

Title

Fragment binding to the Nsp3 macrodomain of SARS-CoV-2 identified through crystallographic screening and computational docking

Authors

Marion Schuller^{1†}, Galen J. Correy^{2†}, Stefan Gahbauer^{3†}, Daren Fearon^{4†}, Taiasean Wu^{5,6}, Roberto Efraín Díaz^{2,7}, Iris D. Young^{2,8}, Luan Carvalho Martins⁹, Dominique H. Smith¹⁰, Ursula Schulze-Gahmen⁸, Tristan W. Owens⁸, Ishan Deshpande⁸, Gregory E. Merz⁸, Aye C. Thwin⁸, Justin T. Biel⁸, Jessica K. Peters⁸, Michelle Moritz⁸, Nadia Herrera⁸, Huong T. Kratochvil⁸, QCRG Structural Biology Consortium^{8†}, Anthony Aimon⁴, James M. Bennett¹¹, Jose Brandao Neto⁴, Aina E. Cohen¹², Alexandre Dias⁴, Alice Douangamath⁴, Louise Dunnett⁴, Oleg Fedorov¹¹, Matteo P. Ferla¹³, Martin Fuchs¹⁴, Tyler J. Gorrie-Stone⁴, James M. Holton^{15,16,12}, Michael G. Johnson¹⁷, Tobias Krojer^{11,18}, George Meigs^{15,16}, Ailsa J. Powell⁴, Johannes Gregor Matthias Rack¹, Victor L. Rangel^{11,18,19}, Silvia Russi¹², Rachael E. Skyner⁴, Clyde A. Smith¹², Alexei S. Soares²⁰, Jennifer L. Wierman¹², Kang Zhu¹, Peter O'Brien²¹, Natalia Jura²², Alan Ashworth¹⁰, John Irwin³, Michael C. Thompson²³, Jason E. Gestwicki^{3,5}, Frank von Delft^{4,11,18,24*}, Brian K. Shoichet^{3*}, James S. Fraser^{2*}, Ivan Ahel^{1*}

Affiliations

¹ Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK.

² Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, CA 94158, USA.

³ Department of Pharmaceutical Chemistry, University of California San Francisco, CA 94158, USA.

⁴ Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0DE, UK.

⁵ Institute for Neurodegenerative Disease, University of California San Francisco, CA 94158, USA.

⁶ Chemistry and Chemical Biology Graduate Program, University of California San Francisco, CA 94158, USA.

⁷ Tetrad Graduate Program, University of California San Francisco, CA 94158, USA.

⁸ Quantitative Biosciences Institute (QBI) Coronavirus Research Group Structural Biology Consortium, University of California San Francisco, CA 94158, USA.

⁹ Biochemistry Department, Institute for Biological Sciences, Federal University of Minas Gerais, Belo Horizonte, Brazil.

¹⁰ Helen Diller Family Comprehensive Cancer, University of California San Francisco, CA 94158, USA.

¹¹ Centre for Medicines Discovery, University of Oxford, South Parks Road, Headington OX3 7DQ, UK.

¹² Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Center, Menlo Park, CA 94025, USA.

¹³ Wellcome Centre for Human Genetics, University of Oxford, Old Road Campus, Oxford OX3 7BN, UK.

¹⁴ National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA.

¹⁵ Department of Biochemistry and Biophysics, University of California San Francisco, CA 94158, USA.

¹⁶ Department of Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

¹⁷ ChemPartner Corporation, South San Francisco, CA 94080, USA.

¹⁸ Structural Genomics Consortium, University of Oxford, Old Road Campus, Roosevelt Drive, Headington OX3 7DQ, UK.

¹⁹ School of Pharmaceutical Sciences of Ribeirao Preto, University of Sao Paulo, São Paulo, Brazil.

²⁰ Photon Sciences, Brookhaven National Laboratory, Upton, NY 11973, USA.

²¹ Department of Chemistry, University of York, Heslington, York YO10 5DD, UK.

²² Department of Cellular and Molecular Pharmacology, University of California San Francisco, CA 94158, USA.

²³ Department of Chemistry and Chemical Biology, University of California Merced, CA 95343, USA.

²⁴ Department of Biochemistry, University of Johannesburg, Auckland Park 2006, South Africa.

† These authors contributed equally

‡ Full listing in the Supplementary Materials

* To whom correspondence should be addressed; Email: ivan.ahel@path.ox.ac.uk, jfraser@fraserlab.com, bshoichet@gmail.com, frank.von-delft@diamond.ac.uk

Abstract

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

Teaser

Massive fragment screening effort provides foundation for the development of macrodomain inhibitors as novel antiviral agents for COVID-19.

MAIN TEXT

Introduction

Macrodomains are conserved protein domains found in all kingdoms of life and in several viruses (1). Viral macrodomains recognize and remove host-derived ADP-ribosylation, a post-translational modification of host and pathogen proteins (2, 3). The innate immune response involves signalling by ADP-ribosylation, which contributes to the suppression of viral replication (3–7). Upon viral infection, ADP-ribosylation is catalyzed by an interferon-induced subset of mammalian ADP-ribosyltransferases (ARTs), collectively termed ‘antiviral poly(ADP-ribosyl) polymerases’ (PARPs) (3, 8). These enzymes transfer the ADP-ribose (‘ADPr’) moiety of NAD⁺ onto target proteins (3, 8). For example, during coronavirus infection, PARP14 stimulates interleukin 4 (IL-4)-dependent transcription, which leads to the production of pro-inflammatory, antiviral cytokines (9). Viral macrodomains, which are found primarily in corona-, alpha-, rubi- and herpes-viruses, can counteract this host defense mechanism via their (ADP-ribosyl)hydrolase activity, contributing to the host-viral arms race for control of cell signalling (10).

Coronaviruses (CoVs) are important pathogens of livestock and humans. Three strains out of seven known to infect humans have caused major outbreaks within the last two decades: the severe acute respiratory syndrome (SARS) coronavirus, causing the SARS epidemic from 2002-2004, the Middle East respiratory syndrome (MERS) coronavirus, causing outbreaks in 2012, 2015 and 2018, and SARS-CoV-2, causing the current COVID-19 pandemic (11). The coronaviral conserved macrodomain (called ‘Mac1’ here; also known as ‘S2-MacroD’ or ‘X domain’) is encoded as part of the non-structural protein 3 (Nsp3), a 200 kDa multi-domain protein (12). While cell culture experiments suggest that SARS Mac1 is dispensable for viral replication in some cell lines (5, 13, 14), animal studies have shown that its hydrolytic activity promotes immune evasion and that it is essential for viral replication and pathogenicity in the host (6, 7). The critical role of macrodomains is further supported by experiments using catalytic null mutations of the murine hepatitis virus (MHV), which render that virus essentially non-pathogenic (5, 6, 13). Collectively, these findings support the idea that SARS-CoV-2 Mac1 is a promising drug target for disrupting the viral life-cycle.

A barrier for macrodomain drug discovery has been the lack of well-behaved inhibitors for this domain. Making matters worse, there are few biochemical assays suitable for screening for such inhibitors. Thus far, PDD00017273, an inhibitor of the poly(ADP-ribose)glycohydrolase (PARG), a macrodomain-type (ADP-ribosyl)hydrolase, remains the only well-characterized inhibitor with convincing on-target pharmacology and selectivity (15). The initial hit was discovered by a homogeneous time-resolved fluorescence (HTRF)-based assay that measures PARG activity, rendering the assay unsuitable for macrodomains that lack this activity (16). A selective allosteric inhibitor targeting PARP14 was identified in an AlphaScreen-based high-throughput screen (HTS) (17). While this inhibitor showed on-target activity in cells, its unique allosteric binding site is difficult to translate to other macrodomains. While potential Mac1 inhibitors have emerged with the advent of SARS-CoV-2 (18), their binding mechanisms and efficacy remain unclear, and the lack of a biochemical assay specific for Mac1 has hindered their development. Furthermore, structures of the new inhibitors bound to Mac1 have not yet been reported, making optimization of initial hits, however promising, difficult.

To address the lack of chemical matter against Mac1, we turned to fragment-based ligand discovery using crystallography as a primary readout (**Fig. 1**). Fragment screens can efficiently address a large and relatively unbiased chemical space (19). Despite typically weak overall affinity, fragments often have high ligand efficiency ($-\Delta G_b/HAC$), and can provide templates for further chemical elaboration into lead-like molecules (20). Crystallography can be used as a primary screening method for fragment discovery (21), and recent automation and processing software at synchrotron radiation sources has made this routinely possible at facilities like the

XChem platform at Diamond Light Source (22–25). As part of Diamond’s contribution toward efforts to combat COVID-19, fragment screening expertise and infrastructure was made immediately available to any users working on SARS-CoV-2 targets (26). Similarly, synchrotron access for essential COVID-19-related research was also made available at the US Department of Energy light sources.

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

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

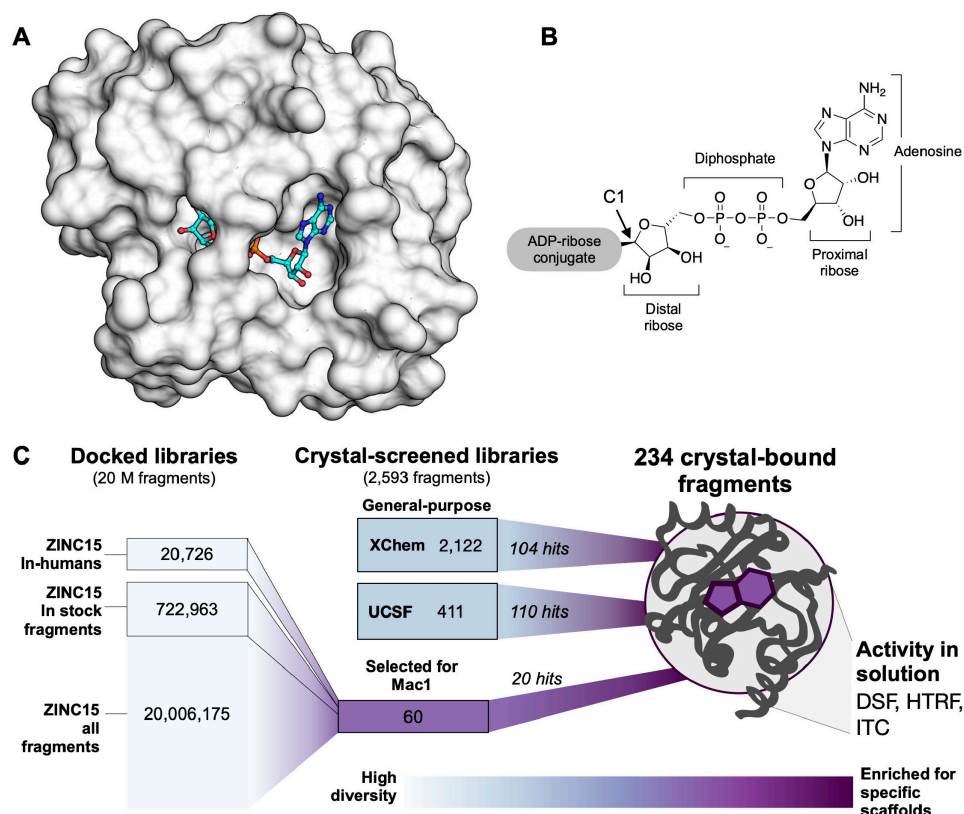


Figure 1. Overview of the fragment discovery approach for SARS-CoV-2 Nsp3 Mac1 presented in this study.

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 20 M 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 a HTRF-based ADPr-peptide displacement assay.

Results

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

We sought a crystal system that enabled consistent ligand soaking for fragment screening and for testing docking predictions. Six Mac1 crystal forms have previously been reported (**Data S1**). Initially, we designed a construct based on PDB entry 6VXS (34). This construct has been reported to crystallize in P1, C2 and P2₁ with either 1 or 2 molecules in the asymmetric unit (ASU) (**Data S1**). This construct crystallized reproducibly in C2 with microseeding and diffracted to a maximum resolution of 0.77 Å (**Data S1**, **Fig. S1**, **Fig. S2A**). This high resolution data yielded electron density maps at true atomic resolution with abundant alternative conformations (**Fig. S1**). The electron density maps also revealed features that are rarely observed in macromolecular crystallography, such as explicit hydrogen atoms, and covalent bond density (**Fig. S1**). Although the active site appears accessible (**Fig. S3B**), efforts to soak ADP-ribose into the crystals were unsuccessful. Additionally, soaking revealed that this crystal form suffers from inconsistent DMSO tolerance (**Fig. S2A**), which is problematic for fragment soaking. In attempts to overcome this problem, we experimented with lysine methylation (35), which

increased DMSO tolerance (**Fig. S2A**), but unfortunately increased occlusion of the active site (**Fig. S3F,G**), and dehydration, which increased DMSO tolerance, at the cost of non-isomorphism (**Fig. S2A,C**).

In parallel, we designed a new Mac1 construct that crystallized in P4₃ with two molecules in the ASU (**Data S1**). This construct crystallized reproducibly with microseeding and diffracted to a maximum resolution of 0.85 Å (**Data S1**). The sequence differences between the two constructs were slight (**Data S1**), yet resulted in substantially different crystal packing (**Fig. S3B-E**). Although the active site of protomer B was obstructed, the active site of protomer A was accessible (**Fig. S3B**), and we were able to soak ADP-ribose into the crystals (**Fig. S4A**). This new structure also revealed a notable difference compared to previously reported Mac1-ADPr structures: the α -anomer of the terminal ribose was observed instead of the β -anomer (**Fig. S4A-D**). Despite this, alignment of ADP-ribose is excellent between all Mac1-ADPr structures (**Fig. S4D**), and the structures are similar overall (**Fig. S4E**). The DMSO tolerance of the P4₃ crystals was excellent (**Fig. S2A**). Accordingly, most of our fragment soaking work proceeded with this construct.

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

Characterization of experimental and virtual screening libraries

Crystal soaking screens at the XChem facility were performed with the P4₃ crystals and a collection of fragment libraries (e.g. Diamond, SGC and iNEXT (DSI)-poised Library including 687 molecules (36) and the EU Open screen containing 968 molecules) totalling 2,122 molecules (see **Data S1** for details). Crystals were screened at the Diamond Light Source. At UCSF, a fragment library composed of Enamine's Essential Fragment library with 320 compounds, augmented by an additional 91 molecules from an in-house library (UCSF_91), was screened against both the P4₃ and C2 crystal forms at the Advanced Light Source (ALS), the Stanford Synchrotron Radiation Lightsource (SSRL) and the National Synchrotron Light Source-II (NSLS-II). On average, molecules across the X-Chem and UCSF collections had molecular weights of 192 ± 47 amu, cLogP values from -1.8 to 3.8, 13 ± 3 heavy atoms, and on average 2 rotatable bonds (**Fig. S5**).

Two fragment libraries were computationally docked against the structure of Mac1 (PDB 6W02): a library of 722,963 fragments "in-stock" at commercial vendors, and the entire ZINC15 fragment library of 20,006,175 mainly make-on-demand fragments that have not been previously synthesized, but can readily be made, available predominantly from Enamine and Wuxi (34). Molecules from the ZINC15 fragment library had molecular weights ≤ 250 amu, cLogP ≤ 3.5 , with an average of 4 rotatable bonds, and typically 4 to 19 heavy atoms (**Fig. S5**). In addition, an "in-human" library of 20,726 drugs, investigational new drugs, and metabolites that have been tested in humans were included into the docking screen, with a view to potential repurposing opportunities. All three sets can be downloaded from ZINC15 (<http://zinc15.docking.org>) (37).

We investigated the fragment libraries for their diversity and their representation of chemotypes likely to bind at the adenine recognition site of Mac1 (**Fig. S5**). Bemis-Murcko (BM) scaffold (38) analysis revealed 179 unique scaffolds in the UCSF libraries, and 809 such scaffolds in the XChem fragment libraries. The in-stock fragment docking library contained 69,244 scaffolds, while 803,333 scaffolds were present in the entire ZINC15 20M fragment collection. Taken together, the experimentally screened libraries contained roughly two compounds per BM scaffold, while the docking libraries contained approximately ten fragments per scaffold, consistent with the expected higher granularity of the docking libraries afforded by their much larger size.

Since adenine-containing compounds are the only structurally characterized binders of Mac1, and fragment libraries are intended to cover a wide chemotype space, we assessed the prevalence of pyrimidines in the libraries. We found pyrimidines in 12 of the 411 fragments in the UCSF libraries, and in 72 of the 2,126

XChem fragments (3.39% of the physically-screened fragments (**Fig. S5**). Pyrimidines were found in 41,531 of the 722,963 (5.74%) in-stock fragments and in 890,199 molecules of the 20,006,175 compound fragment library (4.44%). While the percentages of molecules carrying the pyrimidine substructure were similar between the physical and docked fragments, the absolute numbers in the latter sets were far higher. Aside from bearing a pyrimidine substructure, these subsets were otherwise diverse: among the 890,199 pyrimidine-containing docking fragments, 60,919 distinct BM scaffolds were identified. Adenine itself was present in 5,457 fragments (582 different scaffolds). Furthermore, as ADP-ribose is negatively charged, anionic compounds were considered to exhibit favorable properties to bind to Mac1 by targeting the diphosphate region. Fortuitously, a substantial fraction (35%) of the UCSF fragment libraries is anionic (**Fig. S5**).

Hit rates and Mac1 interaction sites of fragments

Across both crystal forms and facilities, we collected diffraction data for Mac1 crystals soaked with 2,954 fragments (**Data S1**). The diffraction characteristics of the P4₃ crystals were excellent: the average resolution was 1.1 Å, and 98% of crystals diffracted beyond 1.35 Å (**Fig. 2C,E, Fig. S2B**). Although diffraction data was collected for 368 fragments soaked into the C2 crystals at UCSF, data pathologies meant that only 234 datasets could be analysed. The datasets collected from C2 crystals had a mean resolution of 1.4 Å and ranged from 1.0 to 2.2 Å (**Fig. 2A, Fig. S2B**). In total, we identified 234 unique fragments binding to Mac1 using the PanDDA method (**Fig. 2, Data S1, Data S2**) (39). Of these, 221 were identified using P4₃ crystals (hit rate of 8.8%) and 13 using C2 crystals (hit rate of 5.6%). 80% of the fragments were identified in the Mac1 active site, near to or overlapping with the regions occupied by the nucleoside (the adenosine site) or the phospho-ribose (the catalytic site) (**Fig. 2G**). Additional fragments were scattered across the surface of the enzyme, with an enrichment at a distal macrodomain-conserved pocket near lysine 90 (the 'K90 site', 14 fragments) and with many others stabilized by crystal contacts (**Fig. 2B,D,F, Fig. S6**). Coordinates, structure factors, and PanDDA electron density maps for all the fragments have been deposited in the Protein Data Bank (PDB) and are available through the Fragalysis webtool (<https://fragalysis.diamond.ac.uk>).

The unusually high hit rate for the adenosine site in the P4₃ form with the Enamine Essential library (21%) was in contrast to the relatively low hit rate with this library with the C2 form (1.3%). Of the five pairs of fragments identified in both crystal forms, two pairs were identified in the adenine subsite in both crystal forms, two in the adenine subsite in P4₃ crystals but in the K90 site in C2 crystals, and the remaining pair bound to a surface site in the P4₃ crystals and in the K90 site in the C2 crystals (**Data S1**). Additional paired high quality datasets were available for 54 fragments that were bound within the P4₃ crystals, but all showed no density for fragments in the C2 crystals (**Data S1**). It is possible that competition for binding with the N-terminal residues may have contributed to the relatively low hit rate for the C2 form (**Fig. S3F**).

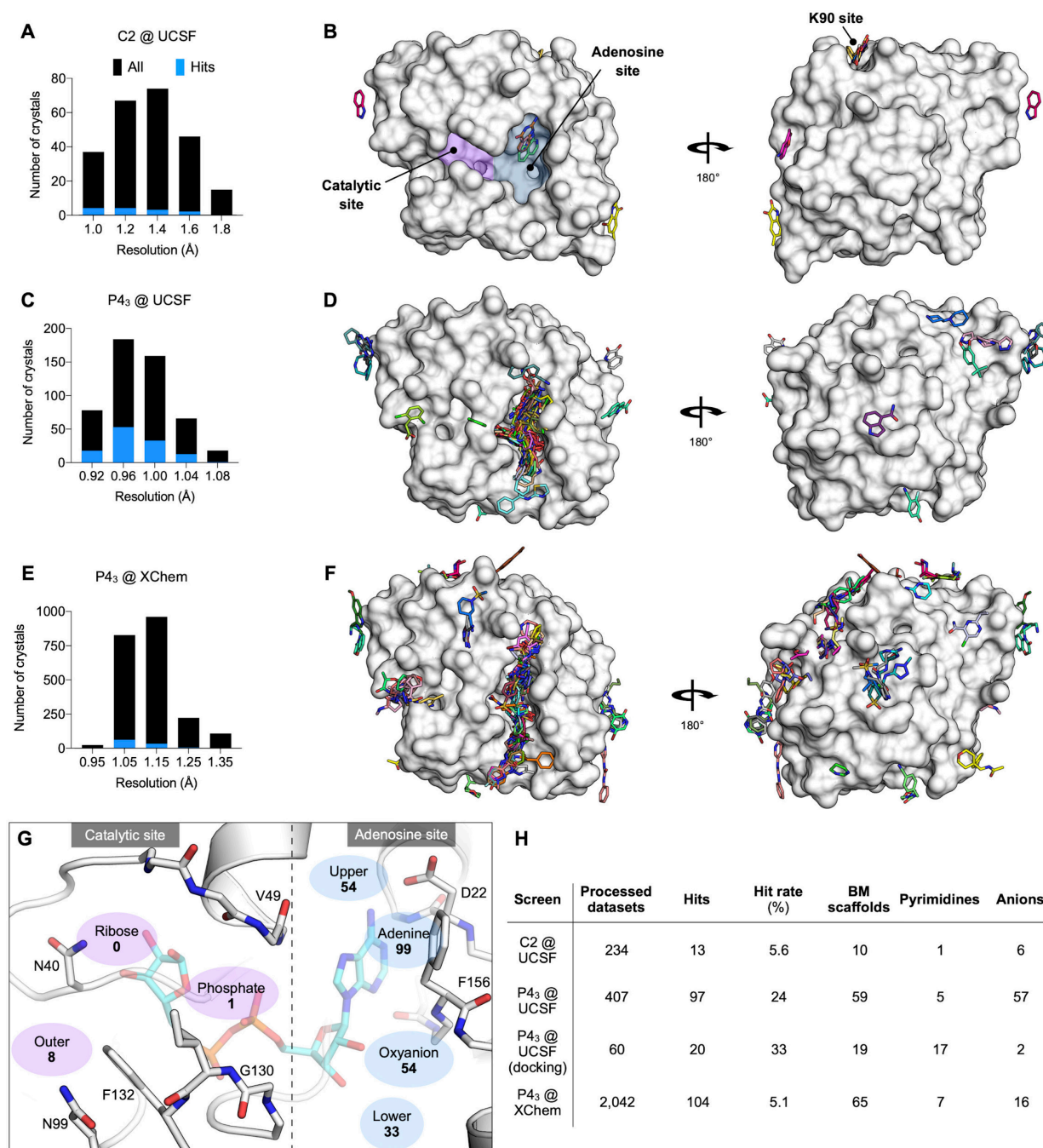


Figure 2. Crystallographic screening identified 234 fragments bound to Mac1.

A,C,E) Histograms showing the resolution of the crystallographic fragment screening data. The resolution of datasets where fragments were identified are shown with blue bars. **B,D,F)** Surface representation of Mac1 with fragments shown as sticks. **G)** The Mac1 active site can be divided based on the interactions made with ADP-ribose. The ‘catalytic’ site recognizes the distal ribose and phosphate portion of the ADP-ribose, and harbours the catalytic residue Asn40 (10). The ‘adenosine’ site recognizes adenine and the proximal ribose. The number of fragments binding in each site is indicated. **H)** Summary of the fragments screened by X-ray crystallography, including the number of Bemis-Murcko (BM) scaffolds and anionic fragments identified as hits in each screen. “Processed datasets” refers to the number of datasets that were analyzed for fragment binding with PanDDA. Out of the datasets collected for 2,954 fragments, 211 (7.1%) were not analyzed due to data pathologies.

