of  
3 the distances shown in **(A)** for all fragment hits. The distances, truncated to 10 Å, are for the closest non-  
4 non-carbon fragment atom. **C)** Stick representation showing all fragments interacting with Asp22-N, Ile23-N or  
5 Ala154-O. The surface is 'sliced' down a plane passing through Asp22. **D)** Structures of the nine unique  
6 motifs that make at least two hydrogen bonds to the adenine site. Colored circles match the interactions  
7 listed in **(A)** and **(B)**. The number of fragments identified for each motif are listed in parentheses. **E)** Example  
8 for the nine structural motifs. The fragment is shown with yellow sticks and the PanDDA event map is shown  
9 as a blue mesh. ADP-ribose is shown as cyan transparent sticks. The apo structure is shown with dark gray  
10 transparent sticks.

1 observed in kinases, with the difference being the engagement of a side chain oxygen rather than the  
2 backbone carbonyl oxygen (**Fig. S8B**) (57). While 7 out of 18 fragments in this class were 4-amino-  
3 pyrimidine derivatives, other moieties were featured, including two 2-amino-thiazole-based fragments and  
4 several purine derivatives (**Supplementary data set 2**). We also observed an unusual adenine-binding mode  
5 with an H-bond formed between Ile23 and N7 instead of N1 (**Fig. 4D,E.II**). The alternative binding mode can  
6 be explained by the N3 substitution of adenine on this fragment, which prevents formation of the canonical  
7 N1-Ile23 H-bond. This pattern of hydrogen bonds to the protein has not been previously observed in adenines  
8 linked through N9 (58).

9  
10 We also observed diverse fragments without adenine-like motifs binding at this site, including succinimides,  
11 amides, thiazoles, diamino-pyridines, pyrazoles, pyrroles, and ureas (**Fig. 4D,E.III-VIII**). These exploited,  
12 separately and together, Asp22 and Ile23, Ala154, and occasionally all three adenine-defining hydrogen-  
13 bonding residues. Several fragments  $\pi$ - $\pi$  stacked with Phe156, while those bearing a urea hydrogen-bonded  
14 with the carboxylate of Asp22 (**Fig. 4D,E.VIII**). These interactions were reproduced by a series of  
15 benzimidazole-based fragments (**Fig. 4D,E.IX**). These classes occupied what might be classified as an  
16 'upper subsite', above that defined by the adenine-ribose axis, and may provide an opportunity to grow  
17 fragments away from the canonical site.

#### 18 *Fragments binding to the oxyanion subsite*

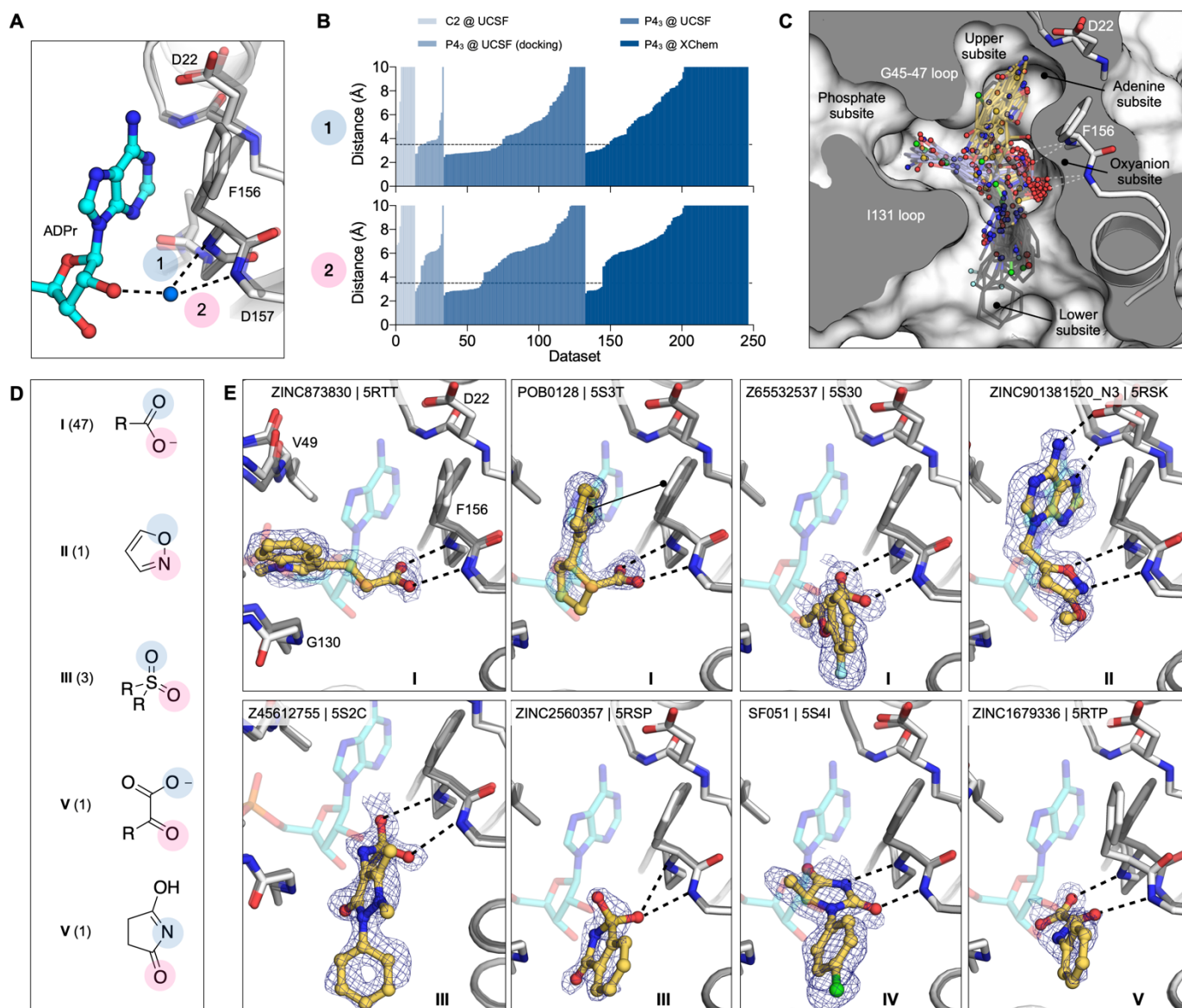
19 As was true of several of the docking hits, 54 fragments formed interactions with an unexpected 'oxyanion  
20 site', defined by the backbone nitrogens of Phe156 and Asp157 adjacent to the adenine site (**Fig. 2G**). As  
21 suggested by its name, most of these fragments (48/56) were anionic (**Supplementary data set 1**).  
22 Intriguingly, the defining backbone nitrogens adopted a similar orientation to those defining the classic  
23 oxyanion hole of serine hydrolases such as acetylcholinesterase and trypsin (**Fig. S8**). In the ADPr-Mac1  
24 structure, the C2 hydroxyl (2'OH) of the proximal ribose interacts with the oxyanion subsite via a bridging  
25 water (**Fig. 5A**). In total, 54 fragments formed at least one H-bond to the oxyanion subsite (**Fig. 5B**). Here  
26 too, the fragments were both geometrically (**Fig. 5C**) and chemically diverse (**Fig. 5D**): orienting groups either  
27 toward the phosphate tunnel, the lower site, or wrapped around toward the upper adenine site, providing  
28 multiple points of departure for further elaboration. Chemically, they interacted with the site using not only a  
29 carboxylate, but also sulfones, and isoxazole,  $\alpha$ -keto acid, and a succinimide (**Fig. 5E**). We suspect that the  
30 presence of the oxyanion subsite explains the higher hit rate for the Enamine Essential library versus the  
31 other crystallographic fragment libraries screened (27% versus 6%), as this library contained a greater  
32 proportion of acids than the others (41% versus 4%) (**Fig. S5**).

#### 33 *Fragments binding to the catalytic and other potential allosteric sites*

34 There were substantially fewer hits against the catalytic site (**Fig. 2G**) compared to the adenosine site (eight  
35 versus >100), though both appear to be accessible (**Fig. S3**). The catalytic site consists of three subsites:  
36 the phosphate tunnel, which is occupied by the diphosphate of the ADPr-diphosphate, the ribose site, which  
37 is occupied by the terminal ribose of the molecule, and the outer site, which sits between Asn40 and Asn99  
38 (**Fig. 6, 2G**). Of the eight fragments binding in the overall catalytic site, seven bound in the outer subsite and  
39 one bound in the phosphate tunnel. Binding to the outer site was often defined by hydrophobic packing  
40 between the Tyr42 and Lys102 side chains, although POB0135 (PDB 5S3W) and POB0128 (PDB 5S3T)  
41 formed a salt bridge to Lys102 (e.g. **Fig. 6A.I**). Interestingly, the latter fragment was also found to bind in the  
42 adenosine-binding site. Other molecules, including Z2234920345 (PDB 5S2L) and Z955123498 (PDB 5S4A)  
43 stabilize an alternative conformation of Lys102 (**Fig. 6A.II**). Three of the fragments, including Z85956652  
44 (PDB 5S2U), positioned a halogen atom in the outer subsite (**Fig. 6A.III**). The only fragment identified in the  
45 phosphate subsite was ZINC84843283 (PDB 5RVI). This fragment was wedged between the Gly47/Ile131

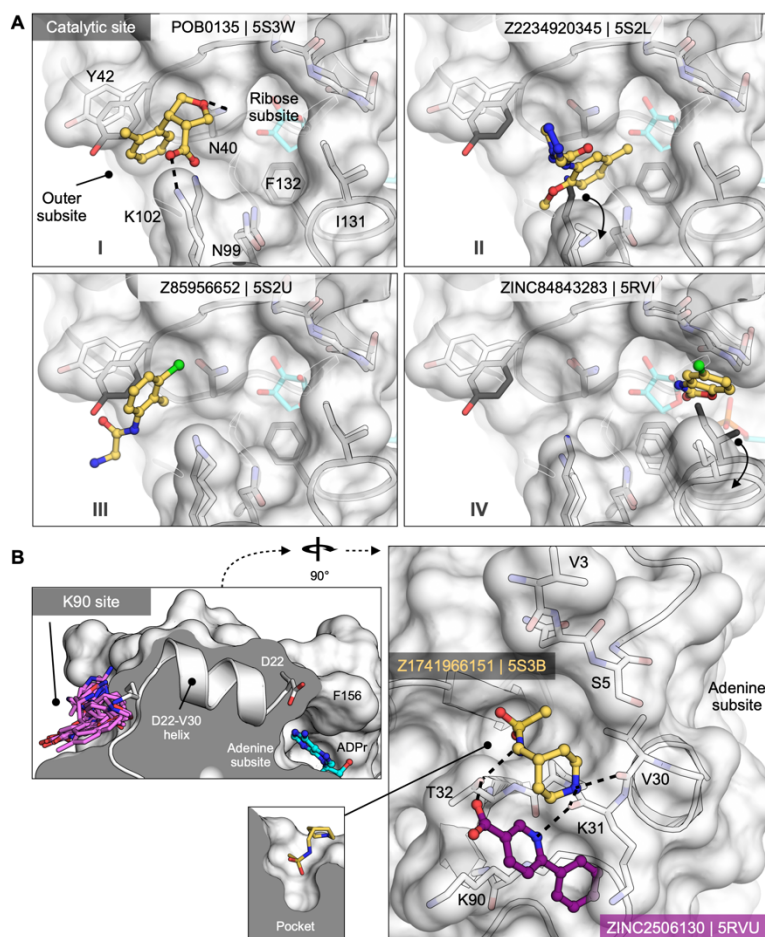


1 loops, and increased the gap between the two loops by 1.6 Å (**Fig. 6A.IV**). The absence of fragments binding  
 2 to the ribose subsite, and the sparsity of fragments in the phosphate subsite, means that designing a Mac1  
 3 inhibitor to occupy the catalytic site will rely more heavily on fragment growing than on fragment merging.



4 **Figure 5. Fragments binding to the oxyanion subsite of the adenosine site.** **A)** Stick representation  
 5 showing the interaction of the adenosine portion of ADP-ribose bound to Mac1. The water molecule bridging  
 6 the ribose moiety and the oxyanion site is shown as a blue sphere. **B)** Plot of the distances highlighted in **(A)**  
 7 for all fragment hits. Distances were calculated as described for Figure 4B. **C)** Stick representation showing  
 8 all fragments interacting with Phe156-N and Asp157-N. Fragments are colored by secondary binding site  
 9 with blue = phosphate, black = lower and yellow = adenine. The surface is “sliced” across a plane passing  
 10 through Phe156 (white surface, grey interior). **D)** Structures of the five unique motifs that bind the oxyanion  
 11 site. **E)** Example for the five structural motifs. Three examples for motif I are shown, where the fragment also  
 12 interacts with the phosphate, adenine or lower subsites. The fragment is shown with yellow sticks and the  
 13 PanDDA event map is shown for reference as a blue mesh. ADPr is shown with transparent cyan sticks. The  
 14 apo structure is shown with transparent gray sticks.





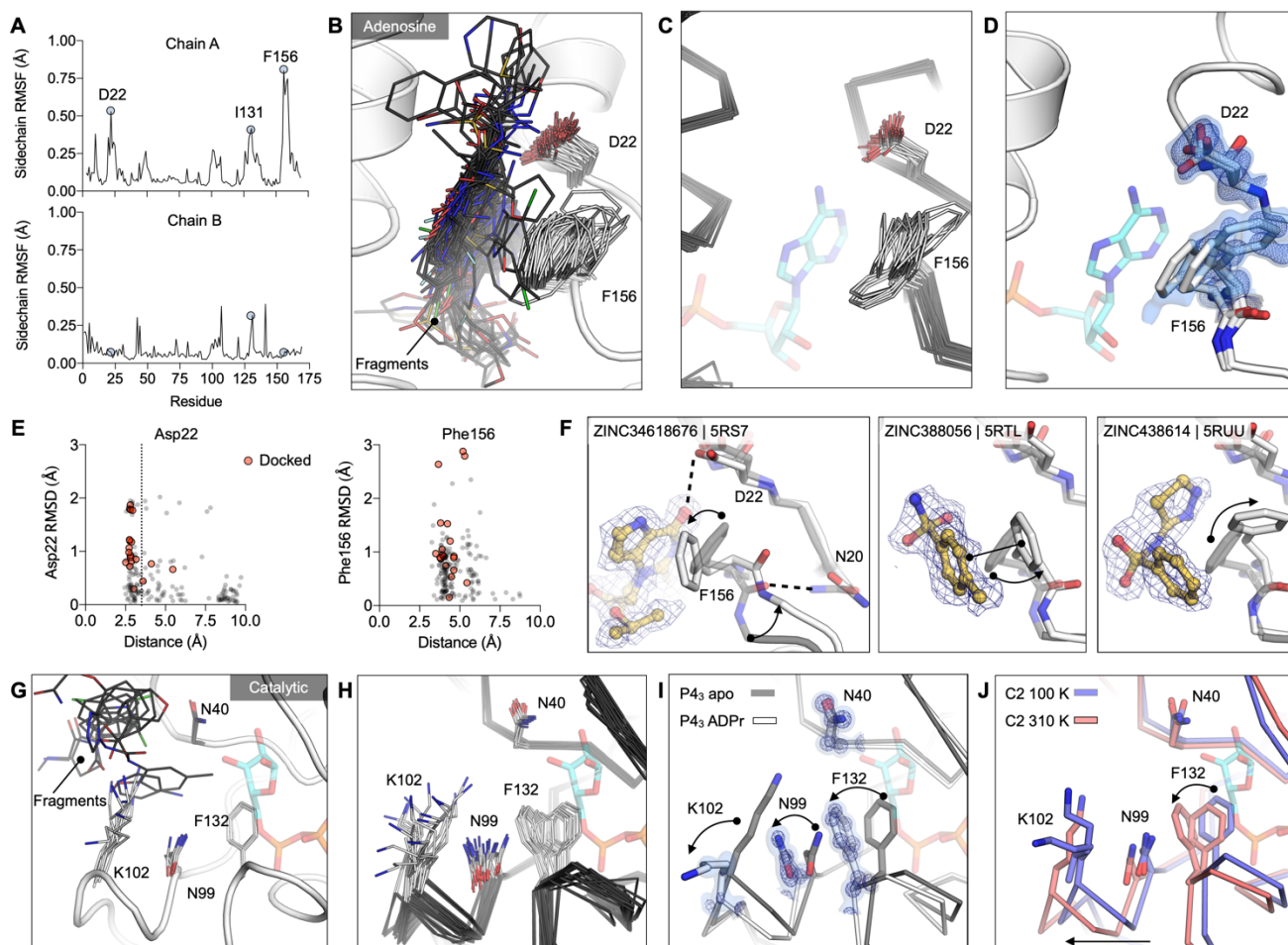
**Figure 6. Fragments targeting the catalytic and potential allosteric sites are sparsely populated compared to the adenosine site.**

**A)** Surface representation showing fragments that bind near the catalytic site. The fragment POB0135 (PDB 5S3W) bridges the gap between Asn40 and Lys102 via a hydrogen bond and a salt bridge, respectively. Although eight fragments bind in the outer site, the fragment POB0135 makes the highest quality interactions. No fragments bind in the ribose subsite. The fragment in ZINC331715 (PDB 5RVI) inserts into the phosphate subsite between Ile131 and Gly47. **B)** Left: the K90 site is connected to the adenosine site via the D22-V30 alpha-helix. Right: surface representation showing two fragments that bind to the K90 site. Hydrogen bonds are shown as dashed black lines. The fragment in Z1741966151 (PDB 5S3B) is partially inserted in a nearby pocket (insert).

3 Both crystallographic screens also identified fragments binding to the 'K90 site', which is formed by a cleft  
 4 between Lys31, Thr32 and Lys90 (**Fig. 6B**). We identified seven fragments from the C2 crystal form and six  
 5 from the P4<sub>3</sub> crystal form; intriguingly, none of the C2-derived fragments were found again when the UCSF  
 6 libraries were rescreened in the P4<sub>3</sub> crystal condition. Although the K90 site is 15 Å from the adenosine  
 7 subsite, it is connected to that subsite via a single alpha-helix (**Fig. 6B**). Although there is no biochemical  
 8 evidence for allosteric communication between these sites, the fragments provide starting points for  
 9 designing chemical probes to test this possibility.

