

The mechanisms of catalysis and ligand binding for the SARS-CoV-2 NSP3 macrodomain from neutron and X-ray diffraction at room temperature

Galen J. Correy¹, Daniel W. Kneller^{2,3}, Gwyndalyn Phillips^{2,3}, Swati Pant^{2,3}, Silvia Russi⁴, Aina E. Cohen⁴, George Meigs^{5,6}, James M. Holton^{4,5,6}, Stefan Gahbauer⁷, Michael C. Thompson⁸, Alan Ashworth⁹, Leighton Coates^{2,3}, Andrey Kovalevsky^{2,3*}, Flora Meilleur^{2,10*}, James S. Fraser^{1*}

Affiliations

¹ Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA 94158, USA

² Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

³ National Virtual Biotechnology Laboratory, US Department of Energy, USA

⁴ Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Center, Menlo Park, CA 94025, USA

⁵ Department of Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

⁶ Department of Biochemistry and Biophysics, University of California San Francisco, CA 94158, USA

⁷ Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA 94158, USA

⁸ Department of Chemistry and Chemical Biology, University of California Merced, CA 95343, USA

⁹ Helen Diller Family Comprehensive Cancer, University of California San Francisco, CA 94158, USA

¹⁰ Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695

* To whom correspondence should be addressed; Email: kovalevskyay@ornl.gov, meilleur@ornl.gov, jfraser@fraserlab.com

Abstract

The NSP3 macrodomain of SARS CoV 2 (Mac1) removes ADP-ribosylation post-translational modifications, playing a key role in the immune evasion capabilities of the virus responsible for the COVID-19 pandemic. Here, we determined neutron and X-ray crystal structures of the SARS-CoV-2 NSP3 macrodomain using multiple crystal forms, temperatures, and pHs, across the apo and ADP-ribose-bound states. We characterize extensive solvation in the Mac1 active site, and visualize how water networks reorganize upon binding of ADP-ribose and non-native ligands, inspiring strategies for displacing waters to increase potency of Mac1 inhibitors. Determining the precise orientations of active site water molecules and the protonation states of key catalytic site residues by neutron crystallography suggests a catalytic mechanism for coronavirus macrodomains distinct from the substrate-assisted mechanism proposed for human MacroD2. These data provoke a re-evaluation of macrodomain catalytic mechanisms and will guide the optimization of Mac1 inhibitors.

1 Introduction

2 Viral infection triggers the release of interferons, a group of secreted cytokines that are key components of the
3 innate immune response (1–4). Among the several hundred interferon-stimulated genes are poly-(ADP-ribose)
4 polymerases (PARPs), a family of enzymes that catalyze the transfer of ADP-ribose from NAD⁺ to proteins as a
5 post-translational modification. Both mono- and poly-ADP-ribosylation are important for signal transduction of
6 the interferon response (5–7). In SARS-CoV-2 infection, ADP-ribosylation is mediated by PARP9 and its binding
7 partner DTX3L (8). To counteract this host defense mechanism, SARS-CoV-2 encodes a macrodomain, Mac1,
8 with mono-(ADP-ribosyl)-hydrolase activity that reverses the protective effect of ADP-ribosylation (9) (**Fig. 1A**).
9 Importantly, an active site mutation that inactivates Mac1 has been shown to render SARS-CoV non-lethal in a
0 mouse model of viral infection (10) and interferon stimulated cells can clear SARS-CoV virus bearing that active
1 site mutation, but not WT Mac1. Accordingly, Mac1 is a potential therapeutic target for small molecule inhibitors.
2 Importantly, the Mac1 active site is highly conserved between SARS-CoV, SARS-CoV-2 and MERS, which
3 raises the possibility that a Mac1 inhibitor could be used as a broad spectrum antiviral against coronaviruses
4 (11).

5
6 The SARS-CoV-2 macrodomain is encoded as a domain of non-structural protein 3 (NSP3), a large multi-domain
7 protein found in all coronaviruses (12) (**Fig. 1B**). Mac1 adopts an $\alpha/\beta/\alpha$ -sandwich fold and binds ADPr through
8 an extensive network of hydrogen bonds in a well-defined cleft (**Fig. 1B/C**) (9, 13). Several conformational
9 changes are stabilized upon ADPr binding, including rotation of Phe132 to accommodate the terminal ribose of
0 ADPr, and two peptide flips to bind the diphosphate portion of ADPr (14). Although the structural determinants
1 of ADPr-binding are well established, the catalytic mechanism of Mac1 is unknown. Based on structural
2 homology to the human macrodomain hMacroD2 (15), a substrate-assisted mechanism has been proposed for
3 viral macrodomains (16), however, this has not been confirmed experimentally. Determining the precise
4 orientations of active site water molecules and the protonation states of key catalytic site residues can
5 discriminate whether the proposed water nucleophile is appropriately orientated for the hydrolysis reaction in the
6 ADPr-bound state or if other catalytic mechanisms should be considered.

7
8 Although X-ray crystallography can provide the high resolution structural information, hydrogen atoms are not
9 normally observed in X-ray diffraction experiments (17). This means that key information may be obscured,
0 including residue protonation states, hydrogen bond networks and water orientations. Neutron crystallography
1 allows hydrogens to be visualized at modest resolutions (<2.5 Å) (18–21), which can reveal details missing from
2 X-ray structures. In practice, hydrogen is typically exchanged for deuterium in protein samples for neutron
3 crystallography, because protons have a small negative neutron scattering length, as well as a large incoherent
4 cross section, whereas deuterium has a comparatively large positive neutron scattering length and small
5 incoherent cross section (22). The low brightness of neutron sources means that the volume of crystals required
6 is 100-1000-fold larger than crystals that are suitable for experiments at synchrotron light sources (23). Because
7 protein crystals rarely grow to this size, relatively few drug targets have been structurally characterized by
8 neutron crystallography (24–26).

9
0 In addition to informing catalytic mechanisms, neutron crystallography can augment structure-based drug design,
1 which depends upon an accurate model of the shape, electrostatic potential, solvation and flexibility of the target
2 site (27–29). We recently screened more than 2500 fragments against Mac1 using X-ray crystallography (14).
3 Structures of 234 fragments bound to Mac1 were determined, including almost 200 in the active site. This screen
4 has given the community an extensive set of chemically diverse starting points for fragment-based ligand
5 discovery against Mac1. As expected based on their low molecular weight, the fragments bound Mac1 weakly
6 (14). Efforts have been assisted by the development of Mac1-specific biophysical and biochemical assays (8,

30, 31). The task of converting fragments into potent and selective Mac1 inhibitors will be helped by a more detailed understanding of the physical chemical features of the Mac1 active site.

Here, we report a series of Mac1 crystal structures determined using neutron and X-ray diffraction (**Fig. 1D**). Two neutron structures of the apo enzyme were determined from distinct crystal forms grown at pH 6.5 and 9.5, and a neutron structure of the ADPr-Mac1 complex was determined at pH 6.5. Protonation states and active site hydrogen bond networks were characterized in the three structures, revealing the molecular basis for functionally relevant flexibility in active site loops and proton locations relevant to the catalytic mechanism. We mapped water networks in the active site, including water orientations from neutron diffraction, and showed that the networks were robust to changes in crystal packing, temperatures and pH. The water networks reorganize upon ADPr binding, with a network of tightly bound water molecules acting as protein-ligand bridges. Finally, we show how these same tightly bound water molecules are co-opted by fragment binding, with implications for inhibitor design. These results advance our knowledge about the structure and function of viral macrodomains, and provide a structural resource to guide the design of macrodomain inhibitors as SARS-CoV-2 antiviral therapeutics.

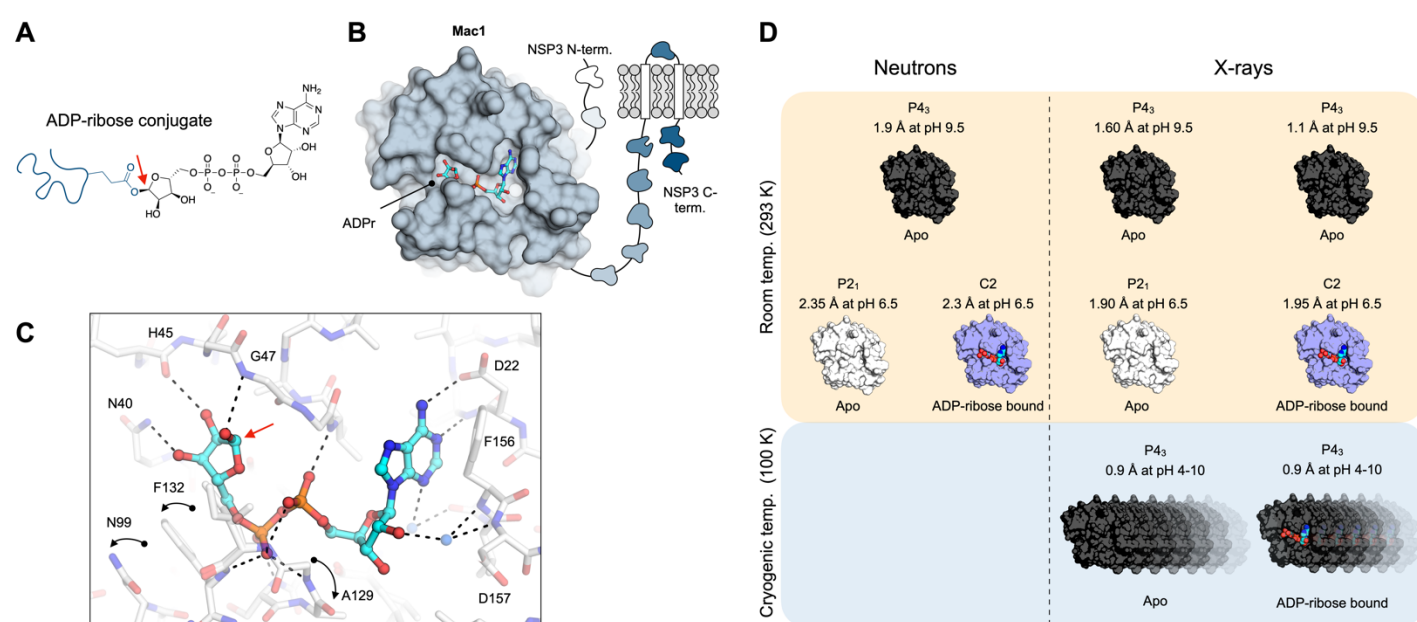


Fig. 1. The NSP3 macrodomain (Mac1) reverses mono-ADP-ribosylation. (A) Chemical structure of ADPr showing the C1' covalent attachment point with a red arrow. (B) Cartoon of the multi-domain NSP3 showing Mac1 with ADPr bound in the active site (PDB code 7KQP). (C) Structure of ADPr bound in the macrodomain active site (PDB code 7KQP) with the changes in protein structure upon ADPr binding indicated with black arrows. (D) Summary of the crystal structures reported in this work.

Results

Room temperature Mac1 structures from distinct crystal forms at different pHs

Mac1 neutron structure at pH 9.5 (1.9 Å, P4₃ crystals)

We grew neutron diffraction quality crystals at pH 9.5 using a Mac1 construct that crystallizes in the P4₃ space group (14) (**Fig. 2A**). Hydrogen/deuterium exchange (HDX) was achieved by growing crystals using solutions

prepared with D₂O, followed by soaking the crystal in a mother liquor prepared with D₂O. Neutron diffraction data to 1.9 Å were collected at room temperature using the Macromolecular Neutron Diffractometer (MaNDi) at the Oak Ridge National Laboratory (32–34) (**Table S1**). To enable joint neutron/X-ray refinement in Phenix (35), we collected a 1.6 Å X-ray diffraction dataset from the same crystal using a rotating anode X-ray source (**Table S1**). To account for differences in HDX at different protein sites, both hydrogen and deuterium were modeled at exchangeable positions, and deuterium occupancies were automatically refined (35). After several cycles of refinement and model building, the $R_{\text{work}}/R_{\text{free}}$ values were 22.6/27.5% for the neutron data and 12.4/15.7% for X-ray data. In total, 203 D₂O molecules were modeled across the two molecules in the asymmetric unit (ASU) (**Fig. 2B**). There was weak electron density for *N*-cyclohexyl-2-aminoethanesulfonic acid (CHES) in the adenosine site, with the CHES sulfonate interacting with the backbone nitrogens of Phe156/Asp157 (the oxyanion subsite), mimicking the bridging interactions between a water molecule and the proximal ribose of ADPr (**Fig. S1**). The atomic coordinates and structure factor amplitudes have been deposited in the Protein Data Bank with the accession code 7TX3.

Mac1 X-ray structure at pH 9.5 (1.1 Å, P₄₃ crystals)

To obtain more structural information about Mac1 at room temperature, we collected higher resolution diffraction data using the P₄₃ crystals and synchrotron radiation (**Table S1**). The sensitivity of the P₄₃ crystals to radiation damage was assessed by collecting four datasets from similarly sized crystals (~0.1 mm³), with the absorbed X-ray dose varied from 73 to 539 kGy. There was no evidence for general or specific radiation damage, as measured by per-image resolution, R_{CP} , unit-cell parameters and side-chain B-factors (**Fig. S2**). The X-ray dose and dataset resolution were correlated, with the highest dose dataset having the highest resolution (1 Å, truncated to 1.1 Å to achieve ~100% completeness). Consistent with radiation damage-free structures, the refined coordinates and B-factors from the four datasets were close to identical (C α root-mean-square fluctuation (RMSF) < 0.1 Å, Pearson correlation coefficient (r) for C α B-factors >0.98) (**Fig. S2**). The atomic coordinates and structure factor amplitudes from the four datasets have been deposited in the PDB with the accession codes 7TWF, 7TWG, 7TWH, 7TWI. Based on having similarly high resolution, but slightly lower C α and side chain B-factors, the structure refined from the 290 kGy dataset was used for subsequent structural analysis ($R_{\text{work}}/R_{\text{free}}$ = 9.94/11.5%).