Docking hits mimic the adenine recognition pattern

Docking the entire (20 million) ZINC15 fragment library, after calibration of docking parameters using different control calculations (see Methods) (37, 40), was completed in just under 5 hours of elapsed time on 500 cores. The 20,006,175 fragments were sampled in over 4.4 trillion complexes. Top-ranked molecules were inspected for their ability to form hydrogen bonds similar to adenine (e.g. with the side chain of Asp22 and with the backbones of Ile23 and Phe156), while molecules with internal molecular strain or unsatisfied hydrogen bond donors were deprioritized. Ultimately, we selected 54 fragments from the entire ZINC15 fragment library screen, 9 of which were immediately available for purchase from Enamine and 33 of 45 make-on-demand molecules were successfully synthesized *de novo*. Furthermore, 8 fragments were purchased from the ZINC15 in-stock fragment library screen, and an additional 10 compounds were sourced based on the ‘in-human’ library docking (**Data S1**).

Of the 60 molecules tested for complex formation by crystal soaking, 20 were observed with unambiguous electron density in complex with Mac1 (**Data S1**). Here too, the crystals diffracted to exceptionally high resolution, between 0.94 and 1.01 Å. The predicted docking poses typically superposed well on the observed crystallographic results (Hungarian method root mean square deviations (41) ranging from 1-to-5 Å) and 19 out of the 20 docking hits bound to the adenine subsite of the Mac1, as targeted by docking (**Fig. 3, Fig. S7**).

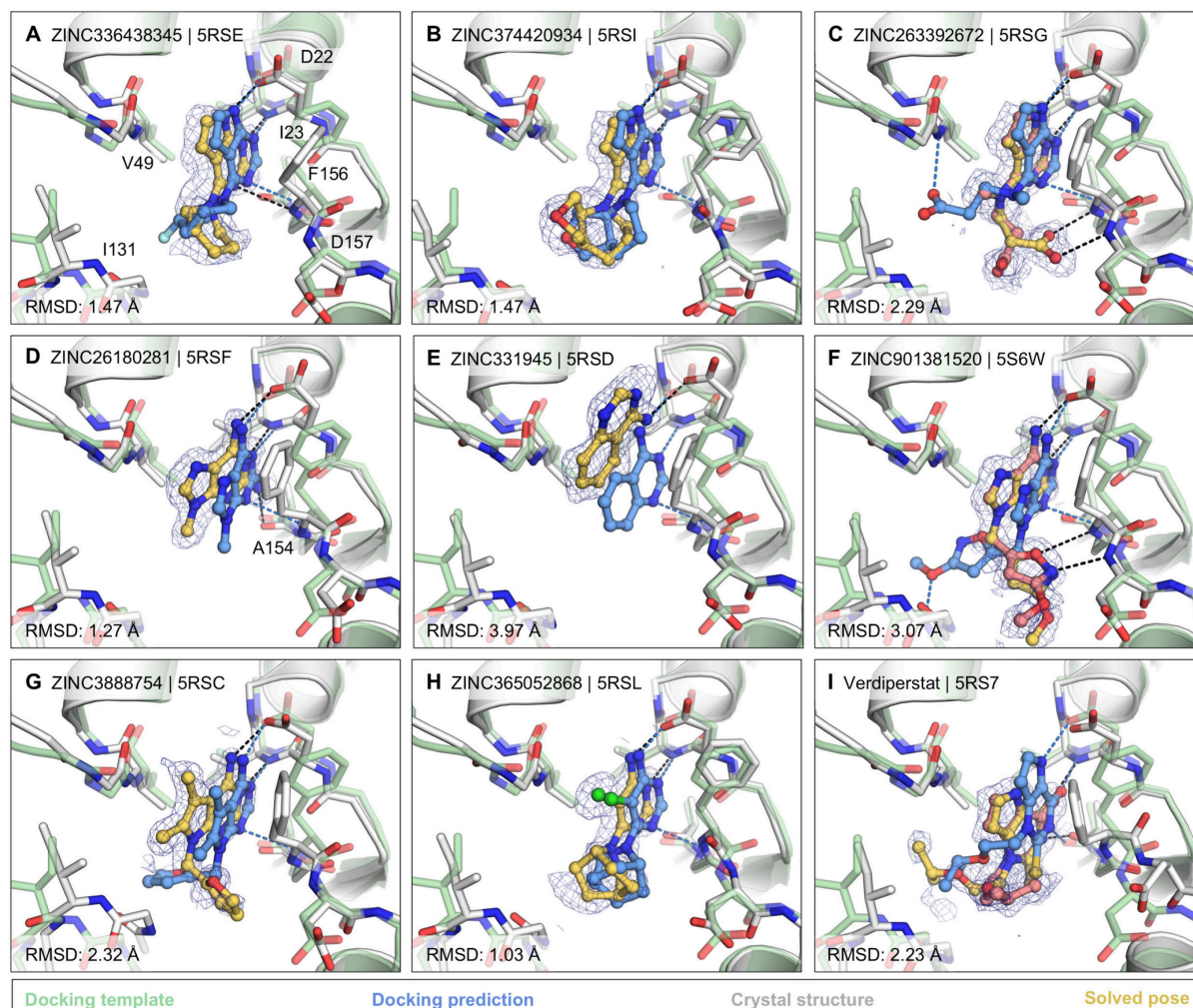


Figure 3. Docking hits confirmed by high-resolution crystal structures.

The protein structure (PDB 6W02) (34), prepared for virtual screens is shown in green, predicted binding poses are shown in blue, the crystal protein structures are shown in grey, the solved fragment poses are shown in yellow, with alternative conformations shown in light pink. PanDDA event maps are shown as a blue mesh. **Event maps were calculated prior to ligand modeling, and the maps are free from model bias towards any ligand (39).** Protein-ligand hydrogen bonds predicted by docking or observed in crystal structures are colored light blue or black, respectively. Hungarian RMSD values are presented between docked and crystallographically determined ligand poses (binding poses for additional docking hits are shown in **Fig. S7**).

The most commonly observed scaffold among the docking hits was 7H-pyrrolo(2,3-d)pyrimidine occupying the adenine-binding subsite (**Fig. 3A-C** and **Fig. S7A,B**). This ring system is typically hydrogen bonded with Asp22, Ile23 and Phe156. Fragments with this scaffold **usually demonstrated** high fidelity between the docking results and the high resolution structures (RMSD 1.5 - 2.3 Å). **For RMSD values >2 Å, indicating noticeable deviations between docking and crystallography (42), visual inspection of docked and solved poses still revealed correct predictions of orientation and key interactions for most fragments in the targeted binding subsite (e.g. Fig. 3C,F,G).** Different substituents can be attached to this headgroup e.g. piperidine, adding a hydrophobic segment to the scaffold (e.g. ZINC336438345 (PDB 5RSE)), occupying most of the adenosine binding site as shown in **Fig. 3A,B** and **Fig. S7A,B**. In addition to hydrophobic variations, ZINC263392672 (PDB 5RSG) attaches an anionic substituent to the pyrrolopyrimidine scaffold, offering additional hydrogen bonds within the binding pocket (**Fig. 3C**). Interestingly, while docking predicted the carboxylic acid of compound ZINC263392672 to insert into the phosphate binding tunnel, forming a hydrogen bond to Val49, the crystal structure instead revealed hydrogen bonds to the backbone amines of Phe156 and Asp157 which we defined as the 'oxyanion' subsite within the adenosine site. Interactions with this backbone-defined oxyanion subsite were also observed for many other hits from both the docking and the crystallographically screened libraries (e.g. **Fig. 3F**, **Fig. S7E**).

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

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

In addition to hydrogen-bonding with residues involved in the adenine recognition of ADP-ribose, several docking hits hydrogen bond to the backbone carbonyl group of Ala154 (**Fig. 3D,I**, **Fig. S7G**), revealing an intriguing accessory polar contact within this subsite. While most residues surrounding the adenosine-binding

site adopted similar conformations in the fragment-bound crystal structures as in the ADPr-bound structure used for docking (PDB 6W02) (34), Asp22 and Phe156 adopted multiple, alternative conformations. In most fragment-bound crystal structures, Phe156 rotated by approximately 90°, enabling improved face-to-face π - π stacking against the aromatic moieties in the bound fragments (**Fig. 3C-G**). However, the docking template orientation of Phe156 was retained for other pyrimidine-containing fragment-bound crystal structures (**Fig. 3B,H**).

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

Analysis of key interactions between Mac1 and fragments from the crystallographic screens

Fragments binding to the adenine subsite

While docking was successful in targeting the adenine binding subsite, crystallographic fragment screening has the advantages of being binding site agnostic and has the potential to identify novel chemotypes at multiple sites. In total, crystallographic screening identified 99 adenine-subsite binding fragments that form subsets of the three hydrogen bonds found between Mac1 and ADP-ribose (**Fig. 4A-C**). Fragments that formed at least two hydrogen bonds to the adenine subsite were separated into nine classes based on the number, nature and connectivity of atoms involved in such hydrogen bonding (**Fig. 4D**). The most common class consisted of a 1,3-hydrogen bond donor/acceptor motif (**Fig. 4D,E.I**). This resembles the kinase hinge binding motif, with the difference being the engagement of a side chain oxygen rather than a backbone carbonyl oxygen (**Fig. S8A-B**) (43). While 7 out of 18 fragments in this class were 4-amino-pyrimidine derivatives, other moieties were also found, including two 2-amino-thiazole-based fragments and several purine derivatives (**Data S1**). We also observed an unusual adenine-binding mode with a hydrogen bond formed between Ile23 and N7 instead of N1 (**Fig. 4D,E.II**). The alternative binding mode can be explained by the N3 substitution of adenine on this fragment, which prevents formation of the canonical N1-Ile23 hydrogen bond. This pattern of hydrogen bonds to the protein has not been previously observed in adenines linked through N9 (44).

We also observed diverse fragments without adenine-like motifs binding at this site, including succinimides, amides, thiazoles, diamino-pyridines, pyrazoles, pyrroles, and ureas (**Fig. 4D,E.III-VIII**). These exploited, separately and together, Asp22 and Ile23, Ala154, and occasionally all three adenine-defining hydrogen-bonding residues. Several fragments π - π stacked with Phe156, while those bearing a urea hydrogen-bonded with the carboxylate of Asp22 (**Fig. 4D,E.VIII**). These interactions were reproduced by a series of benzimidazole-based fragments (**Fig. 4D,E.IX**). These classes occupied what might be classified as an 'upper' subsite, above that defined by the adenine-ribose axis (**Fig. 2G**), and may provide an opportunity to grow fragments away from the canonical site.

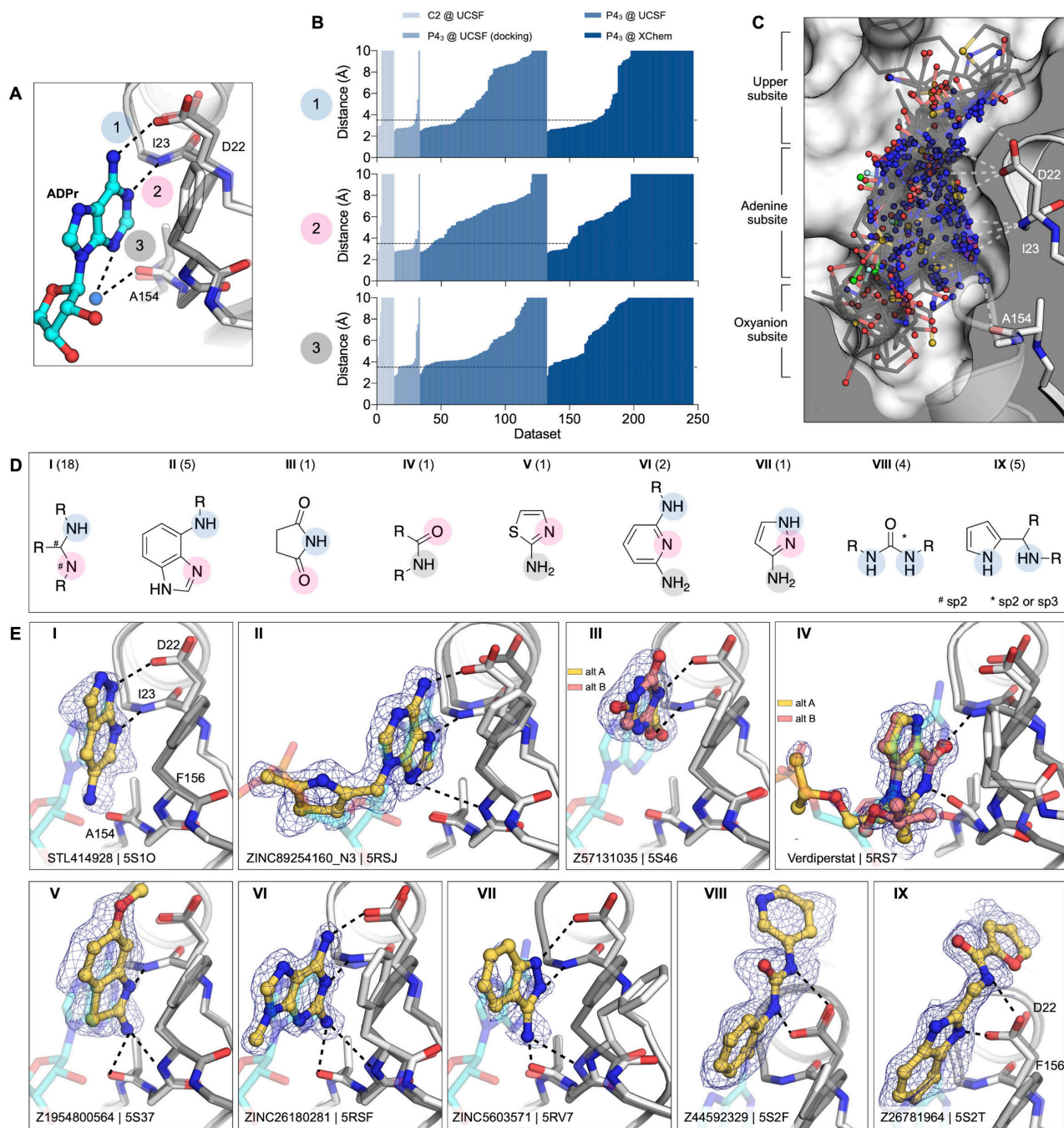


Figure 4. Fragments binding to the adenine subsite.

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

Fragments binding to the oxyanion subsite

In total, we identified 54 fragments that formed interactions with an unexpected “oxyanion” subsite, defined by the backbone nitrogens of Phe156 and Asp157 adjacent to the adenine subsite (**Fig. 2G**, **Fig. 5A**). As suggested by its name, most of these fragments (48/54) were anionic (**Data S1**). Intriguingly, the defining backbone nitrogens adopted a similar orientation to those defining the classic oxyanion hole of serine hydrolases such as acetylcholinesterase (**Fig. S8D-F**). In the Mac1-ADPr structure, the C2 hydroxyl (2'OH) of the proximal ribose interacts with the oxyanion subsite via a bridging water (**Fig. 5A**). In total, 54 fragments formed at least one hydrogen bond to the oxyanion subsite (**Fig. 5B**). Here too, the fragments were both geometrically (**Fig. 5C**) and chemically diverse (**Fig. 5D**): orienting groups either toward the phosphate tunnel, the lower site, or wrapped around toward the upper adenine subsite, providing multiple opportunities for further elaboration. Chemically, they interacted with the site using not only a carboxylate, but also sulfones, and isoxazole, α -keto acid, and a succinimide (**Fig. 5E**). We suspect that the presence of the oxyanion subsite explains the higher hit rate for the Enamine Essential library versus the other crystallographic fragment libraries screened (27% versus 6%), as the former had a greater proportion of acids than the others (41% versus 4%) (**Fig. S5**).

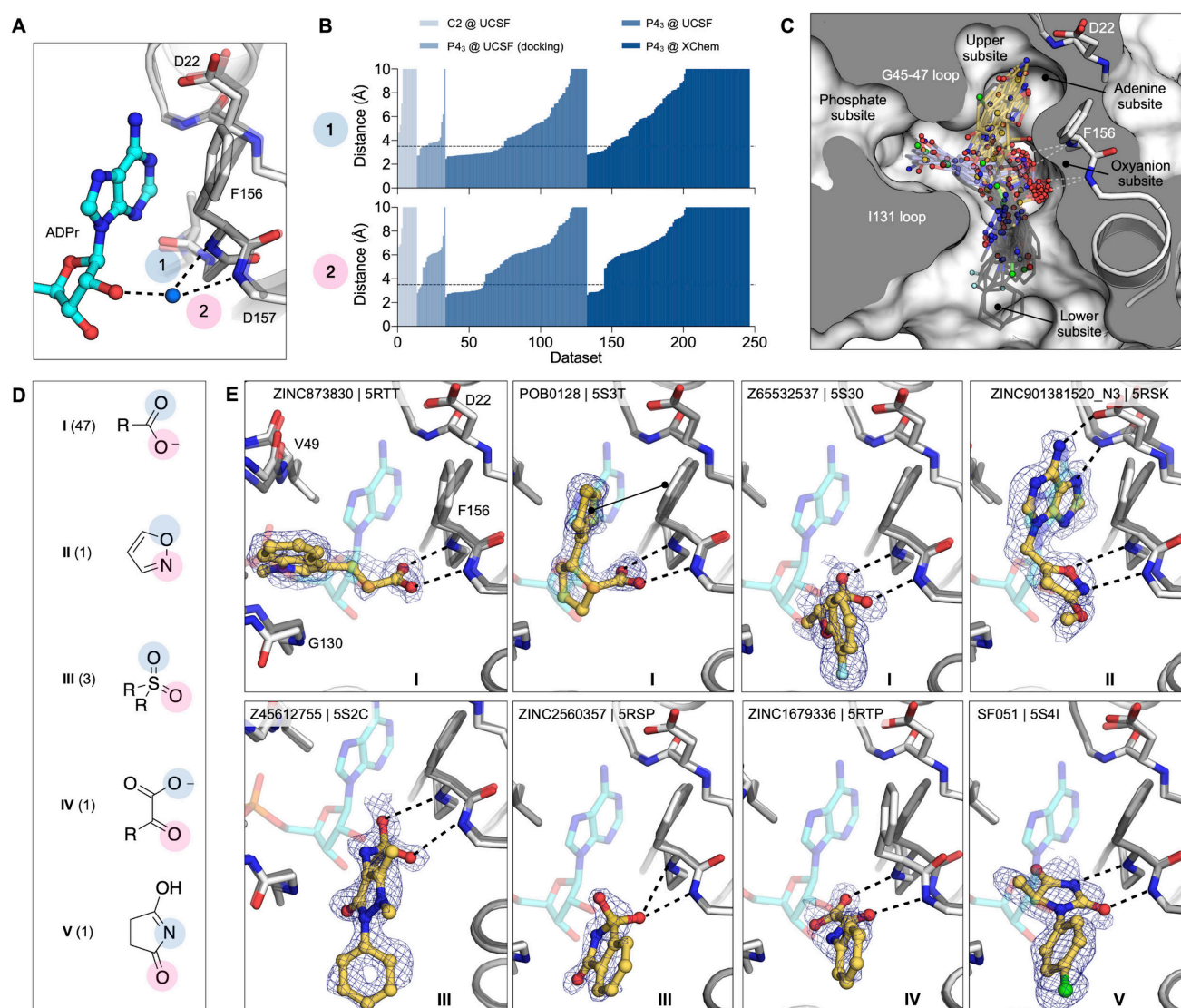


Figure 5. Fragments binding to the oxyanion subsite.

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

Fragments binding to the catalytic and other potential allosteric sites

There were substantially fewer hits against the catalytic site (**Fig. 2G**) versus the adenosine site (eight versus >100), though both appear to be accessible (**Fig. S3B**). The catalytic site consists of three subsites: the phosphate tunnel, which is occupied by the diphosphate of ADP-ribose, the ribose subsite, which is occupied by the terminal ribose of the molecule, and the outer subsite, which sits between Asn40 and Asn99 (**Fig. 2G**, **Fig. 6A**). Of the eight fragments binding in the catalytic site, seven bound in the outer subsite and one bound in the phosphate tunnel. Binding to the outer site was often defined by hydrophobic packing between the Tyr42 and Lys102 side chains, although POB0135 (PDB 5S3W) and POB0128 (PDB 5S3T) formed a salt bridge to Lys102 (e.g. **Fig. 6A.I**). Interestingly, the latter fragment was also found to bind in the adenosine site. Other molecules, including Z2234920345 (PDB 5S2L) and Z955123498 (PDB 5S4A) stabilize an alternative conformation of Lys102 (**Fig. 6A.II**). Three of the fragments, including Z85956652 (PDB 5S2U), positioned a halogen atom in the outer subsite (e.g. **Fig. 6A.III**). The only fragment identified in the phosphate subsite was ZINC84843283 (PDB 5RVI). This fragment was wedged between the Gly47/Ile131 loops, and increased the gap between the two loops by 1.6 Å (**Fig. 6A.IV**). The absence of fragments binding to the ribose subsite, and the sparsity of fragments in the phosphate tunnel, means that designing a Mac1 inhibitor to occupy the catalytic site will rely more heavily on fragment growing than on fragment merging.

Both crystallographic screens also identified fragments binding to the ‘K90 site’, which is formed by a cleft between Lys31, Thr32 and Lys90 (**Fig. 6B**). We identified seven fragments from the C2 crystal form and six from the P4₃ crystal form; intriguingly, none of the C2-derived fragments were found again when the UCSF libraries were rescreened in the P4₃ crystal condition. Although the K90 site is 15 Å from the adenosine site, it is connected to that subsite via a single alpha-helix (**Fig. 6B**). Although there is no biochemical evidence for allosteric communication between these sites, the fragments provide starting points for designing chemical probes to test this possibility.

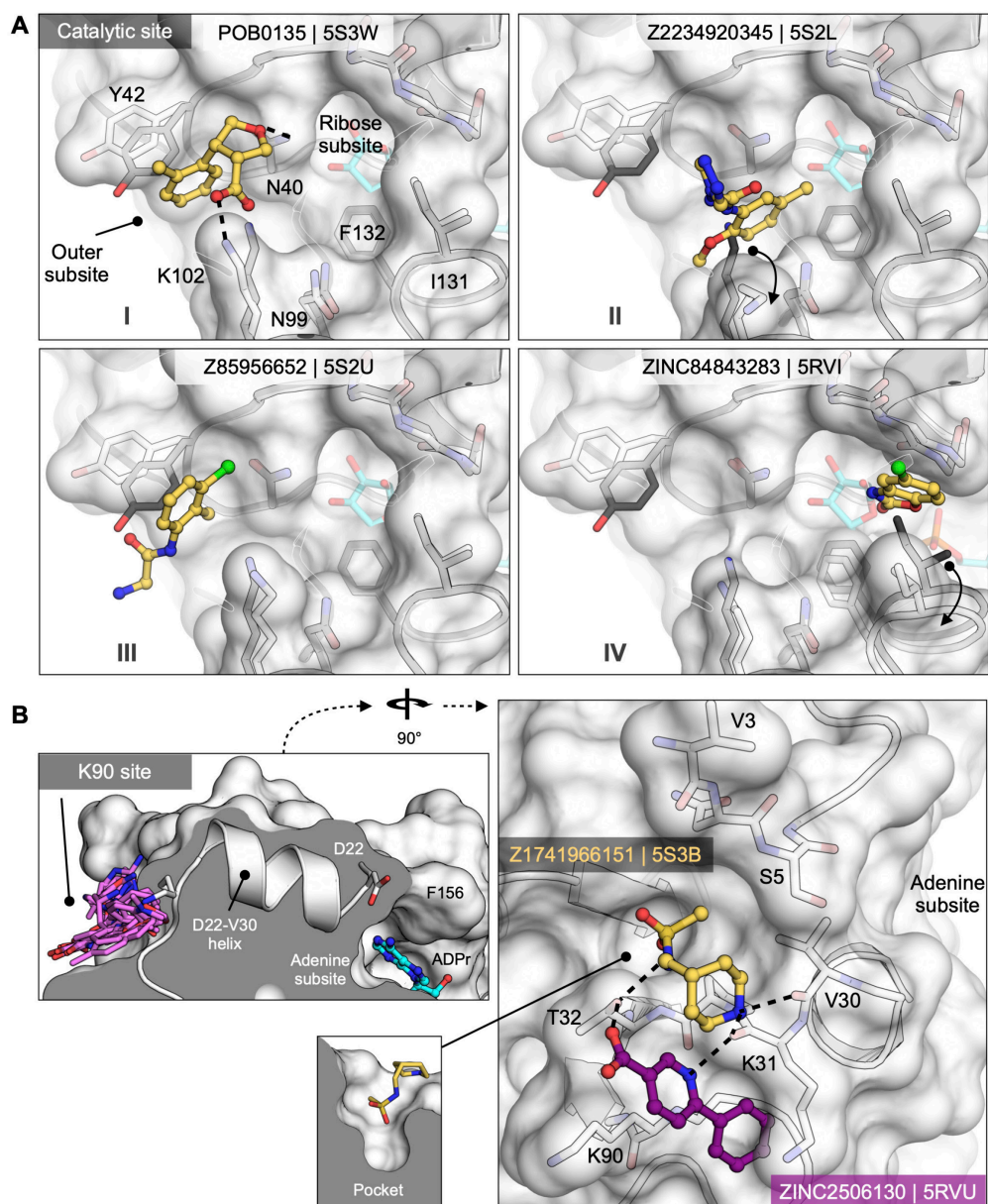


Figure 6. Fragments targeting the catalytic and potential allosteric sites are sparsely populated compared to the adenosine site.