#### 10 Fragment binding exploits protein conformational flexibility

11 To identify Mac1 flexibility associated with molecular recognition, we calculated the root-mean-square  
 12 fluctuation (RMSF) of side-chain atoms across the P4<sub>3</sub> fragment-bound structures. Residues lining the  
 13 adenosine site, especially Asp22 and Phe156, are the most flexible (**Fig. 7A,B**). The flexibility of both  
 14 residues is paralleled in previously reported crystal structures (**Fig. 7C, Fig. S4A**) and also in the 0.77 Å apo  
 15 structure, where multiple alternative conformations are clearly defined in electron density maps (**Fig. 7D** and  
 16 **Fig. S1A-C**). In the high resolution structure, residues 155-159 are modeled as a combination of two distinct  
 17 backbone conformations that diverge substantially at Phe156, which requires three distinct conformations of  
 18 this residue to satisfy the observed density (**Fig. 7D, Fig. S1C**). Despite this flexibility, hydrogen bonds to  
 19 Asp22 are present in many fragments, even docking compounds that were chosen based on interactions  
 20 with a static receptor. Similarly, the flexibility of the aromatic side-chain of Phe156 enables adaptable stacking  
 21 interactions with fragments (**Fig. 7F**), with 46 fragments binding within 4 Å to Phe156. As with Asp22, the  
 22 nature and geometry of these interactions are maintained for many soaked and docked fragments even as  
 23 the residue moves relative to the rest of the protein.



1 **Figure 7. Experimentally observed conformational heterogeneity is sampled by various fragments.**  
 2 **A)** Plots of side-chain RMSF for the 117 fragment structures from the UCSF screen using P4<sub>3</sub> crystals. **B)**  
 3 Stick representation showing all fragments (black sticks) within 3.5 Å of the Asp22 carboxylate and 4 Å of  
 4 of the Phe156 ring (white sticks). **C)** Structural heterogeneity in the previously reported Mac1 structures. **D)** The  
 5 Phe156 side-chain is captured in three conformations in C2 apo structure. Electron density maps (2mF<sub>O</sub>-  
 6 DF<sub>C</sub>) are contoured at 0.5 σ (blue surface) and 1 σ (blue mesh). For reference, ADP-ribose is shown with  
 7 blue sticks. **E)** Plots of side-chain RMSD for Asp22 and Phe156 from the Mac1 apo structure as a function of  
 8 of ligand-protein distance. Structures were aligned by their C<sub>α</sub> atoms, before RMSDs were calculated for the  
 9 Asp22 carboxylate and the Phe156 aromatic carbons. **F)** Fragment binding exploits preexisting  
 10 conformational heterogeneity in the Phe156 side-chain. The apo structure is shown with dark transparent  
 11 gray sticks in each panel and the conformational changes are annotated with arrows. **G)** Stick representation  
 12 showing all fragments (black sticks) in the outer subsite of the catalytic site. **H)** Conformational heterogeneity  
 13 of residues in the catalytic site of the previously reported Mac1 crystal structures. **I)** ADP-ribose binding  
 14 induces a coupled conformational change in the Phe132, Asn99 and Lys102 side-chain, as well as a 2 Å  
 15 shift in the Phe132 loop. Electron density maps (2mF<sub>O</sub>-DF<sub>C</sub>) are contoured at 1.5 σ (blue surface) and 4 σ  
 16 (blue mesh). **J)** Mac1 structures determined at 100 K and 310 K.

17  
 18 In contrast to the adenosine site, little conformational heterogeneity is observed at the catalytic site, with only  
 19 minimal changes in Lys102 and Tyr42 conformations (**Fig. 7G**). Still, even in this site, there is more  
 20 conformational heterogeneity observed in previously published structures (**Fig. 7H, Fig. S4A**). In particular,  
 21 a network of flexible side-chains encompassing Phe132, Asn99, and Lys102 is stabilized in a distinct  
 22 conformation upon ADP-ribose binding (**Fig. 7I**). To further probe the flexibility of the Phe132-Asn99-Lys102  
 23 network, we determined structures of Mac1 using the C2 crystal at human physiological temperature (37°C,  
 24 310 K) to 1.5 Å resolution (**Fig. 7J**). As observed in other systems (59, 60), we noticed that the cryogenic

1 structure appeared more compact than the structure at higher temperatures. Specifically, we observed  
2 substantial loop displacements near the ribose-binding pocket of the active site, which are coupled to a global  
3 hinge-bending motion involving correlated motion of helices about the central  $\beta$ -sheet (**Fig. S4G**). The  
4 structure at physiological temperature more closely resembles the structure with ADPr-bound, with the  
5 backbone adopting a more open conformation (**Fig. 7J**). However, the side-chain rotamers of Phe132, Asn99  
6 and Lys102 do not undergo the larger rearrangements. This temperature-dependent change in the width of  
7 the active site cleft can provide alternative, potentially more relevant, conformations for future ligand-  
8 discovery efforts targeting the catalytic site around the distal ribose.

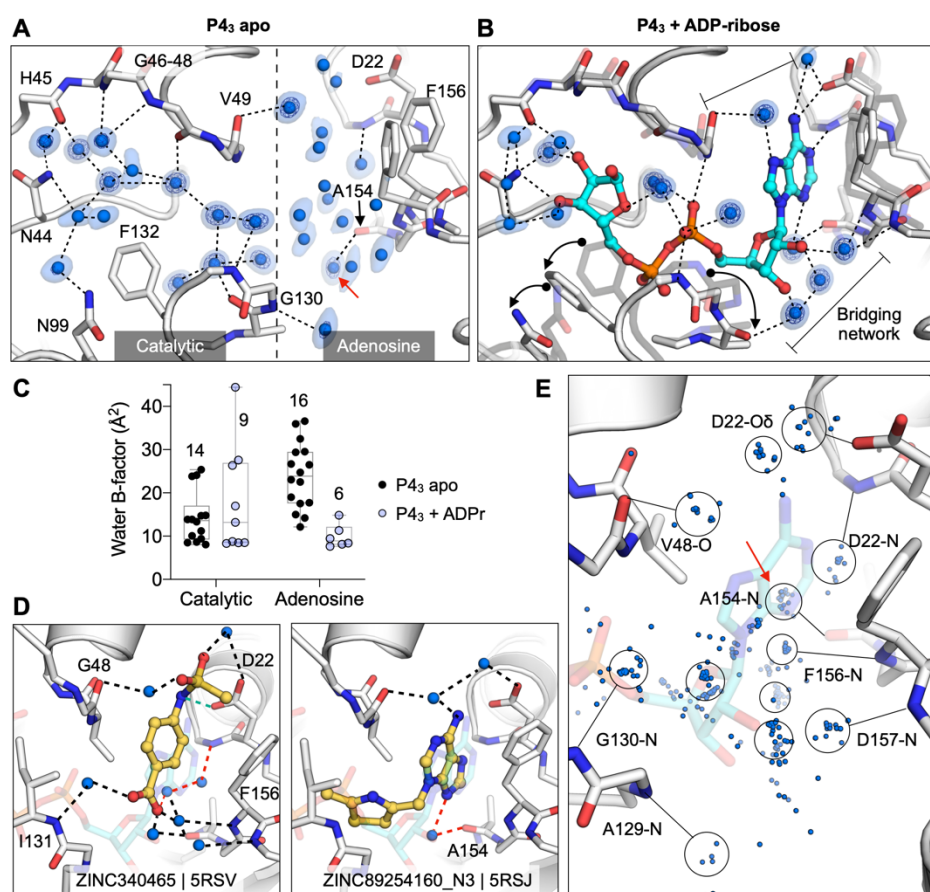
### 9 Changes in water networks upon fragment binding

10 To assess the role of water networks in fragment binding, we examined changes in water networks upon  
11 ADP-ribose binding. In the 0.85 Å P4<sub>3</sub> apo structure, the catalytic site contains 14 water molecules arranged  
12 in an ordered network that connects the Gly47 loop and the Ile131 loop, with an arc formed around the  
13 Phe132 side-chain (**Fig. 8A**). In contrast, waters were more disordered in the adenosine site, with more  
14 diffuse electron density and with waters modelled with higher B-factors (**Fig. 8E**). Upon ADP-ribose binding,  
15 six waters were displaced from the catalytic site and the water network was disrupted (**Fig. 8B**). This  
16 disruption is partly caused by altered conformation of the Phe132 and Asn99 side-chains, which break the  
17 network between residues Asn40 and Asn99. Conversely, the network in the adenosine site was stabilized  
18 in the Mac1-ADPr complex (**Fig. 8B**). The average B-factor decreased from 24 to 10, and two networks  
19 connect the phosphate tunnel with the adenine/oxyanion subsites (**Fig. 8C**). Although the adenine moiety  
20 only forms two direct hydrogen bonds to protein, it has four additional contacts via bridging water molecules  
21 (**Fig. 8B**). Similar bridging waters were observed for fragments binding in the adenosine site including  
22 ZINC2055 (PDB 5RSV), which forms only one direct H-bond to the protein, but has an extensive H-bond  
23 network via water molecules (**Figure 8D**). Visualizing all water molecules within 3.5 Å of fragment atoms  
24 shows clusters near protein hydrogen bond acceptors and donors (**Fig. 8E**). Of particular interest is the  
25 cluster near the backbone carbonyl of Ala154 (**Fig. 8**). This site is occupied by a water molecule in the ADPr-  
26 Mac1 structure, and is bridged by adenine derivatives such as ZINC340465 (PDB 5RSJ) (**Fig. 8D**). In  
27 addition, four fragments occupy this site directly (**Fig. 4E**), including the C2-amino-substituted adenine  
28 present in ZINC26180281 (PDB 5RSF, **Fig. 3D**). Extending fragments to displace the water molecules at  
29 other frequently populated sites could help to quantify the contribution of water networks to Mac1, and to  
30 provide a test set for computational methods that seek to exploit solvent dynamics for ligand optimization  
31 (61–63).

### 32 Characterization of in-solution binding of fragment hits

33 To buttress the crystallographic studies, selected compounds were biophysically screened using differential  
34 scanning fluorimetry (DSF) (64–66), isothermal titration calorimetry (ITC) (67–69), and a homogeneous time-  
35 resolved fluorescence (HTRF) ADPr-peptide displacement assay (**Fig. 9, Supplementary Table 1,**  
36 **Supplementary data set 2**). Because of their ready availability in useful amounts, most of these experiments  
37 focused on the docking hits. For DSF, in agreement with previous reports for this domain (24), we observed  
38 substantial elevation of the apparent melting temperature ( $T_{m_a}$ ) of Nsp3 Mac1 upon addition of its known  
39 binding partner, ADP-ribose (**Fig. 9C,D,G**). When tested in concentration-response from 0.188 to 3 mM, 10  
40 of 54 docked fragments also induced small, but statistically significant and dose-responsive  $T_{m_a}$  elevation  
41 (**Fig. 9C,D,G, Supplementary Table 1, Supplementary data set 2**). All 10 of these were also observed to  
42 bind in the crystallographic studies, providing relatively good agreement between these assays. However,  
43 the correlation was incomplete, as the remaining fragments observed in crystallography either decreased the  
44  $T_{m_a}$  or had no significant effect (**Supplementary Table 1**). Despite this, the results suggest that DSF is a  
45 rapid way to corroborate the fragment hits from crystallography and docking.





**Figure 8. Water networks in the active site are displaced as well as used by fragments for bridging interactions. A)** Water networks in the apo enzyme (P<sub>4</sub><sub>3</sub> crystal form). Waters are shown as blue spheres, with electron density contoured at 5  $\sigma$  (blue mesh) and 1.5  $\sigma$  (blue surface). Hydrogen bonds are shown as dashed lines (distances are 2.6-3 Å). **B)** Water networks in the Mac1-ADPr complex. ADP-ribose is shown as cyan sticks. Conformational changes upon ADP-ribose binding are highlighted with black arrows. **C)** Comparison of crystallographic B-factors of water molecules in the catalytic site and adenosine site. The range and 95% confidence interval are shown. **D)** Examples of the role of water networks in fragment binding. Right: ZINC340465 (PDB 5RSV) makes a single hydrogen bond to the protein (green dashed line), but makes five hydrogen bonds

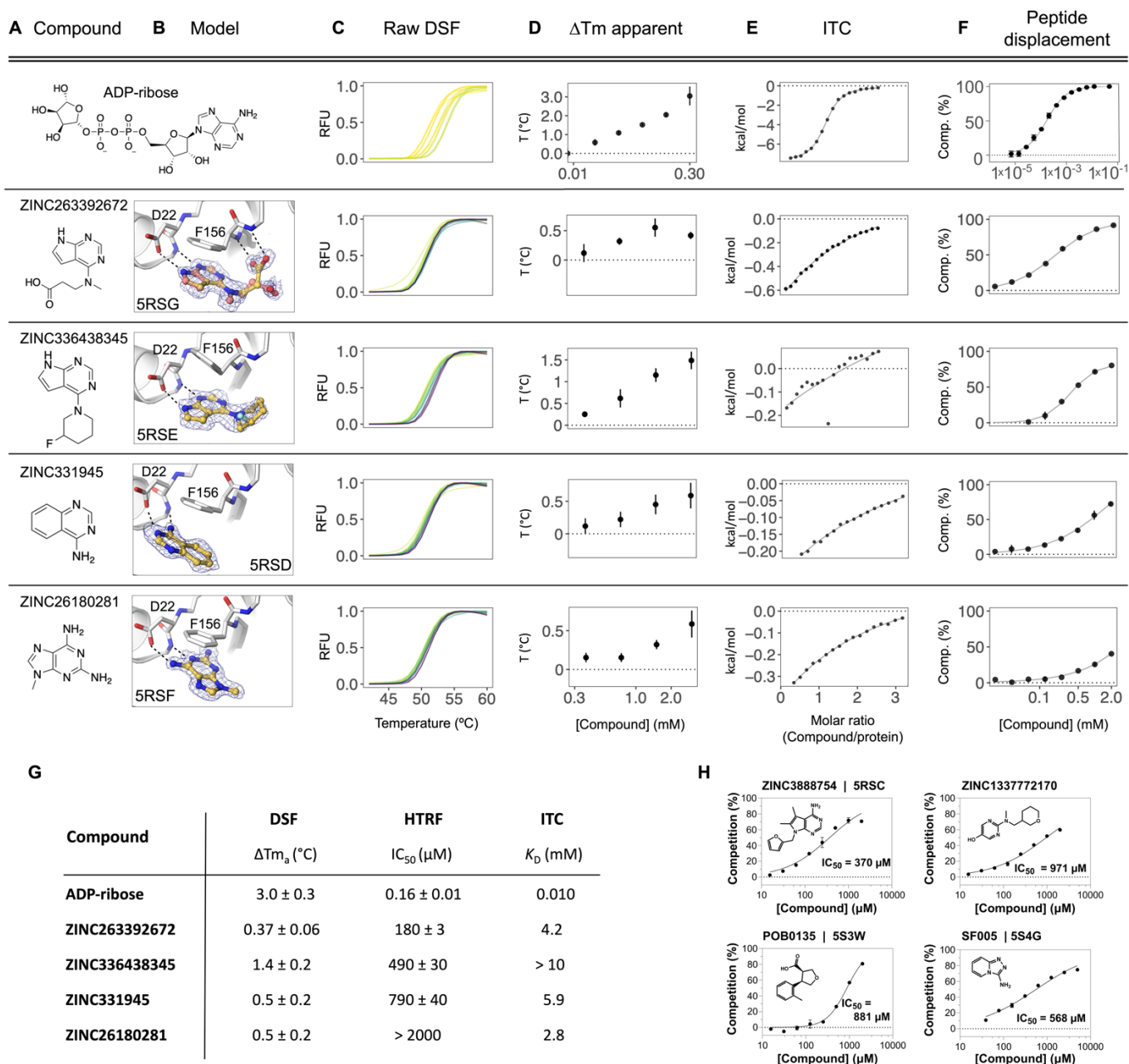
via water molecules. Although few fragments hydrogen bond directly to the backbone oxygen of Ala154 (Figure 4E), several fragments interact with this residue via bridging water molecules (red dashed line) including ZINC89254160\_N3 (PDB 5RSJ). **E)** Plot showing all water molecules that lie within 3.5 Å of a non-carbon fragment atom. Water molecules are shown as blue spheres, with the major clusters circled. The cluster in a red circle bridges fragments and the Ala154 backbone oxygen.

1  
2 To identify fragments with the most promising binding affinity for optimization, we tested the 19  
3 crystallographically observed docking hits using ITC. Due to their small size, most of these fragments have  
4 low binding affinity and release little heat upon binding versus ADP-ribose. Thus, we only observed reliable  
5 thermodynamic measurements for four of the 19 fragments. These could be fit to a 1:1 binding stoichiometry  
6 with affinities in the low mM range (Fig. 9E, Supplementary Table 1, Supplementary data set 2),  
7 consistent with the DSF results. Furthermore, the compounds measured by ITC that released the greatest  
8 amount of heat also induced the most significant T<sub>m</sub> shifts in DSF, highlighting the complementarity between  
9 these two biophysical assays.