Alignment of the 1.1 Å X-ray structure with the 1.9/1.6 Å neutron/X-ray structure shows that the protein coordinates are nearly identical (C α root-mean-square deviation (RMSD) <0.05 Å, Pearson r >0.93 for C α B-factors) (**Fig. S3**). Although the crystals used for X-ray diffraction were grown in the same conditions as the neutron crystal (100 mM CHES pH 9.5, 34% PEG 3000), there was no evidence for CHES in the active site. This may be due to slight differences in how the crystals were handled. There were slightly more water molecules modeled in the 1.1 Å structure compared to the jointly refined neutron/X-ray structure (360 versus 317). Comparison of the waters between the two structures shows that their positions are conserved (**Fig. S3**). Of the 317 waters modeled in the 1.9/1.6 Å neutron/X-ray structure, 257 had a matching water within 0.5 Å in the 1.1 Å X-ray structure. The waters that were not conserved tended to have higher B-factors, indicating increased disorder (**Fig. S3**).

Mac1 neutron structure at pH 6.5 (2.3 Å, P₂₁ crystals)

To investigate structural differences in Mac1 across different crystal forms and different pHs, we also grew neutron-quality crystals at pH 6.5 using a construct that crystallized in the P₂₁ space group (36) (**Fig. 2A/B**). The difference between the two constructs was slight: the P₂₁ construct contained a Gly-Glu sequence immediately following the TEV cleavage site, while the P₄₃ construct had Ser-Met at the equivalent position. Hydrogen/deuterium exchange was performed in a similar manner to the P₄₃ crystals, and diffraction data was collected to 2.35 Å at MaNDi (**Table S1**). As with the P₄₃ crystals, an X-ray dataset was collected with the same

crystal (**Table S1**), and joint X-ray/neutron refinement was performed with phenix.refine. The final $R_{\text{work}}/R_{\text{free}}$ values were 17.7/26.1% for the neutron data and 16.6/22.5% for X-ray data. Overall, the structure was similar to both monomers of the $P4_3$ structure ($C\alpha$ RMSD <0.2 Å). Although MES was bound at full occupancy in the previously reported X-ray structure using the same construct and similar crystallization conditions (36), there was no evidence for MES binding in the new structure. The atomic coordinates and structure factor amplitudes have been deposited in the PDB with the accession code 7TX4.

ADPr-bound Mac1 neutron structure at pH 6.5 (2.3 Å, C2 crystal)

To better understand the molecular basis for ADPr recognition by Mac1, we co-crystallized ADPr with the $P2_1$ construct (**Fig. 2A/B**). Unlike the $P4_3$ construct, where crystal packing prevents co-crystallization with ADPr (14), co-crystals with the $P2_1$ construct grew readily. Neutron diffraction data was collected to 2.3 Å using the IMAGINE beamline at the High Flux Isotope Reactor (HFIR) at ORNL (37) and an X-ray dataset was collected to 1.95 Å using the same crystal. The space group of the co-crystal was C2, indicating that binding of ADPr had perturbed crystal packing. We performed joint neutron/X-ray refinement of the ADPr-Mac1 data using phenix.refine, with $R_{\text{work}}/R_{\text{free}}$ values of 18.7/26.2% for the neutron data and 13.3/17.4% for X-ray data. There was clear positive density for ADPr bound in the Mac1 active site in the $mF_o - DF_c$ electron density map (**Fig. S1**), and the ADPr occupancy was refined to 93%. The terminal ribose of ADPr is bound in the β -configuration, with the C1'' hydroxyl hydrogen bonded to the backbone nitrogen of Gly48. Alignment of the Mac1-ADPr neutron/X-ray structure with the three apo structures shows that the structures are highly similar ($C\alpha$ RMSD <0.3 Å) (**Fig. S3**). Structural differences include the peptide flips of Gly48 and Ala129 that allow two new hydrogen bonds with the diphosphate portion of ADPr, and a coupled conformational change in the Phe132 and Asn99 side-chains to accommodate the terminal ribose of ADPr (**Fig. S1**). These changes have been documented across several crystal forms (9, 13, 36, 38). The main difference with the ADPr-bound structure obtained with the $P4_3$ crystal form is that crystal packing interactions prevent the flip of Gly48, forcing ADPr to adopt the α configuration (14). The atomic coordinates and structure factor amplitudes for the ADPr-bound neutron/X-ray structure have been deposited in the PDB with the accession code 7TX5.

Features of deuteration between the different crystal structures

Next, we examined whether the refined deuterium occupancy in the neutron structures could reveal information about Mac1 structure and dynamics (39–41). Although deuterium occupancy was refined at all exchangeable positions, we restricted our analysis to backbone amide deuteriums. For the $P4_3$ crystal, the average backbone nitrogen deuterium occupancy was 78 and 76% for the two monomers in the ASU, and there was good agreement between the refined deuterium occupancies (Pearson $r = 0.87$) (**Fig. S4**). Because the monomers have unique lattice contacts, this suggests that the differences in HDX are caused by structural features of Mac1, not merely differences in crystal packing. Although the average backbone nitrogen deuterium occupancy in the $P2_1$ and C2 crystals were slightly lower (at 69% and 57%, respectively), there was still good agreement between the refined backbone amide deuterium occupancy (**Fig. S4**). In all four monomers, backbone deuteriums with low occupancy tended to be located on the buried β -strands (**Fig. 2C, Fig. S4**), which is consistent with the protected nature of backbone amides in β -sheets (39). The protective effect was pronounced in the solvent exposed β -strand that includes residues 80–84, where the refined D occupancies alternate between high and low, depending on whether the backbone amide is solvent exposed or not (**Fig. S4**). Compared to the $P4_3$ crystals, there were more positions that were refined with 0% occupancy in the ADPr-bound structure determined using C2 crystals (**Fig. S4**). This included a cluster of residues on the two α -helices composed of residues 50–70 (**Fig. 2D**). This may reflect increased protection from HDX caused by stabilization of the Gly46–48 loop in the ADPr-bound structure. However, comparing deuteration between crystal forms is complicated by differences in crystal packing, different pH of crystallization and differences in deuteration procedure. For example, the α -helix composed of residues 138–145 has relatively low HDX in the $P2_1$ and C2 crystals, which may be connected to

substantial differences in local crystal packing (**Fig. S4**). Despite the possible connection between lattice contacts and HDX, the refined deuterium occupancies were only weakly correlated with crystallographic B-factors (**Fig. S4**).

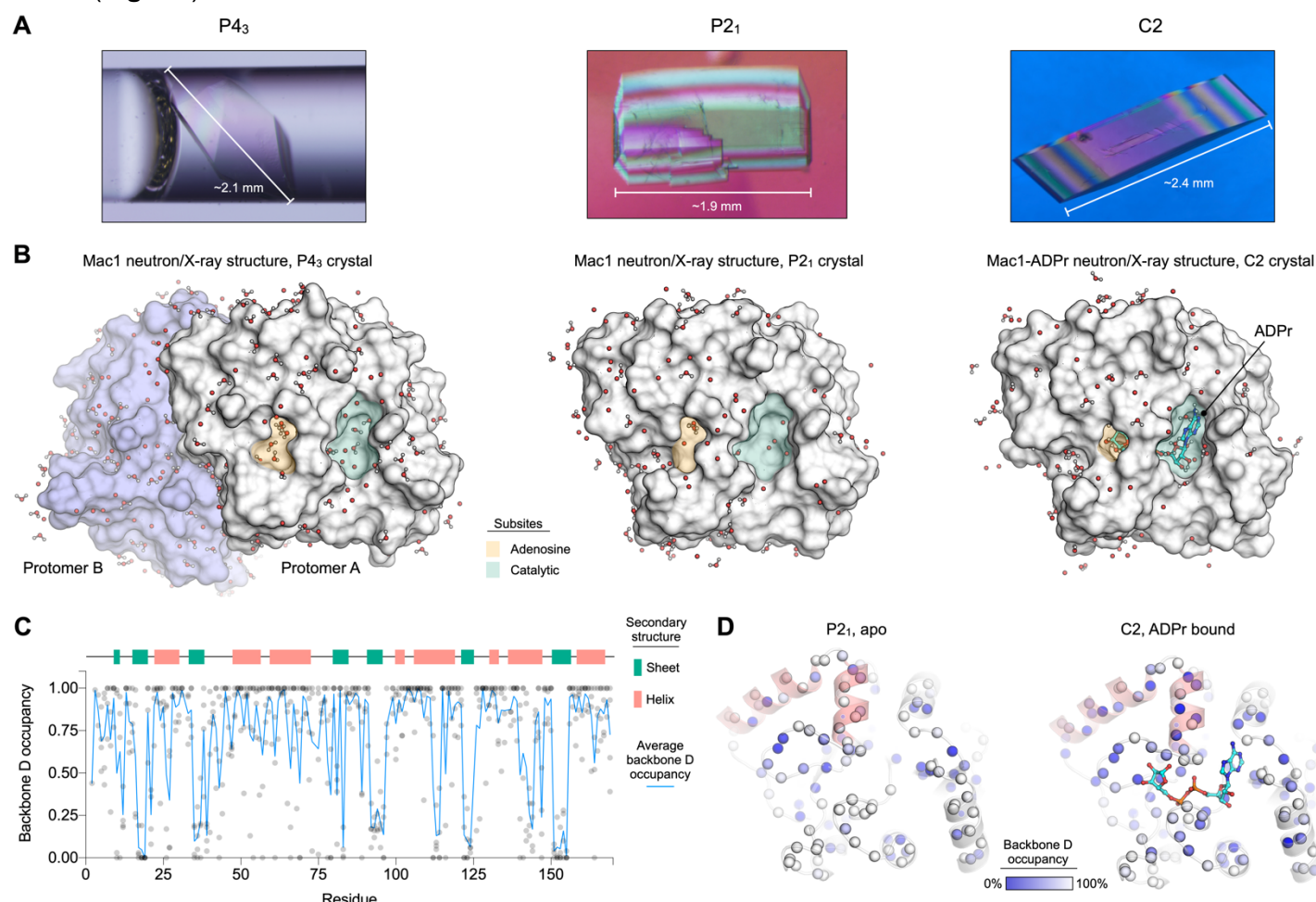


Fig. 2. Crystal structures of Mac1 determined using neutron diffraction. **(A)** Neutron quality Mac1 crystals grown in the P4₃, P2₁ and C2 space groups. The P4₃ crystal is shown in the quartz capillary used for data collection, while the P2₁ and C2 crystals are shown prior to mounting. **(B)** Molecular surface showing Mac1 neutron structures. D₂O molecules within 5 Å of the protein surface are shown with sticks/spheres. The adenosine and catalytic sites are shaded green and yellow respectively. **(C)** Plot showing the occupancy of backbone amide deuterium atoms in the three Mac1 crystal structures. **(D)** Backbone amide deuterium occupancy mapped onto the P2₁ and C2 Mac1 structures. Backbone nitrogens are shown with blue spheres, colored by backbone D occupancy (blue=0%, white=100%). The helices composed of residues 50-70 are shaded red. The average backbone D occupancy for these helices was 82% in the P2₁ structure and 52% in the C2 structure.

Assignment of Mac1 protonation states at pH 6.5 and 9.5

Histidines

Histidine can adopt multiple protonation states and tautomers at physiological pH, which allows it to play diverse roles in protein structure and function. Mac1 contains four histidines that are exposed to the bulk solvent, and two that are buried (**Fig. 3A**). We assigned histidine protonation states based on F_O-F_C neutron scattering length (NSL) difference density maps (**Fig. 3B/C**, **Fig. S5**). The histidine protonation states were the same across both crystal forms, except for histidines 86, 119 and 138, which were singly protonated in the P4₃ structure, but doubly protonated in the P2₁ structure (**Fig. 3C**). The importance of experimentally determining protonation states is

highlighted by the mismatch between the protonation states predicted by the Reduce program (42) and the states assigned using neutron crystallography (**Fig. 3C**). Although the protonation states of the two buried histidines, His45 and His94, were assigned correctly by Reduce, there was a discrepancy in the protonation states for the remaining histidines in at least one of the P4₃ protomers (**Fig. 3C**). To test whether histidine protonation states could be assigned from the previously published high resolution X-ray structures of Mac1 (14), we calculated F_O-F_C electron density difference maps from the 0.86 Å X-ray structure determined using the P4₃ crystals (PDB code 7KQO), and the 0.77 Å structure determined using C2 crystals (PDB code 7KR0). Although the protonation state of His45 and His94 could be assigned based on the electron density maps, the protonation states of the other residues were ambiguous (**Fig. S5**). This highlights the power of neutron crystallography for assigning histidine protonation states, even when X-ray data is available at ultra-high resolution.