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 subsite, 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 by 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).

Fragment binding exploits protein conformational flexibility

To identify Mac1 flexibility associated with molecular recognition, we calculated the root-mean-square fluctuation (RMSF) of side-chain atoms across the P4₃ fragment-bound structures. Residues lining the adenosine site, especially Asp22 and Phe156, are the most flexible (**Fig. 7A,B**). The flexibility of both residues is paralleled in previously reported crystal structures (**Fig. 7C**) and also in the 0.77 Å apo structure, where multiple alternative conformations are clearly defined in electron density maps (**Fig. 7D** and **Fig. S1A-C**). In the ultra-high resolution structure, residues 155-159 are modeled as a combination of two distinct backbone conformations that diverge substantially at Phe156, which requires three distinct conformations of this residue to satisfy the observed density (**Fig. 7D, Fig. S1C**). Despite this flexibility, hydrogen bonds to Asp22 are present in many fragments, including docking compounds that were chosen based on interactions with a static receptor (**Fig. 7E**). Similarly, the flexibility of the aromatic side-chain of Phe156 enables adaptable stacking interactions with fragments (**Fig. 7E,F**), with 46 fragments binding within 4 Å of Phe156. As with Asp22, the nature and geometry of these interactions are maintained for many soaked and docked fragments even as the residue moves relative to the rest of the protein.

In contrast to the adenosine site, little conformational heterogeneity is observed at the catalytic site, with only minimal changes in Lys102 and Tyr42 conformations (**Fig. 7G**). Still, even in this site, there is more conformational heterogeneity observed in previously published structures (**Fig. 7H**). In particular, a network of flexible side-chains encompassing Phe132, Asn99, and Lys102 is stabilized in a distinct conformation upon ADP-ribose binding (**Fig. 7I**). To further probe the flexibility of the Phe132-Asn99-Lys102 network, we determined structures of Mac1 using the C2 crystal at human physiological temperature (37°C, 310 K) to 1.5 Å resolution (**Fig. 7J, Data S1**). As observed in other systems (45, 46), we noticed that the cryogenic structure appeared more compact than the structure at higher temperatures. Specifically, we observed substantial loop displacements near the ribose-binding pocket of the active site, which are coupled to a global hinge-bending motion involving correlated motion of helices about the central β -sheet (**Fig. S4F,G**). The structure at physiological temperature more closely resembles the structure with ADPr-bound, with the backbone adopting a more open conformation (**Fig. 7J**). However, the side-chain rotamers of Asn99 and Lys102 do not undergo the larger rearrangements. This temperature-dependent change in the width of the active site cleft can provide alternative, potentially more relevant, conformations for future ligand-discovery efforts targeting the catalytic site around the distal ribose.

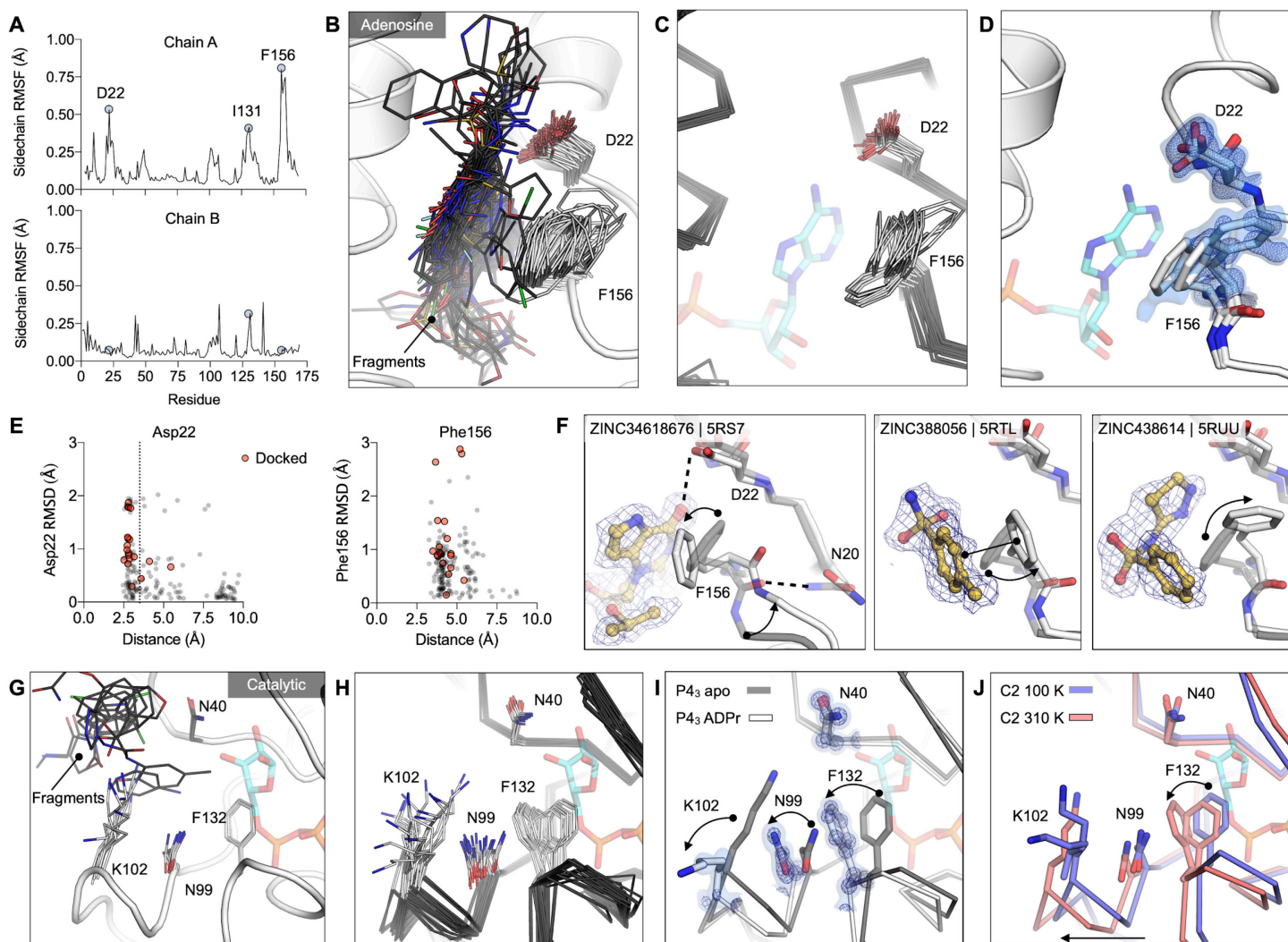


Figure 7. Experimentally observed conformational heterogeneity is sampled by various fragments.

A) Plots of side-chain RMSF for the 117 fragment structures from the UCSF screen using P4₃ crystals. **B**) Stick representation showing all fragments (black sticks) within 3.5 Å of the Asp22 carboxylate and 4 Å of the Phe156 ring (white sticks). **C**) Structural heterogeneity in the previously reported Mac1 structures. **D**) The Phe156 side-chain is captured in three conformations in the C2 apo structure. Electron density maps (2mF_O-DF_C) are contoured at 0.5 σ (blue surface) and 1 σ (blue mesh). For reference, ADP-ribose is shown with blue sticks. **E**) Plots of side-chain RMSD for Asp22 and Phe156 from the Mac1 apo structure as a function of ligand-protein distance. Structures were aligned by their Cα atoms, before RMSDs were calculated for the Asp22 carboxylate and the Phe156 aromatic carbons. **F**) Fragment binding exploits preexisting conformational heterogeneity in the Phe156 side-chain. The apo structure is shown with dark transparent gray sticks in each panel and the conformational changes are annotated with arrows. **G**) Stick representation showing all fragments (black sticks) in the outer subsite of the catalytic site. **H**) Conformational heterogeneity of residues in the catalytic site of the previously reported Mac1 crystal structures. **I**) ADP-ribose binding induces a coupled conformational change in the Phe132, Asn99 and Lys102 side-chains, as well as a 2 Å shift in the Phe132 loop. Electron density maps (2mF_O-DF_C) are contoured at 1.5 σ (blue surface) and 4 σ (blue mesh). **J**) Mac1 structures determined at 100 K and 310 K using C2 crystals.

Changes in water networks upon fragment binding

To assess the role of water networks in fragment binding, we first examined changes in water networks upon ADP-ribose binding. In the 0.85 Å P4₃ apo structure, the catalytic site contains 14 water molecules arranged

in an ordered network that connects the Gly47 loop and the Ile131 loop, with an arc formed around the Phe132 side-chain (**Fig. 8A**). In contrast, waters were more disordered in the adenosine site, with more diffuse electron density and higher B-factors (**Fig. 8A,C**). Upon ADP-ribose binding, five waters were displaced from the catalytic site and the water network was disrupted (**Fig. 8B**). This disruption is partly caused by altered conformation of the Phe132 and Asn99 side-chains, which break the network between residues Asn40 and Asn99. Conversely, the network in the adenosine site was stabilized in the Mac1-ADPr complex (**Fig. 8B**). The average B-factor decreased from 24 to 10 Å², and two networks connect the phosphate tunnel with the adenine/oxyanion subsites (**Fig. 8C**). Although the adenine moiety only forms two direct hydrogen bonds to protein, it has four additional contacts via bridging water molecules (**Fig. 8B**). Similar bridging waters were observed for fragments binding in the adenosine site including ZINC340465 (PDB 5RSV), which forms only one direct hydrogen bond to the protein, but has an extensive hydrogen-bond network via water molecules (**Fig. 8D**). Visualizing all water molecules within 3.5 Å of fragment atoms shows clusters near protein hydrogen bond acceptors and donors (**Fig. 8E**). Of particular interest is the cluster near the backbone carbonyl of Ala154. This site is occupied by a water molecule in the Mac1-ADPr structure, and is bridged by adenine derivatives such as ZINC340465 (PDB 5RSJ) (**Fig. 8D**). In addition, five fragments occupy this site directly (**Fig. 4A,D**), including the C2-amino-substituted adenine present in ZINC89254160_N3 (PDB 5RSJ, **Fig. 3D**). Extending fragments to displace the water molecules at other frequently populated sites could help to quantify the contribution of water networks to Mac1-binding, and to provide a test set for computational methods that seek to exploit solvent dynamics for ligand optimization (47, 48).

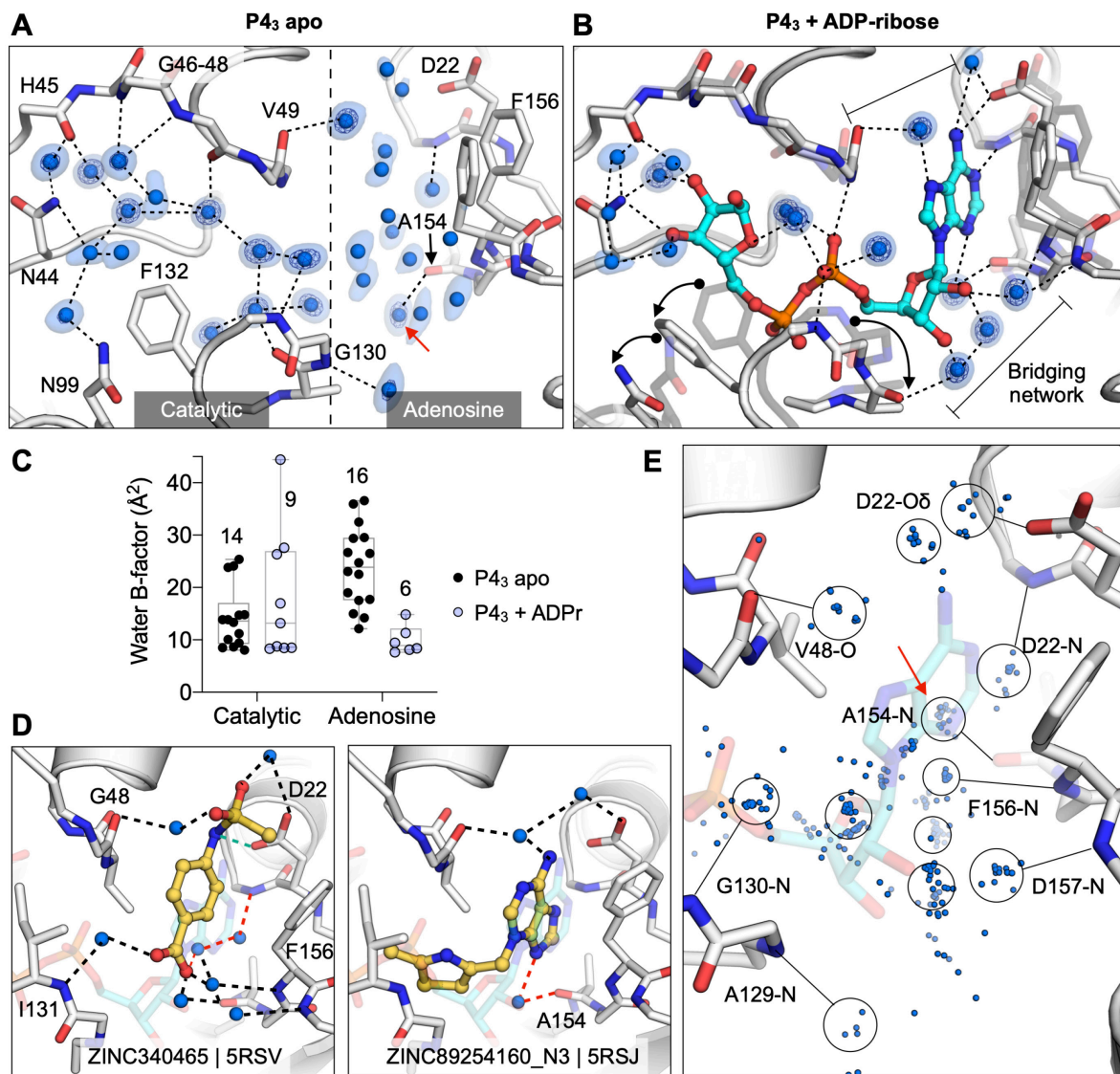


Figure 8. Water networks in the active site are displaced as well as used by fragments for bridging interactions.

A) Water networks in the apo enzyme (P4₃ crystal form). Waters are shown as blue spheres, with electron density contoured at 5.0 σ (blue mesh) and 1.5 σ (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. Left: ZINC340465 (PDB 5RSV) forms a single hydrogen bond to the protein (green dashed line), but forms five hydrogen bonds via water molecules. Right: although few fragments hydrogen bond directly to the backbone oxygen of Ala154, 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.

Solution binding of fragment hits

To buttress the crystallographic studies, selected compounds were biophysically screened using differential scanning fluorimetry (DSF), isothermal titration calorimetry (ITC), and a homogeneous time-resolved fluorescence (HTRF) ADPr-peptide displacement assay (**Fig. 9, Data S1, Data S2**). Because of their ready availability in useful amounts, most of these experiments focused on the docking hits. For DSF, in agreement with previous reports for this enzyme (18), we observed substantial elevation of the apparent melting temperature (T_{m_a}) upon addition of ADP-ribose (**Fig. 9C,D,G**). When tested in concentration-response from 0.188 to 3 mM, 10 of 54 docked fragments also induced small, but statistically significant and dose-responsive T_{m_a} elevation (**Fig. 9C,D,G, Data S1, Data S2**). All 10 of these were also observed to bind in the crystallographic studies, providing relatively good agreement between these assays. However, the correlation was incomplete, as the remaining fragments observed by crystallography either decreased the T_{m_a} or had no significant effect (**Data S1**).

To identify fragments with the most promising binding affinity for optimization, we tested the 19 crystallographically observed docking hits using ITC. Due to their small size, most of these fragments have low binding affinity and release little heat upon binding versus ADP-ribose. Thus, we only observed reliable thermodynamic measurements for four of the 19 fragments. These could be fit to a 1:1 binding stoichiometry with affinities in the low mM range (**Fig. 9E, Data S1, Data S2**), consistent with the DSF results. Furthermore, the compounds measured by ITC that released the greatest amount of heat also induced the most significant T_{m_a} shifts in DSF.

Finally, we tested 57 docking-derived fragments and 18 crystallographic hits from the XChem library in an HTRF-based peptide displacement assay, which monitors displacement of a fluorescently labeled ADPr-conjugated peptide from the active site of Mac1 (**Fig. 9F,G, Data S1, Data S2**). Eight of 57 docking hits (14%) and three of 18 crystallographic hits (17%) inhibited the enzyme with IC_{50} values between 180 μ M - 1 mM, with the most potent fragment being the docking-derived ZINC263392672 (PDB 5RSG) with an IC_{50} of 183 μ M in this assay. Only five (ZINC3888754 (PDB 5RSC), ZINC331945 (PDB 5RSD), ZINC263392672 (PDB 5RSG), ZINC336438345 (PDB 5RSE) and ZINC6180281(PDB 5RSF), **Fig. 3**) of the 10 docking hits that stabilized Mac1 as measured by DSF were inhibitory in the ADPr-peptide displacement assay. Interestingly, two docking hits that were not identified as binders by DSF or crystallography, ZINC1337772170 (IC_{50} = 971 μ M) and pterin (IC_{50} = 784 μ M), were found to be inhibitors in the peptide displacement assay (**Fig. 9H**). This result might be explained by the use of a detergent in the peptide displacement assay that could increase compound solubility. With its ability to detect inhibition of Mac1, the ADPr-peptide displacement assay proved to be a sensitive and complementary strategy for further characterization of the fragment hits obtained from the docking and

crystallographic screens. Assuming that the HTRF-based peptide displacement assay produced the most reliable inhibition data, we estimated ligand efficiencies from IC_{50} values for hits for which we obtained reasonable dose-response curves. ADP-ribose, with an IC_{50} of 161 nM and 36 heavy atoms, has a ligand efficiency (LE) of 0.26 kcal/mol per non-hydrogen atom. The docking hits ZINC3888754 (PDB 5RSC, LE=0.26), ZINC336438345 (PDB 5RSE, LE=0.28), ZINC263392672 (PDB 5RSG, LE=0.32) and ZINC331945 (PDB 5RSD, LE=0.38) reveal similar or slightly improved ligand efficiencies, while the highest LE was calculated for the XChem library hit SF005 (PDB 5S4G, **Fig. 9H**), with 0.44 kcal/mol per heavy atom.

In summary, all crystallographically confirmed docking hits were tested using three complementary in-solution binding techniques - DSF, ITC, and an HTRF-based peptide displacement assay (**Fig. S9, Data S1, Data S2**). ZINC336438345 (PDB 5RSE), ZINC331945 (PDB 5RSD), ZINC263392672 (PDB 5RSG) and ZINC26180281 (PDB 5RSF) were the only four fragment hits for which binding data could be obtained by all three techniques (**Fig. 9**). All of these fragments have key hydrogen bonds in the adenine subsite and π - π stack with Phe156. Furthermore, ZINC263392672 (PDB 5RSG) interacts via its carboxyl group with the oxyanion subsite of Mac1. Finally, we note that crystallography, DSF, ITC all monitor binding, but do not measure function. The peptide displacement assay is thus of particular value for fragment characterization, since it measures specific displacement of an analog of the natural Mac1 substrate.

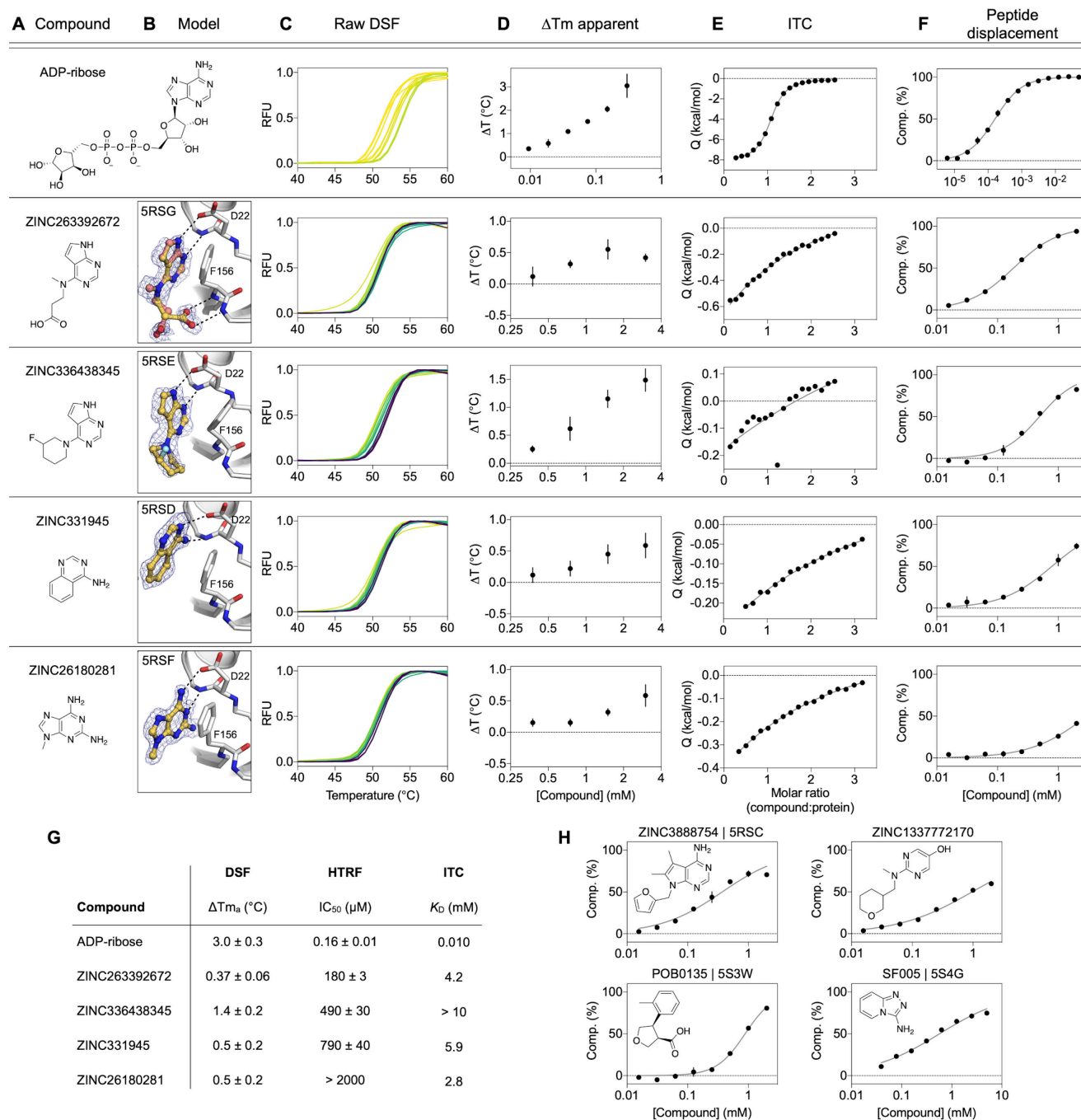


Figure 9. Biophysical corroboration of solution binding of crystallographic fragment hits by DSF, ITC and ADPr-peptide displacement assay.

Top panel (**A-F**) shows performance of the most potent fragment hits in DSF, ITC, and ADPr-peptide displacement assay compared to ADP-ribose. **C,D**) Normalized raw DSF RFU data demonstrates canonical unfolding curves and minimal compound-associated curve shape aberrations. T_m elevation reveals Mac1 stabilization through fragment binding. Gradient color scale: 0 mM = yellow; 3 mM = purple. **E**) Integrated heat peaks measured by ITC as a function of binding site saturation. The black line represents a non-linear least squares (NLLS) fit using a single-site binding model. **F**) Peptide displacement assay measures ADPr-peptide displacement (i.e. % competition) from Mac1 by ligand. **G**) Summary of solution binding data for fragments from

top panel. ΔTm_a are given for the highest compound concentration in this assay. **H)** Additional fragment hits showing Mac1 peptide competition.