10  
11 Finally, we tested 57 docking-derived fragments and 18 crystallographic hits from the XChem library in a  
12 HTRF-based peptide displacement assay, which monitors displacement of a fluorescently labeled ADPr-  
13 conjugated peptide from the active site of Mac1 by the fragment (Fig. 9F,G, Supplementary Table 1,  
14 Supplementary data set 2). Eight of 57 docking hits (14%) and three of 18 crystallographic hits (17%)  
15 inhibited the enzyme with IC<sub>50</sub> values generally between 200  $\mu$ M-1 mM, with the most potent fragment being  
16 the docking-derived ZINC263392672 (PDB 5RSG) with an IC<sub>50</sub> of 183  $\mu$ M in this assay. Only five  
17 (ZINC3888754 (PDB 5RSC), ZINC331945 (PDB 5RSD), ZINC263392672 (PDB 5RSG), ZINC336438345  
18 (PDB 5RSE) and ZINC6180281(PDB 5RSF), Fig. 3) of the 10 docking hits that stabilized the macrodomain



1 observed by DSF proved also inhibitory in the ADPr-peptide displacement assay. Interestingly, two docking  
 2 hits which were not identified as binders by DSF or crystallography, ZINC133772170 ( $IC_{50} = 971 \mu M$ ) and  
 3 pterin ( $IC_{50} = 790 \mu M$ ), were found to be macrodomain inhibitors in this assay (**Fig. 9H**). This result might be



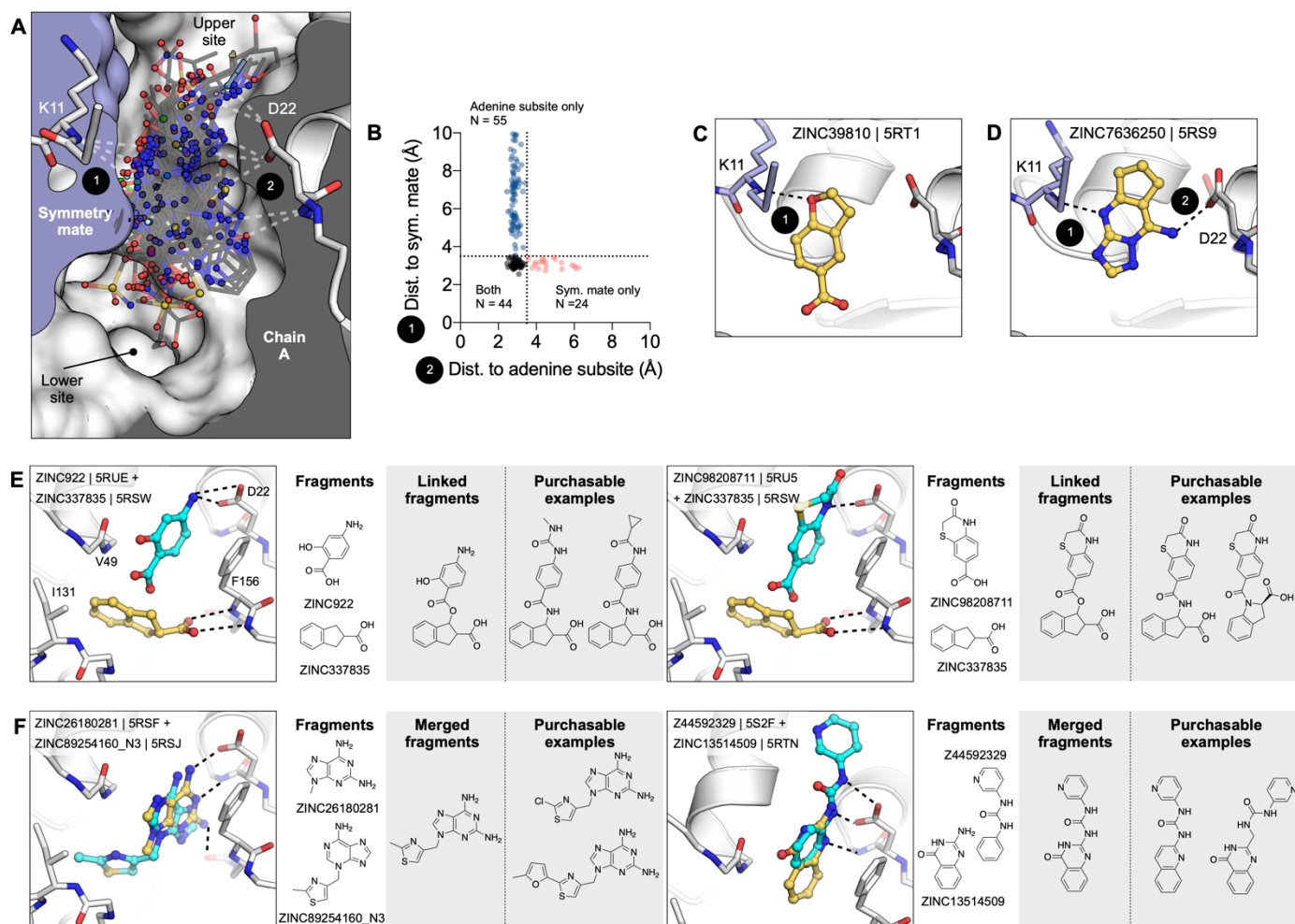
4 **Figure 9. Biophysical corroboration of solution binding of crystallographic fragment hits by DSF, ITC**  
 5 **and ADPr-peptide displacement assay.** Top panel (**A-F**) shows performance of the most potent fragment  
 6 hits in DSF, ITC, and ADPr-peptide displacement assay compared to ADP-ribose. **C,D**) Normalized raw DSF  
 7 RFU demonstrates canonical unfolding curves and minimal compound-associated curve shape aberrations.  
 8  $T_m$  elevation reveals Mac1 stabilization through fragment binding. Gradient color scale: 0 mM = yellow; 3  
 9 mM = purple. **E**) Integrated heat peaks measured by ITC as a function of binding site saturation. The black  
 10 line represents a non-linear least squares (NLLS) fit using a single-site binding model. **F**) Peptide  
 11 displacement assay measures ADPr-peptide displacement (i.e. % competition) from Mac1 by ligand. **G**)  
 12 Summary of solution binding data for fragments from top panel.  $\Delta T_m$  are given for the highest compound  
 13 concentration in this assay. **H**) Additional fragment hits showing Mac1 peptide competition.

1 explained with the use of a detergent-based buffer in the peptide displacement assay that could increase  
2 compound solubilization for what appear to be low-solubility compounds. With its ability to detect specific  
3 inhibition of the macrodomain, this ADPr-peptide displacement assay proved to be a sensitive and  
4 complementary strategy for further characterization of the fragment hits obtained from the docking and  
5 crystallographic screens. Assuming that the HTRF-based peptide displacement assay produced the most  
6 reliable affinity data, we estimated ligand efficiencies from  $IC_{50}$  values for hits for which we obtained  
7 reasonable dose-response curves. ADP-ribose with an  $IC_{50}$  of 161 nM and 36 heavy atoms indicates a ligand  
8 efficiency (LE) of 0.26 kcal/mol per non-hydrogen atom. The docking hits ZINC3888754 (PDB 5RSC,  
9 LE=0.26), ZINC336438345 (PDB 5RSE, LE=0.28), ZINC263392672 (PDB 5RSG, LE=0.32) and  
10 ZINC331945 (PDB 5RSD, LE=0.38) reveal similar or slightly improved ligand efficiencies, while the highest  
11 LE was calculated for the XChem library hit SF005 (PDB 5S4G, **Fig. 9H**) with 0.44 kcal/mol per heavy atom.  
12

13 In summary, all crystallographically confirmed docking hits were tested using three complementary in-  
14 solution binding techniques - DSF, ITC, and an HTRF-based peptide displacement assay (**Fig. S9**,  
15 **Supplementary Table 1, Supplementary data set 2**) (70). ZINC336438345 (PDB 5RSE), ZINC331945  
16 (5RSD), ZINC263392672 (5RSG) and ZINC26180281 (5RSF) were the only four fragment hits for which  
17 binding data could be obtained by all three techniques (**Fig. 9**). All of these fragments have key hydrogen  
18 bonds in the adenine site and display  $\pi$ - $\pi$  stacking with Phe156. Furthermore, ZINC263392672 (PDB 5RSG)  
19 can additionally interact via its carboxyl group with the oxyanion site of Mac1. Finally, we note that  
20 crystallography, DSF, ITC all monitor binding, but do not measure function. The peptide displacement assay  
21 is thus of particular value for fragment characterization, since it measures specific displacement of an analog  
22 of the natural Mac1 substrate.

### 23 Opportunities for fragment linking and merging to optimize Mac1 inhibitors

24 Typically, one might be reluctant to speculate on optimization from fragment structures alone, but the  
25 unusually large number of structures, perhaps supports some cautious inference here. Prior to modifying,  
26 linking, or merging fragments, it is important to consider the crystalline environment. In the macrodomain  
27 crystal, the active site forms a bipartite enclosed pocket with a symmetry mate (**Fig. 10A**). In particular, 24  
28 fragments only hydrogen bond to Lys11 of the symmetry mate, and do not hydrogen bond with any residues  
29 in the adenosine site, indicating that these molecules should not be considered for fragment elaboration  
30 (**Fig. 10C,D**). Based on the solved binding poses of remaining compounds, fragment pairs were linked into  
31 hypothetical scaffolds. These were used as templates to search the make-on-demand chemical space of the  
32 Enamine REAL database employing the smallworld similarity (<http://sw.docking.org>) and arthor substructure  
33 (<http://arthor.docking.org>) search engines (**Fig. 10E,F**) (71). In a second approach, fragments with  
34 overlapping binding poses were merged into larger scaffolds, e.g. the purine of ZINC89254160\_N3 (PDB  
35 5RSJ) interacting in the adenine binding pocket was replaced by ZINC26180281 (PDB 5RSF) adding an  
36 additional hydrogen bond to Ala154 (**Fig. 10F**). Whereas it remains speculative whether the suggested linked  
37 or merged molecules are indeed active against the macrodomain, the scaffolds observed here, and the key  
38 interactions they make with the enzyme, indicate a fruitful chemical space to further explore. Naturally, many  
39 of the fragments described here also merit investigation by alternative fragment growing or analoging  
40 strategies.



1 **Figure 10. Fragments bridging multiple adenosine subsites provide direct merging opportunities. A)**  
2 Sliced view of the adenosine site (white surface, grey interior) and a symmetry mate (blue surface and  
3 interior) showing the deep pocket created by crystal packing in the P4<sub>3</sub> crystals. The 66 fragments that H-  
4 bond with the Lys11 backbone nitrogen are shown as sticks. **B)** Plot showing distances between the  
5 symmetry mate (Lys11-N) and the adenosine subsite (Asp22-Oδ, Ile23-N, Ala154-O) for all fragments iden-  
6 tified in the adenosine site. Dashed lines show the 3.5 Å cut-off used to classify H-bonds. **C)** An example showing  
7 one of the 24 fragments that bound in the adenosine site, yet only formed an H-bond with the sym-  
8 metry mate. **D)** An example of one of the fragments that bridged the 9-11 Å gap between the adenosine subsite and  
9 the symmetry mate. **I, J)** Opportunities for fragment linking and merging. Adjacent or overlapping fragments  
10 were initially merged into a single new compound. Examples of readily available make-on-demand  
11 compounds are shown.

## 12 DISCUSSION

13 Three key observations emerge from this study. Most noteworthy is the sheer number and the unusually high  
14 resolution of the 234 fragment-bound Mac1 structures reported here, including 192 fragments identified in  
15 the active site. The fragments cover both stereotyped interactions (such as adenine-like hydrogen bonding  
16 to the Asp22 side chain/ Ile23 backbone and stacking interaction with Phe156) as well as chemical diversity  
17 exploiting flexibility in the active site (for instance targeting of the oxyanion-subsite). This abundance and  
18 diversity afford robust starting points for future elaboration into lead-like molecules. Second, the high fidelity  
19 of docked poses to the subsequent crystallographic results supports the use of docking to explore the  
20 adenine recognition site, and importantly, demonstrates an ability of docking to prioritize fragments in this  
21 case, an approach still debated in the field. Finally, with 234 diverse fragment structures determined, it should

1 be possible to exploit the fortuitous juxtaposition of pairs of fragments to design joined ligands that combine  
2 the affinities of both, leading to robust starting points with the low micromolar affinity needed for initiating hit-  
3 to-lead optimization of inhibitors. One clear strategy involves extending molecules bound to the adenine site  
4 and with biophysically measurable binding affinities into the phosphate and ribose recognition regions.

5  
6 In contrast to the large number of chemically diverse hits binding to the adenine site, the lack of fragments  
7 bound to the catalytic site is notable and may inform models of how the ADP-ribosylated peptide may bind  
8 to the macrodomain. This result also suggests that this is not likely to be a productive region for growing  
9 fragments which are binding in the adenine site. This paucity of fragments is especially surprising given that  
10 three crystal environments (the A and B chains in the P4<sub>3</sub> crystal and the C2 crystal) were screened and that  
11 the site appears accessible in all lattices (**Fig. S2**). The two major models for peptide-macrodomain  
12 interactions are either that the peptide binds along the widened cleft defined by Tyr42 and Lys102, or that it  
13 extends into solution through the flexible Gly46-48 loop (72). Indeed, we observe fragments that bind in both  
14 locations (**Fig. 6A**). Regardless of the binding mode, which could be distinct depending on the identity of the  
15 modified residue and target substrate, the lack of binding at this site suggests that the binding energy comes  
16 almost exclusively from the ADP-ribose and not from the amino acids on the ADPr-conjugated protein. This  
17 hypothesis is also supported by the fact that Mac1 can hydrolyse a wide range of ADP-ribosylated substrates  
18 (2, 3). Docking of larger 'lead-like' molecules, perhaps enabled by the expanded catalytic site revealed by  
19 the physiological temperature structure, and detailed description of solvent may help to identify molecules  
20 exploiting this site.

21  
22 The success of the fragment docking campaign contrasts, perhaps, with expectations of the field that  
23 fragments have too few functional-group handles to accurately dock or prioritize (73, 74). Not only were hit  
24 rates high (33%), so too was the fidelity of most docking poses to the crystallographic results. Even judged  
25 by potency, the most active fragment to emerge from this study, the 183  $\mu$ M inhibitor ZINC263392672 (PDB  
26 5RSG) (**Fig. 3 and 9**), was a docking hit. Also, it was the docking hits that were most readily available for  
27 such functional testing, as they were sourced as 10 mg of dry powder, while the crystallographic screening  
28 compounds were often in short supply. This is a purely mechanical advantage of docking, and it is  
29 counterbalanced by the small numbers tested versus the crystallographic screens; still, having substantial  
30 material to work with is a pragmatic advantage. Admittedly, weaknesses also emerged from the docking.  
31 Intriguingly, the oxyanion site that featured so prominently among the crystallographic screening hits were  
32 not to be found among the docking predictions. This gap reflects both a failure of the docking scoring function  
33 to prioritize anions binding to this site (as they were at least sampled), and to some extent a failure of the  
34 docking group to pick the few molecules that did dock well to this site as likely candidates. More broadly, as  
35 we docked against a single rigid structure of the protein, the subsequent conformational changes that the  
36 protein underwent, and the changes in the water network, were not captured in the docking predictions, and  
37 this was sometimes reflected in the larger RMSD differences between predicted and observed fragment  
38 poses. These caveats, important as they are, should not obscure a central observation from this study: the  
39 docking hit rate was not only high, but the hits were typically right for the right reasons; this may be something  
40 to build on for the field.

41  
42 From the docked compounds, the most promising hits identified by in-solution binding experiments were also  
43 crystallographically confirmed. However, as expected, the majority of soaking hits did not show appreciable  
44 activity in the orthogonal biophysical assays within the tested concentration range (up to 10 mM in ITC,  
45 **Supplementary Table 1**). The macrodomain ADPr-peptide displacement assay also identified two docking  
46 hits not previously observed in soaking (ZINC1337772170 and pterin), which suggests that the crystal  
47 environment limited the ability of some fragments to bind. Yet, between solution experiments good consensus



1 across all assays was observed for ZINC263392672 (PDB 5RSG), ZINC336438345 (PDB 5RSE) and  
2 ZINC331945 (PDB 5RSD). While we are aware that obtaining high-quality binding data remains particularly  
3 challenging to determine for weak binders such as fragments, the dose-response results obtained in the  
4 complementary assays for many of the identified hits provided convincing evidence for their true binding to  
5 the macrodomain. The inconsistency of fragment binding is apparent when comparing fragments that  
6 resulted in high quality data sets in both the P4<sub>3</sub> and C2 crystal systems. Surprisingly, only five of 59 possible  
7 fragments were observed in both systems, with three fragments binding with equivalent poses in the adenine  
8 subsite. This observation points to the value of having multiple measurements, and even multiple crystal  
9 systems when they are available, in fragment-based drug discovery approaches.

10  
11 Overall, this study has three main implications for inhibitor discovery of SARS-CoV-2 Nsp3 macrodomain,  
12 and for antiviral efforts targeting this domain more broadly. First, we describe not only the first new chemical  
13 matter for this target, but map its hot spots at high resolution. This provides a template for future inhibitor  
14 discovery and development against this enzyme. Such efforts will need to navigate selectivity over human  
15 macrodomains but also ATP-binding proteins including kinases (**Fig. S8**) and consider breadth across other  
16 viral macrodomains (16) (**Fig. S4**). Second, the specific fragments that we describe may lend themselves  
17 directly to optimization: several examples are discussed explicitly, amenable to make-on-demand chemistry  
18 (**Fig. 10**); and the 234 structures should provide inspiration for countless other molecules. Finally, important  
19 technical advances emerged from this study: a crystal form that lends itself to ready structure determination,  
20 the creation of a reliable peptide-displacement assay for the macrodomain, and evidence supporting the  
21 ability of structure-based screening, such as molecular docking, to predict effective fragments. The ability to  
22 position explicit hydrogen atoms experimentally, and to resolve electron density on a subatomic scale, makes  
23 Mac1 an attractive candidate for in-depth computational dissection of its catalytic mechanism using  
24 approaches that integrate both classical and quantum calculations. Taken together, these advances will  
25 speed progress throughout the community to help validate this target and create effective antivirals.

## 1 MATERIALS AND METHODS

### 2 Fragment libraries

3 We screened 2,126 molecules from the XChem facility at Diamond Light Source and 411 molecules from  
4 UCSF against Mac1 crystallised in P4<sub>3</sub> crystal form (**Supplementary data set 1**). The fragment library at  
5 XChem combined molecules from multiple fragment libraries: the Diamond, SGC and iNEXT (DSI)-poised  
6 Library (768 molecules, (34)), the Edelris fragment collection (280 molecules, (75)), the MiniFragments Probing  
7 Library (80 molecules, (76)), the FragLites collection (31 compounds, (77)), the PepLite library (25 molecules,  
8 (38)), the SpotFinder (~100 compounds, (78)), and the York3D (106 molecules, (79)) and the EU Open  
9 screen (969 molecules, (52)).

10

11 The UCSF fragment library was composed of Enamine's Essential Fragment library (320 compounds) and  
12 91 additional compounds from an in-house library (UCSF\_91). To assemble the UCSF\_91 library, we  
13 selected topologically diverse molecules having over 10,000 commercially available analogs in at least three  
14 points of substitution, allowing for rapid and extensive analog-by-catalog without having to resort to flask  
15 synthesis. We picked molecules that were also Bemis-Murcko scaffolds (80), stripped of acyclic terminal  
16 substituents. We thought simple, unsubstituted frameworks would be easier to optimize by adding chemical  
17 matter during analoging. From among these, we prioritized by eye scaffolds with various ring sizes and  
18 combinations including fused rings, spiro systems, with linkers of varying lengths between rings, in an attempt  
19 to sample a diverse range of compact shapes and properties. We added anions where the anionic moiety  
20 was a small acyclic substituent on the scaffold, again picking by eye for shape diversity. We chose molecules  
21 with 11-21 heavy atoms, with molecular weights between 200-300 amu and with a logP < 2.5 for solubility.  
22 Physical properties of all screened libraries are shown in **Fig. S5**.

23

24 Analyses of scaffolds and specific chemotypes in the employed chemical libraries are shown in **Fig. S5E**.  
25 Bemis-Murcko (BM) scaffold analysis was performed with the molinspiration mib engine  
26 (<http://www.molinspiration.com>) (81). Pyrimidines were identified using RDKit (<http://www.rdkit.org>) (82) and  
27 molecular charges at pH 7.4 were approximated using ChemAxon Jchem version 2019.15  
28 (<http://www.chemaxon.com>) to identify anionic fragments (83).

### 29 C2 crystals at UCSF

#### 30 *Protein expression and purification*

31 SARS-CoV-2 Nsp3 Mac1 (residues 2-170) was cloned into a pET22b(+) expression vector with an N-terminal  
32 His<sub>6</sub>-tag and a TEV protease recognition site for removal of the tag (GenScript; Piscataway, NJ). In addition,  
33 a short linker (Asn-Ala-Gly) was included between the TEV recognition site and the Mac1 gene  
34 (**Supplementary data set 1**). To express Mac1, plasmid DNA was transformed into BL21(DE3) *E. coli*. After  
35 overnight growth on LB agar supplemented with carbenicillin (100 ug/ml), starter cultures (10 ml LB) were  
36 grown at 37°C for 8 hours. Large scale cultures (1 l TB) were grown at 37°C until an optical density of 0.8.  
37 Cultures were cooled at 4°C for 15 minutes, before protein expression was induced with 1 mM IPTG, and  
38 the cultures were shaken at 20°C for 12 hours. Cells were collected by centrifugation and frozen at -80°C.