Next, we examined the hydrogen bonding networks around the buried histidines. His45 is singly protonated on the Nε2 nitrogen, and hydrogen bonds to the Ser65 side chain and the backbone carbonyl of Gln62 (**Fig. 3D**). The neutron structure of Mac1 with ADPr bound shows that this network extends to the active site via a D₂O molecule, which forms a hydrogen bond with the C2'' hydroxyl of the terminal ribose (**Fig. 3D**). It has previously been shown that mutation of His45 to an alanine substantially decreases Mac1 activity (16). This water mediated network provides a potential link between His45 and Mac1 activity: His45 could participate either through stabilization of the terminal ribose in a catalytically productive conformation, as proposed previously(15), or through the action of His45 as a general base in the hydrolysis reaction. The other buried histidine, His94, participates in an extensive hydrogen bond network connecting the core of the Mac1 to its surface (**Fig. 3E**). Asn37 is hydrogen bonded to the singly protonated His94, and the network extends to His86 via a D₂O molecule and the Glu64 backbone (**Fig. 3E**). The residues comprising these two networks are conserved in the SARS-CoV and MERS macrodomains (16).

Cysteines

The pK_a of the cysteine thiol (8.5) means that it often acts as a nucleophile in enzyme-catalyzed hydrolysis reactions. Mac1 has three buried cysteines, however, none are located in the active site, and they are unlikely to play a direct role in ADPr hydrolysis (**Fig. 3A**). To determine the protonation states of the Mac1 cysteines in the P4₃ and P2₁ crystal structures, we modeled the protonated and deprotonated forms and calculated F_O-F_C NSL density difference maps (**Fig. S5**). Assigning the cysteine protonation states was difficult. There was a large positive peak placing the Cys81 thiol deuterium atom 3.1 Å from the Tyr68 aromatic ring (**Fig. 3F**), which is consistent with a rare thiol-aromatic bond (43). However, this peak was weaker or absent in protomer B and the P2₁/C2 structures (**Fig. S5**). There was a similar difficulty locating the Cys143 thiol deuterium atom: a positive peak in the F_O-F_C NSL density map was present in protomer B of the P4₃ structure, but absent in the other structures. These discrepancies may be due to incomplete HDX because the cysteines are buried, and/or the different resolutions of the neutron structures.

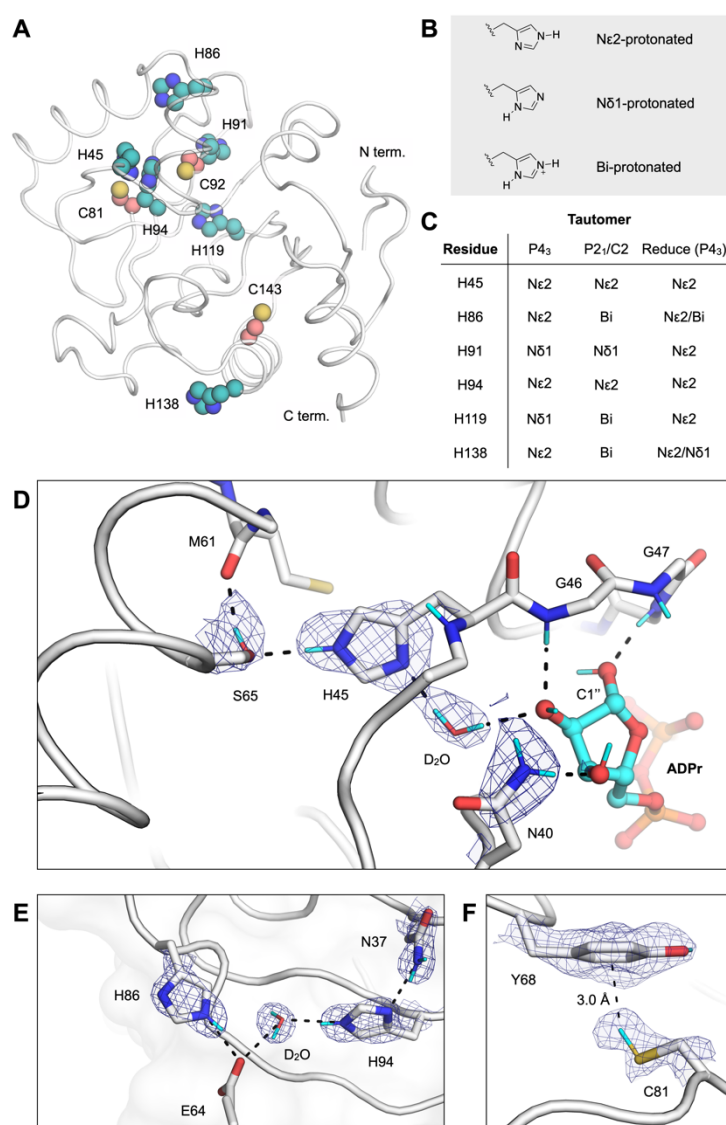


Fig. 3. Protonation states of Mac1 histidine residues assigned by neutron diffraction. **(A)** The location of histidine (teal spheres) and cysteine (salmon spheres) residues mapped onto the Mac1 structure (PDB code 7KQO). **(B)** Chemical structures showing the three possible protonation states of histidine. **(C)** Histidine protonation states assigned based on NSL density maps (maps are shown in **Fig. S5**). The tautomers assigned to the high resolution P4₃ X-ray structure (PDB code 7KQO) by the program Reduce are also shown. **(D)** NSL density maps reveal the hydrogen bond network connecting His45 and ADPr in the C2 structure (PDB code 7TX5). The protein is shown with a white cartoon/stick representation and the 2mF_O-DF_C NSL density map is shown with blue mesh (contoured at 2.5 σ). **(E)** An extensive hydrogen bond network connects the Asn37 side chain with a surface histidine (His86). The P4₃ structure (protomer A, PDB code 7TX3) and the corresponding 2mF_O-DF_C NSL density map is shown with blue mesh (contoured at 2.5 σ). **(F)** An aromatic-thiol bond was observed between Tyr68 and Cys81 in protomer A of the P4₃ structure (PDB code 7TX3). The 2mF_O-DF_C NSL density map is shown with blue mesh (contoured at 1 σ).

7 Protein flexibility and hydrogen bond networks in Mac1

To better understand the structural basis for ADPr recognition in Mac1, we examined hydrogen bond networks in the functionally important loops that surround the active site. First, we analyzed variation across Mac1 crystal structures by aligning structures determined at cryogenic and room temperature from several different crystal forms (**Fig. 4A/B**). The resulting ensemble shows that structure variation is mainly restricted to two regions around the active site: the loop at that positions Phe156 in the adenosine binding site (the Phe156 loop), and the two loops that position Phe132 in the catalytic site (the Ile131 loop and the Lys102 loop). Next, we compared structural variation across structures determined at cryogenic temperature and ambient temperature (**Fig. 4C**). The C α RMSF was highly correlated (Pearson $r = 0.93$), indicating that structural variation was independent of data collection temperature. The C α B-factors were also highly correlated between the two temperatures (Pearson $r = 0.92$) (**Fig. 4C**). The region with the highest C α B-factors at both temperatures was the Lys102 loop, which is consistent with the structural variation seen across different crystal forms (**Fig. 4B**).

9 Asn40 network

The need for flexibility to allow substrate binding is often balanced by the need to stabilize a substrate for catalysis (44, 45). In Mac1, Asn40 plays an important role in substrate stabilization, based on its absolute

conservation across evolutionarily diverse macrodomains, and loss of macrodomain function when Asn40 is mutated to alanine (46). In the ADPr-bound Mac1 neutron structure, Asn40 forms a hydrogen bond with the C3'' hydroxyl of ADPr, with the side chain orientation clearly defined in the NSL density map (**Fig. 4G**). The Asn40 side chain is preorganized to make this interaction, based on the similarity in the side chain conformation in the apo and ADPr-bound structures (**Fig. 4D/G, Fig. S1**). This preorganization is achieved by hydrogen bonds to the backbone of Tyr42 and Lys44, which is in turn stabilized by hydrogen bonds to the side chain of Glu62 (**Fig. 4D/G**). This network is conserved at 100 and 293 K, across different Mac1 crystal forms and between the SARS-CoV and MERS macrodomains (16). This extensive network stabilizes Asn40, while allowing the adjacent glycine-rich loop (Gly46-48) to be flexible during ADPr binding.

Ile131/Lys102 network

Next, we examined hydrogen bond networks in the Ile131/Lys102 loops, which undergo a series of conformational changes upon ADPr binding (**Fig. 1C**). The orientation of Asn99 is clearly defined in NSL density maps in the three neutron structures (**Fig. 4E/H**). The side chain nitrogen is oriented towards the catalytic site, while the side chain oxygen is hydrogen bonded to the backbone nitrogen of Asn101. In the P4₃ structure, the nearby Asn101 forms a hydrogen bond with the backbone of Ile131, connecting the two loops. This network is disrupted in the P2₁ crystal form, where Asn101 forms a hydrogen bond with a symmetry mate. Overall, this network is conserved between the 100 K and 293 K P4₃ X-ray structures, with only a slight shift in the Lys102 loop (**Fig. S6**). However, there is substantial variation in this loop in the ensemble of Mac1 determined using different crystal forms, and this variation matches the conformational change that occurs upon ADPr binding (**Fig. S6**). The Lys102 and Ile131 loop movement is correlated, with inter-loop hydrogen bonds maintained across the ensemble. In the SARS-CoV macrodomain, an arginine and aspartic acid are present at the position equivalent to Asn99 and Asn101. Although this removes a hydrogen bond between the Lys102 loop and the Ile131 loop, it introduces a new salt bridge to constrain the Asp98 side chain (equivalent to Asn99 in Mac1). Taken together, analysis of the Ile131/Lys102 loops shows how a key flexible region of Mac1 is supported by inter-loop hydrogen bond networks. This flexibility may be important for accommodating a wide range of ADP-ribosylated substrates.

Phe156 network

The adenosine portion of ADPr is recognised by hydrogen bonds to Asp22 and Ile23, and a water mediated contact with the Phe156/Asp157 backbone (**Fig. 1C**). The Phe156 loop is stabilized by two hydrogen bonds that connect the Asn20 side-chain to the Val155/Phe156 backbone (**Fig. 4F/I**). The orientation of the Asn20 side-chain is clearly defined in the NSL density maps obtained from both the P4₃ and C2 crystals (**Fig. 4F/I**). Although multiple conformations of the Phe156 side-chain were seen in the cryogenic P4₃ structure, only a single conformation was modeled in the 293 K X-ray structure (**Fig. S6**). However, in the Mac1 ensemble obtained from different crystal forms, the Phe156 side-chain is structurally diverse, matching the diversity seen for fragment-bound structures (14). Both SARS-CoV and MERS have an asparagine at the position equivalent to Phe156 and a cysteine at the position equivalent to Asn20 (**Fig. S6**). This indicates a degree of tolerance of the residues at this position, given that all three enzymes have similar ADPr binding properties (9). One consequence of this amino acid change is the potential for π - π stacking interactions at this site. Although the adenine portion of ADPr is too far to form π - π stacking interactions, we previously observed several fragments making close contacts with the Phe156 side-chain (14). An inhibitor that relies too heavily on stacking interactions with Phe156 may therefore have limited use as a broad-spectrum antiviral against coronaviruses.