Opportunities for fragment linking and merging to optimize Mac1 inhibitors

Typically, one might be reluctant to speculate on optimization from fragment structures alone, but the unusually large number of structures perhaps supports some cautious inference here. Prior to modifying, linking, or merging fragments, it is important to consider the crystalline environment. In the P4₃ crystal form, the active site forms a bipartite enclosed pocket with a symmetry mate (**Fig. 10A,B**). In particular, 24 fragments only hydrogen bond to Lys11 of the symmetry mate, and not with any residues in the adenosine site, indicating that these molecules should not be considered for fragment elaboration (**Fig. 10C,D**). Based on the binding poses of remaining compounds, fragment pairs were linked into hypothetical scaffolds. These were used as templates to search the make-on-demand chemical space of the Enamine REAL database employing the Smallworld similarity (<http://sw.docking.org>) and Arthor substructure (<http://arthor.docking.org>) search engines (**Fig. 10E,F**) (49). In a second approach, fragments with overlapping binding poses were merged into larger scaffolds, e.g. the purine of ZINC89254160_N3 (PDB 5RSJ) interacting in the adenine binding subsite was replaced by ZINC26180281 (PDB 5RSF) adding an additional hydrogen bond to Ala154 (**Fig. 10F**). Whereas it remains speculative whether the suggested linked or merged molecules are indeed active against Mac1, the scaffolds observed here, and the key interactions they make with the enzyme, indicate a fruitful chemical space to further explore. Naturally, many of the fragments described here also merit investigation by alternative fragment growing or analoging strategies.

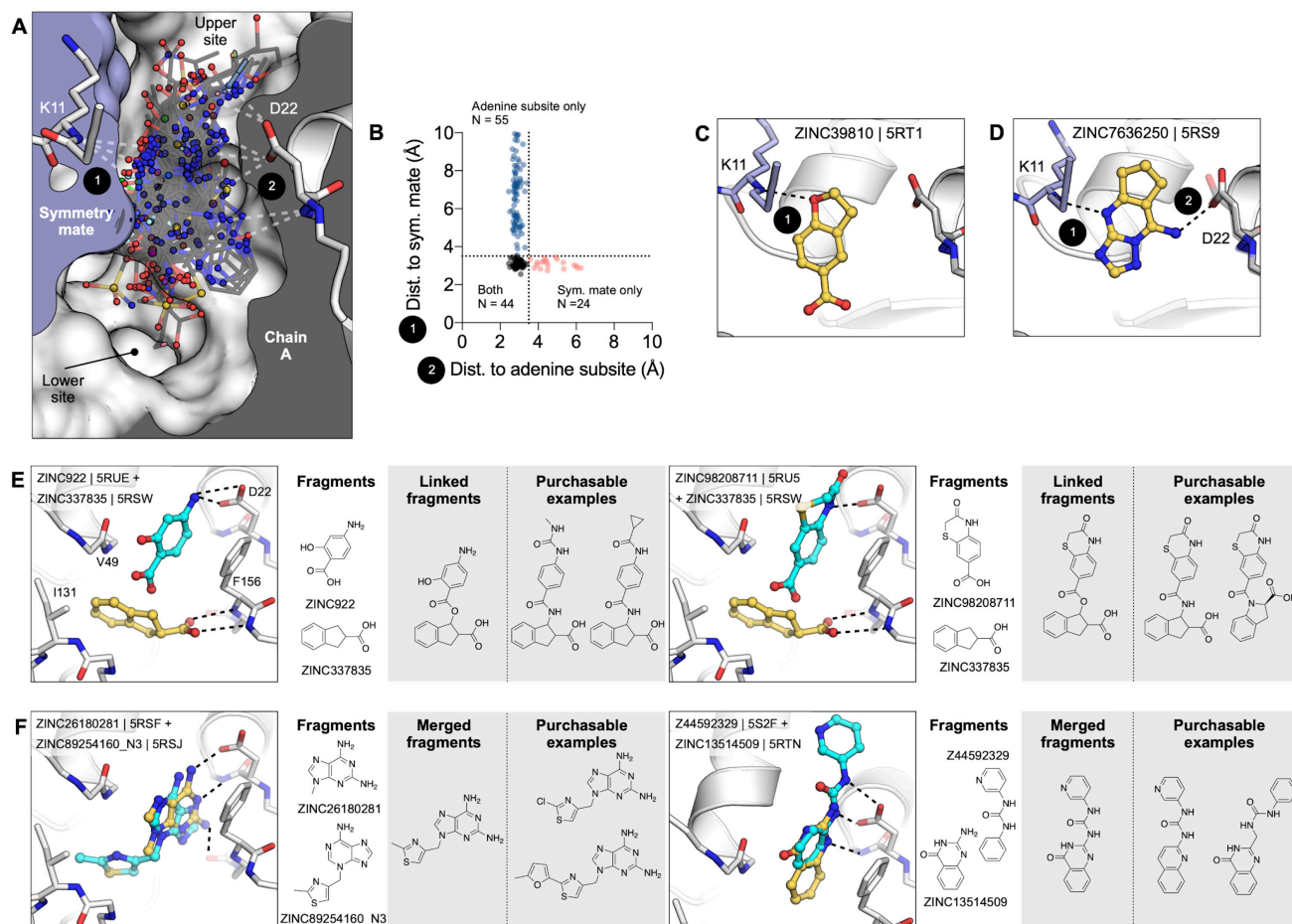


Figure 10. Fragments bridging multiple adenosine sites provide direct merging opportunities.

A) Sliced view of the adenosine site (white surface, grey interior) and a symmetry mate (blue surface and interior) showing the deep pocket created by crystal packing in the P4₃ crystals. The 66 fragments that hydrogen bond with the Lys11 backbone nitrogen are shown as sticks. **B)** Plot showing distances between the symmetry mate (Lys11-N) and the adenine subsite (Asp22-Oδ, Ile23-N, Ala154-O) for all fragments identified in the adenosine site. Dashed lines show the 3.5 Å cut-off used to classify hydrogen -bonds. **C)** An example showing one of the 24 fragments that bound in the adenosine site, yet only formed a hydrogen bond with the symmetry mate. **D)** An example of one of the fragments that bridged the 9-11 Å gap between the adenine subsite and the symmetry mate. **E, F)** Opportunities for fragment linking and merging. Adjacent or overlapping fragments were initially merged into a single new compound. Examples of readily available make-on-demand compounds are shown.

Discussion

Three key observations emerge from this study. Most noteworthy is the sheer number and the unusually high resolution of the 234 fragment-bound Mac1 structures, including 192 fragments identified in the active site. The fragments cover both stereotypical interactions (such as adenine-like hydrogen bonding to the Asp22 side chain/ Ile23 backbone and stacking interaction with Phe156) as well as diverse and unusual chemotypes that exploit active site flexibility (for instance by targeting the oxyanion-subsite). This abundance and diversity afford multiple starting points for future elaboration into lead-like molecules. Second, the high fidelity of docked poses to the subsequent crystallographic results supports the use of docking to explore the adenine recognition site, and importantly, demonstrates an ability of docking to prioritize fragments, at least for this target, something still debated in the field. Finally, with 234 diverse fragment structures determined, it should be possible to exploit the fortuitous juxtaposition of fragment pairs to design joined ligands that combine the affinities of both, leading to inhibitors with the low micromolar affinity needed for hit-to-lead optimization. One clear strategy involves extending molecules bound to the adenine subsite and with biophysically measurable binding affinities into the phosphate and ribose recognition regions.

In contrast to the large number of chemically diverse hits binding to the adenine subsite, the lack of fragments bound to the catalytic site is notable and may inform models of how ADP-ribosylated peptides bind to Mac1. The paucity of fragments is especially surprising given that three crystal environments (the A and B chains in the P4₃ crystal and the C2 crystal) were screened and that the site appears accessible in all lattices (**Fig. S2B**). The two major models for peptide-macrodomein interactions are either that the peptide binds along the widened cleft defined by Tyr42 and Lys102, or that it extends into solution through the flexible Gly46-48 loop (50). Indeed, we observe fragments that bind in both locations (**Fig. 6A**). Regardless of the binding mode, which could be distinct depending on the identity of the modified residue and target substrate, the lack of binding at this site suggests that the binding energy comes mostly from the ADP-ribose and not from the amino acids on the ADPr-conjugated protein. This hypothesis is also supported by the fact that Mac1 can hydrolyse a wide range of ADP-ribosylated substrates (2, 51). Docking of larger 'lead-like' molecules, perhaps enabled by the expanded catalytic site revealed by the physiological temperature structure, and detailed description of solvent, may help to identify molecules exploiting this site.

The success of the fragment docking campaign contrasts, perhaps, with expectations of the field that fragments have too few functional-group handles to accurately dock or prioritize (52). Not only were hit rates high (33%), so too was the fidelity of most docking poses to the crystallographic results. Even judged by potency, the most active fragment to emerge from this study, the 183 μM inhibitor ZINC263392672 (PDB 5RSG) (**Fig. 3, Fig. 9**), was a docking hit. Also, it was the docking hits that were most readily available for such functional testing, as they were sourced in 10 mg amounts, while the crystallographic screening compounds were often in short supply. This is a purely mechanical advantage of docking, and it is counterbalanced by the small numbers tested versus the crystallographic screens; still, having substantial material to work with is a pragmatic advantage. Admittedly, weaknesses also emerged from the docking. Intriguingly, the oxyanion site that featured so

prominently among the crystallographic screening hits were not to be found among the docking predictions. This gap reflects both a failure of the docking scoring function to prioritize anions binding to this site (as they were at least sampled), and to some extent a failure of the docking group to pick the few molecules that did dock well to this site as likely candidates. More broadly, as we docked against a single rigid structure of the protein, the subsequent conformational changes that the protein underwent, and the changes in the water network, were not captured in the docking predictions, and this was sometimes reflected in the larger RMSD differences between predicted and observed fragment poses (**Fig. 3**). These caveats, important as they are, should not obscure a central observation from this study: the docking hit rate was not only high, but the hits were typically right for the right reasons; this may be something to build on for the field.

From the docked compounds, the most promising hits identified by in-solution binding experiments were also crystallographically confirmed. However, as expected, the majority of hits from crystallography did not show appreciable activity in the orthogonal biophysical assays within the tested concentration range (up to 10 mM in ITC, **Data S1**). The macrodomain ADPr-peptide displacement assay also identified two docking hits not previously observed in soaking (ZINC1337772170 and pterin), which suggests that the crystal environment limited the ability of some fragments to bind. Yet, between solution experiments good consensus was observed for ZINC263392672 (PDB 5RSG), ZINC336438345 (PDB 5RSE) and ZINC331945 (PDB 5RSD). While we are aware that obtaining high-quality binding data remains particularly challenging for weak binders such as fragments, the dose-response results obtained in the complementary assays for many of the identified hits provided convincing evidence for their true binding to Mac1. The inconsistency of fragment binding to different crystal systems of the same protein is apparent when comparing fragments that resulted in high quality data sets in both the P4₃ and C2 crystal systems. Surprisingly, only five of 59 possible fragments were observed in both systems, with three fragments binding with equivalent poses in the adenine subsite. This observation points to the value of having multiple measurements, and even multiple crystal systems when they are available, in fragment-based drug discovery approaches.

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

Materials and Methods

Fragment libraries

We screened 2,122 molecules from the XChem facility at Diamond Light Source against the Mac1 P4₃ crystal form, and 411 molecules from UCSF against the C2 and P4₃ crystal forms (**Data S1**). The fragment library at XChem combined molecules from multiple fragment libraries: the Diamond, SGC and iNEXT (DSI)-poised Library (687 molecules (36)), the Edelris fragment collection (132 molecules), the MiniFragments Probing Library (80 molecules (53)), the FragLites collection (31 compounds (54)), the PepLite library (22 molecules (26)), the SpotFinder library (96 compounds), the York3D library (106 molecules (55)) and the EU Open screen (968 molecules). The UCSF fragment library was composed of Enamine's Essential Fragment library (320 compounds) and 91 additional compounds from an in-house library (UCSF_91). To assemble the UCSF_91 library, we selected topologically diverse molecules having over 10,000 commercially available analogs in at least three points of substitution, allowing for rapid and extensive analog-by-catalog without having to resort to flask synthesis. We picked molecules that were also Bemis-Murcko scaffolds (38), stripped of acyclic terminal substituents. We thought simple, unsubstituted frameworks would be easier to optimize by adding chemical matter during analoging. From among these, we prioritized by eye scaffolds with various ring sizes and combinations including fused rings, spiro systems, with linkers of varying lengths between rings, in an attempt to sample a diverse range of compact shapes and properties. We added anions where the anionic moiety was a small acyclic substituent on the scaffold, again picking by eye for shape diversity. We chose molecules with 11-21 heavy atoms, with molecular weights between 200-300 amu and with a logP < 2.5 for solubility. Physical properties of all screened libraries are shown in **Fig. S5**.

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

C2 crystals at UCSF

Protein expression and purification

SARS-CoV-2 Nsp3 Mac1 (residues 2-170) was cloned into a pET22b(+) expression vector with an N-terminal His₆ tag and a TEV protease recognition site for removal of the tag (GenScript). In addition, a short linker (Asn-Ala-Gly) was included between the TEV recognition site and the Mac1 gene (**Data S1**). To express Mac1, plasmid DNA was transformed into BL21(DE3) *E. coli*. After overnight growth on lysogeny broth (LB) agar supplemented with carbenicillin (100 ug/ml), starter cultures (10 ml LB) were grown at 37°C for 8 hours. Large scale cultures (1 l terrific broth (TB)) were grown at 37°C until an optical density of 0.8. Cultures were cooled at 4°C for 15 minutes, before protein expression was induced with 1 mM isopropyl 1-thio-D-galactopyranoside (IPTG), and the cultures were shaken at 20°C for 12 hours. Cells were collected by centrifugation and frozen at -80°C.

All purification steps were performed at 4°C using an AKTA FPLC system (Cytiva). Cells were resuspended in Ni-NTA binding buffer (50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10 mM imidazole, 5% glycerol, 2 mM βME supplemented with 5 units/ml TurboNuclease (Sigma, T4330)) and lysed by sonication. Cell debris was collected by centrifugation and the lysate was applied to a 5 ml HisTrap HP column (Cytiva, 17524802). The column was washed with 25 ml binding buffer followed by 25 ml 5% Ni-NTA elution buffer (50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 500 mM imidazole, 5% glycerol, 2 mM β-ME), and then eluted with 100% elution buffer. Eluted protein was exchanged into TEV reaction buffer (50 mM Tris (pH 8.0), 100 mM NaCl, 1 mM DTT and 1%

glycerol) using a HiPrep 26/10 desalting column (Cytiva, 17508701). To cleave the His₆ tag, Mac1 was diluted to 1.5 mg/ml using TEV reaction buffer and incubated with recombinant TEV protease (56) at a 1:20 ratio (Mac1:TEV) for 16 hours at 4°C. Cleaved Mac1 was separated from the uncleaved protein and TEV protease by re-running the sample over a HisTrap HP column (pre-equilibrated with TEV reaction buffer) and collecting the flow-through. The flow-through was supplemented with 10 mM DTT and concentrated to 2.5 ml using a 10 kDa molecular weight cut-off (MWCO) centrifugal concentrator (Amicon, UFC901024). The sample was further purified by size-exclusion chromatography (SEC) using a HiLoad 16/600 Superdex 75 pg column (Cytiva, 28989333) equilibrated with SEC buffer (20 mM Tris (pH 8.0), 150 mM NaCl, 5% glycerol, 2 mM DTT). Eluted fractions were concentrated to 15 mg/ml and flash-frozen in liquid nitrogen, and stored at -80°C. Protein used for ITC was purified in the same manner, but the SEC was run with 150 mM NaCl, 20 mM Tris (pH 8.0). Protein was concentrated to 10.8 mg/ml prior to flash freezing in liquid nitrogen and storage at -80°C.

Crystallization

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

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

Crystal dehydration and fragment soaking

Fragments were added to crystallization drops using acoustic dispensing with an Echo 650 liquid handler (Labcyte) (23). Two libraries were soaked at UCSF: the Enamine Essential fragment library (Enamine, 320 fragments), and the UCSF_91 library (91 fragments) (**Data S1**). To limit DMSO-induced crystal damage, fragments were targeted to crystallization drops as far away from crystals as possible (23). Initial DMSO tolerance tests indicated that the C2 crystals were sensitive, rapidly disintegrating upon soaking with 10% DMSO (**Fig. S2B**). To enhance DMSO tolerance, 300 nl of a solution containing 35% PEG 4000, 100 mM Tris (pH 8.5) and 100 mM sodium acetate was added to drops containing crystals using the Echo. Plates were resealed and incubated at 19°C for 6 hours. Fragment solutions (120 nl, 10% of the drop volume) were added using the Echo, and plates were re-sealed and incubated at 20°C for 3-8 hours. Crystals were vitrified directly from crystallization drops without additional cryoprotection.

Lysine methylation

Lysine methylation is a routine strategy for altering the crystallization properties of a protein (35). All reagents were added with the protein on ice and incubation steps were performed at 4°C with gentle shaking. First, 20 mg Mac1 was exchanged into lysine methylation buffer (50 mM HEPES (pH 7.5), 150 mM NaCl, 5% glycerol) using a HiPrep 26/10 desalting column. The protein was diluted to 1 mg/ml with lysine methylation buffer, and 400 µl 1 M dimethylamine borane (DMAB, prepared in water) (Sigma, 180238) and 800 µl 1 M formaldehyde (prepared in water) (Sigma, F8775) were added to initiate the methylation reaction. The reaction

was left to proceed for 2 hours, then 400 μ l 1 M DMAB and 800 μ l 1 M formaldehyde was added. After an additional 2 hours, 200 μ l 1 M DMAB was added and the reaction was left for a further 16 hours. To consume any remaining formaldehyde, and to cleave any intermolecular disulfide bonds, 2.5 ml of 1 M glycine (prepared in water) and 2.5 ml of 50 mM DTT (prepared in water) was added and the reaction was incubated for an additional 2 hours. Next, the sample was concentrated to 2.5 ml using a 10 kDa MWCO concentrator, and purified by SEC. The methylated protein was concentrated to 15 mg/ml before flash freezing in liquid nitrogen and storage at -80°C.

To test the extent of lysine methylation, the purified sample was analysed by liquid chromatography-mass spectrometry (LC-MS), using a Waters Acquity LC connected to a Waters TQ detector with electrospray ionization. The sample was separated on a C4 column held at 40°C using water with 0.1% formic acid as solvent A and acetonitrile with 0.1% formic acid as solvent B. After sample injection (5 μ l at 10 μ M diluted in 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 min. Then, a linear gradient elution was run for 6.5 min to 95% solvent B. Finally, an isocratic elution was run with 95% solvent B for 2 min. The flow rate was 0.2 ml/min.

Crystallization of methylated Mac1

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

Ultra high resolution data collection, refinement and modelling

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

The data sets were individually indexed and integrated with XDS (57). During data processing, we merged the high and low resolution datasets from multiple crystals in different orientations to maximize our coverage of reciprocal space given a square detector surface. A low-resolution cutoff of 2.5 Å was applied to the high-resolution (high energy) data sets, because this cutoff simultaneously excludes potentially overlapping reflections at low scattering angles and allows for a significant number of shared observations between high and low resolution data sets, which facilitates robust scaling. Scaling and merging were performed using XSCALE, and the merged intensities were converted to structure factor magnitudes using XDSCONV (57).

We calculated phases by the method of molecular replacement, using the program Phaser (58) and a previous structure of Mac1 (PDB 6WCF) as the search model. The model was manually adjusted in Coot (59) to fit the electron density map calculated from molecular replacement, followed by automated refinement of coordinates, atomic displacement parameters, and occupancies using phenix.refine (60) with optimization of restraint weights. Following two initial rounds of iterative model building and refinement using the aforementioned strategy, we began introducing additional parameters into the model, enabled by the extraordinarily high resolution of our diffraction data. First we implemented anisotropic atomic displacement parameters for heavy atoms (C,N,O,S), followed by refinement of explicit hydrogen atom positions. During early rounds of model building, we noticed

mF_o-DF_c difference density peaks appearing between heavy atom positions, suggesting that we are able to resolve covalent bonding densities (**Fig. S1E**). Indeed, atomic refinement that included a model for inter-atomic scatterers (IAS) (61) was able to account for these densities and reduce the free-R value by approximately 0.0043 (0.43%). Although the refined atomic coordinates do not differ significantly based on the inclusion or exclusion of IAS, the maximum-likelihood estimation of the phase error calculated by phenix.refine is 0.49° less when the IAS are included, suggesting an improvement in map quality (which may indirectly improve the model by aiding in subsequent manual interpretation of electron density features). Final refinement was performed without geometry or ADP weights (unrestrained).

Data collection at physiological temperature, refinement and modelling

We used a low-dose X-ray data collection strategy to acquire diffraction data from macrodomain crystals (C2 crystal form) at human physiological temperature (37°C, 310 K), which is the temperature most relevant to studies of SARS-CoV-2 infection. Using this strategy, we acquired data sets using an X-ray exposure of only 50 kGy - less than 1% of the total dose used at 100 K, which is essential to mitigate the rapid rate of radiation damage at 310 K compared to 100 K. The lower overall X-ray dose resulted in data with a lower overall resolution, extending to 1.5 Å.

Diffraction data from multiple crystals were merged using xia2 (62), implementing DIALS (63) for indexing and integration, and Aimless (64) for scaling and merging. We calculated phases by the method of molecular replacement, using the program Phaser (58) and our high resolution 100K structure as the search model. The model was manually adjusted in Coot to fit the electron density map calculated from molecular replacement, followed by automated refinement of coordinates, atomic displacement parameters, and occupancies using phenix.refine (60) with optimization of restraint weights.

Fragment data collection, refinement and modelling

Diffraction data was collected at ALS beamline 8.3.1 and SSRL beamlines 12-1 and 12-2. The data collection strategy is summarized in **Data S1**. Fragment datasets were indexed, integrated and scaled using XDS (57) run through xia2 (62). Based on the space group and unit cell dimensions, six crystal forms were present (**Fig. S2C**). For each of the three C2 isoforms with one molecule in the ASU (isoform A, B and C), a single, high resolution dataset was selected to create a representative model for each isoform. Phases were obtained via molecular replacement with Phaser (58), using the ultra-high resolution C2 coordinates as the search model (PDB 7KR0). Coordinates were refined with iterative rounds of manual model building in Coot and refinement with phenix.refine (60). Default refinement parameters were used, except five refinement macrocycles carried out per iteration and water molecules were automatically added to peaks in the $2mF_o-DF_c$ electron density map higher than 3.5 σ . The minimum model-water distance was set to 1.8 Å and a maximum model-water distance to 6 Å. For later rounds of refinement, hydrogens were added to riding positions using phenix.ready_set, and B-factors were refined anisotropically for non-hydrogen and non-water atoms. Although these datasets were obtained from crystals soaked with fragments, there was no evidence for fragment binding in the mF_o-DF_c difference density maps, therefore the datasets were deemed acceptable as representative DMSO-only models for each isoform.

For the fragment datasets, molecular replacement was performed with Phaser (58) and initial refinement with Refmac (65), both run through the DIMPLE pipeline (66). The search model used for molecular replacement was selected to match the isoform of the dataset. Waters were included in the initial refinement by changing the HOH records in the PDB file to WWW. After refinement, waters were stripped from models and electron density maps were analyzed for fragment binding using PanDDA (39). Electron density maps from 31 datasets were used to calculate the background electron density map for the A isoform, and 24 datasets were used for isoforms B and C (**Data S1**). Datasets selected for background map calculation had the highest resolution and lowest R_{free}

values. After PanDDA was run with default parameters, the threshold used to classify a hit was decreased by adjusting the Z-map analysis settings (contour_level = 2, min_blob_volume = 5, min_blob_z_peak = 2.5). Although there was a substantial increase in false positives, the decreased threshold allowed an additional seven fragments to be identified. Fragments were modelled into PanDDA event maps with COOT, using restraints generated by phenix.elbow from a SMILES string (67). Changes in protein conformation and solvation were also modeled. Because PanDDA can identify fragments binding with low occupancies, any changes in protein coordinates will have similar, low occupancies. If un-restrained refinement is performed on these low occupancy models, changes supported by PanDDA event maps are often reverted to the ground state model. In the past, this has been overcome by refining both ground-state (apo) and changed-state (fragment bound) structures simultaneously, with the changed state coordinates restrained. However, these multi-state models can be difficult to interpret. As an alternative, we modeled and refined the changed-state only. To prevent reversion of the model into ground state density, coordinate refinement was switched off after fragments were modelled. Hydrogens were added with phenix.ready_set, waters were updated automatically and B-factors were refined anisotropically for non-hydrogen and non-water atoms. After one round of refinement, waters added into ground state electron density were removed. This was achieved by aligning the DMSO-only model to the refined model, and removing any water molecules within 2.2 Å of the DMSO-only model. A final round of refinement was performed without updating water molecules.