39

40 All purification steps were performed at 4°C using an AKTA FPLC system (Cytiva). Cells were resuspended  
41 in Ni-NTA binding buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 5% glycerol, 2 mM βME  
42 supplemented with 5 units/ml TurboNuclease (Sigma, T4330)) and lysed by sonication. Cell debris was  
43 collected by centrifugation and the lysate was applied to a 5 ml HisTrap HP column (Cytiva, 17524802). The  
44 column was washed with 25 ml binding buffer followed by 25 ml 5% Ni-NTA elution buffer (50 mM Tris-HCl

1 pH 8.0, 500 mM NaCl, 500 mM imidazole, 5% glycerol, 2 mM  $\beta$ -ME), and then eluted with 100% elution  
2 buffer. Eluted protein was exchanged into TEV reaction buffer (50 mM Tris pH 8.0, 100 mM NaCl, 1 mM DTT  
3 and 1% glycerol) using a HiPrep 26/10 desalting column (Cytiva, 17508701). To cleave the His<sub>6</sub> tag, Mac1  
4 was diluted to 1.5 mg/ml using TEV reaction buffer and incubated with recombinant TEV protease (84) at a  
5 1:20 ratio (Mac1:TEV) for 16 hours at 4°C. Cleaved Mac1 was separated from the uncleaved protein and  
6 TEV protease by re-running the sample over a HisTrap HP column (pre-equilibrated with TEV reaction buffer)  
7 and collecting the flow-through. The flow-through was supplemented with 10 mM DTT and concentrated to  
8 2.5 ml using a 10 kDa molecular weight cut-off (MWCO) centrifugal concentrator (Amicon, UFC901024). The  
9 sample was further purified by size-exclusion chromatography (SEC) using a HiLoad 16/600 Superdex 75  
10 pg column (Cytiva, 28989333) and SEC buffer (20 mM Tris pH 8.0, 150 mM NaCl, 5% glycerol, 2 mM DTT).  
11 Eluted fractions were concentrated to 15 mg/ml and flash-frozen in liquid nitrogen, and stored at -80°C.  
12 Protein used for ITC was purified in the same manner, but SEC was run with 150 mM NaCl, 20 mM Tris pH  
13 8.0. Protein was concentrated to 10.8 mg/ml prior to flash freezing in liquid nitrogen and storage at -80°C.

#### 14 *Crystallization*

15 Crystals were grown at 19°C using sitting-drop vapor diffusion with a reservoir solution containing 100 mM  
16 Tris pH 8.5, 100 mM sodium acetate and 28% PEG 4000. Crystallization drops were set up with 200 nl  
17 protein and 200 nl reservoir. Initially, crystals were grown in MRC 2-well plates (SwissCI, MRC96TUV) with  
18 a reservoir volume of 40  $\mu$ l. Crystals grew to a maximum size after 1-2 days, and were vitrified in liquid  
19 nitrogen without additional cryoprotection. For diffraction experiments at physiological temperatures, crystals  
20 were mounted using ALS-style goniometer bases (Mitegen, GB-B3S) and sealed with plastic capillary and  
21 vacuum grease (Mitegen, RT-T1). The capillary contained 4  $\mu$ l reservoir solution to prevent crystal  
22 dehydration.

23 Fragment soaking was performed using crystals grown with SwissCI 3-well plates (SwissCi, 3W96T-UVP).  
24 Microseeding was required to achieve consistent nucleation. Several large crystals grown in 100 mM Tris pH  
25 8.5, 100 mM sodium acetate and 28% PEG 4000 were transferred to a drop containing 5  $\mu$ l seed storage  
26 buffer (100 mM Tris pH 8.5, 100 mM sodium acetate, 32% PEG 4000, 2 mM DTT) on a silicon coverslip  
27 (Hampton Research, HR3-233). Crystals were crushed using a flattened glass rod and transferred to 200  $\mu$ l  
28 of seed storage buffer, before being serially diluted 1:10 with seed storage buffer. Consistent nucleation was  
29 achieved with seeds at a 1:100 dilution, with crystallization drops containing 200 nl reservoir, 100 nl seed  
30 stock and 300 nl protein with 30  $\mu$ l in each reservoir.

#### 31 *Crystal dehydration and fragment soaking*

32 Fragments were added to crystallization drops using acoustic dispensing with an Echo 650 liquid handler  
33 (Labcyte) (35). Two libraries were soaked at UCSF: the Enamine Essential fragment library (Enamine, 320  
34 fragments), and the UCSF\_91 library (91 fragments) (**Supplementary data set 1**). To limit DMSO-induced  
35 crystal damage, fragments were targeted to crystallization drops as far away from crystals as possible (35).  
36 Initial DMSO tolerance tests indicated that the C2 crystals were sensitive, rapidly disintegrating upon soaking  
37 with 10% DMSO (**Fig. S2**). To enhance DMSO tolerance, 300 nl of a solution containing 35% PEG 400, 100  
38 mM Tris pH 8.5 and 50 mM sodium acetate was added to drops containing crystals using the Echo. Plates  
39 were resealed and incubated at 19°C for 6 hours. Fragment solutions (120 nl, 10% of the drop volume) were  
40 added using the Echo, and plates were re-sealed and incubated at 20°C for 3-8 hours. Crystals were vitrified  
41 directly from crystallization drops without additional cryoprotection..

42



## 1 *Lysine methylation*

2 Lysine methylation is a routine strategy for altering the crystallization properties of a protein (51). All reagents  
3 were added with the protein on ice and incubation steps were performed at 4°C with gentle shaking. First,  
4 20 mg Mac1 was exchanged into lysine methylation buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 5%  
5 glycerol) using a HiPrep 26/10 desalting column (Cytiva, 17508701). The protein was diluted to 1 mg/ml with  
6 lysine methylation buffer, and 400 µl 1 M dimethylamine borane (DMAB, prepared in water) (Sigma, 180238)  
7 and 800 µl 1 M formaldehyde (prepared in water) (Sigma, F8775) were added to initiate the methylation  
8 reaction. The reaction was left to proceed for 2 hours, then 400 µl 1 M DMAB and 800 µl 1 M formaldehyde  
9 was added. After an additional 2 hours, 200 µl 1 M DMAB was added and the reaction was left for a further  
10 16 hours. To consume any remaining formaldehyde, and to cleave any intermolecular disulfide bonds, 2.5  
11 ml of 1 M glycine (prepared in water) and 2.5 ml of 50 mM DTT (prepared in water) was added and the  
12 reaction was incubated for an additional 2 hours. Next, the sample was concentrated to 2.5 ml using a 10  
13 kDa MWCO concentrator, and purified by SEC. The methylated protein was concentrated to 15 mg/ml before  
14 flash freezing in liquid nitrogen and storage at -80°C.

15  
16 To test the extent of lysine methylation, the purified sample was analysed by liquid chromatography-mass  
17 spectrometry (LC-MS), using a Waters Acquity LC connected to a Waters TQ detector with electrospray  
18 ionization. The sample was separated on a C4 column held at 40°C using water with 0.1% formic acid as  
19 solvent A and acetonitrile with 0.1% formic acid as solvent B. After sample injection (5 µl at 10 µM diluted in  
20 150 mM NaCl, 20 mM Tris pH 8.0), an isocratic elution was run with 95% solvent A and 5% solvent B for 1.5  
21 min. Then, a linear gradient elution was run for 6.5 min to 95% solvent B. Finally, an isocratic elution was run  
22 with 95% solvent B for 2 min. The flow rate was 0.2 ml/min.

## 23 *Crystallization of methylated Mac1*

24 Crystals grew readily in the same conditions as the non-methylated protein (100 mM Tris pH 8.5, 100 mM  
25 sodium acetate, 28% PEG 4000). Consistent nucleation was achieved using microseeding with the same  
26 protocol as the non-methylated protein. Crystallization drops were set up with 100 nl reservoir, 100 nl seed  
27 stocks and 200 nl protein using SwissCI 3-well plates. The methylated crystals displayed increased DMSO  
28 tolerance, so DMSO/fragment soaks were performed directly with 40 nl DMSO (10% of the drop volume).

## 29 *Ultra high resolution data collection, refinement and modelling*

30 To measure the diffraction at such high resolution, we employed a multi-pass, multi-crystal data collection  
31 strategy. We collected ultra high resolution X-ray diffraction data for Mac1 (C2 crystal form) by performing  
32 sequential high-energy (17000 eV) and low-energy (11111 eV) runs to accurately measure reflection  
33 intensities at high and low scattering angles respectively. The same data collection strategy (wedge,  
34 oscillation angle, exposure) was implemented for multiple crystal specimens, each held in different  
35 orientations relative to the X-ray beam and phi rotation axis.

36  
37 The data sets were individually indexed and integrated with XDS (85). During data processing, we merged  
38 the high and low resolution datasets from multiple crystals in different orientations to maximize our coverage  
39 of reciprocal space given a square detector surface. A low-resolution cutoff of 2.5 Å was applied to the high-  
40 resolution (high energy) data sets, because this cutoff simultaneously excludes potentially overlapping  
41 reflections at low scattering angles and allows for a significant number of shared observations between high  
42 and low resolution data sets, which facilitates robust scaling. Scaling and merging were performed using  
43 XSCALE, and the merged intensities were converted to structure factor magnitudes using XDSCONV (85).  
44 We calculated phases by the method of molecular replacement, using the program Phaser (86) and a  
45 previous structure of Mac1 (PDB 6WCF) as the search model. The model was manually adjusted in Coot to

1 fit the electron density map calculated from molecular replacement, followed by automated refinement of  
2 coordinates, atomic displacement parameters, and occupancies using phenix.refine (87) with optimization of  
3 restraint weights. Following two initial rounds of iterative model building and refinement using the  
4 aforementioned strategy, we began introducing additional parameters into the model, enabled by the  
5 extraordinarily high resolution of our diffraction data. First we implemented anisotropic atomic displacement  
6 parameters for heavy atoms (C,N,O,S), followed by refinement of explicit hydrogen atom positions. During  
7 early rounds of model building, we noticed mF<sub>o</sub>-DF<sub>c</sub> difference density peaks appearing between heavy  
8 atom positions, suggesting that we are able to resolve covalent bonding densities (**Fig. S1**). Indeed, atomic  
9 refinement that included a model for inter-atomic scatterers (IAS) was able to account for these densities  
10 (88). Final refinement was performed without geometry or ADP weights (unrestrained).

#### 11 *Data collection at physiological temperature, refinement and modelling*

12 We used a low-dose X-ray data collection strategy to acquire diffraction data from macrodomain crystals (C2  
13 crystal form) at human physiological temperature (37°C, 310 K), which is the temperature most relevant to  
14 studies of SARS-CoV infection. Using this strategy, we acquired data sets using an X-ray exposure of only  
15 50 kGy - less than 1% of the total dose used at 100K, which is essential to mitigate the rapid rate of radiation  
16 damage at 310 K compared to 100 K. The lower overall X-ray dose resulted in data with a lower overall  
17 resolution, extending to 1.5 Å.

18  
19 Diffraction data from multiple crystals were merged using xia2 (89), implementing DIALS (90) for indexing  
20 and integration, and Aimless (91) for scaling and merging. We calculated phases by the method of molecular  
21 replacement, using the program Phaser (86) and our high resolution 100K structure as the search model.  
22 The model was manually adjusted in Coot (92) to fit the electron density map calculated from molecular  
23 replacement, followed by automated refinement of coordinates, atomic displacement parameters, and  
24 occupancies using phenix.refine (87) with optimization of restraint weights.

#### 25 *Fragment data collection, refinement and modelling*

26 Diffraction data was collected at ALS beamline 8.3.1 and SSRL beamlines 12-1 and 12-2. The data collection  
27 strategy is summarized in **Supplementary data set 1**. Fragment datasets were indexed, integrated and  
28 scaled using XDS (85) run through xia2 (89). Based on the space group and unit cell dimensions, six crystal  
29 forms were present (**Fig. S2**). For each of the three C2 isoforms with one molecule in the ASU (isoform A, B  
30 and C), a single, high resolution dataset was selected to create a representative model for each isoform.  
31 Phases were obtained via molecular replacement with Phaser (86), using the ultra-high resolution C2  
32 coordinates as the search model (PDB 7KR0). Coordinates were refined with iterative rounds of manual  
33 model building in Coot (92) and refinement with phenix.refine (87). Default refinement parameters were used,  
34 except five refinement macrocycles carried out per iteration and water molecules were automatically added  
35 to peaks in the 2mF<sub>o</sub>-DF<sub>c</sub> electron density map higher than 3.5  $\sigma$ . The minimum model-water distance was  
36 set to 1.8 Å and a maximum model-water distance to 6 Å. For later rounds of refinement, hydrogens were  
37 added to riding positions using phenix.ready\_set, and B-factors were refined anisotropically for non-hydrogen  
38 and non-water atoms. Although these datasets were obtained from crystals soaked with fragments, three  
39 was no evidence for fragment binding in the mF<sub>o</sub>-DF<sub>c</sub> difference density maps, therefore the datasets were  
40 deemed acceptable as representative DMSO-only models for each isoform.

41  
42 For the fragment datasets, molecular replacement was performed with Phaser (86) and initial refinement with  
43 Refmac (93), both run through the DIMPLE pipeline (91). The search model used for molecular replacement  
44 was selected to match the isoform of the dataset. Waters were included in the initial refinement by changing  
45 the HOH records in the PDB file to "WWW". After refinement, waters were stripped from models and electron

1 density maps were analyzed for fragment binding using PanDDA (54). Electron density maps from 31  
2 datasets were used to calculate the background electron density map for the A isoform, and 24 datasets  
3 were used for isoforms B and C (**Supplementary data set 1**). Datasets selected for background map  
4 calculation had the highest resolution and lowest  $R_{\text{free}}$  values. After PanDDA was run with default parameters,  
5 the threshold used to classify a hit was decreased by adjusting the Z-map analysis settings (contour\_level =  
6 2, min\_blob\_volume = 5, min\_blob\_z\_peak = 2.5). Although there was a substantial increase in false  
7 positives, the decreased threshold allowed an additional seven fragments to be identified. Fragments were  
8 modelled into PanDDA event maps with COOT, using restraints generated by phenix.elbow from a SMILES  
9 string. Changes in protein conformation and solvation were also modeled. Because PanDDA can identify  
10 fragments binding with low occupancies, any changes in protein coordinates will have similar, low  
11 occupancies. If un-restrained refinement is performed on these low occupancy models, changes supported  
12 by PanDDA event maps are often reverted to the ground state model. In the past, this has been overcome  
13 by refining both ground-state (apo) and changed-state (fragment bound) structures simultaneously, with the  
14 changed state coordinates restrained. However, these multi-state models can be difficult to interpret. As an  
15 alternative, we modeled and refined the changed-state only. To prevent reversion of the model into ground  
16 state density, coordinate refinement was switched off after fragments were modelled. Hydrogens were added  
17 with phenix.ready\_set, waters were updated automatically and B-factors were refined anisotropically for non-  
18 hydrogen and non-water atoms. After one round of refinement, waters added into ground state electron  
19 density were removed. This was achieved by aligning the DMSO-only model to the refined model, and  
20 removing any water molecules within 2.2 Å of the DMSO-only model. A final round of refinement was  
21 performed without updating water molecules.

22 P4<sub>3</sub> crystals at UCSF

### 23 *Protein expression and purification*

24 The C2 sequence in pET22b(+) was converted into the P4<sub>3</sub> sequence by removal of Glu170 and replacement  
25 of the N-terminal Asn-Ala-Gly-Glu motif with a methionine. Additionally, a Ser-Ser-Gly-Val-Asp-Leu-Gly-Thr  
26 linker was introduced between the His<sub>6</sub> tag and the TEV recognition sequence (**Supplementary data set 1**).  
27 All cloning steps were performed by PCR with overlapping primers and Gibson assembly (94). Protein was  
28 purified using the same protocol as the C2 protein, except that after SEC, the protein was concentrated to  
29 40 mg/ml prior to flash freezing in liquid nitrogen.

### 30 *Crystallization*

31 Initially, crystals were grown by hanging-drop vapour diffusion with a reservoir solution containing 34% PEG  
32 3000 and 100 mM CHES pH 9.5. Screens were performed using pre-greased VDX plates (Hampton  
33 Research, HR3-142) with 0.5 ml reservoir solution in each well. Crystallization drops were set up on silicon  
34 coverslips (Hampton Research, HR3-233) with 2 µl Mac1 at 10 mg/ml and 2 µl reservoir. Crystals grew after  
35 2-4 days at 19°C. As with the C2 crystals, microseeding was required to achieve consistent nucleation. Seed  
36 stocks were prepared as described previously, except the seed storage buffer used was 35% PEG 3000,  
37 100 mM CHES pH 9.5 and 2 mM DTT. Crystals for fragment soaking were grown using SwissCI 3-well sitting  
38 drop plates with reservoirs containing 30 µl 28% PEG 3000, 100 mM CHES pH 9.5. Crystallization drops  
39 were set-up with 100 nl reservoir solution, 100 nl seed stocks (1:100,000 dilution) and 200 nl Mac1 at 40  
40 mg/ml. Crystals were grown at 19°C and reached a maximum size after 24 hours.

### 41 *Fragment and ADP-ribose soaking*

42 Fragment soaks were performed using the same protocol as the C2 crystals, with soak times between 2-6  
43 hours. ADPr soaks were performed similarly, except that ADP-ribose was prepared in water to 100 mM, and



1 crystals were soaked with 80 nL ADPr (20 mM final concentration). Crystals were vitrified directly after soaking  
2 using a Nanuq cryocooling device (Mitegen).

### 3 *Fragment data collection, processing, modelling and refinement*

4 Diffraction data was collected at ALS beamline 8.3.1, SSRL beamline 12-1 and NSLS-II beamline 17-ID-2.  
5 The data collection summary is summarized in **Supplementary data set 1**. Fragment datasets were indexed,  
6 integrated and scaled using XDS (85) and merged with Aimless (91). In addition to the fragment soaks, we  
7 collected diffraction data for 40 crystals soaked only with DMSO. To generate a DMSO-only model, a single  
8 high resolution dataset was selected and phases were obtained by molecular replacement using the 0.77 Å  
9 C2 structure as a search model. Refinement and model building was performed as described previously for  
10 the C2 crystals. The fragment datasets were prepared for PanDDA analysis using the DIMPLE pipeline.  
11 Fragments were identified using PanDDA, with the background electron density map generated using 35  
12 DMSO-only datasets (**Supplementary data set 1**). As with the analysis of C2 electron density maps,  
13 PanDDA was re-run with a decreased Z-map threshold (contour\_level = 2.5, min\_blob\_volume = 5,  
14 min\_blob\_z\_peak = 2.5). This strategy identified an additional 24 fragments. Fragment modeling and  
15 refinement was carried out using the same protocol as the experiment with C2 crystals.