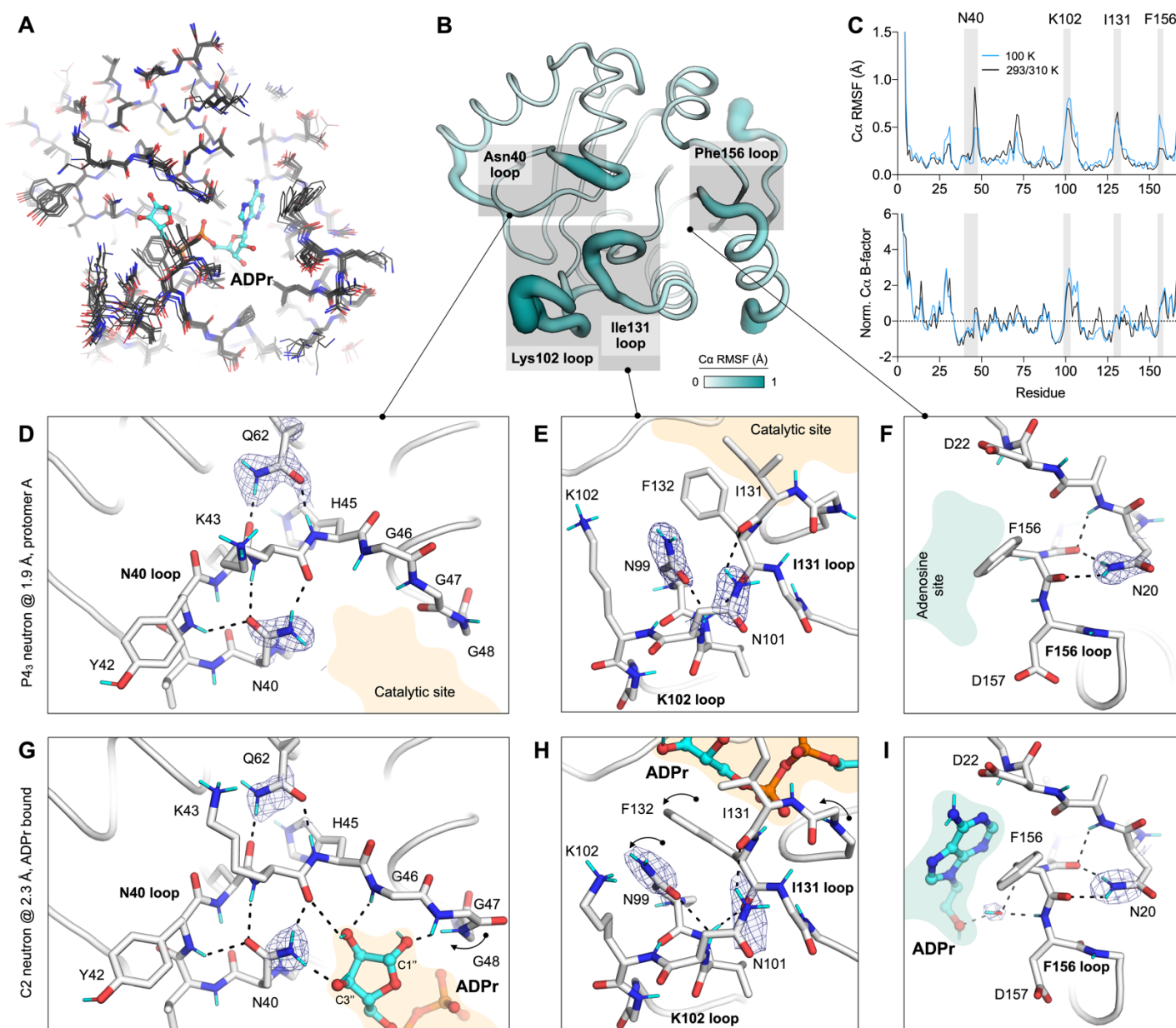


Fig. 4. Protein flexibility and hydrogen bond networks in the Mac1 active site. **(A)** Alignment of P4₃, P2₁ and C2 Mac1 neutron and X-ray crystal structures determined at 100/293/310 K (PDB codes 7KQO, 7KQP, 7KR0, 7KR1, 7TWH, 7TX3, 7TX4, 7TX5). **(B)** Ca RMSF calculated from structures shown in **(A)** mapped onto the Mac1 structure. **(C)** Top: plot showing Ca RMSF of Mac1 structures determined at 100 K (black line) and those at 293/310 K (blue line). Bottom: plot showing Ca B-factors from Mac1 structure. B-factors were normalized by Z-score as described in the Methods section. **(D,E,F)** Active site hydrogen bond networks assigned based on NSL density maps for the P4₃ neutron structure (protomer A, PDB code 7TX3). The protein is shown with a white stick/cartoon representation and the 2mF_o-DF_c NSL density map is shown with a blue mesh (contoured at 2 σ around the asparagine/glutamine side chains). Hydrogen bonds (<3.5 Å) are shown with dashed black lines. **(G, H, I)** Same as **(D,E,F)**, but showing the hydrogen bond networks in the C2 ADPr co-crystal structure (PDB code 7TX5). The NSL density maps are shown with a blue mesh (contoured at 3, 2 and 2.5 σ in **G, H** and **I**).

Water networks in Mac1

Water networks play an important role in protein-ligand recognition (47, 48), however their contribution to ligand binding is often difficult to quantify. Here, the combination of ultra-high resolution cryogenic X-ray structures (0.77 and 0.85 Å), a high-resolution room temperature structure (1.1 Å), structures determined using different crystal

forms (P4₃, P2₁ and C2), high resolution neutron structures (1.9 Å and 2.3 Å), and more than 200 fragment bound structures offers an exceptional opportunity to study the role of water in macrodomain structure, function and ligand binding.

Active site water networks at room temperature

To examine water structure within the Mac1 active site, we first mapped water networks in the 1.1 Å room temperature X-ray structure determined using P4₃ crystals (**Fig. 5A/B**). There were 19 active site water molecules in total: 14 in the catalytic site and five in the adenosine site. The catalytic site network encircles the Phe132 side chain, and is larger and more interconnected compared to the adenosine site network (**Fig. 5A**). To assess the tolerance of water networks to crystal packing, we compared the networks between the two protomers in the P4₃ ASU (**Fig. 5C**). Although the adenine site of protomer B is occupied by the N-terminal residues of a symmetry mate, the water positions and B-factors were similar (RMSD = 0.40 Å, Pearson r for B-factors = 0.42). Next, we compared the water network from the 1.9/2.3 Å X-ray/neutron structure determined using the P2₁ crystal. The water network was similar (RMSD = 0.45 Å), with only W10 and W16 not resolved in the P2₁ structure. Together, these results indicate that the Mac1 active site water network is conserved across different crystal packing environments.

D₂O orientations from neutron diffraction

Although the high resolution Mac1 X-ray crystal structures allow the accurate placement of water molecule oxygens, hydrogens are not visible at this resolution, which means that the orientation of water molecules is unclear. Neutron crystallography yields a more detailed map of water networks by determining the orientation of D₂O molecules. To model D₂O molecules in the neutron structures, deuterium atoms were added to D₂O oxygens and their orientations were refined using the 2F_O-F_C NSL density maps (35). Deuterium atoms were removed from D₂O molecules that lacked density. Although all D₂O molecules in the active site of the P4₃ structure were modeled with deuterium, the lower resolution of the P2₁ structure meant that six D₂O molecules were modeled without deuterium (**Fig. S7**). The quality of water modeling was assessed by calculating the real-space correlation coefficient (RSCC) (49) (**Fig. 5D**). The RSCC values were consistently high, indicating that there was good agreement between D₂O orientations and NSL maps. Next, we assessed the reproducibility of D₂O orientation in refinement by running 100 independent rounds of refinement starting from coordinates where the D₂O deuterium atoms were randomly shifted by 0.5 Å (**Fig. S8**). Deuterium atoms were included for all D₂O molecules in this procedure. The RMSF of deuterium atoms calculated across the 100 structures provides a measure of the reproducibility of D₂O orientations. For D₂O molecules in the P4₃ structure, RMSF values were low, consistent with highly reproducible D₂O orientations (**Fig. S8**). RMSF values for W3, W4, W9, W15 and W18 were comparatively high in the P2₁ structure, supporting their omission from the final model. To measure the agreement between equivalent D₂O molecules across the two P4₃ protomers, we calculated the distance between the average deuterium coordinates from the 100 structures (**Fig. 5E**). The D1/D2 positions for W3, W4, W6, W11 and W14 were all within 0.6 Å on average, consistent with conserved D₂O orientations between the P4₃ protomers. Based on B-factor and coordination number, W6 and W14 are the most ordered waters in the active site, and this matches their conserved orientation across the three Mac1 monomers (**Fig. 5F**).

Water networks in Mac1 are conserved from 100 K to 293 K

It is well established that crystal cryocooling causes crystal lattice contraction and remodeling of water networks at crystal packing interfaces (50–52). In the P4₃ crystal form of Mac1, lattice contraction upon cryocooling was modest, with a 2% decrease in unit cell volume from 293 K to 100 K (PDB codes 7TWH versus 7QKO). To investigate the effect of cryocooling on Mac1 solvation, we compared the overall water structure in the 1.1 Å P4₃ structure with the previously published 0.85 Å structure determined at 100 K (PDB code 7QKO). Although there

were substantially fewer waters modeled in the room temperature structure relative to the cryogenic structure (360 versus 745), the positions and B-factors of the waters were highly conserved (**Fig. 5G-I**). Of the 360 waters modeled in the room temperature structure, 240 had a matching water within 0.5 Å in the cryogenic structure. Waters that were not conserved tended to be second-shell waters (i.e. they were more than 3.5 Å from protein atoms) and also tended to have higher B-factors. The resolution gap between the two structures (0.25 Å) is not large enough to explain the difference in the number of waters modeled (53). In line with previous observations (52), it seems likely that this difference is due to increased solvent disorder at room temperature. However, unlike previous observations (52), the apparent increase in solvent disorder is not coupled to increased protein disorder, based on the highly correlated disorder in the structures determined at cryogenic and room temperature (**Fig. 4C**, **Fig. S3**). Next, we compared active site water networks at 293 and 100 K (**Fig. 5I/J**). Water positions and B-factors were highly conserved (RMSD = 0.38, Pearson r for B-factors = 0.62). Unsurprisingly, the positions of waters with lower B-factors tended to be more similar.

Mac1 active site water networks are conserved from pH 4-10

Based on the similarity of the Mac1 structures determined using crystals grown at pH 6.5 and 9.5, Mac1 tolerates a broad pH range. To extend this analysis to a wider pH range and higher resolution, whilst eliminating the impact of different crystal packing, we performed a pH-shift experiment with the P4₃ crystals. Briefly, crystals grown at pH 9.5 were soaked for four hours in a buffer composed of citric acid, HEPES and CHES, with pH values from 3 to 10 (54). There was no diffraction from crystals soaked at pH 3, despite the crystals appearing undamaged. From pH 4 to 10, crystals diffracted to better than 0.8 Å (data were truncated to 0.9 Å to achieve ~100% completeness) (**Table S1**). Based on diffraction data recorded from at least three crystals at each pH, there was a small but significant pH-dependent change in unit cell lengths (**Fig. S9**). This was most pronounced for the c-axis, which contracted 0.8% from pH 4 to 10. The unit cell changes may be linked to changes in the protonation states of His91 and His119 (**Fig. S10**). These residues are located at the interface between the two protomers in the P4₃ ASU, and there was a pH dependent shift in the side chain χ_1 and χ_2 dihedral angles (**Fig. S10**). Surprisingly, we observed multiple buffer components binding to the oxyanion subsite (**Fig. S1**). Binding was pH dependent, with density for citrate between pH 4-6, weak density for acetate at pH 7 and 8, and weak density for CHES at pH 9 and 10. There was also weak density for CHES binding in an alternative site at pH 5 and 6. Although acetate was not added to the soaking buffer, it is a known contaminant of HEPES buffer (55). Binding of the buffer components in the oxyanion subsite led to substantial conformational changes in the Phe156 loop (**Fig. S9**), mirroring the changes seen across different crystal forms (**Fig. 4A**) and upon fragment binding (14). There was an increase in backbone disorder from pH 4 to 10 for the α -helices composed of residues 20-31 and 156-169, as well as the Lys102 loop (**Fig. S9**). Surprisingly, we observed a low occupancy alternative conformation of the Ile131 loop in protomer B at pH 4. This alternative conformation was associated with an alternative water network and shared some similarity to the ADPr-bound conformation (**Fig. S1**), although there was no flip in Ala129. Overall, the active site water network was highly similar across the entire pH range (**Fig. 5L**, **Fig. S9**).

Re-organization of water networks upon ADPr binding

To investigate the role of water in ADPr binding, we compared the active site water networks in the previously published apo and ADPr-bound structures determined using P4₃ crystals (PDB code 7KQO and 7KQP) (**Fig. 6A/B**). In total, 11 waters are displaced by ADPr binding, either through direct displacement by ADPr (W4-5, W9-13, W15-16, W18) or through displacement by a shift in protein conformation (W3, displaced by Phe132). Of the remaining waters, six form bridging hydrogen bonds between ADPr and Mac1 (W6, W7, W8, W14, W17, W19). These waters tend to have more protein contacts and lower B-factors compared to the displaced waters (**Fig. 5B/C**). A notable exception is W19, which connects the oxyanion subsite to the C2' hydroxyl and is