P4₃ crystals at UCSF

Protein expression and purification

The C2 sequence in pET22b(+) was converted into the P4₃ sequence by removal of Glu170 and replacement of the N-terminal Asn-Ala-Gly-Glu motif with a methionine. Additionally, a Ser-Ser-Gly-Val-Asp-Leu-Gly-Thr linker was introduced between the His₆ tag and the TEV recognition sequence (**Data S1**). All cloning steps were performed by PCR with overlapping primers and Gibson assembly (68). Protein was purified using the same protocol as the C2 protein, except that after SEC, the protein was concentrated to 40 mg/ml prior to flash freezing in liquid nitrogen.

Crystallization

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

Fragment and ADP-ribose soaking

Fragment soaks were performed using the same protocol as the C2 crystals, with soak times between 2-6 hours. ADP-ribose soaks were performed similarly, except that ADP-ribose was prepared in water to 100 mM, and crystals were soaked with 80 nl ADPr (20 mM final concentration). Crystals were vitrified directly after soaking using a Nanuq cryocooling device (Mitegen).

Fragment data collection, processing, modelling and refinement

Diffraction data was collected at ALS beamline 8.3.1, SSRL beamline 12-1 and NSLS-II beamline 17-ID-2. The data collection strategy is summarized in **Data S1**. Fragment datasets were indexed, integrated and scaled using XDS (57) and merged with Aimless (64). In addition to the fragment soaks, we collected diffraction data for 40 crystals soaked only with DMSO. To generate a DMSO-only model, a single high resolution dataset was selected and phases were obtained by molecular replacement using the 0.77 Å C2 structure as a search model (PDB 7KR0). Refinement and model building was performed as described previously for the C2 crystals. The fragment datasets were prepared for PanDDA analysis using the DIMPLE pipeline(39, 66). Fragments were identified using PanDDA, with the background electron density map generated using 35 DMSO-only datasets (**Data S1**). As with the analysis of C2 electron density maps, PanDDA was re-run with a decreased Z-map threshold (contour_level = 2.5, min_blob_volume = 5, min_blob_z_peak = 2.5). This strategy identified an additional 24 fragments. Fragment modeling and refinement was carried out using the same protocol as the experiment with C2 crystals.

P4₃ crystals at Oxford/XChem

Protein expression and purification

SARS-CoV-2 Nsp3 Mac1 (residues 3-169) was cloned into a pNIC28-Bsa4 expression vector which adds an N-terminal His₆-tag and a TEV protease recognition site for removal of the tag. For expression of protein used for crystallisation, the constructs was transformed into the *E. coli* Rosetta strain BL21(DE3)-R3 and cells were grown at 37°C in LB medium (Miller) supplemented with 50 µg/ml of kanamycin and 35 µg/ml of chloramphenicol. After reaching an OD₆₀₀ of 0.5–0.6, the temperature was lowered to 18°C prior to induction of protein expression overnight by adding 0.5 mM IPTG. Harvested cells were resuspended in lysis buffer (50 mM HEPES (pH 7.5), 500 mM NaCl, 5% glycerol, 20 mM imidazole, 10 mM βME, cOmplete EDTA-free protease inhibitors (Roche)) and stored at -20°C until purification. For protein purification, pellets were gently thawed in lukewarm water and lysed by high-pressure homogenisation. DNA was digested using Benzonase. Proteins were purified by immobilised metal affinity chromatography (IMAC) using Ni-Sepharose resin (GE Healthcare) and eluted stepwise in binding buffer containing 40–500 mM imidazole. A high salt wash with 1 M NaCl was combined with the first elution step including 40 mM imidazole. Removal of the His₆ tag was carried out by addition of recombinant TEV protease during overnight dialysis into buffer without imidazole, followed by purification on a second IMAC column and finally by size-exclusion chromatography (SEC) (Superdex 75, GE Healthcare) in a buffer consisting of 20 mM HEPES (pH 8.0), 250 mM NaCl and 2 mM DTT. Macrodomein protein used for HTRF assay was not subjected to TEV cleavage and purified after the IMAC step by SEC in a buffer consisting of 25 mM HEPES (pH 7.4), 300 mM NaCl, 5% glycerol and 0.5 mM TCEP. Proteins were characterised by SDS-PAGE, then flash frozen in liquid nitrogen and stored at -80°C until required.

Crystallographic fragment screening

SARS-CoV-2 Nsp3 Mac1 was concentrated to a final concentration of 47 mg/ml and *apo* crystals were grown in crystallization solution containing 100 mM CHES (pH 9.5) and 30% PEG 3000. Fragments were soaked into crystals as previously described (23) by adding dissolved compounds directly to the crystallisation drops using an ECHO liquid handler (final concentration 10% DMSO); drops were incubated for approximately 1-3 hours prior to mounting and flash freezing in liquid nitrogen.

Data was collected at the beamline I04-1 at 100 K and automatically processed with Diamond Light Source's auto-processing pipelines using XDS (57) and either xia2 (62) or DIALS (63) with the default settings. Most Mac1 data processed to a resolution of approximately 1.1 Å. Further analysis was performed with XChemExplorer (24), electron density maps were generated with Dimple (66) and ligand-binding events were identified using PanDDA (39). Ligands were modelled into PanDDA-calculated event maps using Coot (59), restraints were calculated with AceDRG (69), and structures were refined with BUSTER (70). Coordinates, structure factors and PanDDA event maps for the structures discussed are deposited in the Protein Data Bank. Data collection and refinement statistics are summarised in **Data S1**.

Molecular Docking Screens

Docking was performed against the crystal structure of SARS-CoV-2 Nsp3 Mac1 bound to ADP-ribose (PDB 6W02 (34)). Chain B and all water molecules except for HOH324, HOH344, HOH384, and HOH406 were removed. These water molecules were included in the docking template structure since they were buried within the ADP-ribose binding site and formed bridging hydrogen bonds between ADP-ribose and the protein. The protein structure in complex with ADP-ribose and the four selected water molecules was capped at N- and C-termini and prepared for docking following the prepwizard protocol in Maestro (Schrödinger; (71)). Accordingly, protons were added using Epik and protonation states were optimized with PropKa at pH 7. Finally, the structure was energetically minimized using the OPLS3e force field (71). The maximum heavy-atom deviation from the initial structure was 0.3 Å (71).

Docking was performed with DOCK3.7 using pre-calculated scoring grids for rapid evaluation of docked molecules (72). AMBER united atom charges (73) were assigned to the minimized protein structure and water molecules. Partial atomic charges of backbone amide hydrogen atoms for residues Ile23 and Phe156 were increased by 0.2 elementary charge units without changing the net charge of the residues, as described previously (29). The low dielectric constant of the protein environment was extended outwards from the protein surface by 1.9 Å using spheres generated by SPHGEN. Electrostatic potentials at the ligand-binding pocket were calculated by numerical solution of the Poisson-Boltzmann equation using QNIFFT (74), scoring grids for van der Waals potentials were generated with CHEMGRID. Ligand desolvation scoring grids were calculated by SOLVMAP (75), and the volume of the low protein dielectric was extended out 0.4 Å from the protein surface, as described previously (40). Since we specifically targeted the adenosine binding site of the full ADP-ribose binding pocket, atomic coordinates of adenosine rather than the whole ADP-ribose molecule were used to generate 45 matching spheres, representing favorable positions for placing ligand atoms with docking (72).

As ADP-ribose was the only known ligand for Mac1 when we started the docking campaign, the generated scoring grids and matching spheres were judged for their ability to place and score adenosine, adenine and ribose at the adenosine binding site of the ligand binding pocket compared to 250 property-matched decoys, generated following the DUDE-Z method (76). Decoys share similar physical properties as the control molecules but are topologically different, hence unlikely to ligate the binding pocket. Furthermore, an "extrema" set (76) of approximately 500,000 molecules including anionic, neutral and cationic compounds with molecular weights ranging from 250-350 Da was screened to ensure similar enrichments for monovalent anions and neutral molecules. We note that the lack of experimentally confirmed ligands for the macrodomain did not allow exhaustive control calculations.

Virtual compound libraries were downloaded from ZINC15 (www.zinc15.docking.org) (37). From the set of 722,963 in-stock fragments, 696,092 compounds were successfully docked, exploring on average 2,355 orientations and 63 conformations per compound in the binding pocket. Roughly 58 billion complexes were

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

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

Fragment linking and merging

Fragment mergers and linkers were generated using Fragmenstein (<https://github.com/matteoferla/Fragmenstein>), a python module that automatically joins fragments or places compounds based on fragments in way that is as faithful to the positions of the fragments as possible in a conformation that is energy acceptable. For merging, using RDKit (77), rings are temporarily collapsed into pseudo-atoms, one-to-one spatial overlapping atoms are identified, pseudo-atoms expanded with appropriate bonds to nearby atoms and various chemical corrections applied. For the constrained energy minimization, Pyrosetta is used (78). Interactive online summary of mergers was made at <https://michelangelo.sgc.ox.ac.uk> (79).

Differential Scanning Fluorimetry (DSF)

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

DSF buffer was prepared by adding 10 µl of SYPRO Orange (Thermo Scientific, S6650) to 10 ml buffer (50 mM Tris HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.01% Triton X-100), for a final dye concentration of 5X (10 µM) SYPRO Orange. A compound plate (see above) was resuspended by the addition of 20 µl of DSF buffer, and set aside for 20 minutes in the dark. Purified Mac1 (P4₃ construct expressed at UCSF) was diluted to 10 µM in DSF buffer, and 2 µl of either protein solution or protein-free buffer was added to each well a 384-well white PCR plate (Axygen, PCR-384-LC480WNFBC) using an E1 ClipTip P125 electronic pipette. 8 µl of

resuspended compound was transferred to each well of the protein- and buffer-containing PCR plate using an Opentrons OT-2 liquid handling system, yielding the following final conditions: 2 μ M Mac1, 5X (10 μ M) SYPRO Orange, 3% DMSO, 0.1-3 mM fragments, and 0.1-1 mM ADP-ribose. The PCR plate was spun briefly in a salad spinner to remove bubbles, and sealed with optically clear film (Applied Biosystems, MicroAmp Optical Adhesive Film, 4311971). In an Analytik Jena qTower 384G qPCR instrument, plate was continuously heated from 25 - 94°C at a rate of 1°C/minute, and fluorescence was measured at each degree in the TAMRA channel (535 nm / 580 nm). 53 of 54 fragments could be tested up to 3 mM without assay interference in these conditions (**Data S1**, **Data S2**). $T_{m,a}$ s were calculated online at DSFworld, using fitting model 2 (80).

Raw DSF data for the Mac1 construct used in this work was characterized by a major transition at 50.8 $\pm$ 0.3°C, with a minor second transition at 67.0 $\pm$ 3.6°C (**Fig. 9C,D**, **Data S1**, **Data S2**); results described refer to the major transition. Significance was defined as compounds with ANOVA p-values < 0.05 for $T_{m,a}$ over the tested concentration regime.

Isothermal Titration Calorimetry (ITC)

All ITC titrations were performed on a MicroCal iTC 200 instrument (GE Healthcare). All reactions were performed in 20 mM Tris (pH 7.5), 150 mM NaCl using 300 - 600 μ M of Mac1 (P4₃ construct expressed at UCSF) at 25°C. Titration of 4 mM ADP-ribose (Sigma-Aldrich, A0752) or 4-10 mM fragment contained in the stirring syringe included a single 0.2 μ l injection, followed by 18 consecutive injections of 2 μ l. Thermodynamic parameters were obtained from a non-linear least squares (NLLS) fit of a single-site binding model in the RITC package (<https://rdrr.io/cran/Ritc>).

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

Fragment inhibitory activity on Mac1 was assessed by the displacement of an ADPr-conjugated biotin peptide from the His₆-tagged Mac1 using HTRF with a Eu³⁺-conjugated anti-His₆ antibody donor and streptavidin-conjugated acceptor. Compounds were dispensed into white ProxiPlate-384 Plus (PerkinElmer) assay plates using an Echo 525 liquid handler (Labcyte). Binding assays were conducted in a final volume of 16 μ l with 12.5 nM Mac1, 400 nM peptide ARTK(Bio)QTARK(Aoa-RADP)S (synthesized by Cambridge Peptides (Birmingham, UK)), 1:125 Streptavidin-XL665 (Cisbio), 1:20000 Anti-His₆-Eu³⁺ cryptate (PerkinElmer) in assay buffer (25 mM HEPES (pH 7.0), 20 mM NaCl, 0.05% BSA, 0.05% Tween20). Assay reagents were dispensed into plates using a Multidrop combi (Thermo Scientific) and incubated at room temperature for 1 h. Fluorescence was measured using a PHERAstar microplate reader (BMG) using the HTRF module with dual emission protocol (A = excitation of 320 nm, emission of 665 nm, and B = excitation of 320 nm, emission of 620 nm). Raw data were processed to give an HTRF ratio (channel A/B $\times$ 10,000), which was used to generate IC₅₀ curves by nonlinear regression using GraphPad Prism v8 (GraphPad Software, CA, USA).

References

1. J. G. M. Rack, D. Perina, I. Ahel, Macrodomains: Structure, Function, Evolution, and Catalytic Activities. *Annu. Rev. Biochem.* **85**, 431–454 (2016).
2. C. Li, Y. Debing, G. Jankevicius, J. Neyts, I. Ahel, B. Coutard, B. Canard, Viral Macro Domains Reverse Protein ADP-Ribosylation. *J. Virol.* **90**, 8478–8486 (2016).
3. A. R. Fehr, S. A. Singh, C. M. Kerr, S. Mukai, H. Higashi, M. Aikawa, The impact of PARPs and ADP-ribosylation on inflammation and host-pathogen interactions. *Genes Dev.* **34**, 341–359 (2020).
4. Y. M. O. Alhammad, A. R. Fehr, The Viral Macrodomain Counters Host Antiviral ADP-Ribosylation. *Viruses*. **12** (2020), doi:10.3390/v12040384.
5. A. R. Fehr, R. Channappanavar, G. Jankevicius, C. Fett, J. Zhao, J. Athmer, D. K. Meyerholz, I. Ahel, S. Perlman, The Conserved Coronavirus Macrodomain Promotes Virulence and Suppresses the Innate Immune Response during Severe Acute Respiratory Syndrome Coronavirus Infection. *MBio*. **7** (2016), doi:10.1128/mBio.01721-16.
6. A. R. Fehr, J. Athmer, R. Channappanavar, J. M. Phillips, D. K. Meyerholz, S. Perlman, The nsp3 macrodomain promotes virulence in mice with coronavirus-induced encephalitis. *J. Virol.* **89**, 1523–1536 (2015).
7. M. E. Grunewald, Y. Chen, C. Kuny, T. Maejima, R. Lease, D. Ferraris, M. Aikawa, C. S. Sullivan, S. Perlman, A. R. Fehr, The coronavirus macrodomain is required to prevent PARP-mediated inhibition of virus replication and enhancement of IFN expression. *PLoS Pathog.* **15**, e1007756 (2019).
8. L. Palazzo, P. Mikolčević, A. Mikoč, I. Ahel, ADP-ribosylation signalling and human disease. *Open Biol.* **9**, 190041 (2019).
9. G. Caprara, E. Prosperini, V. Piccolo, G. Sigismondo, A. Melacarne, A. Cuomo, M. Boothby, M. Rescigno, T. Bonaldi, G. Natoli, PARP14 Controls the Nuclear Accumulation of a Subset of Type I IFN-Inducible Proteins. *J. Immunol.* **200**, 2439–2454 (2018).
10. A. R. Fehr, G. Jankevicius, I. Ahel, S. Perlman, Viral Macrodomains: Unique Mediators of Viral Replication and Pathogenesis. *Trends Microbiol.* **26**, 598–610 (2018).
11. J. Cui, F. Li, Z.-L. Shi, Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192 (2019).
12. J. G. M. Rack, V. Zorzini, Z. Zhu, M. Schuller, D. Ahel, I. Ahel, Viral macrodomains: a structural and evolutionary assessment of the pharmacological potential. *Open Biol.* **10**, 200237 (2020).
13. K. K. Eriksson, L. Cervantes-Barragán, B. Ludewig, V. Thiel, Mouse Hepatitis Virus Liver Pathology Is Dependent on ADP-Ribose-1"-Phosphatase, a Viral Function Conserved in the Alpha-Like Supergroup. *JVI.* **82**, 12325–12334 (2008).
14. A. Putics, W. Filipowicz, J. Hall, A. E. Gorbalenya, J. Ziebuhr, ADP-ribose-1"-monophosphatase: a conserved coronavirus enzyme that is dispensable for viral replication in tissue culture. *J. Virol.* **79**, 12721–12731 (2005).
15. D. I. James, K. M. Smith, A. M. Jordan, E. E. Fairweather, L. A. Griffiths, N. S. Hamilton, J. R. Hitchin, C. P. Hutton, S. Jones, P. Kelly, A. E. McGonagle, H. Small, A. I. J. Stowell, J. Tucker, I. D. Waddell, B. Waszkowycz, D. J. Ogilvie, First-in-Class Chemical Probes against Poly(ADP-ribose) Glycohydrolase (PARG) Inhibit DNA Repair with Differential Pharmacology to Olaparib. *ACS Chem. Biol.* **11**, 3179–3190 (2016).

16. A. I. J. Stowell, D. I. James, I. D. Waddell, N. Bennett, C. Truman, I. M. Hardern, D. J. Ogilvie, A high-throughput screening-compatible homogeneous time-resolved fluorescence assay measuring the glycohydrolase activity of human poly(ADP-ribose) glycohydrolase. *Anal. Biochem.* **503**, 58–64 (2016).
17. M. Schuller, K. Riedel, I. Gibbs-Seymour, K. Uth, C. Sieg, A. P. Gehring, I. Ahel, F. Bracher, B. M. Kessler, J. M. Elkins, S. Knapp, Discovery of a Selective Allosteric Inhibitor Targeting Macrodomain 2 of Polyadenosine-Diphosphate-Ribose Polymerase 14. *ACS Chem. Biol.* **12**, 2866–2874 (2017).
18. R. S. Viridi, R. V. Bavisotto, N. C. Hopper, D. N. Frick, Discovery of Drug-like Ligands for the Mac1 Domain of SARS-CoV-2 Nsp3. *bioRxiv* (2020), doi:10.1101/2020.07.06.190413.
19. M. M. Hann, A. R. Leach, G. Harper, Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **41**, 856–864 (2001).
20. C. W. Murray, D. C. Rees, The rise of fragment-based drug discovery. *Nat. Chem.* **1**, 187–192 (2009).
21. T. Krojer, J. S. Fraser, F. von Delft, Discovery of allosteric binding sites by crystallographic fragment screening. *Curr. Opin. Struct. Biol.* **65**, 209–216 (2020).
22. D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **15**, 605–619 (2016).
23. P. M. Collins, J. T. Ng, R. Talon, K. Nekrosiute, T. Krojer, A. Douangamath, J. Brandao-Neto, N. Wright, N. M. Pearce, F. von Delft, Gentle, fast and effective crystal soaking by acoustic dispensing. *Acta Crystallogr D Struct Biol.* **73**, 246–255 (2017).
24. T. Krojer, R. Talon, N. Pearce, P. Collins, A. Douangamath, J. Brandao-Neto, A. Dias, B. Marsden, F. von Delft, The XChemExplorer graphical workflow tool for routine or large-scale protein-ligand structure determination. *Acta Crystallogr D Struct Biol.* **73**, 267–278 (2017).
25. N. D. Wright, P. Collins, R. Talon, E. Nelson, L. Koekemoer, M. Ye, R. Nowak, J. Newman, J. T. Ng, N. Mitrovich, H. Wiggers, F. von Delft, The Low-Cost, Semi-Automated Shifter Microscope Stage Transforms Speed and Robustness of Manual Protein Crystal Harvesting. *Cold Spring Harbor Laboratory* (2019), p. 2019.12.20.875674.
26. A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C. David Owen, E. Resnick, C. Strain-Damerell, A. Aimon, P. Ábrányi-Balogh, J. Brandão-Neto, A. Carbery, G. Davison, A. Dias, T. D. Downes, L. Dunnett, M. Fairhead, J. D. Firth, S. Paul Jones, A. Keely, G. M. Keserü, H. F. Klein, M. P. Martin, M. E. M. Noble, P. O'Brien, A. Powell, R. Reddi, R. Skyner, M. Snee, M. J. Waring, C. Wild, N. London, F. von Delft, M. A. Walsh, Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease (2020), p. 2020.05.27.118117.
27. R. J. Hall, P. N. Mortenson, C. W. Murray, Efficient exploration of chemical space by fragment-based screening. *Prog. Biophys. Mol. Biol.* **116**, 82–91 (2014).
28. A. Manglik, H. Lin, D. K. Aryal, J. D. McCorvy, D. Dengler, G. Corder, A. Levit, R. C. Kling, V. Bernat, H. Hübner, X.-P. Huang, M. F. Sassano, P. M. Giguère, S. Löber, Da Duan, G. Scherrer, B. K. Kobilka, P. Gmeiner, B. L. Roth, B. K. Shoichet, Structure-based discovery of opioid analgesics with reduced side effects. *Nature*. **537**, 185–190 (2016).
29. J. Lyu, S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O'Meara, T. Che, E. Algaa, K. Tolmachova, A. A. Tolmachev, B. K. Shoichet, B. L. Roth, J. J. Irwin, Ultra-large library docking for discovering new chemotypes. *Nature*. **566**, 224–229 (2019).
30. R. M. Stein, H. J. Kang, J. D. McCorvy, G. C. Glatfelter, A. J. Jones, T. Che, S. Slocum, X.-P. Huang, O. Savych, Y. S. Moroz, B. Stauch, L. C. Johansson, V. Cherezov, T. Kenakin, J. J. Irwin, B. K. Shoichet, B. L. Roth, M. L. Dubocovich, Virtual discovery of melatonin receptor ligands to modulate circadian rhythms.

Nature. **579**, 609–614 (2020).

31. Y. Bian, X.-Q. S. Xie, Computational Fragment-Based Drug Design: Current Trends, Strategies, and Applications. *AAPS J.* **20**, 59 (2018).
32. Y. Chen, B. K. Shoichet, Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.* **5**, 358–364 (2009).
33. D. G. Teotico, K. Babaoglu, G. J. Rocklin, R. S. Ferreira, A. M. Giannetti, B. K. Shoichet, Docking for fragment inhibitors of AmpC beta-lactamase. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7455–7460 (2009).
34. K. Michalska, Y. Kim, R. Jedrzejczak, N. I. Maltseva, L. Stols, M. Endres, A. Joachimiak, Crystal structures of SARS-CoV-2 ADP-ribose phosphatase: from the apo form to ligand complexes. *IUCrJ.* **7**, 814–824 (2020).
35. T. S. Walter, C. Meier, R. Assenberg, K.-F. Au, J. Ren, A. Verma, J. E. Nettleship, R. J. Owens, D. I. Stuart, J. M. Grimes, Lysine methylation as a routine rescue strategy for protein crystallization. *Structure*. **14**, 1617–1622 (2006).
36. O. B. Cox, T. Krojer, P. Collins, O. Monteiro, R. Talon, A. Bradley, O. Fedorov, J. Amin, B. D. Marsden, J. Spencer, F. von Delft, P. E. Brennan, A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.* **7**, 2322–2330 (2016).
37. T. Sterling, J. J. Irwin, ZINC 15--ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).
38. G. W. Bemis, M. A. Murcko, The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* **39**, 2887–2893 (1996).
39. N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M. Deane, F. von Delft, A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat. Commun.* **8**, 15123 (2017).
40. M. M. Mysinger, D. R. Weiss, J. J. Ziarek, S. Gravel, A. K. Doak, J. Karpiak, N. Heveker, B. K. Shoichet, B. F. Volkman, Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5517–5522 (2012).
41. W. J. Allen, R. C. Rizzo, Implementation of the Hungarian algorithm to account for ligand symmetry and similarity in structure-based design. *J. Chem. Inf. Model.* **54**, 518–529 (2014).
42. D. Ramírez, J. Caballero, Is It Reliable to Take the Molecular Docking Top Scoring Position as the Best Solution without Considering Available Structural Data? *Molecules*. **23**, 1038 (2018).
43. L. Xing, J. Klug-Mcleod, B. Rai, E. A. Lunney, Kinase hinge binding scaffolds and their hydrogen bond patterns. *Bioorg. Med. Chem.* **23**, 6520–6527 (2015).
44. A. Narunsky, A. Kessel, R. Solan, V. Alva, R. Kolodny, N. Ben-Tal, On the evolution of protein-adenine binding. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 4701–4709 (2020).
45. J. S. Fraser, H. van den Bedem, A. J. Samelson, P. T. Lang, J. M. Holton, N. Echols, T. Alber, Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16247–16252 (2011).
46. D. A. Keedy, Z. B. Hill, J. T. Biel, E. Kang, T. J. Rettenmaier, J. Brandão-Neto, N. M. Pearce, F. von Delft, J. A. Wells, J. S. Fraser, An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. *Elife*. **7** (2018), doi:10.7554/eLife.36307.