### 16 *P4<sub>3</sub> crystals at Oxford/XChem*

#### 17 *Protein expression and purification*

18 SARS-CoV-2 Nsp3 Mac1 (residues 3-169) was cloned into a pNIC28-Bsa4 expression vector which adds an  
19 N-terminal His<sub>6</sub>-tag and a TEV protease recognition site for removal of the tag. For expression of protein  
20 used for crystallisation, the constructs was transformed into the *E. coli* Rosetta strain BL21(DE3)-R3 and  
21 cells were grown at 37°C in LB medium (Miller) supplemented with 50 µg/ml of kanamycin and 35 µg/ml of  
22 chloramphenicol. After reaching an OD<sub>600</sub> of 0.5–0.6, the temperature was lowered to 18°C prior to induction  
23 of protein expression overnight by adding 0.5 mM isopropyl 1-thio-D-galactopyranoside. Harvested cells  
24 were resuspended in lysis buffer (50 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol, 20 mM imidazole, 10  
25 mM βME, cOmplete EDTA-free protease inhibitors (Roche)) and stored at -20°C until purification. For protein  
26 purification, pellets were gently thawed in lukewarm water and lysed by high-pressure homogenisation. DNA  
27 was digested using Benzonase. Proteins were purified by immobilised metal affinity chromatography (IMAC)  
28 using Ni-Sepharose resin (GE Healthcare) and eluted stepwise in binding buffer containing 40–500 mM  
29 imidazole. A high salt wash with 1 M NaCl was combined with the first elution step including 40 mM imidazole.  
30 Removal of the hexahistidine tag was carried out by addition of recombinant TEV protease during overnight  
31 dialysis into buffer without imidazole, followed by purification on a second IMAC column and finally by size-  
32 exclusion chromatography (SEC) (Superdex 75, GE Healthcare) in a buffer consisting of 20 mM HEPES (pH  
33 8.0), 250 mM NaCl and 2 mM DTT. Macrodomein protein used for HTRF assay was not subjected to TEV  
34 cleavage and purified after the IMAC step by SEC in a buffer consisting of 25 mM HEPES pH 7.4, 300 mM  
35 NaCl, 5% glycerol and 0.5 mM TCEP. Proteins were characterised by SDS-PAGE, then flash frozen in liquid  
36 nitrogen and stored at -80°C until required.

#### 37 *Crystallographic fragment screening*

38 SARS-CoV-2 Nsp3 Mac1 was concentrated to a final concentration of 47 mg/ml and *apo* crystals were grown  
39 in crystallization solution containing 100 mM CHES pH 9.5 and 30% PEG3000. Fragments were soaked into  
40 crystals as previously described (35) by adding dissolved compounds directly to the crystallisation drops  
41 using an ECHO liquid handler (final concentration 10% DMSO); drops were incubated for approximately 1-3  
42 hours prior to mounting and flash freezing in liquid nitrogen.

43

1 Data was collected at the beamline I04-1 at 100K and automatically processed with Diamond Light Source's  
2 auto-processing pipelines using XDS (85) and either xia2 (95), autoPROC (96) or DIALS (90) with the default  
3 settings. Most Nsp3 macrodomain data processed to a resolution of approximately 1.1 Å. Further analysis  
4 was performed with XChemExplorer, (36), electron density maps were generated with Dimple (97) and  
5 ligand-binding events were identified using PanDDA (54). Ligands were modelled into PanDDA-calculated  
6 event maps using Coot (98), restraints were calculated with AceDRG (99) or GRADE (100), and structures  
7 were refined with BUSTER (101). Coordinates, structure factors and PanDDA event maps for the structures  
8 discussed are deposited in the Protein Data Bank. Data collection and refinement statistics are summarised  
9 in the Supplementary Data File 1.

## 10 Molecular Docking Screens

11 Docking was performed against the crystal structure of SARS-CoV-2 Nsp3 Mac1 bound to ADP-ribose (PDB  
12 6W02). Chain B and all water molecules except for HOH324, HOH344, HOH384, and HOH406 were  
13 removed. These water molecules were included in the docking template structure since they were buried  
14 within the ADP-ribose binding site and formed bridging hydrogen bonds between ADP-ribose and the protein.  
15 The protein structure in complex with ADP-ribose and the 4 selected water molecules was capped at N- and  
16 C-termini and prepared for docking following the prepwizard protocol in Maestro (Schrödinger; (102)).  
17 Accordingly, protons were added using Epik and protonation states were optimized with PropKa at pH 7.  
18 Finally, the structure was energetically minimized using the OPLS3e force field, thereby, the maximum  
19 heavy-atom deviation from the initial structure was 0.3 Å (102).

20  
21 Docking was performed with DOCK3.7 using pre-calculated scoring grids for rapid evaluation of docked  
22 molecules (103). AMBER united atom charges (104) were assigned to the minimized protein structure and  
23 water molecules. Partial atomic charges of backbone amides for residues Ile23 and Phe156 were increased  
24 by 0.2 elementary charge units without changing the net charge of the residues, as described previously  
25 (105). The low dielectric constant of the protein environment was extended outwards from the protein surface  
26 by 1.9Å using spheres generated by SPHGEN. Electrostatic potentials at the ligand-binding pocket were  
27 calculated by numerical solution of the Poisson-Boltzmann equation using QNIFFT (106), scoring grids for  
28 van der Waals potentials were generated with CHEMGRID. Ligand desolvation scoring grids were calculated  
29 by SOLVMAP (107), thereby, the volume of the low protein dielectric was extended out 0.4 Å from the protein  
30 surface, as described previously (55).

31  
32 Since we specifically targeted the adenosine binding site of the full ADP-ribose binding pocket, atomic  
33 coordinates of adenosine rather than the whole ADP-ribose molecule were used to generate 45 matching  
34 spheres, representing favorable positions for placing ligand atoms with docking (103).

35 As ADP-ribose was the only known ligand for the viral macrodomain when we started the docking campaign,  
36 the generated scoring grids and matching spheres were judged for their ability to place and score adenosine,  
37 adenine and ribose at the adenosine binding site of the ligand binding pocket compared to 250 property-  
38 matched decoys, generated following the DUD-EZ method (108). Decoys share similar physical properties  
39 as the control molecules but are topologically different, hence unlikely to ligate the binding pocket.  
40 Furthermore, an "extrema" set (108) of approximately 500,000 molecules including anionic, neutral and  
41 cationic compounds with molecular weights ranging from 250-350 Da was screened to ensure similar  
42 enrichments for monovalent anions and neutral molecules. We note that the lack of experimentally confirmed  
43 ligands for the macrodomain did not allow exhaustive control calculations.

44  
45 Virtual compound libraries were downloaded from ZINC15 ([www.zinc15.docking.org](http://www.zinc15.docking.org)) (53). From the set of  
46 722,963 in-stock fragments, 696,092 compounds were successfully docked, exploring on average 2,355

1 orientations and 63 conformations per compound in the binding pocket. Roughly 58 billion complexes were  
2 sampled in 88 core hours, or roughly 10 minutes on a 500 core cluster. Screening the entire 20 million  
3 ZINC15 fragment library resulted in the evaluation of ca. 4.4 trillion complexes within 2,342 core hours, or  
4 4.7 hours on 500 cores. In that screen, 19,130,798 compounds were scored and sampled in ca. 2,145  
5 orientations and 180 conformations each. From the relatively small “in-human” library, containing 20,726  
6 molecules, 17,362 compounds were scored, and sampling was increased to roughly 16,615 orientations per  
7 compound. 84 billion complexes were evaluated within 27 core hours.

8  
9 Compounds with DOCK scores < -20 (top 500,000 compounds from the entire fragment screen), were  
10 subsequently filtered for those with strained conformations, and inspected for their ability to form hydrogen  
11 bonds to residues Asp22, Ile23, Gly48, Val49, Gly130 or Phe156. Compounds with unsatisfied hydrogen  
12 bond donors or more than three unsatisfied hydrogen bond acceptors were deprioritized. From both fragment  
13 screens, 17 in-stock compounds (8 selected from the ZINC15 in-stock library docking screen) were  
14 purchased, and 45 make-on-demand fragments were ordered of which 33 were successfully synthesized,  
15 both from Enamine. The following compounds were selected from the “in-human” collection docking screen  
16 and purchased from different vendors: Pterin (Sigma-Aldrich, P1132), Verdiperstat (MedChem Express, HY-  
17 17646), Kinetin (Cayman Chemical, 20712), Irsogladine (Cayman Chemical, 30223), Diaveridine (Cayman  
18 Chemical, 29427), N6-Benzyladenine (Cayman Chemical, 21711), PP2 (Cayman Chemical, 13198),  
19 Temozolomide (Cayman Chemical, 14163), Chrysophanol (Cayman Chemical, 19870), Isoxanthopterin  
20 (Cayman Chemical, 17564).

## 21 Fragment linking and merging

22 Fragment mergers and linkers were generated using Fragmenstein  
23 (<https://github.com/matteoferla/Fragmenstein>) (109), a python module that automatically joins inspiration hits  
24 or places compounds based on inspiration hits in way that is as faithful to the positions of the inspiration hits  
25 as possible in a conformation that is energy acceptable. For merging, using RDKit (110) rings are temporarily  
26 collapsed into pseudo-atoms, one-to-one spatial overlapping atoms are identified, pseudo-atoms expanded  
27 with appropriate bonds to nearby atoms and various chemical corrections applied. For the constrained energy  
28 minimisation, Pyrosetta is used (111). Interactive online summary of mergers was made in  
29 <https://michelangelo.sgc.ox.ac.uk> (112, 113).

## 30 Purity and structure determination of initial fragment samples ZINC901381520, ZINC82473428 and 31 ZINC89254160 provided by Enamine

32 Samples of ZINC901391520, ZINC82473428 and ZINC89254160 obtained from Enamine were expected to  
33 be N<sub>9</sub>-alkylated isomers but electron density of compounds in X-ray co-crystal structures indicated these  
34 samples were N<sub>3</sub>-alkylated isomers instead (ZINC901391520\_N3, ZINC82473428\_N3 and  
35 ZINC89254160\_N3, see **Fig. S7I-L**). The original samples of ZINC901391520, ZINC82473428 and  
36 ZINC89254160 used in fragment screening by X-ray crystallography were analyzed by HPLC-MS and <sup>1</sup>H  
37 NMR to confirm sample purity and corroborate structure.

38  
39 There is no reported characterization data to be used as reference for structure confirmation for N<sub>9</sub>- or N<sub>3</sub>-  
40 alkylated compounds ZINC901391520 and ZINC89254160. The N<sub>9</sub>-alkylated structure ZINC82473428 is a  
41 previously prepared compound with tabulated NMR data reported by Rad *et al.* (114).

42 A re-supplied sample of ZINC901391520 from a new batch synthesized at Enamine was confirmed by <sup>1</sup>H  
43 NMR to be >95% purity and a different isomer than the original sample of ZINC901391520. The X-ray crystal  
44 structure of this fragment in complex with Mac1 revealed the fragment to be N<sub>9</sub>-alkylated isomer (**Fig. S7I**).



1 Original samples of ZINC901391520, ZINC82473428 and ZINC89254160 from Enamine used in fragment  
2 screen were evaluated for purity by HPLC on an Agilent 1200 Binary SL system with diode array detection  
3 and mass spectrometric detection on an Agilent 6135B Quadrupole system in electrospray ionization mode  
4 (positive ion detection). One of two HPLC Methods A or B were used to determine sample purity using mobile  
5 phase linear gradients of acetonitrile with 0.1% TFA in water with 0.1% TFA detailed below at 1.000 ml/min  
6 flow rate through a Phenomenex Gemini 3 mm C18 110 Å LC column (4.6 mm dia. x 150 mm length).

7  
8 HPLC Method A mobile phase gradient: Gradient time points (minutes): 1.0-1.5-10.5-11.0-12.5-13.0-15.0; %  
9 acetonitrile at gradient time points: 5-5-20-95-95-5-5

10  
11 HPLC Method B mobile phase gradient: Gradient time points (minutes): 1.0-7.0-8.0-10.0-10.5-12.0; %  
12 acetonitrile at gradient time points: 5-30-95-95-5-5

### 13 *NMR Experiments for Samples ZINC901391520, ZINC82473428 and ZINC89254160*

14 Original samples of ZINC901391520, ZINC82473428 and ZINC89254160 from Enamine used in the  
15 fragment screen were dissolved in  $d_6$ -DMSO and analyzed by  $^1\text{H}$  and  $^{13}\text{C}$  NMR on a Bruker 400 MHz  
16 instrument with Avance III electronics. Data was obtained at ambient temperature (ca. 25°C) collecting 64  
17 scans for proton experiments and 1024 scans for carbon experiments. Raw data was processed and reports  
18 created using ACD Spectrus software.

#### 19 *Original sample ZINC901391520*

20 A sample of 5.5 mg ZINC901391520 was dissolved in 0.75 ml  $d_6$ -DMSO for NMR analysis and from this  
21 solution 50  $\mu\text{l}$  was diluted in 0.45 ml acetonitrile to make up the analytical sample for HPLC-MS using HPLC  
22 Method A. The sample chromatogram from HPLC revealed a single peak with UV absorbance at both 214  
23 and 254 nm at  $t_R = 5.272$  minutes. Aside from a very strong UV<sub>214</sub> peak at  $t_R = 2.00$  minutes attributed to  
24 DMSO co-solvent in the sample, no other peaks were observed at these UV wavelengths and sample purity  
25 estimated >98% based on UV peak area.  $^1\text{H}$  NMR (400 MHz,  $d_6$ -DMSO, 25 °C)  $\delta$  ppm 8.49 (s, 1H), 7.91-  
26 8.26 (br d, 2H), 7.76 (s, 1H), 6.30 (s, 1H), 5.63 (s, 2H), 3.86 (s, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $d_6$ -DMSO, 25 °C)  
27  $\delta$  ppm 172.05, 167.57, 155.01, 152.46, 149.39, 143.55, 120.18, 94.85, 57.13, 44.32. LRMS (ESI<sup>+</sup>) for peak  
28 at  $t_R = 5.272$  minutes: observed  $m/z = 247.3$  [MH]<sup>+</sup> for C<sub>10</sub>H<sub>10</sub>N<sub>6</sub>O<sub>2</sub> exact mass = 246.09.

#### 29 *Second batch sample ZINC901391520*

30 A sample was dissolved in 0.75 ml  $d_6$ -DMSO for NMR analysis.  $^1\text{H}$  NMR (400 MHz,  $d_6$ -DMSO, 25 °C)  $\delta$  ppm  
31 8.24 (s, 1H), 8.16 (s, 1H), 7.31 (br s, 2H), 6.22 (s, 1H), 5.49 (s, 2H), 3.86 (s, 3H).

#### 32 *Sample ZINC82473428*

33 A sample of 3.9 mg ZINC82473428 was dissolved in 0.75 ml  $d_6$ -DMSO for NMR analysis and from this  
34 solution 50  $\mu\text{l}$  was diluted in 0.45 ml acetonitrile to make up the analytical sample for HPLC-MS using HPLC  
35 Method B. The sample chromatogram from HPLC revealed a single peak with UV absorbance at both 214  
36 and 254 nm at  $t_R = 3.766$  minutes. Aside from a very strong UV<sub>214</sub> peak at  $t_R = 2.00$  minutes attributed to  
37 DMSO cosolvent in the sample no other peaks were observed at these UV wavelengths and sample purity  
38 estimated >98% based on UV peak area.  $^1\text{H}$  NMR (400 MHz,  $d_6$ -DMSO, 25 °C)  $\delta$  ppm 8.31 (s, 1H), 8.01 (br  
39 s, 2H), 7.86 (s, 1H), 4.44 (dd,  $J=13.18, 3.39$  Hz, 1H), 4.31-4.40 (m, 1H), 4.20-4.30 (m, 1H), 3.75-3.87 (m,  
40 1H), 3.58-3.70 (m, 1H), 1.93-2.07 (m, 1H), 1.75-1.92 (m, 2H), 1.58-1.73 (m, 1H).  $^{13}\text{C}$  NMR (101 MHz,  $d_6$ -  
41 DMSO, 25 °C)  $\delta$  ppm 154.78, 151.53, 149.56, 144.30, 75.35, 67.24, 52.54, 40.44, 28.23, 25.03. LRMS (ESI<sup>+</sup>)  
42 for peak at  $t_R = 3.766$  minutes: observed  $m/z = 220.3$  [MH]<sup>+</sup> for C<sub>10</sub>H<sub>13</sub>N<sub>5</sub>O exact mass = 219.11.

1 Reported NMR data for compound ZINC82473428\_N9 from Rad et al., 2015 (114):  $^1\text{H}$  NMR (400 MHz,  $d_6$ -  
2 DMSO, 25 °C)  $\delta$  ppm 7.91 (s, 1H), 7.83 (s, 1H), 7.01 (br s, 2H), 3.87-3.99 (m, 3H), 3.34-3.52 (m, 2H), 1.30-  
3 1.54 (complex m, 4H).  $^{13}\text{C}$  NMR (101 MHz,  $d_6$ -DMSO, 25 °C)  $\delta$  ppm 156.6, 152.9, 149.2, 144.7, 117.2, 80.6,  
4 67.9, 57.8, 29.1, 25.1.

#### 5 *Sample ZINC89254160*

6 A sample of 3.2 mg ZINC89254160 was dissolved in 0.75 ml  $d_6$ -DMSO for NMR analysis and from this  
7 solution 50  $\mu\text{l}$  was diluted in 0.45 ml acetonitrile to make up the analytical sample for HPLC-MS using HPLC  
8 Method A. The sample chromatogram from HPLC revealed a major peak and a minor peak with UV  
9 absorbances at both 214 and 254 nm: major peak  $t_R$  = 6.530 minutes and minor peak  $t_R$  = 6.751 minutes.  
10 Relative peak area calculated as percentage of combined UV peak area at 254 nm was 93.3% major peak  
11 and 6.7% minor peak (corresponds to ca. 14:1 ratio). Aside from a very strong UV<sub>214</sub> peak at  $t_R$  = 2.00 minutes  
12 attributed to DMSO cosolvent in the sample no other peaks were observed at these UV wavelengths.  
13 Tabulated NMR data reported here for major peaks only.  $^1\text{H}$  NMR (400 MHz,  $d_6$ -DMSO, 25 °C)  $\delta$  ppm 8.47  
14 (s, 1H), 7.95 (br s, 2H), 7.73 (s, 1H), 7.47 (s, 1H), 5.55 (s, 2H), 2.60 (s, 3H).  $^{13}\text{C}$  NMR (101 MHz,  $d_6$ -DMSO,  
15 25 °C)  $\delta$  ppm 166.21, 154.93, 152.47, 149.63, 149.51, 143.63, 120.43, 117.69, 48.08, 18.66. LRMS (ESI<sup>+</sup>)  
16 for major peak at  $t_R$  = 6.530 minutes: observed  $m/z$  = 247.3 [MH]<sup>+</sup> for C<sub>10</sub>H<sub>10</sub>N<sub>6</sub>S exact mass = 246.07. LRMS  
17 (ESI<sup>+</sup>) for minor peak at  $t_R$  = 6.751 minutes: observed  $m/z$  = 247.3 [MH]<sup>+</sup> for C<sub>10</sub>H<sub>10</sub>N<sub>6</sub>S exact mass = 246.07.  
18 Major peak and minor peak have the same observed mass peak in LRMS and are presumed to be different  
19 N-alkylated isomers.