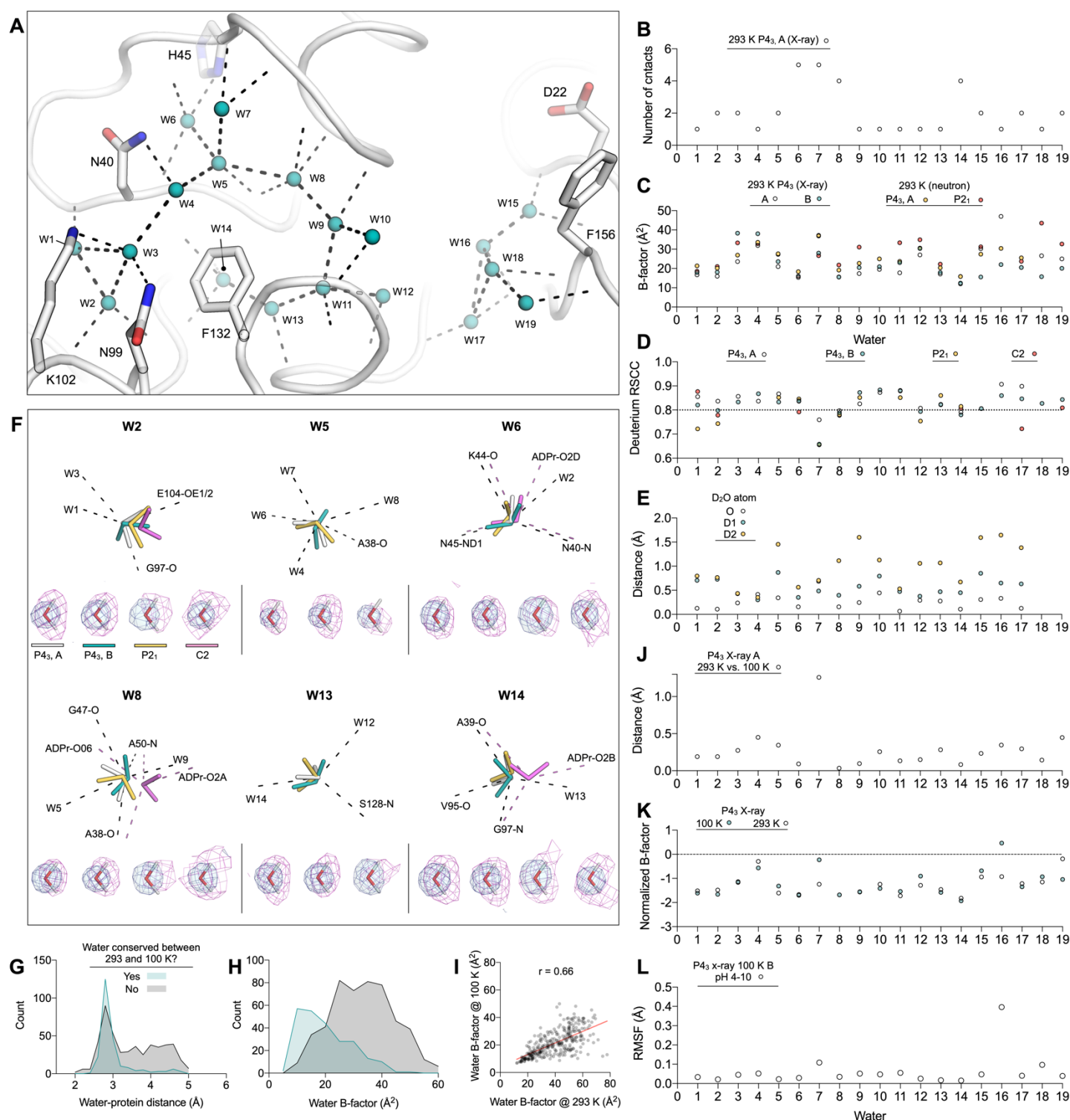


Fig. 5. Mac1 active site water networks are robust to changes in crystal packing, temperature and pH. **(A)** Water network in the Mac1 active site from the 1.1 Å P4₃ X-ray structure obtained at room temperature (PDB code 7TWH). Waters were considered part of the active site network if they were within 3.5 Å of an active site hydrogen bond acceptor/donor. For clarity, only the side chains of selected residues are shown. The protein is shown with white cartoon/sticks, the waters are shown as teal spheres, and hydrogen bonds with dashed black lines. **(B)** Plot showing the number of contacts for water molecules shown in **(A)**. **(C)** Plot showing B-factors for the active site H₂O/D₂O molecules in the room temperature P4₃ X-ray structure and the P2₁ and C2 neutron structures. Solvent molecules are numbered according to **(A)**. **(D)** Real-space correlation coefficients (RSCC) for active site D₂O molecules in the P4₃, P2₁ and C2 neutron structures calculated using the 2mF_o-DF_c NSL density map. A line is drawn at an RSCC=0.8, which has previously been used as a threshold for assessing

whether a D₂O is correctly oriented (49). (E) Variation in water orientations across 100 independent rounds of refinement. The plot shows distances between the average deuterium position of protomer A and B of the P₄₃ structure. For comparison, the D₂O oxygen distances between the A and B protomers are shown. (F) Selected active site D₂O molecules showing the 2mF_O-DF_C NSL density map (purple mesh contoured at 2 σ) and 2mF_O-DF_C electron density map (blue mesh/surface contoured at 2 σ). Hydrogen bonds are shown with dashed black lines. For the C2 structure, ADPr specific hydrogen bonds are shown with pink dashed lines. (G) Histogram showing water protein-distances in the 0.85 Å P₄₃ X-ray structure determined at 100 K (PDB code 7KQO). Waters were grouped based on whether a matching water molecule within 0.5 Å was found in the P₄₃ structure determined at 293 K (PDB code 7TWH). The histogram was generated with a bin width of 0.2 Å. (H) Same as (G) but showing B-factors for all water molecules in the 100 K P₄₃ structure. The histogram was generated with a bin width of 5 Å². (I) Scatter plot showing correlation between water B-factor at 100 and 293 K. The red line shows a linear fit of the data using GraphPad Prism. (J) Plot showing distances between H₂O oxygen atoms between protomer A of the P₄₃ X-ray structures determined at 293 and 100 K (PDB codes 7TWH and 7KQO). (K) Plot showing the B-factors of active site H₂O molecules in the P₄₃ X-ray structures determined at 293 and 100 K. B-factors were normalized by Z-score using the B-factors from all the H₂O molecules in a structure. (L) Plot showing the RMSF of active site water molecules calculated across the seven structures determined from pH 4 to 10 using the P₄₃ crystal form at 100 K. Because of pH-dependent binding of buffer components in the active site of protomer A, only the waters in protomer B are shown.

stabilized in the ADPr-bound structure (**Fig. 6A/C**). There were minor adjustments in the bridging waters that bind ADPr: W6, W7, W8, W14 and W17 all shift by between 0.3-1 Å (**Fig. 6B**). In addition to the water displacement and reorganization, two new waters are found in the ADPr-bound structure. W_{OXY} extends the hydrogen bonding network near the oxyanion subsite, and W_{PHOS} is hydrogen bonded to the terminal ribose and adjacent phosphate of ADPr (**Fig. 6A**). In total, the re-organization of water networks upon ADPr binding fits a model where weakly bound waters (as measured by number of protein contacts and B-factors) are displaced, while tightly bound waters are recruited to form bridging ligand-protein interactions.

Next, we examined hydrogen bond networks in the Mac1-ADPr in the C2 co-crystal system using neutron diffraction. Although the ADPr used for co-crystallization was hydrogenated, HDX occurred when the crystal was soaked in the deuterated buffer. This is supported by positive peaks in the F_O-F_C NSL density map calculated prior to modeling ADPr deuterium atoms (**Fig. 6D**). In the neutron structure, the terminal ribose adopts the β -configuration, with rearrangement of the Gly46-48 loop accommodating the C1'' hydroxyl (**Fig. 6E**). This configuration contrasts with the P₄₃ structure, where crystal packing interactions preclude rearrangement of the Gly46-48 loop, which forces the terminal ribose to adopt the α -configuration (**Fig. 6E**) (14). Despite the different configurations of the terminal ribose, and rearrangement of the Gly46-48 loop, the ADPr binding pose was very similar between the C2 and P₄₃ structures (RMSD = 0.34 Å, excluding the C1'' hydroxyl). Likewise, the positions of bridging D₂O molecules were highly conserved (RMSD for W6, W8, W14, W17 = 0.46 Å). Although there was a 1.2 Å shift in the position of W19, the hydrogen bonding network with the oxyanion subsite is maintained. Of the ADPr-specific bridging water molecules (W_{OXY} and W_{PHOS}), only W_{OXY} was identified in the C2 structure. The absence of W_{PHOS} is due to crystal packing interactions: a close contact between Ala69 of a symmetry mate forces the Ile131 side chain to rotate and occupy the W_{PHOS} position (**Fig. S1**).

Refinement of the positions of ADPr deuteriums, and the orientations of bridging D₂O molecules, allows the hydrogen bond networks that underpin ADPr binding to be examined (**Fig. 6F**). In the catalytic site, W8 forms hydrogen bonds with the C1'' hydroxyl, the backbone oxygen of Ala38, the backbone nitrogen of Ala50 and one of the ADPr phosphate oxygens. This highly coordinated D₂O is appropriately oriented to act as a water nucleophile in the putative substrate-assisted hydrolysis mechanism (16), with the D₂O oxygen oriented towards the C1'' carbon. Also in the catalytic site, W6 forms a bridging interaction between the deprotonated N δ 1 of His45 and the C2'' hydroxyl of ADPr. The nearby Asn40, which plays a key role in Mac1 function (16), forms a 3 Å

hydrogen bond with the C3' hydroxyl. In the adenosine site, W19 acts as a bridge between the oxyanion subsite and the C2' hydroxyl (**Fig. 6F**). The second bridging water in this site, W17, forms hydrogen bonds to the backbone oxygen of Ala154, the backbone nitrogen of Leu126, and is within hydrogen bonding distance of both the N3 nitrogen and the O4' oxygen of ADPr. In summary, the structure of ADPr bound to Mac1 determined by neutron diffraction has revealed the precise configuration of hydrogen bond networks in the active site, and highlights the role of bridging water molecules in ADPr binding.

ADPr binding is conserved from pH 4-10

The high resolution X-ray structures used to investigate changes in water networks upon ADPr binding were obtained using crystals grown at pH 9.5 (14). To investigate ADPr binding across a broader pH range, we repeated the pH-shift experiment, this time soaking crystals in a buffer supplemented with 20 mM ADPr. Similarly to the apo pH-shift experiment, the diffraction properties of the crystals at pH 4-10 were excellent, with reflections recorded beyond 0.8 Å at all pHs. Data were truncated to 0.9 Å to achieve 100% completeness, with the exception of the dataset at pH 9, which was recorded at 17 keV resulting in ~100% completeness at 0.84 Å (**Table S1**). There was clear density for ADPr in the active site at all pHs, with the refined occupancy exceeding 95% (**Fig. S1**). Like the apo crystals, we observed a pH-dependent contraction along the c-axis of the unit cell, although the a/b-axis expansion was absent (**Fig. S9**). Compared to the apo structures, there was only a slight increase in disorder going from pH 4 to 10, consistent with stabilization of protomer A by ADPr binding. At pH 4, there was weak density for an alternative conformation of His45 in protomer A (**Fig. S10**). The conformational change involved an 8 Å shift in the His45 side chain, with an adjustment in the backbone torsion angles of Gly46/47 and rearrangement of water molecules to occupy the cavity left by His45 (**Fig. S10**). The change in His45 conformation is likely induced by protonation at pH 4, indicating that the doubly protonated tautomer is sterically incompatible with the buried position. Despite the conformational change in this active site loop, the coordinates and B-factors of ADPr were highly similar across the entire pH range (**Fig. S9**), as were the positions and B-factors of active site water molecules (**Fig. 6G**). These pH-shift experiments show that Mac1 can bind ADPr from pH 4 to 10 in the P4₃ crystals.

Role of water networks in fragment binding

To investigate the role of bridging water molecules in the binding of non-native ligands to Mac1, we analyzed water networks in the previously published structures of 232 fragments bound to Mac1 (14). The fragment structures were determined as part of a large-scale screen using X-ray crystallography, and included 196 fragments bound in the Mac1 active site. Our previous analysis revealed that in some cases, fragment binding occurred through extensive water-mediated interactions (14). Here, we extend this analysis to include the newly identified water networks from the room temperature X-ray and neutron crystal structures. First, we created a map of bridging water density in the adenosine site using multi-Gaussian spreading (56) (**Fig. 7A**). Waters were classified as bridging if they were within 3.5 Å of both a fragment hydrogen bond donor/acceptor and a protein hydrogen bond donor/acceptor. Of the 196 fragments binding in the Mac1 active site, 130 fragments had at least one bridging water, with an average of 2.1 bridging waters per fragment. To quantify the extent of clustering around apo waters in the active site, we measured the distance between bridging waters and the nearest apo water (**Fig. 7B**). Clusters with at least 10 waters were found near W10/W12, and W15-19. The largest cluster was near W17, with 64 waters within 2 Å and 42 waters within 0.2 Å. W17 sits at the base of the adenine subsite, and is involved in bridging interactions with ADPr (**Fig. 6F**). Out of the 64 fragments that hydrogen bond to W17, the fragment atom was an oxygen hydrogen bond donor in 59 of them. The remaining five contained either nitrogen hydrogen bond donors or acceptors. Visualization of the fragment atoms that hydrogen bond to W17 shows that this interaction is highly flexible (**Fig. 7D**). Hydrogen bonds between W17 and fragment atoms were distributed from 2.6-3.5 Å, and over a wide range of angles (**Fig. 7E-G**). The flexibility in bridging interactions with W17 may explain the large number of carboxylic acid containing fragments that bound to the oxyanion

subsite (14). Of the 64 fragments interacting with W17, 55 were carboxylic acids. In particular, W17 allows carboxylic acids to interact with the oxyanion subsite in several distinct ways, based on bond connectivity and carboxylic acid orientation (**Fig. 7H**). Bridging waters were also identified beyond the adenosine site, with three nearby clusters visible in the water density map (labeled WB1, WB2 and WB3 in **Fig. 7A**).