47. A. S. Bayden, D. T. Moustakas, D. Joseph-McCarthy, M. L. Lamb, Evaluating Free Energies of Binding and Conservation of Crystallographic Waters Using SZMAP. *J. Chem. Inf. Model.* **55**, 1552–1565 (2015).
48. D. Cappel, W. Sherman, T. Beuming, Calculating Water Thermodynamics in the Binding Site of Proteins - Applications of WaterMap to Drug Discovery. *Curr. Top. Med. Chem.* **17**, 2586–2598 (2017).
49. J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield, R. A. Sayle, ZINC20-A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *J. Chem. Inf. Model.* (2020), doi:10.1021/acs.jcim.0c00675.
50. G. Jankevicius, M. Hassler, B. Golia, V. Rybin, M. Zacharias, G. Timinszky, A. G. Ladurner, A family of macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nat. Struct. Mol. Biol.* **20**, 508–514 (2013).
51. D. Munnur, E. Bartlett, P. Mikočević, I. T. Kirby, J. G. M. Rack, A. Mikoč, M. S. Cohen, I. Ahel, Reversible ADP-ribosylation of RNA. *Nucleic Acids Res.* **47**, 5658–5669 (2019).
52. B. Lamoree, R. E. Hubbard, Current perspectives in fragment-based lead discovery (FBLD). *Essays Biochem.* **61**, 453–464 (2017).
53. M. O'Reilly, A. Cleasby, T. G. Davies, R. J. Hall, R. F. Ludlow, C. W. Murray, D. Tisi, H. Jhoti, Crystallographic screening using ultra-low-molecular-weight ligands to guide drug design. *Drug Discov. Today*. **24**, 1081–1086 (2019).
54. D. J. Wood, J. D. Lopez-Fernandez, L. E. Knight, I. Al-Khawaldeh, C. Gai, S. Lin, M. P. Martin, D. C. Miller, C. Cano, J. A. Endicott, I. R. Hardcastle, M. E. M. Noble, M. J. Waring, FragLites-Minimal, Halogenated Fragments Displaying Pharmacophore Doublets. An Efficient Approach to Druggability Assessment and Hit Generation. *J. Med. Chem.* **62**, 3741–3752 (2019).
55. T. D. Downes, S. P. Jones, H. F. Klein, M. C. Wheldon, M. Atobe, P. S. Bond, J. D. Firth, N. S. Chan, L. Waddelove, R. E. Hubbard, D. C. Blakemore, C. De Fusco, S. D. Roughley, L. R. Vidler, M. A. Whatton, A. J.-A. Woolford, G. L. Wrigley, P. O'Brien, Design and Synthesis of 56 Shape-Diverse 3D Fragments. *Chemistry*. **26**, 8969–8975 (2020).
56. J. E. Tropea, S. Cherry, D. S. Waugh, Expression and purification of soluble His(6)-tagged TEV protease. *Methods Mol. Biol.* **498**, 297–307 (2009).
57. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
58. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
59. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
60. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
61. P. V. Afonine, R. W. Grosse-Kunstleve, P. D. Adams, V. Y. Lunin, A. Urzhumtsev, On macromolecular refinement at subatomic resolution with interatomic scatterers. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 1194–1197 (2007).
62. G. Winter, xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2009).
63. G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero,

- M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter, G. Evans, DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr D Struct Biol.* **74**, 85–97 (2018).
64. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
 65. G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
 66. M. Wojdyr, R. Keegan, G. Winter, A. Ashton, DIMPLE- a pipeline for the rapid generation of difference maps from protein crystals with putatively bound ligands. *Acta Crystallographica Section A Foundations of Crystallography.* **69** (2013), pp. s299–s299.
 67. N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1074–1080 (2009).
 68. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison 3rd, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods.* **6**, 343–345 (2009).
 69. F. Long, R. A. Nicholls, P. Emsley, S. Gra  ulis, A. Merkys, A. Vaitkus, G. N. Murshudov, AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr D Struct Biol.* **73**, 112–122 (2017).
 70. G. Bricogne, Direct phase determination by entropy maximization and likelihood ranking: status report and perspectives. *Acta Crystallogr. D Biol. Crystallogr.* **49**, 37–60 (1993).
 71. G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234 (2013).
 72. R. G. Coleman, M. Carchia, T. Sterling, J. J. Irwin, B. K. Shoichet, Ligand pose and orientational sampling in molecular docking. *PLoS One.* **8**, e75992 (2013).
 73. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, P. Weiner, A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).
 74. K. Gallagher, K. Sharp, Electrostatic contributions to heat capacity changes of DNA-ligand binding. *Biophys. J.* **75**, 769–776 (1998).
 75. M. M. Mysinger, B. K. Shoichet, Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* **50**, 1561–1573 (2010).
 76. R. M. Stein, Y. Yang, T. E. Balius, M. J. O'Meara, J. Lyu, J. Young, K. Tang, B. K. Shoichet, J. J. Irwin, Property-unmatched decoys in docking benchmarks. *Understanding virtual solvent through large-scale ligand discovery*, 114 (2020).
 77. P. Tosco, N. Stiefl, G. Landrum, Bringing the MMFF force field to the RDKit: implementation and validation. *J. Cheminform.* **6**, 37 (2014).
 78. S. Chaudhury, S. Lyskov, J. J. Gray, PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics.* **26**, 689–691 (2010).
 79. M. P. Ferla, A. T. Pagnamenta, D. Damerell, J. C. Taylor, B. D. Marsden, MichelaNglo: sculpting protein views on web pages without coding. *Bioinformatics.* **36**, 3268–3270 (2020).

80. T. Wu, J. Yu, Z. Gale-Day, A. Woo, A. Suresh, M. Hornsby, J. E. Gestwicki, Three Essential Resources to Improve Differential Scanning Fluorimetry (DSF) Experiments. *Cold Spring Harbor Laboratory* (2020), p. 2020.03.22.002543.
81. A. Morin, B. Eisenbraun, J. Key, P. C. Sanschagrin, M. A. Timony, M. Ottaviano, P. Sliz, Cutting edge: Collaboration gets the most out of software. *Elife*. **2**, e01456 (2013).
82. M. N. S. Rad, S. Behrouz, E. Zarenezhad, N. Kaviani, Highly efficient protocol for one-pot N-alkylation of nucleobases using alcohols in bmim[Br]: a rapid route to access acyclic nucleosides. *Journal of the Iranian Chemical Society*. **12** (2015), pp. 1603–1612.

Acknowledgements

General: We thank all the staff of Diamond Light Source for providing support and encouragement which allowed us to carry out this work during the COVID-19 lockdown. We also thank Gemma Davison, Selma Dorman, James Sanderson, Matthew Martin, Mike Waring and Martin Noble (CRUK Newcastle Drug Discovery Unit, Newcastle University), Thomas Downes, Paul Jones, Hanna Klein, James Firth (York University), David Bajusz and Gyorgy Keseru (Hungarian Academy of Sciences) for providing fragment libraries. We also acknowledge EU-OPENSOURCE ERIC for providing its fragment library for the presented scientific work. EU-OPENSOURCE ERIC has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement No 823893 (EU- OPENSOURCE-DRIVE). Structural biology applications used at UCSF were compiled and configured by SBGrid (81). X-ray diffraction data processed at UCSF used the Wynton high-performance compute cluster.

Funding: This work was supported by NIH GM123159, NSF Rapid 2031205, and a TMC Award from the UCSF Program for Breakthrough Biomedical Research, funded in part by the Sandler Foundation (to J. Fraser); by the Wellcome Trust (grants 101794 and 210634), BBSRC (BB/R007195/1) and Cancer Research UK (C35050/A22284) (to I. Ahel); NIH R35GM122481 and DARPA HR0011-19-2-0020 (to B. Shoichet); GM071896 (to J. Irwin). The crystallographic screen at Oxford was supported by the XChem facility at Diamond Light Source (proposal ID MX27001). The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA Darmstadt Germany, MSD, Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome [106169/ZZ14/Z]. Beamline 8.3.1 at the Advanced Light Source is operated by the University of California Office of the President, Multicampus Research Programs and Initiatives grant MR-15-328599, NIH (R01 GM124149 and P30 GM124169), Plexxikon Inc., and the Integrated Diffraction Analysis Technologies program of the US Department of Energy Office of Biological and Environmental Research. The crystallographic fragment screens at UCSF were carried out using beamlines at the Advanced Light Source, the Stanford Synchrotron Radiation Lightsource and the National Synchrotron Light Source II. The Advanced Light Source (Berkeley, CA) is a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the US Department of Energy under contract number DE-AC02-05CH11231, Office of Basic Energy Sciences. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). Extraordinary SSRL operations were supported in part by the DOE Office of Science through the National Virtual Biotechnology Laboratory, a consortium of DOE national laboratories focused on response to COVID-19, with funding provided by the Coronavirus CARES Act. This research used beamline 17-ID-2 of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. The Center for BioMolecular Structure (CBMS) is primarily supported by the National Institutes of Health, National Institute of General Medical Sciences (NIGMS) through a Center Core P30 Grant (P30GM133893), and by the DOE Office of Biological and Environmental Research (KP1605010). The QCRG Structural Biology Consortium has received support from: Quantitative Biosciences Institute, Defense Advanced Research Projects Agency HR0011-19-2-0020 (to D. Agard. and K. Verba; B. Shoichet PI), FastGrants COVID19 grant (K. Verba PI), Laboratory For Genomics Research (O. Rosenberg PI) and Laboratory for Genomics Research LGR-ERA (R. Stroud PI). R. Díaz and T. Wu were supported by NSF GRFP. R. Díaz is a Howard Hughes Medical Institute Gilliam Fellow. I. Young was supported by NIH F32GM133129. M. Ferla was supported

by the Wellcome Trust 203141/Z/16/Z and the NIHR Biomedical Research Centre Oxford. L. Martins was supported by CAPES-PrInt 88887.364628/2019-00.

Authors contributions: M. Schuller designed and cloned the construct that yielded the P4₃ crystals at Oxford/XChem; expressed, purified and establishing crystallization conditions for P4₃ crystals at Oxford/XChem; assisted with data processing/analysis at XChem; set up the HTRF functional assay and performed data interpretation; prepared the manuscript. G. Correy cloned, expressed and purified the P4₃ construct at UCSF; crystallized, performed fragment soaking, vitrified crystals, collected X-ray diffraction data and processed data for fragment screens at UCSF; modelled, refined and analyzed fragment structures at UCSF; purified, methylated and crystallised the C2 construct; performed ADP-ribose soaks; refined the apo and ADP-ribose bound structures determined using the P4₃ crystals; prepared the manuscript. S. Gahbauer performed docking screens against Mac1; performed the chemoinformatic analysis of fragment libraries; assisted with fragment-linking and -merging; prepared the manuscript. D. Fearon crystallized, prepared samples, collected X-ray diffraction data, refined and analyzed fragment structures at XChem; prepared the manuscript. T. Wu performed and analyzed DSF experiments; prepared the manuscript. R. Díaz designed the construct that yielded the C2 crystals; expressed the C2 construct; performed and analyzed ITC experiments; prepared the manuscript. I. Young collected X-ray diffraction data and processed the diffraction data for crystals screened at UCSF. L. Martins assisted with docking screens against Mac1. D. Smith assisted with DSF experiments. U. Schulze-Gahmen crystallized the C2 construct at UCSF. T. Owens purified the C2 construct at UCSF. I. Deshpande collected X-ray diffraction data for C2 crystals at UCSF. G. Merz purified the C2 construct at UCSF. A. Thwin purified the C2 construct at UCSF. J. Biel supported the fragment soaking experiments performed at UCSF; supported the fragment modeling and refinement at UCSF. J. Peters purified C2 construct at UCSF. M. Moritz purified C2 construct at UCSF. N. Herrera supported crystallization of the C2 construct at UCSF. H. Kratochvil supported crystallization of the C2 construct at UCSF. QCRG Structural Biology Consortium provided infrastructure and support for experiments performed at UCSF. A. Aimon prepared samples and collected X-ray diffraction data at XChem. J. Bennett set up the HTRF functional assay; analyzed and collected data for the HTRF assay. J. Neto collected X-ray diffraction data at XChem. A. Cohen supported X-ray diffraction experiments at the SSRL. A. Dias collected X-ray diffraction data at XChem. A. Douangamath crystallized, prepared samples, collected X-ray diffraction data, refined and analyzed fragment structures at XChem. L. Dunnett collected X-ray diffraction data at XChem. O. Fedorov set up the HTRF functional assay; analyzed and collected data for the HTRF assay. M. Ferla performed fragment merging and linking. M. Fuchs supported X-ray diffraction experiments at the NSLS-II. T. Gorrie-Stone deposited fragments structures at XChem. J. Holton supported X-ray diffraction experiments at the ALS. M. Johnson analyzed and assigned structures to the N3- and N9-alkylated fragments. T. Krojer analyzed structural data at XChem. G. Meigs supported X-ray diffraction experiments at the ALS. A. Powell collected X-ray diffraction data at XChem. J. Rack provided feedback on the manuscript. V. Rangel refined fragment structures at XChem. S. Russi supported X-ray diffraction experiments at the SSRL. R. Skyner deposited fragments structures at XChem. C. Smith supported X-ray diffraction experiments at the SSRL. A. Soares supported X-ray diffraction experiments at the NSLS-II. J. Wierman supported X-ray diffraction experiments at the SSRL. K. Zhu expressed and crystallized the P4₃ construct at XChem. P. O'Brien designed and provided the York3D fragment library. N. Jura supervised work. A. Ashworth supervised work; prepared the manuscript. J. Irwin designed the UCSF_91 fragment library; prepared the manuscript. M. Thompson vitrified C2 crystals at UCSF; collected X-ray diffraction data for C2 crystals at UCSF; refined the ultra-high-resolution structure determined using C2 crystals; prepared the manuscript. J. Gestwicki supervised work. F. Delft provided the XChem facility; prepared the manuscript; supervised work. B. Shoichet guided and evaluated the docking work; prepared the manuscript; supervised work. J. Fraser supervised work; prepared the manuscript; arranged funding. I. Ahel supervised work; prepared the manuscript; arranged funding.

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

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

Supplementary materials for “Fragment binding to the Nsp3 macrodomain of SARS-CoV-2 identified through crystallographic screening and computational docking”

This PDF file includes:

Supplementary Text
Figs. S1 to S9

Other Supplementary Materials for this manuscript include the following:

Data  and 

Supplementary Text

Purity and structure determination of fragments ZINC901391520, ZINC82473428 and ZINC89254160 from Enamine

Samples of ZINC901391520, ZINC82473428 and ZINC89254160 obtained from Enamine were expected to be N₉-alkylated isomers but electron density of the fragments in X-ray structures indicated these samples were N₃-alkylated isomers instead (ZINC901391520_N3, ZINC82473428_N3 and ZINC89254160_N3, see **Fig. S7I-L**). The original samples of ZINC901391520, ZINC82473428 and ZINC89254160 used in fragment screening by X-ray crystallography were analyzed by HPLC-MS and ¹H NMR to confirm sample purity and corroborate structure. There is no reported characterization data to be used as reference for structure confirmation for N₉- or N₃-alkylated compounds ZINC901391520 and ZINC89254160. The N₉-alkylated structure ZINC82473428 is a previously prepared compound with tabulated NMR data reported by Rad *et al.* (82).

A re-supplied sample of ZINC901391520 from a new batch synthesized at Enamine was confirmed by ¹H NMR to be >95% purity and a different isomer than the original sample of ZINC901391520. The X-ray crystal structure of this fragment in complex with Mac1 revealed the fragment to be N₉-alkylated isomer (**Fig. S7I**).

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

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

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

NMR experiments for samples ZINC901391520, ZINC82473428 and ZINC89254160

Original samples of ZINC901391520, ZINC82473428 and ZINC89254160 from Enamine used in the fragment screen were dissolved in *d*₆-DMSO and analyzed by ¹H and ¹³C NMR on a Bruker 400 MHz instrument with Avance III electronics. Data was obtained at ambient temperature (ca. 25°C) collecting 64 scans for proton experiments and 1024 scans for carbon experiments. Raw data was processed and reports created using ACD Spectrus software.

Original sample ZINC901391520

A sample of 5.5 mg ZINC901391520 was dissolved in 0.75 ml *d*₆-DMSO for NMR analysis and from this solution 50 µl was diluted in 0.45 ml acetonitrile to make up the analytical sample for HPLC-MS using HPLC Method A. The sample chromatogram from HPLC revealed a single peak with UV absorbance at both 214 and 254 nm at *t*_R = 5.272 minutes. Aside from a very strong UV₂₁₄ peak at *t*_R = 2.00 minutes attributed to DMSO co-solvent in the sample, no other peaks were observed at these UV wavelengths and sample purity estimated >98% based on UV peak area. ¹H NMR (400 MHz, *d*₆-DMSO, 25°C) δ ppm 8.49 (s, 1H), 7.91-8.26 (br d, 2H), 7.76 (s, 1H), 6.30 (s, 1H), 5.63 (s, 2H), 3.86 (s, 3H). ¹³C NMR (101 MHz, *d*₆-DMSO, 25°C) δ ppm 172.05, 167.57, 155.01, 152.46, 149.39, 143.55, 120.18, 94.85, 57.13, 44.32. LRMS (ESI⁺) for peak at *t*_R = 5.272 minutes: observed *m/z* = 247.3 [MH]⁺ for C₁₀H₁₀N₆O₂ exact mass = 246.09.

Second batch sample ZINC901391520

A sample was dissolved in 0.75 ml d_6 -DMSO for NMR analysis. ^1H NMR (400 MHz, d_6 -DMSO, 25°C) δ ppm 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).

Sample ZINC82473428

A sample of 3.9 mg ZINC82473428 was dissolved in 0.75 ml d_6 -DMSO for NMR analysis and from this solution 50 μl was diluted in 0.45 ml acetonitrile to make up the analytical sample for HPLC-MS using HPLC Method B. The sample chromatogram from HPLC revealed a single peak with UV absorbance at both 214 and 254 nm at $t_R = 3.766$ minutes. Aside from a very strong UV₂₁₄ peak at $t_R = 2.00$ minutes attributed to DMSO cosolvent in the sample no other peaks were observed at these UV wavelengths and sample purity estimated >98% based on UV peak area. ^1H NMR (400 MHz, d_6 -DMSO, 25°C) δ ppm 8.31 (s, 1H), 8.01 (br 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, 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}C NMR (101 MHz, d_6 -DMSO, 25°C) δ ppm 154.78, 151.53, 149.56, 144.30, 75.35, 67.24, 52.54, 40.44, 28.23, 25.03. LRMS (ESI⁺) for peak at $t_R = 3.766$ minutes: observed $m/z = 220.3$ [MH]⁺ for C₁₀H₁₃N₅O exact mass = 219.11.

Reported NMR data for compound ZINC82473428_N9 from Rad et al., 2015 (82): ^1H NMR (400 MHz, d_6 -DMSO, 25°C) δ 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), 1.30-1.54 (complex m, 4H). ^{13}C NMR (101 MHz, d_6 -DMSO, 25°C) δ ppm 156.6, 152.9, 149.2, 144.7, 117.2, 80.6, 67.9, 57.8, 29.1, 25.1.

Sample ZINC89254160

A sample of 3.2 mg ZINC89254160 was dissolved in 0.75 ml d_6 -DMSO for NMR analysis and from this solution 50 μl was diluted in 0.45 ml acetonitrile to make up the analytical sample for HPLC-MS using HPLC Method A. The sample chromatogram from HPLC revealed a major peak and a minor peak with UV absorbances at both 214 and 254 nm: major peak $t_R = 6.530$ minutes and minor peak $t_R = 6.751$ minutes. Relative peak area calculated as percentage of combined UV peak area at 254 nm was 93.3% major peak and 6.7% minor peak (corresponds to ca. 14:1 ratio). Aside from a very strong UV₂₁₄ peak at $t_R = 2.00$ minutes attributed to DMSO cosolvent in the sample no other peaks were observed at these UV wavelengths. Tabulated NMR data reported here for major peaks only. ^1H NMR (400 MHz, d_6 -DMSO, 25°C) δ ppm 8.47 (s, 1H), 7.95 (br s, 2H), 7.73 (s, 1H), 7.47 (s, 1H), 5.55 (s, 2H), 2.60 (s, 3H). ^{13}C NMR (101 MHz, d_6 -DMSO, 25°C) δ ppm 166.21, 154.93, 152.47, 149.63, 149.51, 143.63, 120.43, 117.69, 48.08, 18.66. LRMS (ESI⁺) for major peak at $t_R = 6.530$ minutes: observed $m/z = 247.3$ [MH]⁺ for C₁₀H₁₀N₆S exact mass = 246.07. LRMS (ESI⁺) for minor peak at $t_R = 6.751$ minutes: observed $m/z = 247.3$ [MH]⁺ for C₁₀H₁₀N₆S exact mass = 246.07.

Major peak and minor peak have the same observed mass peak in LRMS and are presumed to be different N-alkylated isomers.

Conclusions based on HPLC-MS and NMR characterization of samples ZINC901391520, ZINC82473428 and ZINC89254160

HPLC-MS data confirmed that samples ZINC901391520 and ZINC82473428 are single compounds >98% purity with mass peak corresponding to either N₉- or N₃-alkylated isomers. Both ^1H and ^{13}C NMR data corroborated initial samples ZINC901391520 and ZINC82473428 are >98% single compound. The very high purity determined for these two samples rules out the possibility that the structures determined by X-ray crystallography were the result of trace amounts of the alternative isomer in the samples. For ZINC89254160, HPLC-MS data confirmed that there was a 13:1 ratio of isomers in this sample and it is possible that the X-ray crystal structure of Mac1 obtained with ZINC89254160 was the result of protein complexed to trace/minor amounts of the alternative isomer (N3-alkylated).

The NMR data obtained for sample ZINC82473428 used in crystallographic fragment screen does not match NMR data reported in the literature for the N₉-alkylated ZINC82473428 and thus this sample is presumed not to

be N₉-alkylated isomer. NMR data is not sufficient to unambiguously assign N₃- or N₉-alkylated structures for ZINC901391520, ZINC82473428 or ZINC89254160 and the unambiguous structure assignment of ZINC901391520, ZINC82473428 and ZINC89254160 as N₃-alkylated isomers in this work was provided by the electron density observed for these fragments in the Mac1 X-ray crystal structures.

The crystal structure of Mac1 with ZINC400552187 additionally revealed the N₃-alkylated structure instead of the requested N₉-alkylated form. Using DSF and ITC, ZINC901391520, ZINC82473428, ZINC89254160, ZINC400552187 were initially screened as the N₃-alkylated isomer (ZINC901391520_N3 (PDB 5RSK), ZINC82473428_N3 (PDB 5RVF), ZINC89254160_N3 (PDB 5RSJ), ZINC400552187_N3 (PDB 5RVG)). In addition the N₉-alkylated ZINC901391520 (PDB 5S6W) was tested in DSF and the peptide-competition assay (HTRF) (**Data S1**).