#### 20 *Conclusions Based on HPLC-MS and NMR Characterization of Samples ZINC901391520, ZINC82473428* 21 *and ZINC89254160*

22 HPLC-MS data confirmed that samples ZINC901391520 and ZINC82473428 are single compounds >98%  
23 purity with mass peak corresponding to either N<sub>9</sub>- or N<sub>3</sub>-alkylated isomers. Both  $^1\text{H}$  and  $^{13}\text{C}$  NMR data  
24 corroborated initial samples ZINC901391520 and ZINC82473428 are >98% single compound. The very high  
25 purity determined for these two samples strongly rules out the possibility that X-ray co-crystals obtained were  
26 the result of protein complexed to trace amount of alternative isomer in the samples. For ZINC89254160,  
27 HPLC-MS data confirmed that there was a 13:1 ratio of isomers in this sample and it is possible the X-ray  
28 co-crystal obtained with ZINC89254160 was the result of protein complexed to trace/minor amounts of the  
29 alternative isomer (N<sub>3</sub>-alkylated).

30  
31 The NMR data obtained for sample ZINC82473428 used in crystallographic fragment screen does not match  
32 NMR data reported in the literature for the N<sub>9</sub>-alkylated ZINC82473428 and thus this sample is presumed  
33 not to be N<sub>9</sub>-alkylated isomer. NMR data is not sufficient to unambiguously assign N<sub>3</sub>- or N<sub>9</sub>-alkylated  
34 structures for ZINC901391520, ZINC82473428 or ZINC89254160 and the unambiguous structure  
35 assignment of ZINC901391520, ZINC82473428 and ZINC89254160 as N<sub>3</sub>-alkylated isomers in this work  
36 was provided by the electron density observed for these fragments in X-ray co-crystal structures obtained.  
37 The crystal structure with ZINC400552187 additionally revealed the N<sub>3</sub>-alkylated structure instead of the  
38 requested N<sub>9</sub>-alkylated form. Using DSF and ITC, ZINC901391520, ZINC82473428, ZINC89254160,  
39 ZINC400552187 were initially screened as the N<sub>3</sub>-alkylated isomer (ZINC901391520\_N3 (PDB 5RSK),  
40 ZINC82473428\_N3 (PDB 5RVF), ZINC89254160\_N3 (PDB 5RSJ), ZINC400552187\_N3 (PDB 5RVG)). In  
41 addition the N<sub>9</sub>-alkylated ZINC901391520 (PDB 5S6W) was tested in DSF and the peptide-competition assay  
42 (HTRF) (**Supplementary Table 1**).

43

## 1 Differential Scanning Fluorimetry (DSF)

### 2 *Compound handling*

3 Compounds were dissolved in DMSO to a final concentration of 100 mM, and placed in a 384-well Echo  
4 source plate (Labcyte PP0200). Using a LabCyte Echo. Each compound was dispensed into a 384-well  
5 storage plate (Greiner BioOne 781280) in five stock concentrations in two-fold serial dilutions (compounds:  
6 6.25-100 mM; ADP-ribose: 0.625-10 mM) and a final volume of 750 nL in triplicate. Two identical plates were  
7 created, with the second plate used to provide protein-free controls for all tested conditions. Echo dispensing  
8 instructions were created by an in-house script ([https://gestwickilab.shinyapps.io/echo\\_layout\\_maker/](https://gestwickilab.shinyapps.io/echo_layout_maker/)).

### 9 *Experimental set-up*

10 DSF buffer was prepared by adding 10  $\mu$ l of SYPRO Orange (Thermo Scientific, Product #S6650) to 10 ml  
11 buffer (50 mM Tris HCl pH 7.50, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.01% Triton X-100), for a final dye  
12 concentration of 5X (10  $\mu$ M) SYPRO Orange. A compound plate (see above) was resuspended by the  
13 addition of 20  $\mu$ l of DSF buffer, and set aside for 20 minutes in the dark. Purified Nsp3 Mac1 domain was  
14 diluted to 10  $\mu$ M in DSF buffer, and 2  $\mu$ l of either protein solution or protein-free buffer was added to each  
15 well a 384-well white PCR plate (Axygen PCR-384-LC480WNFBC, Lot #23517000) using an E100 ClipTip  
16 p125 Matrix Pipette. 8  $\mu$ l of resuspended compound was transferred to each well of the protein- and buffer-  
17 containing PCR plate using an Opentrons OT-2 liquid handling system, yielding the following final conditions:  
18 2  $\mu$ M Nsp3 Mac1, 5X (10  $\mu$ M) SYPRO Orange, 3% DMSO, 0.1-3 mM fragments, and 0.1-1 mM ADP-ribose.  
19 The PCR plate was spun briefly in a salad spinner to remove bubbles, and sealed with optically clear film  
20 (Applied Biosystems, MicroAmp Optical Adhesive Film, Product #4311971). In an Analytik Jena qTower  
21 384G qPCR instrument, plate was continuously heated from 25 - 94°C at a rate of 1°C/minute, and  
22 fluorescence was measured at each degree in the TAMRA channel (535 nm / 580 nm). 53 of 54 fragments  
23 could be tested up to 3 mM without assay interference in these conditions (**Supplementary data set 2,**  
24 **Supplementary Table 1**).

25  
26 Raw DSF data for the Nsp3 Mac1 construct used in this work was characterized by a major transition at 50.8  
27 +/- 0.3°C, with a minor second transition at 67.0 +/- 3.6°C (**Fig. 9, Supplementary data set 2,**  
28 **Supplementary Table 1**); results described refer to the major transition. Significance was defined as  
29 compounds with ANOVA p-values < 0.005 for  $T_{m_a}$  over the tested concentration regime.

30

## 31 Isothermal Titration Calorimetry (ITC)

32 All ITC titrations were performed on a MicroCal iTC 200 instrument (GE Healthcare, IL). All reactions were  
33 performed in 20 mM Tris pH 7.5, 150 mM NaCl using 300 - 600  $\mu$ M of Mac1 at 25°C. Titration of 4 mM ADP-  
34 ribose (Sigma-Aldrich A0752) or 4-10 mM fragment contained in the stirring syringe included a single 0.2  $\mu$ l  
35 injection, followed by 18 consecutive injections of 2  $\mu$ l. Data analysis of thermograms was analyzed using  
36 one set of binding sites model of the [RITC package](#) to obtain all fitting model parameters for the experiments.

## 37 Homogeneous Time-Resolved Fluorescence (HTRF)-based Peptide Displacement Assay

38 Fragment inhibitory activity on Mac1 was assessed by the displacement of an ADP-ribose-conjugated biotin  
39 peptide from the His<sub>6</sub>-tagged Nsp3 Mac1 domain using HTRF with a Eu<sup>3+</sup>-conjugated anti-His<sub>6</sub> antibody  
40 donor and streptavidin-conjugated acceptor. Compounds were dispensed into ProxiPlate-384 Plus  
41 (PerkinElmer) assay plates using an Echo 525 Liquid Handler (Labcyte). Binding assays were conducted in  
42 a final volume of 16  $\mu$ l with 12.5 nM Nsp3 Mac1 domain, 400 nM peptide ARTK(Bio)QTARK(Aoa-RADP)S,  
43 1:125 Streptavidin-XL665, 1:20000 Anti-His<sub>6</sub>-Eu<sup>3+</sup> cryptate in assay buffer (25 mM HEPES pH7.0, 20 mM



1 NaCl, 0.05% BSA, 0.05% Tween20). Assay reagents were dispensed into plates using a Multidrop combi  
2 (Thermo Scientific) and incubated at room temperature for 1 h. Fluorescence was measured using a  
3 PHERAstar microplate reader (BMG) using the HTRF module with dual emission protocol (A = excitation of  
4 320 nm, emission of 665 nm, and B = excitation of 320 nm, emission of 620 nm). Raw data were processed  
5 to give an HTRF ratio (channel A/B × 10,000), which was used to generate IC<sub>50</sub> curves by nonlinear  
6 regression using GraphPad Prism v8 (GraphPad Software, CA, USA).

## 1 ACKNOWLEDGMENTS

2 J. Fraser is supported by NIH GM123159, NSF Rapid 2031205, and a TMC Award from the UCSF Program  
3 for Breakthrough Biomedical Research, funded in part by the Sandler Foundation. Work in the Ahel laboratory  
4 is funded by the Wellcome Trust (grants 101794 and 210634), BBSRC (BB/R007195/1) and Cancer  
5 Research UK (C35050/A22284). BKS is supported by NIH R35GM122481 and DARPA HR0011-19-2-0020.  
6 The crystallographic screen at Oxford was supported by the XChem facility at Diamond Light Source  
7 (proposal ID MX27001). We thank all the staff of Diamond Light Source for providing support and  
8 encouragement which allowed us to carry out this work during the COVID-19 lockdown. We also thank  
9 Gemma Davison, Selma Dormen, James Sanderson, Matthew Martin, Mike Waring and Martin Noble (CRUK  
10 Newcastle Drug Discovery Unit, Newcastle University), Thomas Downes, Paul Jones, Hanna Klein, James  
11 Firth and Peter O'Brien (York University), David Bajusz and Gyorgy Keseru (Hungarian Academy of  
12 Sciences) for providing fragment libraries. We also acknowledge EU-OPENSOURCE ERIC for providing its  
13 fragment library for the presented scientific work. EU-OPENSOURCE ERIC has received funding from  
14 European Union's Horizon 2020 research and innovation programme under grant agreement No 823893  
15 (EU- OPENSOURCE-DRIVE).The SGC is a registered charity (number 1097737) that receives funds from  
16 AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for  
17 Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766],  
18 Janssen, Merck KGaA Darmstadt Germany, MSD, Novartis Pharma AG, Ontario Ministry of Economic  
19 Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome  
20 [106169/ZZ14/Z]. R. Díaz and T. Wu were supported by NSF GRFP. R. Díaz is a Howard Hughes Medical  
21 Institute Gilliam Fellow. I. Young was supported by NIH F32GM133129. M. Ferla is supported by the  
22 Wellcome Trust 203141/Z/16/Z and the NIHR Biomedical Research Centre Oxford

## 23 QCRG

24 The QCRG Structural Biology Consortium has received support from: Quantitative Biosciences Institute,  
25 Defense Advanced Research Projects Agency HR0011-19-2-0020 (to D.A.A. and K.A.V.; B. Shoichet PI),  
26 FastGrants COVID19 grant (K.A. Verba PI), Laboratory For Genomics Research (O.S. Rosenberg PI) and  
27 Laboratory for Genomics Research LGR-ERA (R.M. Stroud PI).

## 28 Beamlines

29 Beamline 8.3.1 at the Advanced Light Source is operated by the University of California Office of the  
30 President, Multicampus Research Programs and Initiatives grant MR-15-328599, NIH (R01 GM124149 and  
31 P30 GM124169), Plexikon Inc., and the Integrated Diffraction Analysis Technologies program of the US  
32 Department of Energy Office of Biological and Environmental Research. The Advanced Light Source  
33 (Berkeley, CA) is a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the  
34 US Department of Energy under contract number DE-AC02-05CH11231, Office of Basic Energy Sciences.

35  
36 Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported  
37 by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No.  
38 DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of  
39 Biological and Environmental Research, and by the National Institutes of Health, National Institute of General  
40 Medical Sciences (including P41GM103393). Extraordinary SSRL operations were supported in part by the  
41 DOE Office of Science through the National Virtual Biotechnology Laboratory, a consortium of DOE national  
42 laboratories focused on response to COVID-19, with funding provided by the Coronavirus CARES Act.

43  
44 This research used beamline 17-ID-2 of the National Synchrotron Light Source II, a U.S. Department of  
45 Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National

1 Laboratory under Contract No. DE-SC0012704. The Center for BioMolecular Structure (CBMS) is primarily  
2 supported by the National Institutes of Health, National Institute of General Medical Sciences (NIGMS)  
3 through a Center Core P30 Grant (P30GM133893), and by the DOE Office of Biological and Environmental  
4 Research (KP1605010).

## 5 Contributions

6 M. Schuller designed and cloned the construct that yielded the P4<sub>3</sub> crystals at Oxford/XChem; expressed,  
7 purified and establishing crystallization conditions for P4<sub>3</sub> crystals for Oxford/XChem; assisted with data  
8 processing/analysis at XChem; set up the HTRF functional assay and performed data interpretation;  
9 prepared the manuscript. G. Correy cloned, expressed and purified the P4<sub>3</sub> construct at UCSF; crystallized,  
10 performed fragment soaking, vitrified crystal, collected X-ray diffraction data and processed data for fragment  
11 screens at UCSF; modelled, refined and analyzed fragment structures at UCSF; purified, methylated and  
12 crystalized the C2 construct; performed ADP-ribose soaks; refined the apo and ADP-ribose bound structures  
13 determined using the P4<sub>3</sub> crystals; prepared the manuscript. S. Gahbauer performed docking screens against  
14 Mac1; performed the chemoinformatic analysis of fragment libraries; assisted with fragment-linking and -  
15 merging; prepared the manuscript. D. Fearon crystallized, prepared samples, collected X-ray diffraction data,  
16 refined and analyzed fragment structures at XChem; prepared the manuscript. T. Wu performed and  
17 analyzed DSF experiments; prepared the manuscript. R. Díaz designed the construct that yielded the C2  
18 crystals; expressed the C2 construct; performed and analyzed ITC experiments; prepared the manuscript. I.  
19 Young collected X-ray diffraction data and processed the diffraction data for crystals screened at UCSF. L.  
20 Martins assisted with docking screens against Mac1. D. Smith assisted with DSF experiments. U. Schulze-  
21 Gahmen crystallized the C2 construct at UCSF. T. Owens purified the C2 construct at UCSF. I. Deshpande  
22 collected X-ray diffraction data for C2 crystals at UCSF. G. Merz purified the C2 construct at UCSF. A. Thwin  
23 purified the C2 construct at UCSF. J. Biel supported the fragment soaking experiments performed at UCSF;  
24 supported the fragment modeling and refinement at UCSF. J. Peters purified C2 construct at UCSF. M. Moritz  
25 purified C2 construct at UCSF. N. Herrera supported crystallization of the C2 construct at UCSF. H. Kratochvil  
26 supported crystallization of the C2 construct at UCSF. QCRG Structural Biology Consortium provided  
27 infrastructure and support for experiments performed at UCSF. A. Aimon prepared samples and collected X-  
28 ray diffraction data at XChem. J. Bennett set up the HTRF functional assay; analyzed and collected data for  
29 the HTRF assay. J. Neto collected X-ray diffraction data at XChem. A. Cohen supported X-ray diffraction  
30 experiments at the SSRL. A. Dias collected X-ray diffraction data at XChem. A. Douangamath crystallized,  
31 prepared samples, collected X-ray diffraction data, refined and analyzed fragment structures at XChem. L.  
32 Dunnett collected X-ray diffraction data at XChem. O. Fedorov set up the HTRF functional assay; analyzed  
33 and collected data for the HTRF assay. M. Ferla performed fragment merging and linking. M. Fuchs  
34 supported X-ray diffraction experiments at the NSLS-II. T. Gorrie-Stone deposited fragments structures at  
35 XChem. J. Holton supported X-ray diffraction experiments at the ALS. M. Johnson analyzed and assigned  
36 structures to the N3- and N9-alkylated fragments. T. Krojer analyzed structural data at XChem. G. Meigs  
37 supported X-ray diffraction experiments at the ALS. A. Powell collected X-ray diffraction data at XChem. J.  
38 Rack provided feedback on the manuscript. V. Rangel refined fragment structures at XChem. S. Russi  
39 supported X-ray diffraction experiments at the SSRL. R. Skyner deposited fragments structures at XChem.  
40 C. Smith supported X-ray diffraction experiments at the SSRL. A. Soares supported X-ray diffraction  
41 experiments at the NSLS-II. J. Wierman supported X-ray diffraction experiments at the SSRL. K. Zhu  
42 expressed and crystallized the P4<sub>3</sub> construct at XChem. N. Jura supervised work. A. Ashworth supervised  
43 work; prepared the manuscript. J. Irwin designed the UCSF\_91 fragment library; prepared the manuscript.  
44 M. Thompson vitrified C2 crystals at UCSF; collected X-ray diffraction data for C2 crystals at UCSF; refined  
45 the ultra-high-resolution structure determined using C2 crystals; prepared the manuscript. J. Gestwicki  
46 supervised work. F. Delft provided the XChem facility; prepared the manuscript; supervised work. B. Shoichet



1 guided and evaluated the docking work; prepared the manuscript; supervised work. J. Fraser supervised  
2 work; prepared the manuscript; arranged funding. I. Ahel supervised work; prepared the manuscript;  
3 arranged funding.

#### 4 Competing interests

5 N. Jura is a member of the SAB of Turning Point Therapeutics and SUDO Biosciences. A. Ashworth is a co-  
6 founder of Tango Therapeutics, Azkarra Therapeutics, Ovibio Corporation; a consultant for SPARC, Bluestar,  
7 ProLynx, Earli, Cura, GenVivo and GSK; a member of the SAB of Genentech, GLAdiator, Circle and  
8 Cambridge Science Corporation; receives grant/research support from SPARC and AstraZeneca; holds  
9 patents on the use of PARP inhibitors held jointly with AstraZeneca which he has benefitted financially (and  
10 may do so in the future). B. Shoichet and J. Irwin are co-founders of a company, BlueDolphin LLC, that does  
11 fee-for-service docking. J. Fraser is a founder of Keyhole Therapeutics and a shareholder of Relay  
12 Therapeutics and Keyhole Therapeutics. The Fraser laboratory has received sponsored research support  
13 from Relay Therapeutics.

#### 14 QCRG Structural Biology Consortium

15 This work was supported by the QCRG (Quantitative Biosciences Institute Coronavirus Research Group)  
16 Structural Biology Consortium. Listed below are the contributing members of the consortium listed by teams.  
17 Within each team the team leads are italicized (responsible for organization of each team, and for the  
18 experimental design utilized within each team), then the rest of team members are listed alphabetically.