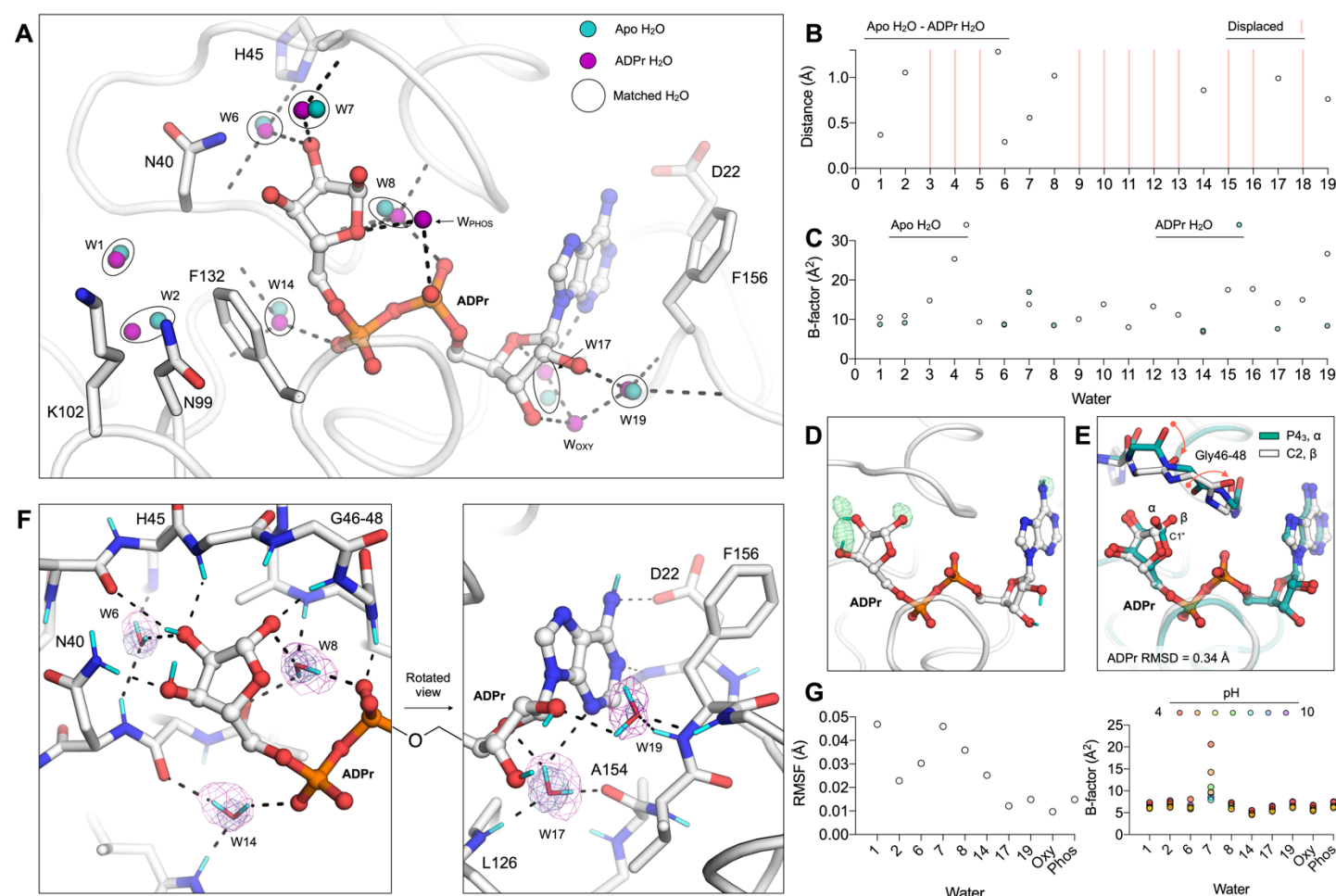


Fig. 6. Reorganization of water networks upon ADPr binding. **(A)** Active site water positions from the structure of ADPr-bound Mac1 determined at 100 K (P4₃ crystal, PDB code 7KQP) compared to the apo structure (PDB code 7KQO). For clarity, only selected side chains of the ADPr-bound structure are shown (white sticks). Hydrogen bonds are shown as dashed black lines. **(B)** Plot showing distances between the water molecules shown in **(A)**. **(C)** Plot showing the B-factors for waters shown in **(A)**. **(D)** Evidence for hydrogen-deuterium exchange in ADPr co-crystallized with Mac1 (C2, PDB code 7TX5). The mFo-DFc NSL density map calculated after joint neutron/X-ray refinement but prior to adding deuterium atoms to ADPr is shown with green mesh (contoured at +3 σ). No density was observed for the C2' and C3' hydroxyl deuteriums. **(E)** Alignment of the ADPr-bound Mac1 structures determined using P4₃ and C2 crystals. The two configurations of the terminal ribose are marked (α and β), and the conformational change required to bind the α configuration in the C2 crystal is shown with red arrows. **(F)** Hydrogen bond networks in the ADPr-bound Mac1 structure determined using neutron diffraction (C2, PDB code 7TX5). The 2mFo-DFc NSL density map is shown for bridging water molecules with a purple mesh (contoured at 2.5 σ for W6, W8, W14 and W17, and at 2 σ for W19). The 2mFo-DFc electron density map is contoured at 2 σ (blue mesh). Deuterium atoms are colored cyan and hydrogen bonds are shown with dashed black lines. **(G)** Left: plot showing the RMSF of active site water molecules calculated across the seven ADPr-bound structures determined from pH 4 to 10. Right: B-factors for water molecules in the ADPr-bound structures from pH 4 to 10.

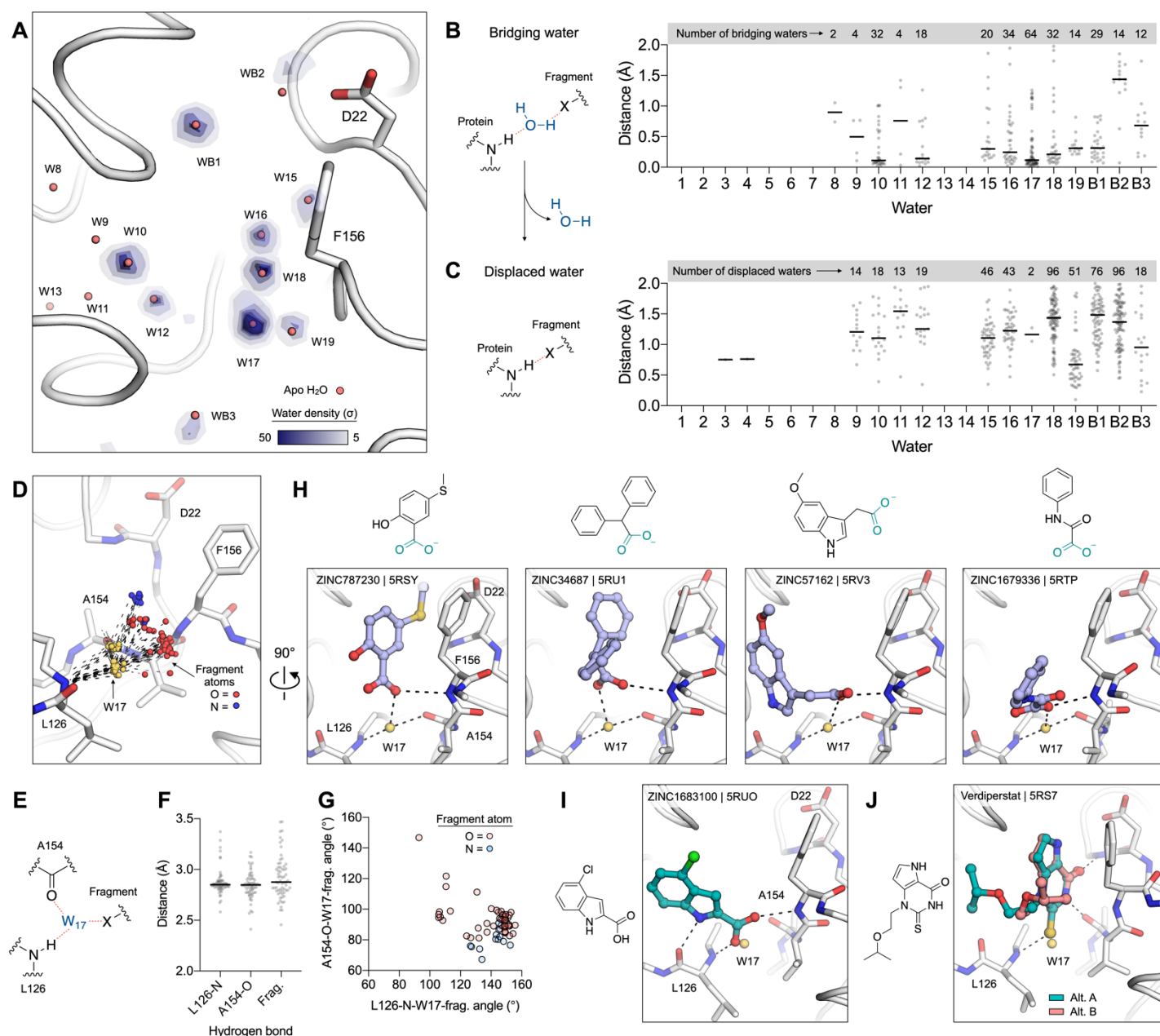


Fig. 7. Water networks mediate fragment binding in the adenosine site of Mac1. (A) Structure of Mac1 showing density of fragment-protein bridging water molecules, calculated using the 232 previously reported fragment structures (14). The P4₃ 100 K structure (PDB code 7KQO) is shown with a white stick/cartoon representation. Water density was calculated with the GROMAPS tool (56) and is contoured from 5 to 50 σ . (B) Left: chemical structure showing an example of a bridging water molecule. Right: distances from bridging waters to apo waters calculated for all fragment structures. (C) Left: chemical structure showing an example of water displacement. Right: the minimum fragment-water molecule for each fragment structure. (D) W17 acts as a bridge to 64 fragments binding in the adenosine site. The Mac1 structure (PDB code 7KQO) is shown, and W17 from the 64 fragments is shown with yellow spheres. The hydrogen bonds to Leu126 and Ala154 are shown with black dashed lines, and the fragment atoms are shown with spheres colored by atom type. (E,F,G) The W17-Leu126/Ala154 hydrogen bonds are conserved across the 64 bridging fragments, whereas the W17-fragment bonds are variable, based on distance (F) and angle (G). (H) W17 mediates diverse interactions with carboxylic-acid containing fragments and the oxyanion subsite. The fragments are shown with blue sticks. (I,J) W17 is displaced by 2 out of the 178 fragments binding in the adenosine site. Two conformations of Verdiperstat were observed.

Next, we examined the displacement of waters from the adenosine network by fragment binding. A water was classified as displaced if a fragment atom was within 2 Å of the water in the apo structure (**Fig. 7C**). On average, each fragment binding in the Mac1 active site displaced 1.6 waters from the active site network. The most frequently displaced waters were W15, W16, W18 and W19, which matches the large number of fragments bound in the adenine and oxyanion subsites (14). Waters that were highly ordered in the apo structure were displaced infrequently or not at all. Notably, W17 was only displaced by two fragments out of 178 binding in the adenosine site (**Fig. 7C**). In both cases, the fragments form extensive hydrogen bond networks to neighboring residues (**Fig. 7I/J**). The indole-containing ZINC1683100 forms hydrogen bonds to the backbone nitrogen/oxygen of Leu129 and to the oxyanion subsite (**Fig. 7I**), while the thioketone of Verdiperstat forms hydrogen bonds with the Leu126 backbone nitrogen and the adenine subsite (**Fig. 7J**). In summary, the perturbation of water networks in the adenosine site by fragment binding is similar to the perturbation seen upon ADPr binding (**Fig. 6A**): waters that are highly ordered in the apo structure are exploited to form bridging interactions in the ligand-bound enzyme. Furthermore, analysis of the fragment-bound structures shows that the majority of waters in the adenosine site network are available for bridging interactions or displacement (**Fig. 7B/C**). This analysis highlights the potential for solvent networks to be exploited in the design of Mac1 ligands, through the introduction of groups that extend bridging hydrogen bond networks or through the targeted displacement of water molecules.

Discussion

The neutron structures reported here advance our understanding of Mac1 structure and function in three main ways. First, the structures allowed the protonation state of histidine residues to be assigned at pH 6.5 and 9.5 in the apo and ADPr-bound enzyme (**Fig. 3C**). Second, the neutron structures helped to identify hydrogen bond networks that control functionally important loop flexibility in and around the active site (**Fig. 4D-I**). Third, the neutron structures enabled the orientation of ordered water molecules to be determined, helping to construct a detailed map of water networks in the Mac1 active site (**Fig. 5, Fig. 6**). These three main findings collectively inform the potential for inhibitor optimization and determination of the catalytic mechanism of Mac1.

The combination of diffraction data collection using different crystal forms, at different temperatures, and at different pHs show that the water networks in the Mac1 active site are surprisingly conserved across a range of perturbations. In contrast, ADPr binding leads to reorganization of the water networks with the displacement of 11 waters, the stabilization of two new waters and small shifts in the remaining waters (**Fig. 6A**). The displacement of weakly bound waters, and the recruitment of tightly bound waters in bridging interactions is consistent with previous large scale surveys of ligand-water-protein interactions (57–59). Although ADPr recognition involves a series of changes in Mac1 conformation that either introduce new polar contacts with ADPr (e.g. the Gly48/Ala129 peptide flips) or remove steric blocks (e.g. the coupled Phe132/Asn99 side chain rotation), small shifts in tightly bound waters are seen in the Mac1-ADPr complex (**Fig. 6B**). This may reflect a general principle governing water-mediated ligand-protein contacts: small shifts in the positions of water molecules leads to ligand stabilization that would be difficult to achieve through direct protein-ligand contacts. Our analysis of bridging water molecules in the structures of ~200 fragments bound in the adenosine site suggests that these principles extend to non-native ligands (**Fig. 7B**). Notably, W17 acts as a bridging water in the ADPr complex, and also forms bridging interactions with 64 fragments.