QCRG Structural Biology Consortium authorship

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

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

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

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

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

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

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

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

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

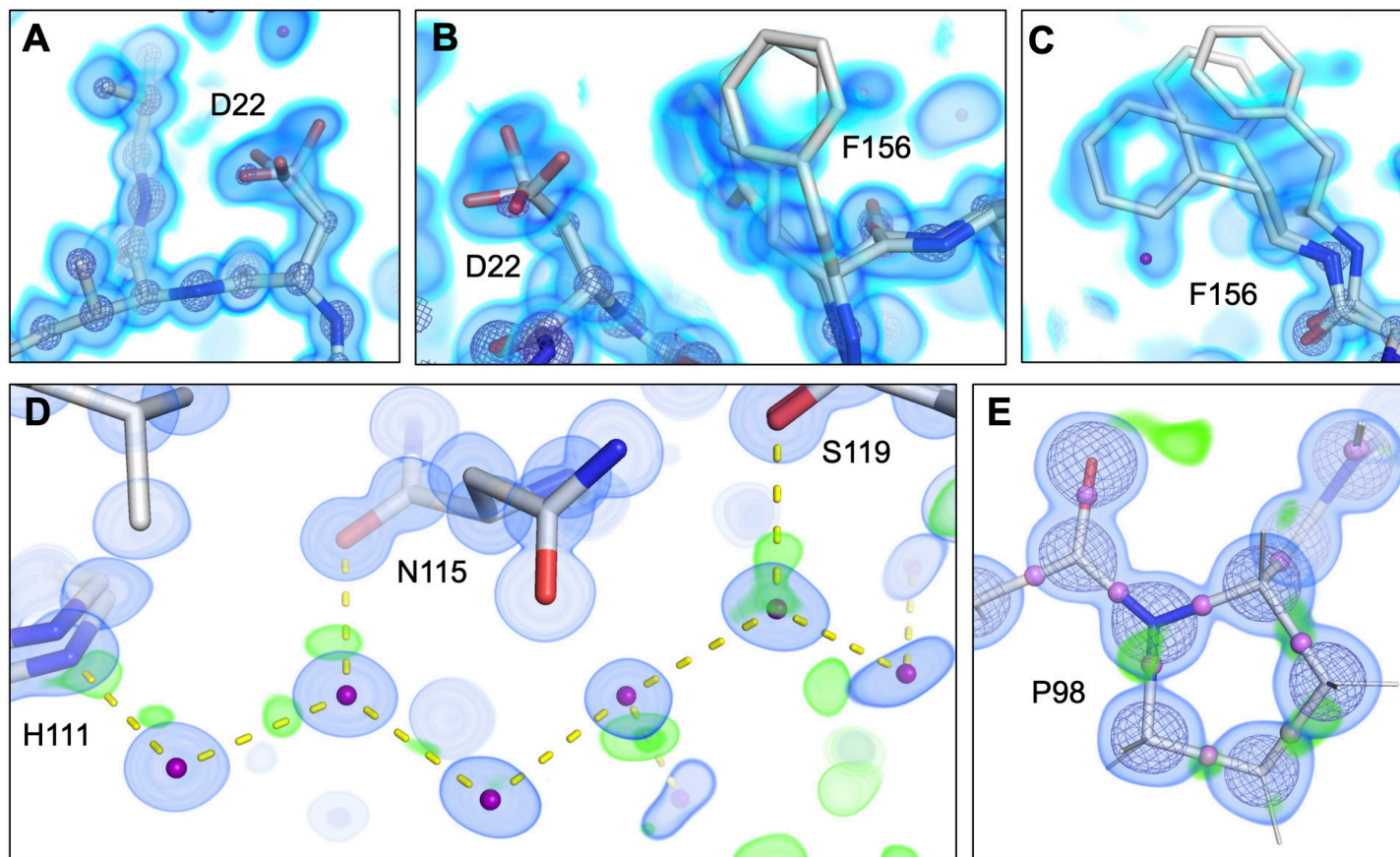


Fig. S1. Ultra-high resolution features in Mac1 electron density maps.

A,B,C) Residues Asp22 and Phe156 both display conformational heterogeneity in high-resolution electron density maps of apo Mac1. The ultra-high resolution nature of the electron density is evident in (A), where atoms belonging to residues Ile23 and Leu24, as well as the backbone, appear as separate spheres at high contour. In (A,B,C) three $2mF_o-DF_c$ maps are shown, contoured at 4.0σ (blue mesh), 1.0σ (blue volume), and 0.5σ (cyan volume). **D)** Mac1 electron density reveals ordered water networks, including difference density that suggests the positions of hydrogen atoms within the water network. In (D), a $2mF_o-Df_c$ map is shown, contoured at 1.5σ (blue volume), and a mF_o-Df_c map is also shown, contoured at 2.5σ (green volume). Hydrogen bonding interactions are depicted as dashed yellow lines. **E)** Two electron density maps are shown for P98, a $2mF_o-Df_c$ map, contoured at 1.5σ (blue volume), and a mF_o-Df_c map, contoured at 2.0σ (green volume). The green peaks of positive difference density between heavy atom positions suggest covalent bonding densities, and we refined a model of Mac1 that included interatomic scatterers (IAS), shown in violet in the figure.

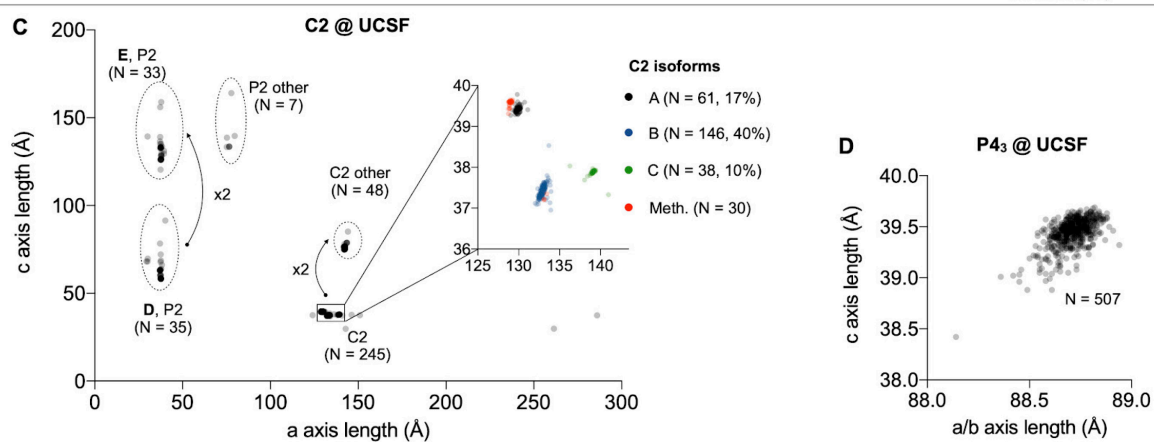
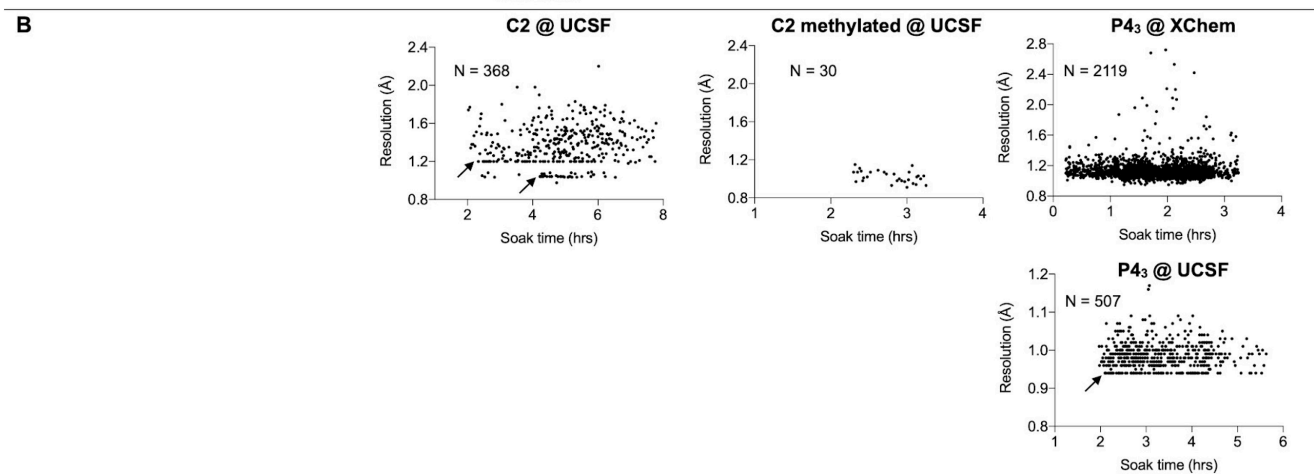
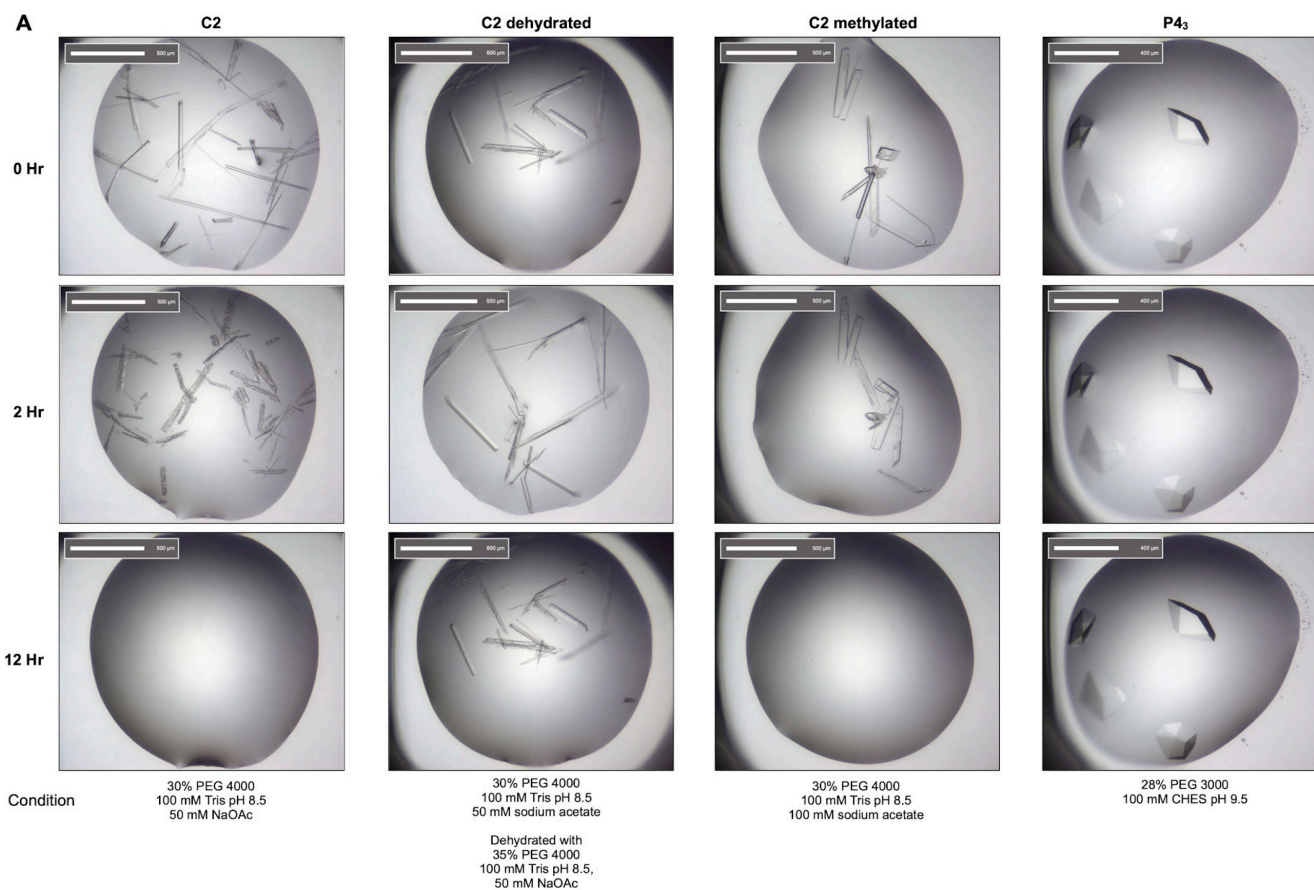


Fig. S2. Comparison of isomorphism and DMSO tolerance of the C2 and P4₃ crystals.

A) Images of crystals after soaking with 10% DMSO for 0, 2 and 12 hours. **B)** Resolution of the three crystal forms as a function of soak time for the datasets collected at XChem and UCSF. The arrows indicate where the measurement of high resolution reflections was limited by the experimental setup. **C)** Multiple isoforms were observed for the C2 crystals after dehydration. Isoforms were distinguished based on differences in the a and c unit cell lengths. Arrows indicate where doubling of the a or c axis occurred. Inset: the majority of the datasets that were indexed in C2 (245, 84%) could be clustered into three isoforms (A, B and C). Of the 30 datasets collected for crystals grown from methylated protein, the majority (28) were similar to the A isoform. **D)** The P4₃ crystals were isomorphous.

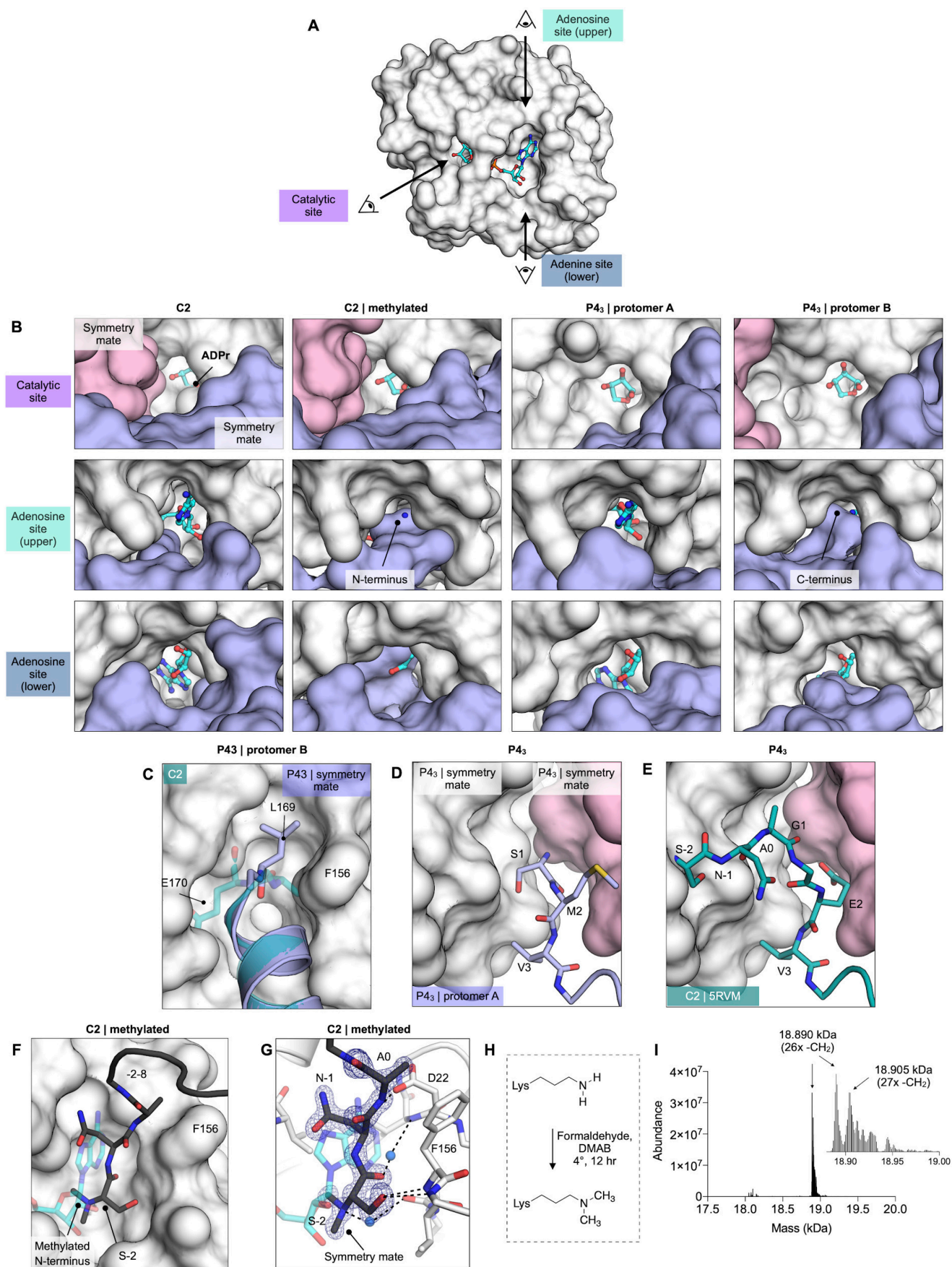


Fig. S3. Crystal packing in Mac1 crystals determines active site accessibility.

A) Active site access in the C2 and P4₃ crystals. Mac1 is shown as a white surface with ADP-ribose bound in the active site shown as cyan sticks. The three access points are indicated with arrows. **B)** Crystal packing defines the three access points. The catalytic site is partially obstructed in the C2 crystals, but open in both protomers of the P4₃ crystals. In both the methylated C2 crystals and protomer B of the P4₃ crystals, the adenosine site is obstructed. **C)** The C-terminal leucine (Leu169) of the P4₃ construct occupies the adenosine site of a symmetry mate. The adenosine site is shown as a white surface and the C-terminal residues with blue sticks/cartoon. The C2 sequence (transparent teal cartoon/sticks) has an additional residue at the C-terminus (Glu170), and is therefore incompatible with the P4₃ crystal packing. **C)** The N-terminal residues of the P4₃ sequence (blue sticks) pack between two symmetry mates (white and pink surface). Compared to the P4₃ sequence, the C2 sequence contains a substitution (Met2Glu) and a three residue insertion (Asn-Ala-Gly). These residues were typically disordered, however, they were resolved in one of the fragment structures (ZINC157088 | 5RVM) (shown aligned to the P4₃ protomer A in **(E)**). Like differences in the C-termini, differences in the N-termini may have contributed to the distinct crystal packing seen for the two Mac1 structures reported in this work. **F)** The adenosine site was obstructed by a symmetry mate in the structure determined from crystals grown using methylated C2 protein. **G)** In the structure of methylated Mac1, the side-chain hydroxyl of Ser-2 occupies the oxyanion subsite. Electron density ($2mF_o - DF_c$) is shown as a blue mesh, contoured at 1.5σ . **H)** Free amines were methylated using formaldehyde and dimethylamine borane (DMAB). The reaction is shown for lysine, however, based on the electron density shown in **(G)**, the N-terminal amine was methylated as well. The methylated amines would be protonated at the pH used to grow crystals (pH 8.5). **I)** LC/MS analysis of methylated Mac1 (C2 construct). The mass spectrum was deconvoluted using MaxEnt1. The major peak (18.89 kDa) is consistent with the methylation of 13 lysine residues ($26 \times -CH_2$). The minor peak (18.905 kDa + 15 Da) suggests that methylation was not 100% complete.

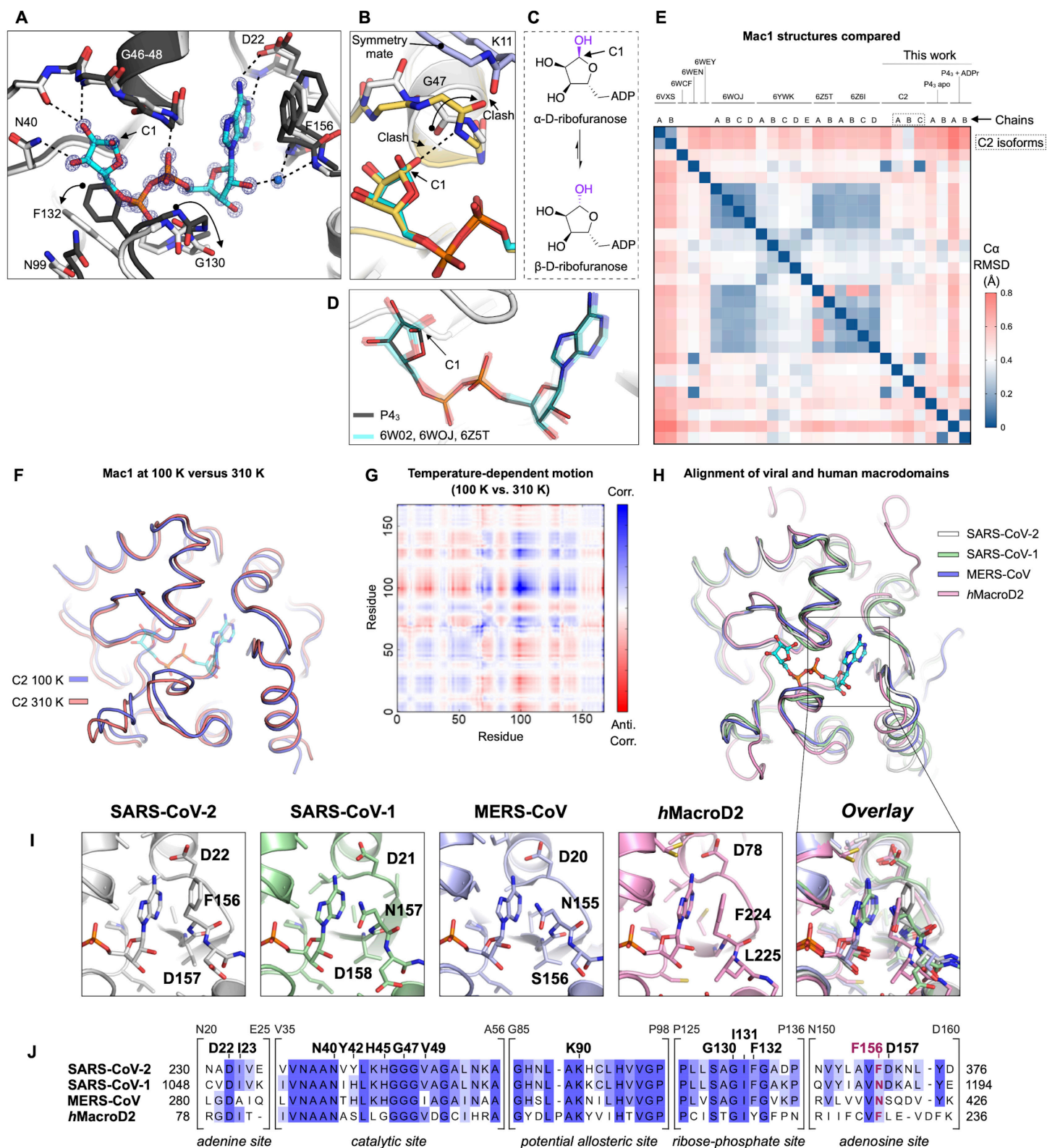
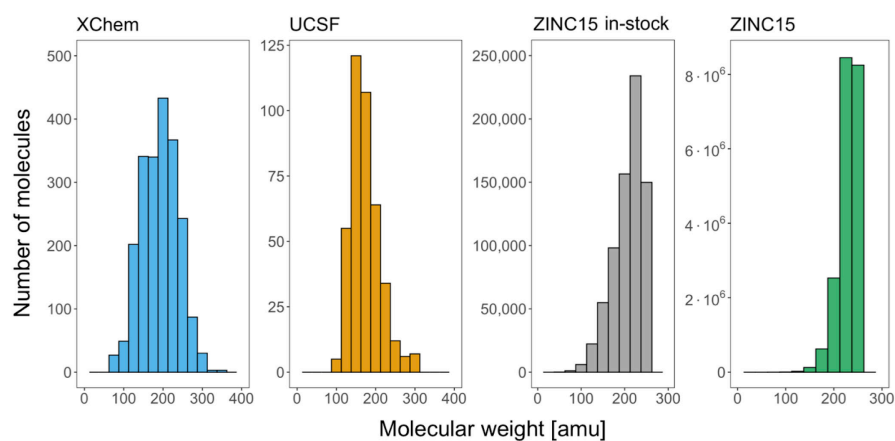


Fig. S4. Structure and sequence comparison of Mac1 with related viral and human macrodomains.