19

20 **Bacterial expression team:** *Amy Diallo, Meghna Gupta, Erron W. Titus*, Jen Chen, Loan Doan, Sebastian Flores,  
21 Mingliang Jin, Huong T. Kratochvil, Victor L. Lam, Yang Li, Megan Lo, Gregory E. Merz, Joana Paulino, Aye C.  
22 Thwin, Zanlin Yu, Fengbo Zhou, Yang Zhang.

23 **Protein purification team:** *Daniel Asarnow, Michelle Moritz, Tristan W. Owens, Sergei Pourmal, Caleigh M.*  
24 *Azumaya, Cynthia M. Chio, Bryan Faust, Meghna Gupta, Kate Kim, Joana Paulino, Jessica K. Peters, Kaitlin*  
25 *Schaefer, Tsz Kin Martin Tsui.*

26 **Crystallography team:** *Nadia Herrera, Huong T. Kratochvil, Ursula Schulze-Gahmen, Iris D. Young*, Justin Biel,  
27 Ishan Deshpande, Xi Liu.

28 **CryoEM grid freezing/collection team:** *Caleigh M. Azumaya, Axel F. Brilot, Gregory E. Merz, Cristina Puchades,*  
29 *Alexandrea N. Rizo, Ming Sun, Julian R. Braxton, Meghna Gupta, Fei Li, Kyle E. Lopez, Arthur Melo, Gregory E.*  
30 *Merz, Frank Moss, Joana Paulino, Thomas H. Pospiech, Jr., Sergei Pourmal, Amber M. Smith, Paul V. Thomas,*  
31 *Feng Wang, Zanlin Yu.*

32 **CryoEM data processing team:** *Axel F. Brilot, Miles Sasha Dickinson, Gregory E. Merz, Henry C. Nguyen,*  
33 *Alexandrea N. Rizo, Daniel Asarnow, Julian R. Braxton, Melody G. Campbell, Cynthia M. Chio, Un Seng Chio,*  
34 *Devan Diwanji, Bryan Faust, Meghna Gupta, Nick Hoppe, Mingliang Jin, Fei Li, Junrui Li, Yanxin Liu, Joana*  
35 *Paulino, Thomas H. Pospiech, Jr., Sergei Pourmal, Smriti Sangwan, Raphael Trenker, Donovan Trinidad, Eric*  
36 *Tse, Kaihua Zhang, Fengbo Zhou.*

37 **Mammalian cell expression team:** *Christian Billesboelle, Melody G. Campbell, Devan Diwanji, Carlos Nowotny,*  
38 *Amber M. Smith, Jianhua Zhao, Caleigh M. Azumaya, Alisa Bowen, Nick Hoppe, Yen-Li Li, Phuong Nguyen,*  
39 *Cristina Puchades, Mali Safari, Smriti Sangwan, Kaitlin Schaefer, Raphael Trenker, Tsz Kin Martin Tsui, Natalie*  
40 *Whitis.*

41 **Infrastructure team:** David Bulkley, Arceli Joves, Almarie Joves, Liam McKay, Mariano Tabios, Eric Tse.

42 **Leadership team:** *Oren S Rosenberg, Kliment A Verba*, David A Agard, Yifan Cheng, James S Fraser, Adam  
43 Frost, Natalia Jura, Tanja Kortemme, Nevan J Krogan, Aashish Manglik, Daniel R. Southworth, Robert M Stroud.

## 1 DATA AND MATERIALS AVAILABILITY

2 All data generated or analyzed during this study are included in this article and its Supplementary Information.  
3 Crystallographic coordinates and structure factors for all structures have been deposited in the Protein Data  
4 Bank with the following accessing codes: 7KR0, 7KR1, 7KQW, 7KQO, 7KQP, 5RVJ, 5RVK, 5RVL, 5RVM,  
5 5RVN, 5RVO, 5RVP, 5RVQ, 5RVR, 5RVS, 5RVT, 5RVU, 5RVV, 5RS7, 5RS8, 5RS9, 5RSB, 5RSC, 5RSD,  
6 5RSE, 5RSF, 5RSG, 5RSH, 5RSI, 5RSJ, 5RSK, 5RSL, 5RSM, 5RSN, 5RSO, 5RSP, 5RSQ, 5RSR, 5RSS,  
7 5RST, 5RSU, 5RSV, 5RSW, 5RSX, 5RSY, 5RSZ, 5RT0, 5RT1, 5RT2, 5RT3, 5RT4, 5RT5, 5RT6, 5RT7,  
8 5RT8, 5RT9, 5RTA, 5RTB, 5RTC, 5RTD, 5RTE, 5RTF, 5RTG, 5RTH, 5RTI, 5RTJ, 5RTK, 5RTL, 5RTM,  
9 5RTN, 5RTO, 5RTP, 5RTQ, 5RTR, 5RTS, 5RTT, 5RTU, 5RTV, 5RTW, 5RTX, 5RTY, 5RTZ, 5RU0, 5RU1,  
10 5RU2, 5RU3, 5RU4, 5RU5, 5RU6, 5RU7, 5RU8, 5RU9, 5RUA, 5RUC, 5RUD, 5RUE, 5RUF, 5RUG, 5RUH,  
11 5RUI, 5RUJ, 5RUK, 5RUL, 5RUM, 5RUN, 5RUO, 5RUP, 5RUQ, 5RUR, 5RUS, 5RUT, 5RUU, 5RUV, 5RUW,  
12 5RUX, 5RUY, 5RUZ, 5RV0, 5RV1, 5RV2, 5RV3, 5RV4, 5RV5, 5RV6, 5RV7, 5RV8, 5RV9, 5RVA, 5RVB,  
13 5RVC, 5RVD, 5RVE, 5RVF, 5RVG, 5RVH, 5RVI, 5S6W, 5S18, 5S1A, 5S1C, 5S1E, 5S1G, 5S1I, 5S1K,  
14 5S1M, 5S1O, 5S1Q, 5S1S, 5S1U, 5S1W, 5S1Y, 5S20, 5S22, 5S24, 5S26, 5S27, 5S28, 5S29, 5S2A, 5S2B,  
15 5S2C, 5S2D, 5S2E, 5S2F, 5S2G, 5S2H, 5S2I, 5S2J, 5S2K, 5S2L, 5S2M, 5S2N, 5S2O, 5S2P, 5S2Q, 5S2R,  
16 5S2S, 5S2T, 5S2U, 5S2V, 5S2W, 5S2X, 5S2Y, 5S2Z, 5S30, 5S31, 5S32, 5S33, 5S34, 5S35, 5S36, 5S37,  
17 5S38, 5S39, 5S3A, 5S3B, 5S3C, 5S3D, 5S3E, 5S3F, 5S3G, 5S3H, 5S3I, 5S3J, 5S3K, 5S3L, 5S3M, 5S3N,  
18 5S3O, 5S3P, 5S3Q, 5S3R, 5S3S, 5S3T, 5S3U, 5S3V, 5S3W, 5S3X, 5S3Y, 5S3Z, 5S40, 5S41, 5S42, 5S43,  
19 5S44, 5S45, 5S46, 5S47, 5S48, 5S49, 5S4A, 5S4B, 5S4C, 5S4D, 5S4E, 5S4F, 5S4G, 5S4H, 5S4I, 5S4J,  
20 5S4K.

## 1 REFERENCES

- 2 1. J. G. M. Rack, D. Perina, I. Ahel, Macrodomains: Structure, Function, Evolution, and Catalytic  
3 Activities. *Annu. Rev. Biochem.* **85**, 431–454 (2016).
- 4 2. C. Li, Y. Debing, G. Jankevicius, J. Neyts, I. Ahel, B. Coutard, B. Canard, Viral Macro Domains  
5 Reverse Protein ADP-Ribosylation. *J. Virol.* **90**, 8478–8486 (2016).
- 6 3. D. Munnur, E. Bartlett, P. Mikolčević, I. T. Kirby, J. G. M. Rack, A. Mikoč, M. S. Cohen, I. Ahel,  
7 Reversible ADP-ribosylation of RNA. *Nucleic Acids Res.* **47**, 5658–5669 (2019).
- 8 4. A. R. Fehr, S. A. Singh, C. M. Kerr, S. Mukai, H. Higashi, M. Aikawa, The impact of PARPs and ADP-  
9 ribosylation on inflammation and host-pathogen interactions. *Genes Dev.* **34**, 341–359 (2020).
- 10 5. Y. M. O. Alhammad, A. R. Fehr, The Viral Macrodomain Counters Host Antiviral ADP-Ribosylation.  
11 *Viruses.* **12** (2020), doi:10.3390/v12040384.
- 12 6. A. R. Fehr, R. Channappanavar, G. Jankevicius, C. Fett, J. Zhao, J. Athmer, D. K. Meyerholz, I. Ahel,  
13 S. Perlman, The Conserved Coronavirus Macrodomain Promotes Virulence and Suppresses the Innate  
14 Immune Response during Severe Acute Respiratory Syndrome Coronavirus Infection. *MBio.* **7** (2016),  
15 doi:10.1128/mBio.01721-16.
- 16 7. K. K. Eriksson, L. Cervantes-Barragán, B. Ludewig, V. Thiel, Mouse Hepatitis Virus Liver Pathology  
17 Is Dependent on ADP-Ribose-1<sup>st</sup>-Phosphatase, a Viral Function Conserved in the Alpha-Like Supergroup.  
18 *JVI.* **82**, 12325–12334 (2008).
- 19 8. A. R. Fehr, J. Athmer, R. Channappanavar, J. M. Phillips, D. K. Meyerholz, S. Perlman, The nsp3  
20 macrodomain promotes virulence in mice with coronavirus-induced encephalitis. *J. Virol.* **89**, 1523–1536  
21 (2015).
- 22 9. M. E. Grunewald, Y. Chen, C. Kuny, T. Maejima, R. Lease, D. Ferraris, M. Aikawa, C. S. Sullivan, S.  
23 Perlman, A. R. Fehr, The coronavirus macrodomain is required to prevent PARP-mediated inhibition of virus  
24 replication and enhancement of IFN expression. *PLoS Pathog.* **15**, e1007756 (2019).
- 25 10. L. Palazzo, P. Mikolčević, A. Mikoč, I. Ahel, ADP-ribosylation signalling and human disease. *Open*  
26 *Biol.* **9**, 190041 (2019).
- 27 11. G. Caprara, E. Prosperini, V. Piccolo, G. Sigismondo, A. Melacarne, A. Cuomo, M. Boothby, M.  
28 Rescigno, T. Bonaldi, G. Natoli, PARP14 Controls the Nuclear Accumulation of a Subset of Type I IFN-  
29 Inducible Proteins. *J. Immunol.* **200**, 2439–2454 (2018).
- 30 12. A. R. Fehr, G. Jankevicius, I. Ahel, S. Perlman, Viral Macrodomains: Unique Mediators of Viral  
31 Replication and Pathogenesis. *Trends Microbiol.* **26**, 598–610 (2018).
- 32 13. T. S. Fung, D. X. Liu, Human Coronavirus: Host-Pathogen Interaction. *Annu. Rev. Microbiol.* **73**, 529–  
33 557 (2019).
- 34 14. J. Cui, F. Li, Z.-L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**,  
35 181–192 (2019).
- 36 15. J. Lei, Y. Kusov, R. Hilgenfeld, Nsp3 of coronaviruses: Structures and functions of a large multi-  
37 domain protein. *Antiviral Res.* **149**, 58–74 (2018).
- 38 16. J. G. M. Rack, V. Zorzini, Z. Zhu, M. Schuller, D. Ahel, I. Ahel, Viral macrodomains: a structural and  
39 evolutionary assessment of the pharmacological potential. *Open Biol.* **10**, 200237 (2020).
- 40 17. A. Putics, W. Filipowicz, J. Hall, A. E. Gorbalenya, J. Ziebuhr, ADP-ribose-1<sup>st</sup>-monophosphatase: a  
41 conserved coronavirus enzyme that is dispensable for viral replication in tissue culture. *J. Virol.* **79**, 12721–  
42 12731 (2005).
- 43 18. T. Kuri, K. K. Eriksson, A. Putics, R. Züst, E. J. Snijder, A. D. Davidson, S. G. Siddell, V. Thiel, J.  
44 Ziebuhr, F. Weber, The ADP-ribose-1<sup>st</sup>-monophosphatase domains of severe acute respiratory syndrome  
45 coronavirus and human coronavirus 229E mediate resistance to antiviral interferon responses. *J. Gen. Virol.*  
46 **92**, 1899–1905 (2011).



- 1 19. M.-P. Egloff, H. Malet, Á. Putics, M. Heinonen, H. Dutartre, A. Frangeul, A. Gruez, V. Campanacci,  
2 C. Cambillau, J. Ziebuhr, Others, Structural and functional basis for ADP-ribose and poly (ADP-ribose)  
3 binding by viral macro domains. *J. Virol.* **80**, 8493–8502 (2006).
- 4 20. D. I. James, K. M. Smith, A. M. Jordan, E. E. Fairweather, L. A. Griffiths, N. S. Hamilton, J. R. Hitchin,  
5 C. P. Hutton, S. Jones, P. Kelly, A. E. McGonagle, H. Small, A. I. J. Stowell, J. Tucker, I. D. Waddell, B.  
6 Waszkowycz, D. J. Ogilvie, First-in-Class Chemical Probes against Poly(ADP-ribose) Glycohydrolase  
7 (PARG) Inhibit DNA Repair with Differential Pharmacology to Olaparib. *ACS Chem. Biol.* **11**, 3179–3190  
8 (2016).
- 9 21. A. I. J. Stowell, D. I. James, I. D. Waddell, N. Bennett, C. Truman, I. M. Hardern, D. J. Ogilvie, A high-  
10 throughput screening-compatible homogeneous time-resolved fluorescence assay measuring the  
11 glycohydrolase activity of human poly(ADP-ribose) glycohydrolase. *Anal. Biochem.* **503**, 58–64 (2016).
- 12 22. M. Schuller, K. Riedel, I. Gibbs-Seymour, K. Uth, C. Sieg, A. P. Gehring, I. Ahel, F. Bracher, B. M.  
13 Kessler, J. M. Elkins, S. Knapp, Discovery of a Selective Allosteric Inhibitor Targeting Macrodomain 2 of  
14 Polyadenosine-Diphosphate-Ribose Polymerase 14. *ACS Chem. Biol.* **12**, 2866–2874 (2017).
- 15 23. D. N. Frick, R. S. Viridi, N. Vuksanovic, N. Dahal, N. R. Silvaggi, Molecular Basis for ADP-Ribose  
16 Binding to the Mac1 Domain of SARS-CoV-2 nsp3. *Biochemistry.* **59**, 2608–2615 (2020).
- 17 24. R. S. Viridi, R. V. Bavisotto, N. C. Hopper, D. N. Frick, Discovery of Drug-like Ligands for the Mac1  
18 Domain of SARS-CoV-2 Nsp3. *bioRxiv* (2020), doi:10.1101/2020.07.06.190413.
- 19 25. K. Michalska, Y. Kim, R. Jedrzejczak, N. I. Maltseva, L. Stols, M. Endres, A. Joachimiak, Crystal  
20 structures of SARS-CoV-2 ADP-ribose phosphatase: from the apo form to ligand complexes. *IUCrJ.* **7**, 814–  
21 824 (2020).
- 22 26. F. Cantini, L. Banci, N. Altincekic, J. K. Bains, K. Dhamotharan, C. Fuks, B. Fürtig, S. L. Gande, B.  
23 Hargittay, M. Hengesbach, M. T. Hutchison, S. M. Korn, N. Kubatova, F. Kutz, V. Linhard, F. Löhr, N. Meiser,  
24 D. J. Pyper, N. S. Qureshi, C. Richter, K. Saxena, A. Schlundt, H. Schwalbe, S. Sreeramulu, J.-N. Tants, A.  
25 Wacker, J. E. Weigand, J. Wöhnert, A. C. Tsika, N. K. Fourkiotis, G. A. Spyroulias, 1H, 13C, and 15N  
26 backbone chemical shift assignments of the apo and the ADP-ribose bound forms of the macrodomain of  
27 SARS-CoV-2 non-structural protein 3b. *Biomol. NMR Assign.* **14**, 339–346 (2020).
- 28 27. A. Sallari Jazzi, K. Mahnam, S. Hossein Hejazi, M. S. Damavandi, P. Sadeghi, M. Zeinalian, F.  
29 Tabesh, S. M. Mirbod, H. Khanahmad, Inhibition of Viral Macrodomain of COVID-19 and Human TRPM2 by  
30 losartan. *other* (2020), , doi:10.20944/preprints202003.0457.v1.
- 31 28. M. M. Hann, A. R. Leach, G. Harper, Molecular complexity and its impact on the probability of finding  
32 leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **41**, 856–864 (2001).
- 33 29. A. L. Hopkins, C. R. Groom, A. Alex, Ligand efficiency: a useful metric for lead selection. *Drug Discov.*  
34 *Today.* **9**, 430–431 (2004).
- 35 30. C. W. Murray, D. C. Rees, The rise of fragment-based drug discovery. *Nat. Chem.* **1**, 187–192 (2009).
- 36 31. T. Krojer, J. S. Fraser, F. von Delft, Discovery of allosteric binding sites by crystallographic fragment  
37 screening. *Curr. Opin. Struct. Biol.* **65**, 209–216 (2020).
- 38 32. D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, Twenty years on: the impact of  
39 fragments on drug discovery. *Nat. Rev. Drug Discov.* **15**, 605–619 (2016).
- 40 33. J. R. Helliwell, E. P. Mitchell, Synchrotron radiation macromolecular crystallography: science and  
41 spin-offs. *IUCrJ.* **2**, 283–291 (2015).
- 42 34. O. B. Cox, T. Krojer, P. Collins, O. Monteiro, R. Talon, A. Bradley, O. Fedorov, J. Amin, B. D. Marsden,  
43 J. Spencer, F. von Delft, P. E. Brennan, A poised fragment library enables rapid synthetic expansion yielding  
44 the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.* **7**, 2322–2330 (2016).
- 45 35. P. M. Collins, J. T. Ng, R. Talon, K. Nekrosiute, T. Krojer, A. Douangamath, J. Brandao-Neto, N.  
46 Wright, N. M. Pearce, F. von Delft, Gentle, fast and effective crystal soaking by acoustic dispensing. *Acta*  
47 *Crystallogr D Struct Biol.* **73**, 246–255 (2017).