The detailed map of active site water molecules presented here will help ligand discovery efforts against Mac1 for three main reasons. First, the water network analysis shows that bridging waters in native and non-native ligands are clustered around the position of water molecules in the apo structure (**Fig. 6A, Fig. 7A**). This correlation can be leveraged by minimizing the energy penalty associated with solvent reorganization. The

involvement of the same waters in bridging interactions with a large collection of chemically diverse fragments is promising for virtual screening efforts against Mac1, because it shows that water positions can be taken directly from the apo structure, which is consistent with previous observations (60). Second, the water network analysis identified buried waters that can be targeted for displacement. This strategy is frequently invoked in ligand optimization efforts (48, 61–63), but it is unclear how broadly applicable this strategy is for optimizing potency. The large number of bridging waters identified here provide an unmatched opportunity to systematically examine the thermodynamics of water displacement (for example, displacement of W17 by the indole-containing ZINC1683100 in **Fig. 7i**). Third, receptor models for virtual screening can be improved by using the experimentally determined water orientations from the Mac1 neutron structures (**Fig. 5F**, **Fig. 6F**). The orientation of water molecules included in receptor models can have a major impact on the success of virtual screening (64). One caveat to using the neutron-determined orientations is that only orientations of highly ordered water molecules could be determined with high confidence (**Fig. 5D/E**). Fortunately, these waters are most frequently involved in bridging interactions. Simulation-based methods might yield an improved model of the orientations of disordered water molecules (65).

We found that the conformational ensemble of Mac1 was unusually robust to temperature and pH perturbation in the crystal. Room temperature crystallography can reveal low occupancy, functionally relevant conformational states of proteins (51, 66–68). The 1.1 Å room temperature structure of Mac1 reported here, in combination with the previously reported 0.85 Å structure determined using the same crystal form at 100 K (14), allowed a detailed analysis of temperature dependent changes in structure and function. Overall, the structure and flexibility at 100 and 293 K were remarkably similar, indicating that unlike several recently reported examples (68–70), Mac1 displays minimal temperature dependent structural changes *in crystallo* (**Fig. 4C**). Similarly, assuming that the crystals have equilibrated to the target pH in the ~4 hour soaks conducted here (71), the pH shift experiments revealed that the Mac1 active site residues and water networks are remarkably robust (**Fig. 6I/J**), and this is matched by the invariance in ADPr binding across the entire pH range (**Fig. 6G/H**). One caveat to the pH-shift experiments is the pH-dependent binding of citrate, acetate and CHES in the oxyanion subsite (**Fig. S1**). Although this complicated the analysis of pH-dependent structural changes, this interference would apply equally to solution-based experiments, and highlights the protein binding potential of commonly used buffers (72). A valuable use of the pH-shift experiment described here would be to investigate the pH-dependent binding of ionizable inhibitors to Mac1. Collectively, these experiments highlight a conspicuous problem related to using crystallography to study protein flexibility: the same tightly packed crystal lattices that help to deliver the high resolution information needed to identify subtle changes in structure may lead to protein stabilization that masks the same structural changes.

We discovered conformational diversity by comparing Mac1 structures across different crystal forms (**Fig. 4C**). Importantly, this ensemble of structures captures most of the structural diversity seen upon ADPr and fragment binding (**Fig. S6**). One notable exception is the Phe132 side chain, which adopts a similar conformation in all of the apo crystal structures, but is stabilized in a different rotamer in the ADPr-bound structure (**Fig. S6**). Surprisingly, the relatively constant position of Phe132 across the apo structures occurs despite substantial variation in the Asn99/Lys102 loops (**Fig. S6**). This stability may help to explain why our fragment screening campaign failed to identify any fragments binding in the terminal ribose site (14): fragments are unable to overcome the energetic penalty required to stabilize the Phe132 side-chain in the ADPr-bound “flipped” conformation. The lack of fragments in this site may also be due to the stability and the associated enthalpic cost of displacing the highly ordered waters in this site (e.g. W13 and W14). While our previous virtual screening campaign focussed solely on identifying fragments binding to the adenosine site, a worthwhile extension would be to use the ADPr-bound conformation as a receptor model to target the catalytic site.

Although the physiologically relevant substrates of Mac1 are unknown, macrodomains are thought to primarily hydrolyse ADPr conjugated to glutamate or aspartate side chains at the C1'' position (**Fig. 1A**) (73). Based on X-ray and neutron crystal structures of the SARS-CoV-2 macrodomain with ADPr bound, two substrate binding modes, distinguishable based on the conformation of the terminal ribose (**Fig. 8A/B**) and leading to different potential catalytic mechanisms are plausible. In the first mode, which matches the conformation seen in the P4₃ crystal structure, the terminal ribose adopts the α configuration, with the conjugated peptide exiting the active site through the gap between the Gly46-48 loop and the Ile131 loop. Although this gap is narrow in the apo structures, it widens by ~ 1 Å in the ADPr-bound structure, and we previously identified a fragment wedged between these loops (PDB code 5RVI) (14). In this binding mode, the proximity of the C1'' carbon to the phosphate groups, and the absence of another suitable general base, implicates the previously proposed substrate-assisted hydrolysis mechanism with either W_{PHOS} or W8 acting as a nucleophile (15, 16) (**Fig. 8C**). If the hydrolysis reaction involves an S_N2 type nucleophilic substitution at the C1'' carbon, then we favor W8 as the water nucleophile, because it can approach the sp³ carbon with a collinear trajectory relative to the peptide leaving group (**Fig. 8C**) and the neutron crystal structure with ADPr-bound shows that the water oxygen is oriented towards the C1'' carbon (**Fig. 6G**). This is in contrast to the water nucleophile proposed by Jankevicius *et al.* (15) in hMacroD2, which is equivalent to W_{PHOS} in Mac1 (**Fig. 8E**).

In the second possible ADPr binding mode, the ribose adopts the β configuration and the terminal ribose is rotated 180° about the C4''-C5'' bond relative to the configuration observed in the P4₃ crystals (**Fig. 8B**). In this binding mode, a substrate-assisted mechanism is unlikely because the C1'' carbon would be >6 Å from the diphosphate. However, several lines of evidence support the nearby His45 as a general base in the hydrolysis mechanism for coronaviral macrodomains (**Fig. 8D**). First, the neutron structures show that His45 has the appropriate orientation and protonation state to act as a general base, with a deprotonated N δ 1 nitrogen and a D₂O molecule hydrogen bonded between His45 and ADPr (**Fig. 3D**). Second, His45 is conserved across all known coronavirus macrodomains, with mutation to alanine eliminating ADPr-hydrolysis activity (16). It should be noted that His45 is absent in alphavirus macrodomains and the human macrodomain hMacroD2, although both macrodomains contain a nearby nucleophilic residue (Asp102 in hMacroD2 and Cys34 in the Chikungunya virus macrodomain) (**Fig. 8E**). Third, the binding mode with the terminal ribose rotated $\sim 180^\circ$ was observed in the structure of the macrodomain from Tylonycteris bat coronavirus HKU4 with NAD⁺ bound (74) (**Fig. 8F**). The nicotinamide group is oriented towards a groove formed between Tyr42 and Lys102. We previously identified several fragments binding in this site (14), and have uncovered conformational flexibility consistent with the site being able to accommodate a range of peptide substrates exiting the Mac1 active site (**Fig. S6**). Moreover, docking of the potential Mac1 substrate PARP-1 into the Mac1 active site places a portion of the PARP-1 peptide in this groove (74). Finally, this binding mode hints at an alternative role for Asn40 in ADPr hydrolysis. Asn40 is strictly conserved across macrodomains and mutation to alanine eliminates activity (10). In the substrate-assisted mechanism, this functional importance has been attributed to the hydrogen bond formed between Asn40 and the C3'' hydroxyl of ADPr. In the alternative mechanism proposed here (**Fig. 8B/D**), Asn40 could have a direct role by stabilizing the peptide leaving group. Discriminating between the two ADPr binding modes would require a structure of Mac1 in complex with a suitable substrate mimic (e.g. a non-hydrolyzable ADP-ribosylated protein or peptide). As has been suggested previously (16), these two possible mechanisms raise the intriguing possibility that distinct mechanisms of ADPr-hydrolysis have evolved from highly similar protein scaffolds.

In summary, we have determined neutron crystal structures of the SARS-CoV-2 NSP3 macrodomain using multiple crystal forms across the apo and ADPr-bound states. We assigned the protonation states of histidine residues, and mapped hydrogen bond networks that control the flexibility of active site loops that are involved in the recognition of ADPr. A comprehensive analysis of water networks in the Mac1 active revealed that the water

networks are conserved across different crystal forms, between 100 and 293 K and from pH 4 to 10. The water networks reorganize upon ADPr binding, and play a key role in the molecular recognition of non-native ligands in the adenosine site. The series of structures presented here increase our understanding of macrodomain catalytic mechanisms and will serve as a resource to guide the design of new Mac1 inhibitors.

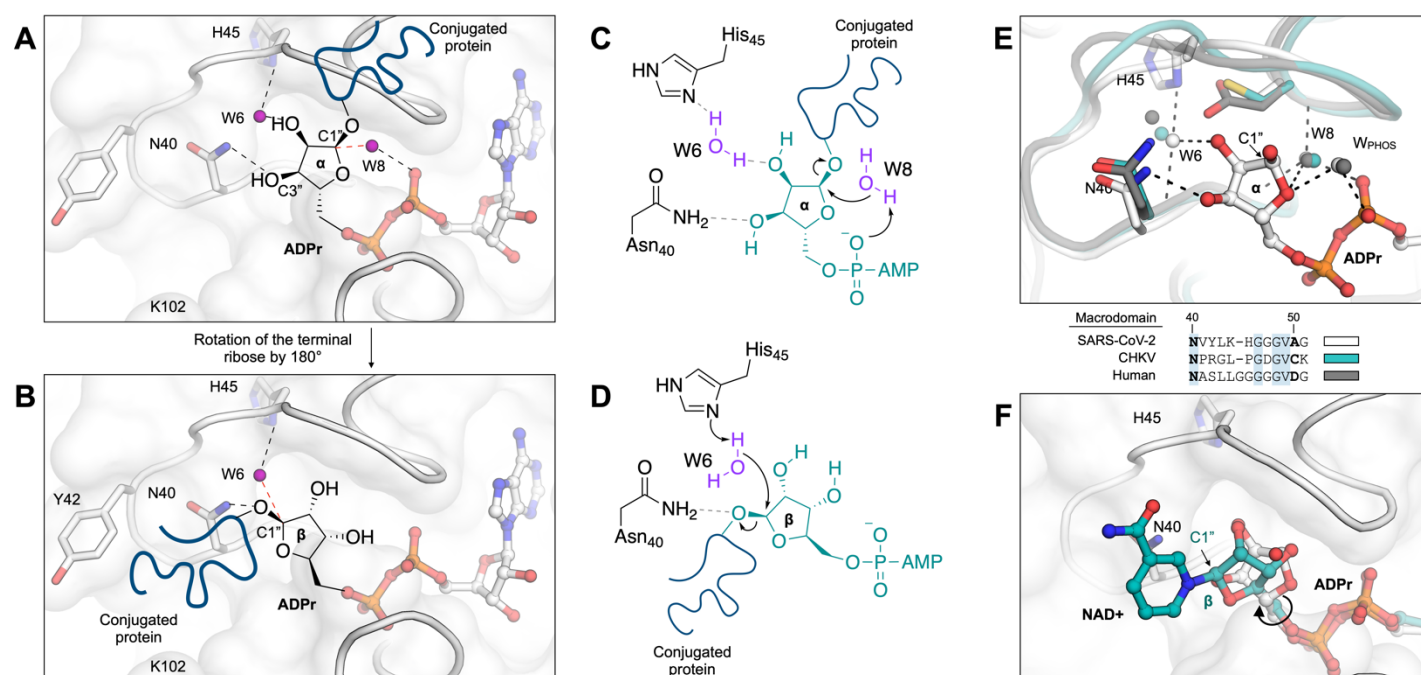


Fig. 8. Mechanism of ADPr-ribose hydrolysis catalyzed by the SARS-CoV-2 NSP3 macrodomain. (A) Composite image showing the ADPr-bound Mac1 structure (white sticks/cartoon/surface, PDB 7KQP). ADPr is shown in the configuration that is compatible with a substrate-assisted mechanism. The terminal ribose adopts the α -configuration and W8 acts as the water nucleophile. For clarity, only the side chains of selected residues are shown. Hydrogen bonds are shown with dashed black lines and the W6-C1'' trajectory is shown with a dashed red line. (B) Same as (A), but showing the β -configuration of the terminal ribose that is compatible with His45 acting as a general base to activate W6 as a nucleophile. (C, D) Chemical structures showing the two possible mechanisms for ADPr hydrolysis. (E) Top: structural alignment of the SARS-CoV-2 macrodomain (PDB code 7KQP), Chikungunya virus macrodomain (PDB code 3GPO) and the human macrodomain hMacroD2 (PDB code 4IQY), all in complex with ADPr. Bottom: protein sequence alignment of residues equivalent to residues 40-51 from the SARS-CoV-2 macrodomain. (F) X-ray crystal structure of the NSP3 macrodomain from the Tylonycteris bat coronavirus HKU4 in complex with NAD⁺ (PDB code 6MEB). The terminal ribose is rotated $\sim 180^\circ$ relative to ADPr. This configuration matches the model proposed in (B/D).

1 Materials and Methods

2 P4₃ crystals

3 Protein expression and purification

4 The gene encoding SARS-CoV-2 NSP3 Mac1 (residues 3-169) was cloned into a pET-22b(+) expression
5 plasmid with a TEV-cleavable N-terminal 6-His tag (Genscript). The protein was expressed and purified as
6 described previously (14).