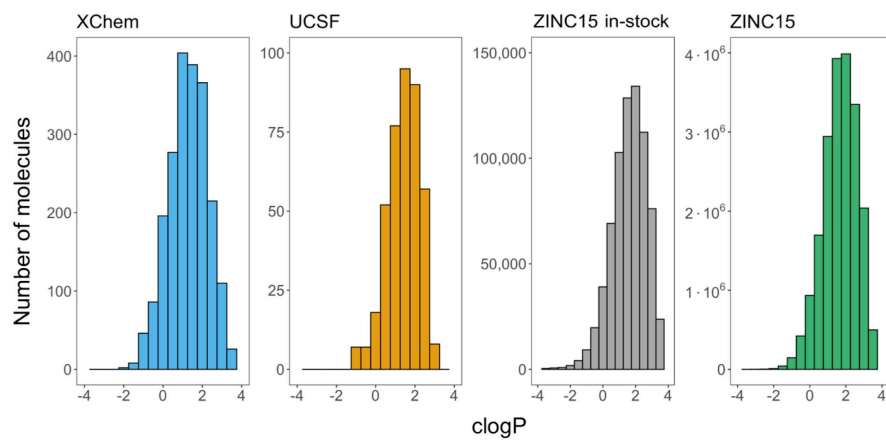
A) The structural changes previously reported to occur upon ADP-ribose binding are captured by the Mac1-ADPr structure determined using P4₃ crystals. The apo P4₃ structure is shown with dark gray sticks, with arrows indicating the changes in protein conformation upon ADP-ribose binding (white sticks). Electron density (2mF_O-DF_C) is contoured at 4 σ (blue mesh). **B)** The α -anomer of the terminal ribose of ADP-ribose was observed in the P4₃ crystal form (cyan and white sticks). In previously reported structures (e.g. PDB 6W02, yellow sticks), a flip in Gly47 allows the β -anomer to bind by removing a steric block (red dashed line) and forming a new hydrogen

bond (black dashed line). However, the Gly47 flip is incompatible with the P4₃ crystal form because it would clash with the Lys11 carbonyl of a symmetry mate (blue sticks). In α -anomer, the anomeric hydroxyl is orientated away from Gly47, and binding can proceed without the peptide flip. **C)** Interconversion between ribose anomers in solution. **D)** Stick representation showing the previously reported Mac1-ADPr structures (cyan sticks) and the new structure determined using P4₃ crystals (grey sticks). The agreement between ADP-ribose is excellent, despite different anomers of the terminal ribose being present (α in the P4₃ structure, β in the previously reported structures). **E)** Heatmap showing the C α RMSD values after C α alignment for 10 previously reported SARS-CoV-2 Mac1 structures (6VXS, 6W02, 6W6Y, 6WCF, 6WEN, 6WOJ, 6WEY, 6YWK, 6Z5T, 6Z6I) and the new structures reported in this work. **F)** Comparison of SARS-CoV-2 Nsp3 Mac1 structures at 100 K (blue) and 310 K (red). The adenosine diphosphoribose ligand shown in the figure (cyan) is modeled according to its position in PDB 6W02. **G)** Correlation plot showing structural differences between the 100 K and 310 K structures. To generate the plot, the 100 K and 310 K structures were aligned and difference vectors were calculated between identical C α atoms in the two structures. The plot shows all pairwise dot products between these difference vectors, revealing the extent to which temperature-dependent changes are correlated across the structure. Positive dot products (positive correlations) are colored blue, while dot products (negative correlations) are shown in red. The pattern of positive and negative correlations is characteristic of a hinge-bending motion. **H)** Alignment of three coronavirus macrodomain structures with a human macrodomain (hMacroD2). ADP-ribose from the SARS-CoV-2 structure is shown with cyan sticks. **I)** Comparison of the adenosine binding site highlighting key residues involved in adenine and fragment interaction. The adenine coordination by Phe156 is unique to SARS-CoV-2 amongst betacoronaviruses and replaced in SARS-CoV-1 (PDB 2FAV) and MERS-CoV (PDB 5HOL) with asparagine. Human macrodomains including MacroD2 (PDB 4IQY) interact with adenine as SARS-CoV-2 Nsp3 Mac1 with a phenylalanine in this position which needs to be considered for achieving inhibitor selectivity for viral over human macrodomains. **J)** Sequence alignment showing conservation of residues in the ADP-ribose, catalytic and potential allosteric sites which are targeted by the fragments. Residue numbers on top refer to the construct residue numbering of SARS-CoV-2 Nsp3 Mac1. Numbers on either end of the alignment are residue numbers in the full-length proteins. The adenine-coordinating Phe156 is highlighted in red.

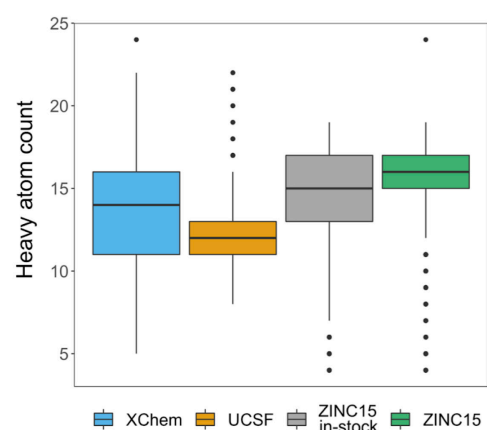
A Molecular weight



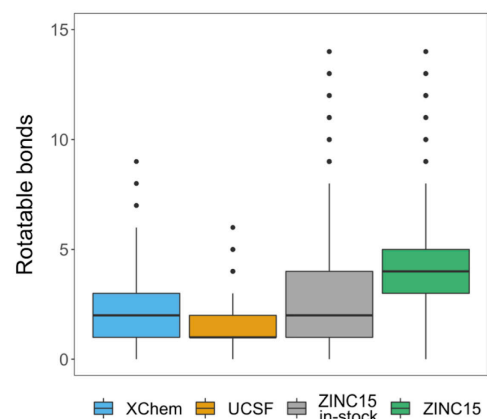
B cLogP



C Number of heavy atoms



D Number of rotatable bonds



E Scaffold and chemotype analysis

Fragment Library	Number of molecules	Distinct Bemis-Murcko scaffolds	Number of pyrimidines	Number of anions
XChem	2,126	809	72 (3.39%)	115 (5.4%)
UCSF	411	179	12 (2.92%)	144 (35%)
ZINC15 (in-stock)	722,963	69,244	41,531 (5.74%)	85,398 (11.8%)
ZINC15	20,006,175	803,333	890,199 (4.44%)	739,184 (3.7%)

Fig. S5. Physical properties, scaffold and chemotype analysis of screened fragment libraries.

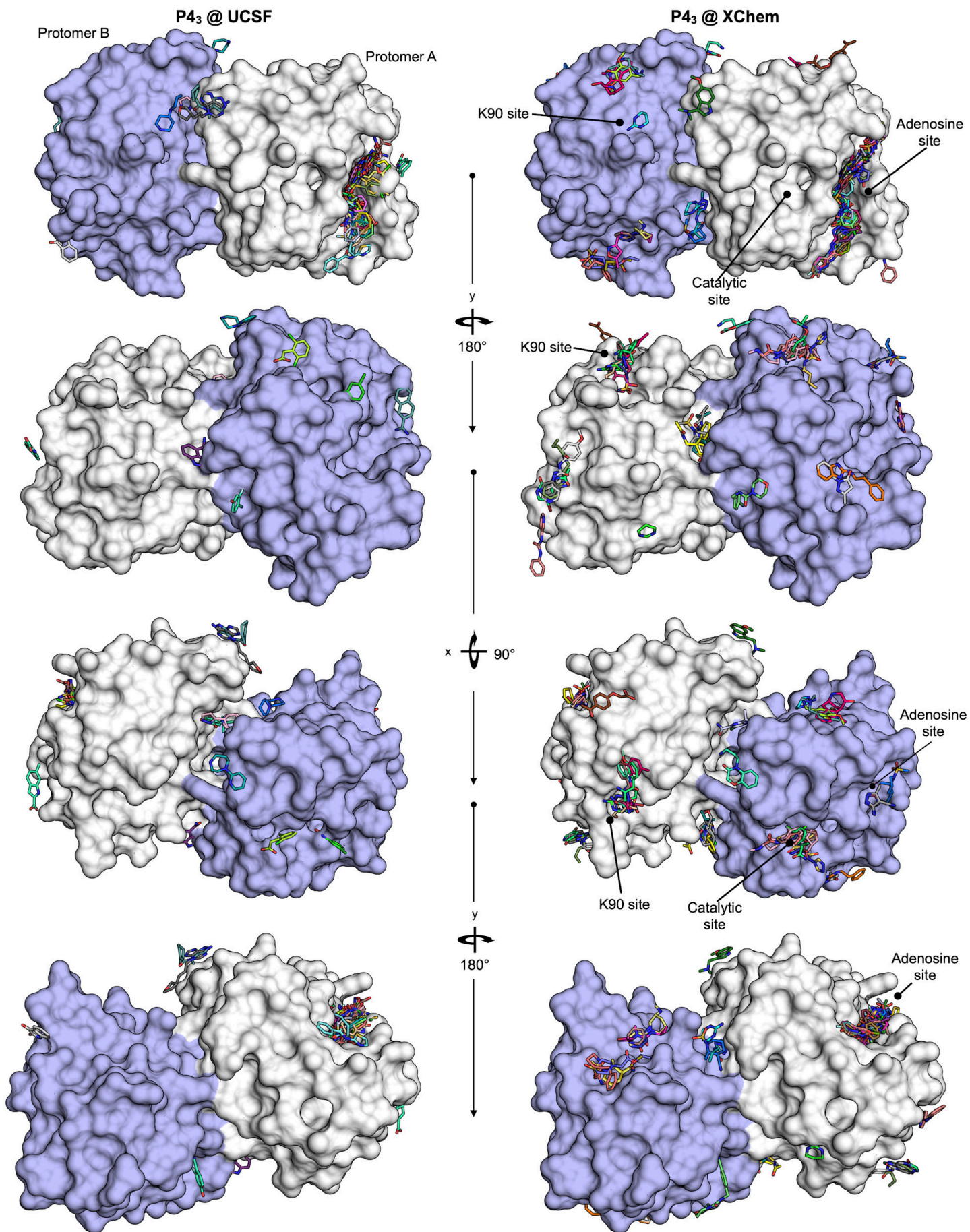


Fig. S6. Overview of fragment binding to protomer A (white surface) and protomer B (blue surface) of the P4₃ crystals.

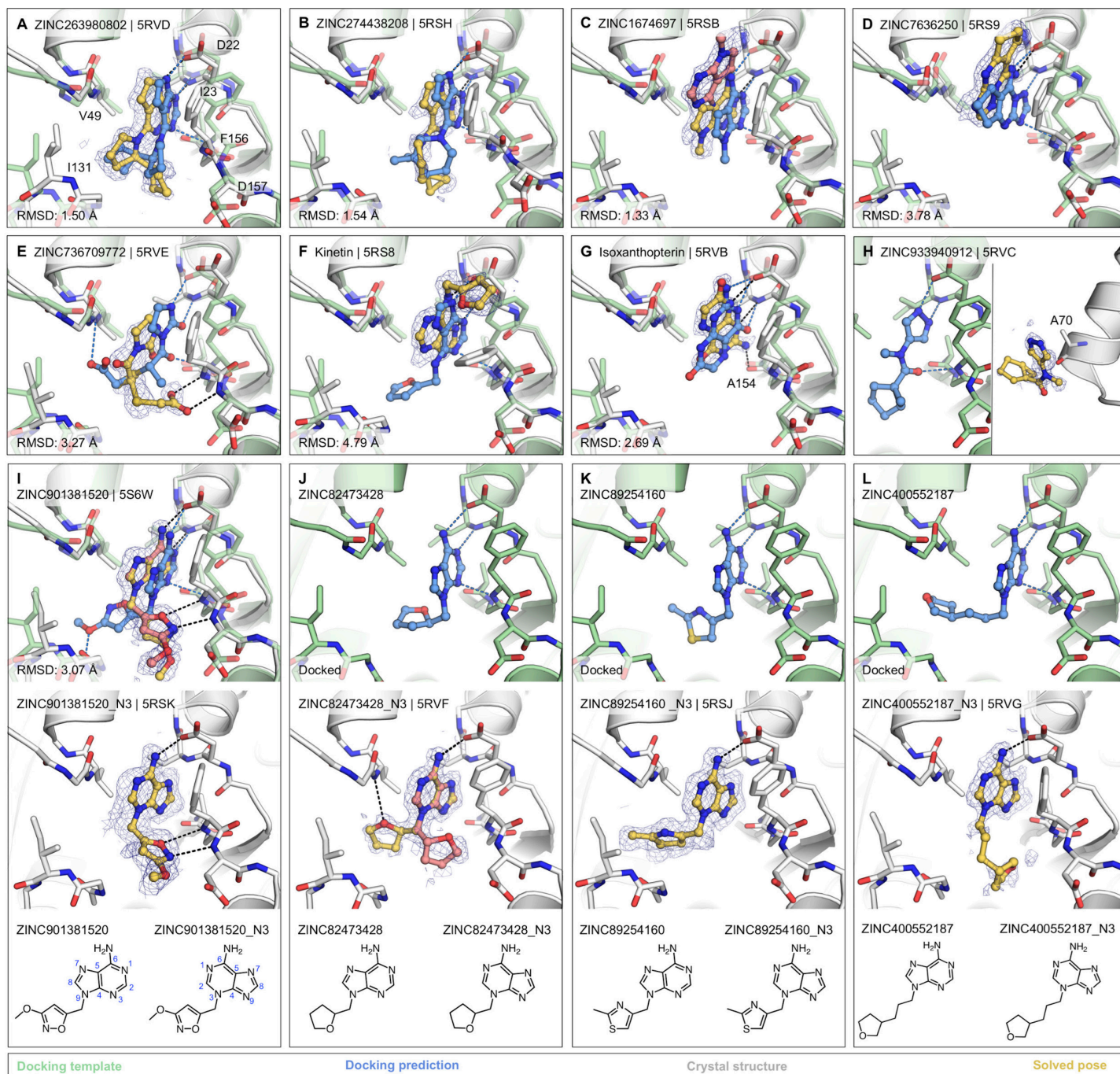


Fig. S7. Additional soaking hits from docking and adenine-N3 vs -N9-alkylated isomers.

The protein structure (PDB 6W02) (34), prepared for virtual screens is shown in green, predicted binding poses are shown in blue, the crystal protein structures are shown in grey, the solved fragment poses are shown in yellow, with alternative conformations shown in light pink. PanDDA event maps are shown as a blue mesh. Protein-ligand hydrogen bonds predicted by docking or observed in crystal structures are colored light blue or

black, respectively. Hungarian RMSD values are presented between docked and crystallographically determined ligand poses.

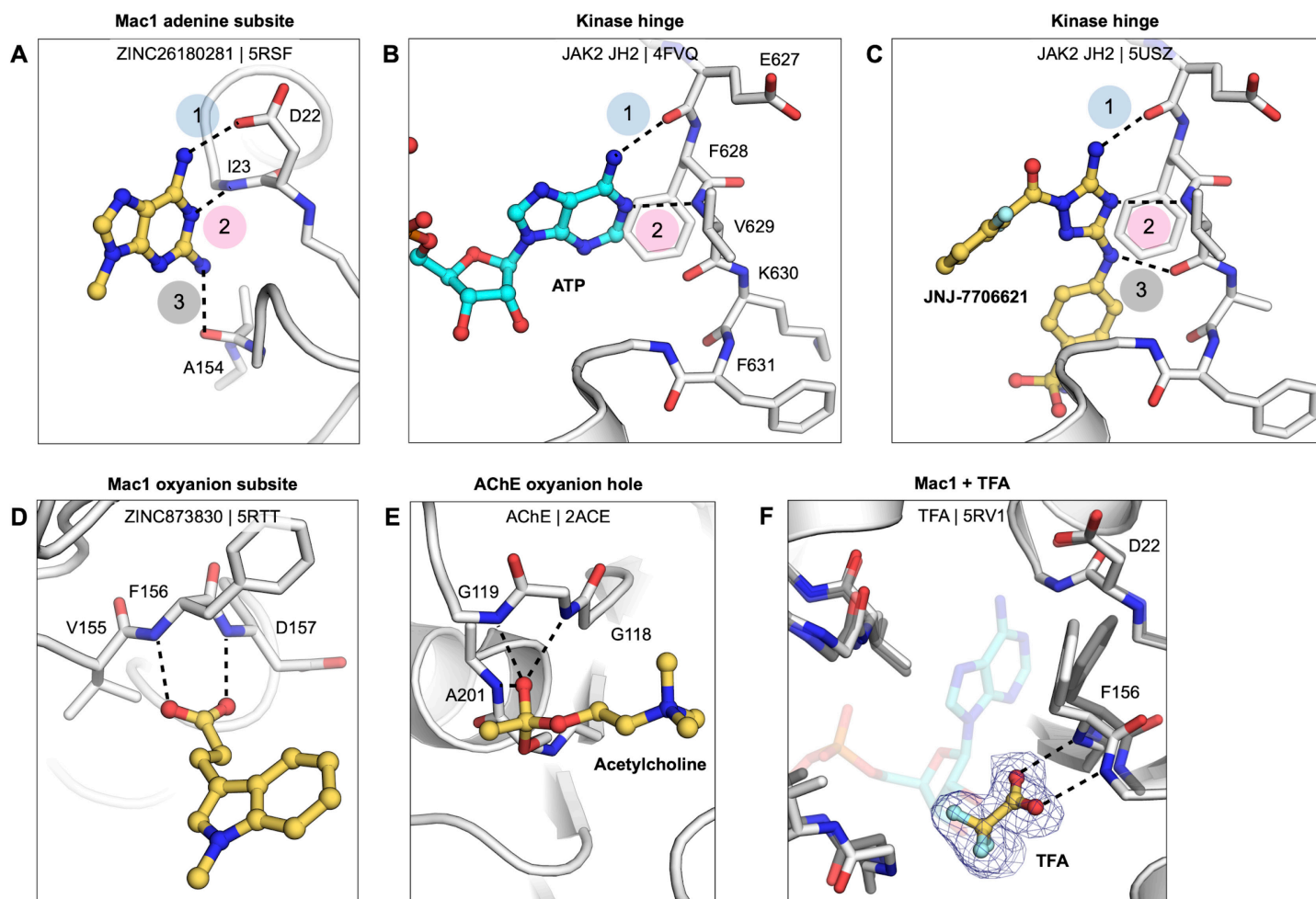


Fig. S8. Mac1 subsites compared to the adenine binding subsite in kinases and the oxyanion binding site in carboxylesterases.

A) Key features of the Mac1 adenine subsite are illustrated by the structure of ZINC26180281 (PDB 5RSF). Hydrogen bonds are formed between the C6-amine of the adenine scaffold and the backbone nitrogen of Ile23, and between N1 of the adenine scaffold and the side-chain carboxylate of Asp22. The C2 amine of ZINC26180281 forms a non-canonical hydrogen bond to the backbone carbonyl oxygen of Ala154. **B)** Adenine recognition is similar in the pseudokinase domain of JAK2, however, the C6-amine forms a hydrogen bond to a backbone carbonyl oxygen rather than a side-chain carboxylate. Adenine binding occurs at the hinge residues that connect the N- and C-terminal lobes of the catalytic domain. Interactions that mimic adenine binding to the hinge residues are conserved in the majority of kinase inhibitors (43). Like ZINC2618028, kinase inhibitors exploit non-canonical hydrogen bonds. The 1,2,4-triazole derived inhibitor shown in **(C)** forms a hydrogen bond to the backbone carbonyl oxygen of Lys630. **D)** The fragment screens against Mac1 identified 47 oxyanions binding to the backbone nitrogens of Phe156 and Asp157. A comparable oxyanion recognition motif is present in acetylcholinesterase (AChE) **(E)**. In AChE, this motif stabilizes negative charge on the oxyanion transition state. **F)** Trifluoroacetic acid (TFA), present as a counter ion for ZINC3860798, was clearly defined in PanDDA event maps binding to the oxyanion subsite. TFA was also observed binding to the oxyanion subsite for fragments ZINC35185198 and ZINC51658946. The docking fragment ZINC263392672 also contained TFA, but no TFA was observed in the oxyanion subsite (PDB 5RSG).

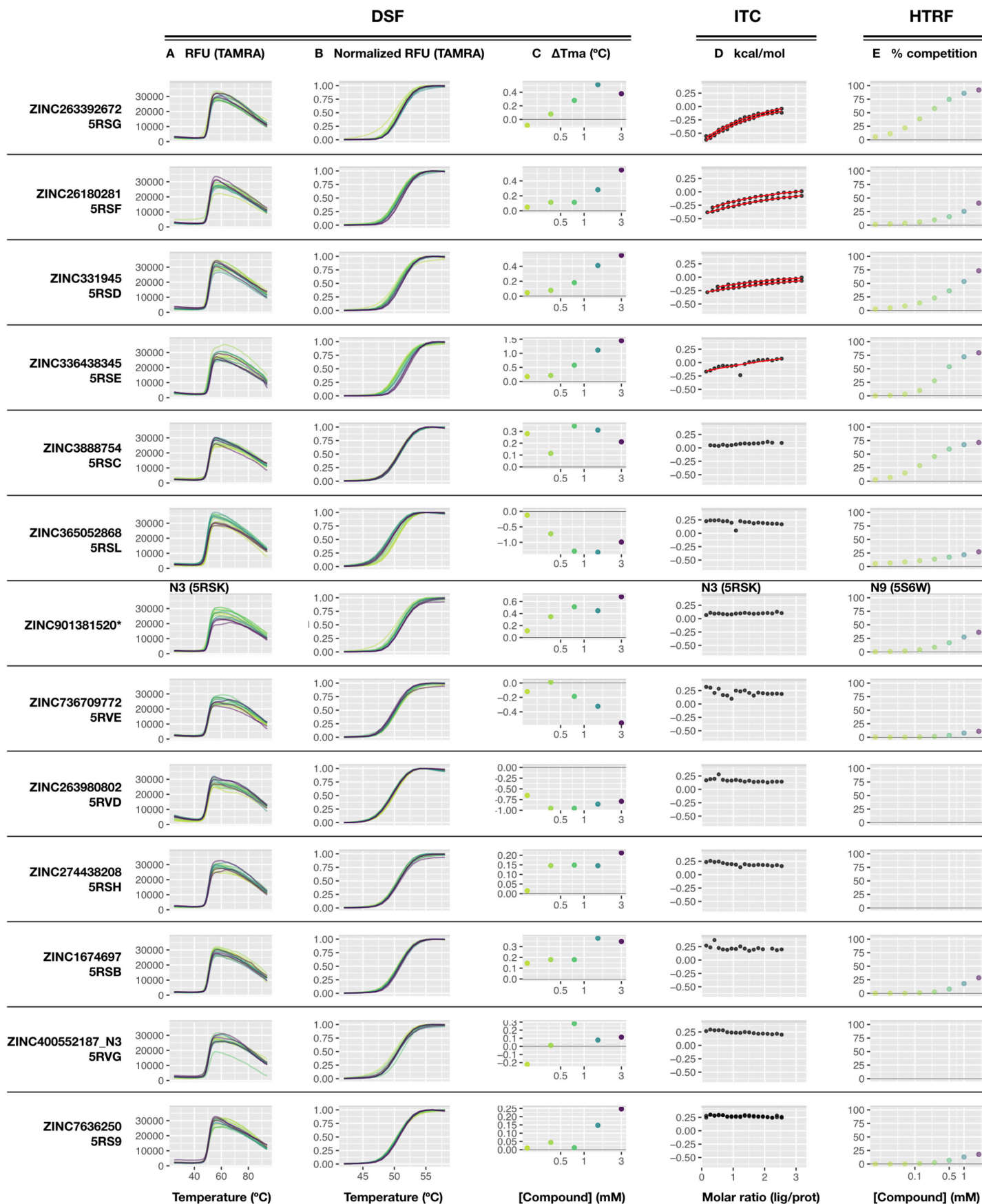


Fig. S9. Comparison of DSF, HTRF, and ITC results for compounds tested in all assays.

A) Raw, un-normalized DSF data for the full measured temperature range (25 - 94 °C) demonstrates the absence of confounding changes in curve shape for all compounds. **B)** Normalized raw DSF data, enlarged to visualize compound-induced thermal shifts. **C)** Changes in T_{m_a} observed in the presence of fragments (0-3 mM fragment). **D)** Integrated heat peaks as a function of binding site saturation shown as black dots. The red line represents a non-linear least squares (NLLS) fit using a single-site binding model. **E)** Dose-response curves showing competition of the fragments with an ADPr-conjugated peptide for Mac1 binding. (*) ZINC901381520_N3 was tested in DSF and ITC, ZINC901381520_N9 was tested in HTRF.

Data S1. (separate file)

Excel spreadsheet with 1) summary of fragment screens, 2) amino acid sequences of constructs used for crystallography, 3) summary of the data collection strategy for the X-ray diffraction experiments, 4) data reduction and refinement statistics for all X-ray crystal structures reported in this work, 5) summary of all X-ray diffraction datasets collected, 6) classification of all fragment hits, and 7) solution binding data for selected fragments.

Data S2. (separate file)

PDF with 1) summary of all fragments binding in the adenosine, catalytic and K90 sites of Mac1, 2) DSF data for all compounds tested, 3) ITC data for all compounds tested, and 4) HTRF peptide-displacement data for all compounds tested.