- 1 36. T. Krojer, R. Talon, N. Pearce, P. Collins, A. Douangamath, J. Brandao-Neto, A. Dias, B. Marsden,  
2 F. von Delft, The XChemExplorer graphical workflow tool for routine or large-scale protein-ligand structure  
3 determination. *Acta Crystallogr D Struct Biol.* **73**, 267–278 (2017).
- 4 37. N. D. Wright, P. Collins, R. Talon, E. Nelson, L. Koekemoer, M. Ye, R. Nowak, J. Newman, J. T. Ng,  
5 N. Mitrovich, H. Wiggers, F. von Delft, The Low-Cost, Semi-Automated Shifter Microscope Stage Transforms  
6 Speed and Robustness of Manual Protein Crystal Harvesting. *Cold Spring Harbor Laboratory* (2019), p.  
7 2019.12.20.875674.
- 8 38. A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C. David Owen, E. Resnick, C. Strain-  
9 Damerell, A. Aimon, P. Ábrányi-Balogh, J. Brandaõ-Neto, A. Carbery, G. Davison, A. Dias, T. D. Downes, L.  
10 Dunnett, M. Fairhead, J. D. Firth, S. Paul Jones, A. Keely, G. M. Keserü, H. F. Klein, M. P. Martin, M. E. M.  
11 Noble, P. O'Brien, A. Powell, R. Reddi, R. Skyner, M. Snee, M. J. Waring, C. Wild, N. London, F. von Delft,  
12 M. A. Walsh, Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease  
13 (2020), p. 2020.05.27.118117.
- 14 39. R. J. Hall, P. N. Mortenson, C. W. Murray, Efficient exploration of chemical space by fragment-based  
15 screening. *Prog. Biophys. Mol. Biol.* **116**, 82–91 (2014).
- 16 40. K. Babaoglu, A. Simeonov, J. J. Irwin, M. E. Nelson, B. Feng, C. J. Thomas, L. Cancian, M. P. Costi,  
17 D. A. Maltby, A. Jadhav, J. Inglese, C. P. Austin, B. K. Shoichet, Comprehensive mechanistic analysis of hits  
18 from high-throughput and docking screens against beta-lactamase. *J. Med. Chem.* **51**, 2502–2511 (2008).
- 19 41. M. Korczynska, D. D. Le, N. Younger, E. Gregori-Puigjané, A. Tumber, T. Krojer, S. Velupillai, C.  
20 Gileadi, R. P. Nowak, E. Iwasa, S. B. Pollock, I. Ortiz Torres, U. Oppermann, B. K. Shoichet, D. G. Fujimori,  
21 Docking and Linking of Fragments To Discover Jumonji Histone Demethylase Inhibitors. *J. Med. Chem.* **59**,  
22 1580–1598 (2016).
- 23 42. A. Manglik, H. Lin, D. K. Aryal, J. D. McCorvy, D. Dengler, G. Corder, A. Levit, R. C. Kling, V. Bernat,  
24 H. Hübner, X.-P. Huang, M. F. Sassano, P. M. Giguère, S. Löber, Da Duan, G. Scherrer, B. K. Kobilka, P.  
25 Gmeiner, B. L. Roth, B. K. Shoichet, Structure-based discovery of opioid analgesics with reduced side  
26 effects. *Nature.* **537**, 185–190 (2016).
- 27 43. W. Wang, M. Wan, D. Liao, G. Peng, X. Xu, W. Yin, G. Guo, F. Jiang, W. Zhong, J. He, Identification  
28 of Potent Chloride Intracellular Channel Protein 1 Inhibitors from Traditional Chinese Medicine through  
29 Structure-Based Virtual Screening and Molecular Dynamics Analysis. *Biomed Res. Int.* **2017**, 4751780  
30 (2017).
- 31 44. J. Lyu, S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O'Meara, T. Che, E. Algaa, K.  
32 Tolmachova, A. A. Tolmachev, B. K. Shoichet, B. L. Roth, J. J. Irwin, Ultra-large library docking for  
33 discovering new chemotypes. *Nature.* **566**, 224–229 (2019).
- 34 45. R. M. Stein, H. J. Kang, J. D. McCorvy, G. C. Glatfelter, A. J. Jones, T. Che, S. Slocum, X.-P. Huang,  
35 O. Savych, Y. S. Moroz, B. Stauch, L. C. Johansson, V. Cherezov, T. Kenakin, J. J. Irwin, B. K. Shoichet, B.  
36 L. Roth, M. L. Dubocovich, Virtual discovery of melatonin receptor ligands to modulate circadian rhythms.  
37 *Nature.* **579**, 609–614 (2020).
- 38 46. X.-P. Huang, J. Karpiak, W. K. Kroeze, H. Zhu, X. Chen, S. S. Moy, K. A. Saddoris, V. D. Nikolova,  
39 M. S. Farrell, S. Wang, T. J. Mangano, D. A. Deshpande, A. Jiang, R. B. Penn, J. Jin, B. H. Koller, T. Kenakin,  
40 B. K. Shoichet, B. L. Roth, Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65.  
41 *Nature.* **527**, 477–483 (2015).
- 42 47. Y. Bian, X.-Q. S. Xie, Computational Fragment-Based Drug Design: Current Trends, Strategies, and  
43 Applications. *AAPS J.* **20**, 59 (2018).
- 44 48. K. Goossens, B. Wroblowski, C. Langini, H. van Vlijmen, A. Cafilisch, H. De Winter, Assessment of  
45 the Fragment Docking Program SEED. *J. Chem. Inf. Model.* **60**, 4881–4893 (2020).
- 46 49. Y. Chen, B. K. Shoichet, Molecular docking and ligand specificity in fragment-based inhibitor  
47 discovery. *Nat. Chem. Biol.* **5**, 358–364 (2009).

- 1 50. D. G. Teotico, K. Babaoglu, G. J. Rocklin, R. S. Ferreira, A. M. Giannetti, B. K. Shoichet, Docking for  
2 fragment inhibitors of AmpC beta-lactamase. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7455–7460 (2009).
- 3 51. T. S. Walter, C. Meier, R. Assenberg, K.-F. Au, J. Ren, A. Verma, J. E. Nettleship, R. J. Owens, D. I.  
4 Stuart, J. M. Grimes, Lysine methylation as a routine rescue strategy for protein crystallization. *Structure.* **14**,  
5 1617–1622 (2006).
- 6 52. EU OPENSREEN - Drive:: DRIVE Startseite, (available at [https://drive.eu-openscreen.eu/drive-](https://drive.eu-openscreen.eu/drive-startseite.html)  
7 [startseite.html](https://drive.eu-openscreen.eu/drive-startseite.html)).
- 8 53. T. Sterling, J. J. Irwin, ZINC 15--ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337  
9 (2015).
- 10 54. N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J.  
11 Shi, C. M. Deane, F. von Delft, A multi-crystal method for extracting obscured crystallographic states from  
12 conventionally uninterpretable electron density. *Nat. Commun.* **8**, 15123 (2017).
- 13 55. M. M. Mysinger, D. R. Weiss, J. J. Ziarek, S. Gravel, A. K. Doak, J. Karpik, N. Heveker, B. K.  
14 Shoichet, B. F. Volkman, Structure-based ligand discovery for the protein-protein interface of chemokine  
15 receptor CXCR4. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5517–5522 (2012).
- 16 56. W. J. Allen, R. C. Rizzo, Implementation of the Hungarian algorithm to account for ligand symmetry  
17 and similarity in structure-based design. *J. Chem. Inf. Model.* **54**, 518–529 (2014).
- 18 57. L. Xing, J. Klug-Mcleod, B. Rai, E. A. Lunney, Kinase hinge binding scaffolds and their hydrogen bond  
19 patterns. *Bioorg. Med. Chem.* **23**, 6520–6527 (2015).
- 20 58. A. Narunsky, A. Kessel, R. Solan, V. Alva, R. Kolodny, N. Ben-Tal, On the evolution of protein-  
21 adenine binding. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 4701–4709 (2020).
- 22 59. J. S. Fraser, H. van den Bedem, A. J. Samelson, P. T. Lang, J. M. Holton, N. Echols, T. Alber,  
23 Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc. Natl.*  
24 *Acad. Sci. U. S. A.* **108**, 16247–16252 (2011).
- 25 60. D. A. Keedy, Z. B. Hill, J. T. Biel, E. Kang, T. J. Rettenmaier, J. Brandão-Neto, N. M. Pearce, F. von  
26 Delft, J. A. Wells, J. S. Fraser, An expanded allosteric network in PTP1B by multitemperature crystallography,  
27 fragment screening, and covalent tethering. *Elife.* **7** (2018), doi:10.7554/eLife.36307.
- 28 61. A. S. Bayden, D. T. Moustakas, D. Joseph-McCarthy, M. L. Lamb, Evaluating Free Energies of  
29 Binding and Conservation of Crystallographic Waters Using SZMAP. *J. Chem. Inf. Model.* **55**, 1552–1565  
30 (2015).
- 31 62. D. Cappel, W. Sherman, T. Beuming, Calculating Water Thermodynamics in the Binding Site of  
32 Proteins - Applications of WaterMap to Drug Discovery. *Curr. Top. Med. Chem.* **17**, 2586–2598 (2017).
- 33 63. B. M. Kearney, M. Schwabe, K. C. Marcus, D. M. Roberts, M. Dechene, P. Swartz, C. Mattos, DRoP:  
34 Automated detection of conserved solvent-binding sites on proteins. *Proteins.* **88**, 152–165 (2020).
- 35 64. F. H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand  
36 interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
- 37 65. M. W. Pantoliano, E. C. Petrella, J. D. Kwasnoski, V. S. Lobanov, J. Myslik, E. Graf, T. Carver, E.  
38 Asel, B. A. Springer, P. Lane, F. R. Salemme, High-density miniaturized thermal shift assays as a general  
39 strategy for drug discovery. *J. Biomol. Screen.* **6**, 429–440 (2001).
- 40 66. A. Simeonov, Recent developments in the use of differential scanning fluorimetry in protein and small  
41 molecule discovery and characterization. *Expert Opin. Drug Discov.* **8**, 1071–1082 (2013).
- 42 67. T. Wiseman, S. Williston, J. F. Brandts, L. N. Lin, Rapid measurement of binding constants and heats  
43 of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131–137 (1989).
- 44 68. G. G. Ferenczy, G. M. Keserű, Thermodynamic profiling for fragment-based lead discovery and  
45 optimization. *Expert Opin. Drug Discov.* **15**, 117–129 (2020).

- 1 69. H. L. Silvestre, T. L. Blundell, C. Abell, A. Ciulli, Integrated biophysical approach to fragment  
2 screening and validation for fragment-based lead discovery. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 12984–  
3 12989 (2013).
- 4 70. J. Schiebel, N. Radeva, S. G. Krimmer, X. Wang, M. Stieler, F. R. Ehrmann, K. Fu, A. Metz, F. U.  
5 Huschmann, M. S. Weiss, U. Mueller, A. Heine, G. Klebe, Six Biophysical Screening Methods Miss a Large  
6 Proportion of Crystallographically Discovered Fragment Hits: A Case Study. *ACS Chem. Biol.* **11**, 1693–1701  
7 (2016).
- 8 71. J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J.  
9 Mayfield, R. A. Sayle, ZINC20-A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *J. Chem.*  
10 *Inf. Model.* (2020), doi:10.1021/acs.jcim.0c00675.
- 11 72. G. Jankevicius, M. Hassler, B. Golia, V. Rybin, M. Zacharias, G. Timinszky, A. G. Ladurner, A family  
12 of macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nat. Struct. Mol. Biol.* **20**, 508–514  
13 (2013).
- 14 73. B. Lamoree, R. E. Hubbard, Current perspectives in fragment-based lead discovery (FBLD). *Essays*  
15 *Biochem.* **61**, 453–464 (2017).
- 16 74. G. Klebe, Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov. Today.* **11**,  
17 580–594 (2006).
- 18 75. D. L. Source, Edelris Keymical (Decommissioned), (available at  
19 [https://www.diamond.ac.uk/Instruments/Mx/Fragment-Screening/New\\_Fragment-Libraries/Edelris-](https://www.diamond.ac.uk/Instruments/Mx/Fragment-Screening/New_Fragment-Libraries/Edelris-Keymical-Library.html)  
20 [Keymical-Library.html](https://www.diamond.ac.uk/Instruments/Mx/Fragment-Screening/New_Fragment-Libraries/Edelris-Keymical-Library.html)).
- 21 76. M. O'Reilly, A. Cleasby, T. G. Davies, R. J. Hall, R. F. Ludlow, C. W. Murray, D. Tisi, H. Jhoti,  
22 Crystallographic screening using ultra-low-molecular-weight ligands to guide drug design. *Drug Discov.*  
23 *Today.* **24**, 1081–1086 (2019).
- 24 77. FragLitesMinimal, Halogenated Fragments Displaying Pharmacophore Doublets. An Efficient  
25 Approach to Druggability Assessment and Hit Generation, , doi:10.1021/acs.jmedchem.9b00304.s001.
- 26 78. RCNS-HAS Medicinal Chemistry Research Group, (available at  
27 [http://medchem.ttk.hu/news\\_archive.html](http://medchem.ttk.hu/news_archive.html)).
- 28 79. T. D. Downes, S. P. Jones, H. F. Klein, M. C. Wheldon, M. Atobe, P. S. Bond, J. D. Firth, N. S. Chan,  
29 L. Waddelove, R. E. Hubbard, D. C. Blakemore, C. De Fusco, S. D. Roughley, L. R. Vidler, M. A. Whatton,  
30 A. J.-A. Woolford, G. L. Wrigley, P. O'Brien, Design and Synthesis of 56 Shape-Diverse 3D Fragments.  
31 *Chemistry.* **26**, 8969–8975 (2020).
- 32 80. G. W. Bemis, M. A. Murcko, The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.*  
33 **39**, 2887–2893 (1996).
- 34 81. Molinspiration Cheminformatics, (available at <http://www.molinspiration.com>).
- 35 82. G. Landrum, RDKit, (available at <http://www.rdkit.org>).
- 36 83. ChemAxon - Software Solutions and Services for Chemistry & Biology, (available at  
37 <http://www.chemaxon.com>).
- 38 84. J. E. Tropea, S. Cherry, D. S. Waugh, Expression and purification of soluble His(6)-tagged TEV  
39 protease. *Methods Mol. Biol.* **498**, 297–307 (2009).
- 40 85. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
- 41 86. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser  
42 crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- 43 87. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T.  
44 C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, Towards automated crystallographic structure  
45 refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).



- 1 88. P. V. Afonine, R. W. Grosse-Kunstleve, P. D. Adams, V. Y. Lunin, A. Urzhumtsev, On macromolecular  
2 refinement at subatomic resolution with interatomic scatterers. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 1194–  
3 1197 (2007).
- 4 89. G. Winter, xia2: an expert system for macromolecular crystallography data reduction. *J. Appl.*  
5 *Crystallogr.* **43**, 186–190 (2009).
- 6 90. G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-  
7 Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter, G. Evans, DIALS: implementation and  
8 evaluation of a new integration package. *Acta Crystallogr D Struct Biol.* **74**, 85–97 (2018).
- 9 91. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B.  
10 Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R.  
11 Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta*  
12 *Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
- 13 92. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol.*  
14 *Crystallogr.* **60**, 2126–2132 (2004).
- 15 93. G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of macromolecular structures by the  
16 maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
- 17 94. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison 3rd, H. O. Smith, Enzymatic  
18 assembly of DNA molecules up to several hundred kilobases. *Nat. Methods.* **6**, 343–345 (2009).
- 19 95. G. Winter, C. M. C. Lobley, S. M. Prince, Decision making in xia2. *Acta Crystallogr. D Biol. Crystallogr.*  
20 **69**, 1260–1273 (2013).
- 21 96. C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, Data  
22 processing and analysis with the autoPROC toolbox. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 293–302  
23 (2011).
- 24 97. M. Wojdyr, R. Keegan, G. Winter, A. Ashton, DIMPLE- a pipeline for the rapid generation of difference  
25 maps from protein crystals with putatively bound ligands. *Acta Crystallographica Section A Foundations of*  
26 *Crystallography.* **69** (2013), pp. s299–s299.
- 27 98. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr.*  
28 *D Biol. Crystallogr.* **66**, 486–501 (2010).
- 29 99. F. Long, R. A. Nicholls, P. Emsley, S. Graæulis, A. Merkys, A. Vaitkus, G. N. Murshudov, AceDRG:  
30 a stereochemical description generator for ligands. *Acta Crystallogr D Struct Biol.* **73**, 112–122 (2017).
- 31 100. GradeCite, (available at <https://www.globalphasing.com/buster/wiki/index.cgi?GradeCite>).
- 32 101. Global Phasing Limited, (available at <http://www.globalphasing.com>).
- 33 102. G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation:  
34 parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–  
35 234 (2013).
- 36 103. R. G. Coleman, M. Carchia, T. Sterling, J. J. Irwin, B. K. Shoichet, Ligand pose and orientational  
37 sampling in molecular docking. *PLoS One.* **8**, e75992 (2013).
- 38 104. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, P. Weiner, A  
39 new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**,  
40 765–784 (1984).
- 41 105. J. Carlsson, L. Yoo, Z.-G. Gao, J. J. Irwin, B. K. Shoichet, K. A. Jacobson, Structure-based discovery  
42 of A2A adenosine receptor ligands. *J. Med. Chem.* **53**, 3748–3755 (2010).
- 43 106. K. Gallagher, K. Sharp, Electrostatic contributions to heat capacity changes of DNA-ligand binding.  
44 *Biophys. J.* **75**, 769–776 (1998).
- 45 107. M. M. Mysinger, B. K. Shoichet, Rapid context-dependent ligand desolvation in molecular docking. *J.*  
46 *Chem. Inf. Model.* **50**, 1561–1573 (2010).

- 1 108. R. M. Stein, Y. Yang, T. E. Balius, M. J. O'Meara, J. Lyu, J. Young, K. Tang, B. K. Shoichet, J. J.  
2 Irwin, Property-unmatched decoys in docking benchmarks. *Understanding virtual solvent through large-scale*  
3 *ligand discovery*, 114 (2020).
- 4 109. M. Ferla, *Fragmenstein* (Github; <https://github.com/matteoferla/Fragmenstein>).
- 5 110. P. Tosco, N. Stiefl, G. Landrum, Bringing the MMFF force field to the RDKit: implementation and  
6 validation. *J. Cheminform.* **6**, 37 (2014).
- 7 111. S. Chaudhury, S. Lyskov, J. J. Gray, PyRosetta: a script-based interface for implementing molecular  
8 modeling algorithms using Rosetta. *Bioinformatics.* **26**, 689–691 (2010).
- 9 112. M. P. Ferla, A. T. Pagnamenta, D. Damerell, J. C. Taylor, B. D. Marsden, Michelangelo: sculpting  
10 protein views on web pages without coding. *Bioinformatics.* **36**, 3268–3270 (2020).
- 11 113. M. F. et al., Michelangelo: sculpting protein views on webpages without coding, (available at  
12 <https://michelangelo.sgc.ox.ac.uk>).
- 13 114. M. N. S. Rad, S. Behrouz, E. Zarenezhad, N. Kaviani, Highly efficient protocol for one-pot N-alkylation  
14 of nucleobases using alcohols in bmim[Br]: a rapid route to access acyclic nucleosides. *Journal of the Iranian*  
15 *Chemical Society.* **12** (2015), pp. 1603–1612.
- 16