7 Crystallization for neutron diffraction

8 Neutron diffraction quality crystals were grown using sitting drop vapor diffusion. Crystallization drops were set
9 up on 22 mm siliconized glass coverslips (Hampton Research, HR3-231) with 65 μ l protein, 60 μ l reservoir (34%
0 PEG 3000, 100 mM CHES pH_{apparent} 9.1 (pD 9.5)) and 5 μ l seed stock. Seed stocks were prepared as described
1 previously (14) and were stored at -80°C. The sitting drops were equilibrated against a 4 ml reservoir solution
2 stored in a small plastic beaker, using two 10 cm plastic petri dishes taped together to create a sealed
3 environment. Crystals reached their maximum size of approximately 2x2x0.5 mm after one week. To increase
4 the deuterium content, crystals were transferred from crystallization drops into an artificial reservoir solution
5 containing 34% PEG 3000 and 100 mM CHES pH 9.5, prepared using D₂O. Because crystals were tightly bound
6 to the coverslips, it was necessary to use a fine needle to dislodge them. Crystals were allowed to soak in the
7 artificial reservoir solution for 24 hours, before being transferred into a 2 mm thin walled quartz capillary
8 (Hampton Research, HR6-150). To keep crystals hydrated, 20 μ l of the D₂O-containing reservoir solution was
9 pipetted into one end, and the capillaries were sealed with wax at both ends (Hampton Research, HR4-328).

0 Neutron/X-ray diffraction data collection, reduction and refinement

1 Neutron time-of-flight diffraction data were collected at room temperature on the MaNDi instrument at the SNS
2 (32, 75). An incident neutron wavelength bandpass of 2-4 Å was used with a $\Delta\phi$ of 10° between frames. A total
3 of 10 diffraction images were collected with an average exposure of 16 h per frame. Exposure was set to 80
4 Coulomb of charge on the SNS target. At 1.4 MW operation, charge accumulates at a rate of approximately 5
5 C/h. Fluctuation in exposure time occurs because the accelerator does not operate 100% of the time. The
6 neutron dataset was reduced using the Mantid package (76) and integrated using three dimensional profile fitting
7 (77). The data were wavelength normalized using LAUENORM from the LAUEGEN suite (78–80).

8
9 Following neutron diffraction data collection, an X-ray dataset was collected on the same crystal at room
0 temperature using a microfocus rotating anode X-ray diffractometer (Rigaku HighFlux HomeLab equipped with
1 a MicroMax-007 HF X-ray generator, Osmic VariMax optics and an Eiger R 4M hybrid photon counting detector).
2 A total of 360 diffraction images were collected with a $\Delta\phi$ of 0.5° and an exposure time of 10 s per frame. X-ray
3 diffraction data were integrated using the CrysAlis Pro software suite (Rigaku Inc.). Diffraction data were then
4 reduced and scaled using the Aimless (81) program from the CCP4 suite (82). The X-ray data collection statistics
5 are shown in Table S1.

6
7 The Mac1 structure was initially refined against the X-ray data only. Phases were obtained by molecular
8 replacement with Phaser (83) using the coordinates of the apo Mac1 as the search model (monomer A of PDB
9 code 7KQO, stripped of alternative conformations and waters). To debias refinement from the coordinates used
0 for molecular replacement, the B-factors of 7KQO were reset to 20 Å² and the coordinates were randomly offset
1 by 0.3 Å using phenix.pdbtools. The structure was refined with phenix.refine (84) using default parameters and

five macrocycles. Manual model building was performed with Coot (85). In the early stages of refinement, waters were added automatically in phenix.refine to peaks in the mF_o-DF_c map 3.5σ or higher. CHES was modeled into the adenosine site with restraints generated using phenix.elbow (86). The occupancy of CHES was set to 50% initially and was refined automatically in subsequent steps. Positive peaks in the mF_o-DF_c electron density map that overlapped with CHES were modeled as partially occupied water molecules. The occupancy of CHES and the partially occupied waters was constrained to 100% in phenix.refine, with the final CHES occupancy equal to 62%. In the later stages of refinement, hydrogens were added to the model using Reduce (42) run through phenix.ready_set. Hydrogens were refined at their riding positions. Cycles of refinement and model building continued until R_{free} no longer decreased. The coordinates were prepared for joint neutron/X-ray refinement by adding deuterium atoms to all exchangeable positions and D_2O molecules with phenix.ready_set. Therefore, exchangeable positions were modeled with both a hydrogen and deuterium using A/B alternate location identifiers, and the deuterium occupancy was automatically refined. Joint refinement was performed using phenix.refine (35) with hydrogens/deuteriums refined individually. Hydrogens/deuterium ADPs were refined and neutron distances were used for hydrogen/deuterium bonds. After five macrocycles of refinement, the $2mF_o-DF_c$ (unfilled) and the mF_o-DF_c NSL density maps were inspected in Coot and modifications made to the coordinates. To assign histidine protonation states, mF_o-DF_c NSL density maps were inspected after refinement with either the doubly protonated or doubly deprotonated forms (**Fig. S5**). Water orientations were automatically refined with phenix.refine. The fit of D_2O molecules into $2mF_o-DF_c$ NSL density maps was assessed by visual inspection, and the orientations of D_2O molecules were adjusted when necessary. Deuteriums were deleted from water molecules that lacked supporting density.

Crystallization for high resolution room temperature X-ray diffraction

Crystals were grown using hanging drop vapor diffusion in 24-well crystallization plates (Hampton, HR3-171). Crystallization drops contained 4 μ l Mac1 at 40 mg/ml, 2 μ l seeds and 2 μ l reservoir consisting of 32% PEG 3000 and 100 mM CHES pH 9.5 with 500 μ l reservoir in each well. Crystals reached a maximum size of approximately 1x1x0.25 mm after four days. Crystals were looped using dual thickness MicroLoops (Mitegen, M5-L18SP) mounted on B3-3 magnetic bases (Mitegen, GB-B3-R). MicroRT tubing (Mitegen, RT-T1) with 5 μ l of reservoir in the end was placed over the crystals to prevent dehydration, and the tubing was sealed to the crystal mount with epoxy.

High resolution room temperature X-ray diffraction data collection, reduction and refinement

X-ray diffraction data was collected at the Stanford Synchrotron Radiation Lightsource (SSRL) using beamline 12-1. The cryojet temperature was set to 293 K and crystals were mounted remotely from an SSRL crystallization plate (Mitegen, M-CP-111-095). To assess the susceptibility of crystals to radiation damage, datasets were collected from four similarly sized crystals with the X-ray dose increased by adjusting the beam transmittance while keeping the other data collection parameters constant (see **Table S1**). The absorbed dose for each dataset was calculated with RADDOSE-3D (87) using the parameters listed in **Table S1**. X-ray diffraction data was indexed, integrated and scaling used XDS (88) and merged with Aimless (81). Data collection statistics are shown in **Table S1**. Phases were obtained by molecular replacement with Phaser using the same search model as the lower resolution X-ray experiment. Refinement was performed with phenix.refine using default settings and five macrocycles at each step. Water molecules were added automatically in phenix.refine using the same settings as the lower resolution X-ray refinement. ADPs were refined isotropically in the first two cycles of refinement and anisotropically in the remaining cycles. After three cycles of refinement and model building, hydrogens were added to the model using Reduce. Hydrogens were refined using a riding model, with their ADPs refined isotropically. In the final two stages of refinement, the weights for the stereochemical and ADP restraints were relaxed (using the $wxc_scale=1.5$ and $wxu_scale=1.5$ flags in phenix.refine). Refinement and model building continued until R_{free} no longer decreased. In the final refinement cycle, automatic water modeling

was switched off, and waters were placed manually into peaks in the mF_o-DF_c electron density map. Coordinates and structure factor intensities were deposited in the PDB with accession codes 7TWF (73 kGy), 7TWG (153 kGy), 7TWH (290 kGy) and 7TWI (539 kGy).

pH-shift experiment

Crystals for the pH shift experiment were grown at 20°C using 96-well sitting drop crystallization plates (Hampton Research, HR3-125). Crystallization drops contained 200 nl protein, 100 nl reservoir solution (28% PEG 3000, 100 mM CHES pH 9.5) and 100 nl seed stock, with 30 µl in the reservoirs. Crystals grew to their maximum size in two days with dimensions approximately 0.4x0.4x0.1 mm. Crystals were transferred from crystallization plates into a buffer containing 22 mM citric acid, 33 mM HEPES and 44 mM CHES (2:3:4 molar ratio of the buffer components) (54) and 28% PEG 3000, with the pH adjusted from 3 to 10. Crystals were soaked for four hours, then vitrified in liquid nitrogen. The soaks with ADPr were performed in the same manner, but the buffers were supplemented with 20 mM ADPr (Sigma, A0752) (from a 500 mM stock prepared in H₂O). Diffraction data were collected at beamline 8.3.1 of the Advanced Light Source (ALS) for 3-6 crystals at each pH. Data collection strategy and statistics are shown in **Table S1** for the best diffracting crystals at each pH. Phases were obtained by molecular replacement using the same search model as the room temperature experiments. Refinement was performed with phenix.refine using a similar protocol to the high resolution room temperature structures. ADPs were refined anisotropically (excluding hydrogens) and the occupancy of all water molecules was refined. Hydrogens were refined using a riding model, and in the final two stages of refinement the weights for stereochemical and ADP restraints were relaxed (using the wxc_scale=2 and wxu_scale=2 flags in phenix.refine). pH dependent binding of buffer components were observed in the apo crystal structures, these were refined with restraints generated by phenix.elbow for CHES or from the internal CCP4 library for citrate and acetate. ADPr binding was confirmed by inspection of mF_o-DF_c electron density maps, with restraints generated by phenix.elbow. Coordinates and structure factor intensities were deposited in the PDB with accession codes 7TWJ (pH 4, apo), 7TWN (pH 5, apo), 7TWO (pH 6, apo), 7TWP (pH 7, apo), 7TWR (pH 8, apo), 7TWQ (pH 9, apo), 7TWS (pH 10, apo), 7TWT (pH 4, ADPr), 7TWV (pH 5, ADPr), 7TWW (pH 6, ADPr), 7TWX (pH 7, ADPr), 7TWY (pH 8, ADPr), 7TX0 (pH 9, ADPr), 7TX1 (pH 10, ADPr). Data reduction and refinement statistics are presented in **Table S1**, with the statistics produced with phenix.table_one.

P2₁/C2 crystals

Protein expression and purification

The gene encoding SARS-CoV-2 2-170 NSP3 Mac1 was cloned into a pET-11a plasmid (Bio Basic) and transformed into *E. coli* (BL21-DE3). A detailed procedure for Mac1 expression and purification has been published elsewhere (36).

Crystallization for neutron diffraction

Purified Mac1 was stored at 4°C and filtered through a 0.2 µm centrifugal filter prior to crystallization trials. Initial crystals grown in a 20 µl drop of Mac1 (~17.5 mg/ml) mixed with 28% PEG 4000, 0.1 M MES pH 6.5 at a 1:1 ratio at 16°C were used to prepare a microseed stock with seed beads (Hampton Research). Large-volume crystals of ligand-free Mac1 were grown using a Hampton nine-well sandwich box set up with 150 µl drops of Mac1 (~19 mg/ml) mixed at a 1:1 ratio with 25% PEG 4000, 0.1 M MES pH 6.5 seeded with 0.2 µl microseeds (1:3000 dilution). Crystals were left to grow undisturbed at 16°C for >2 months resulting in a crystal measuring ~1.9 mm at the longest edge. The crystal was mounted in a fused quartz capillary accompanied by a plug of 27% PEG 4000 prepared in 100% D₂O. For the ADP-ribose complex crystal, Mac1 (~15 mg/ml) was mixed with ADPr at a 1:2 molar ratio, incubated at room-temperature for 30 minutes. The crystallization conditions were

identical to those for ligand-free Mac1 crystal growth. The crystal drops were seed struck with a cat whisker from a small Mac1+ADPr co-crystal previously grown in a 24-well plate using a 24% PEG 4000, 0.1 M MES pH 6.5 well solution, and then the nine-well plate was incubated at 16°C. A Mac1+ADPr crystal measuring ~2.4 mm at the longest edge was mounted in quartz capillary with 27% PEG 4000 prepared in 100% D₂O. At the time of mounting, both crystallization drops were measured with a microelectrode to be pH 6.5.

Neutron/X-ray diffraction data collection and reduction

The full neutron diffraction dataset for ligand-free Mac1 crystal at pH 6.5 was collected using the MaNDi instrument and the data were processed in a similar fashion as for the ligand-free Mac1 crystal at pH 9.5. Neutron diffraction data for the Mac1-ADPr complex were collected at the IMAGINE instrument at the High Flux Isotope Reactor (HFIR) at Oak Ridge National Laboratory (37, 89, 90). The crystal was held stationary at room temperature, and diffraction data for each image were collected for 23 hours using neutrons between 3.3-4.5 Å. During the experiment, the next 23-hour diffraction image was collected after crystal rotation by $\Delta\phi = 10^\circ$. IMAGINE neutron diffraction data were processed using the Daresbury Laboratory LAUE suite program LAUEGEN modified to account for the cylindrical geometry of the detector (79, 91). The program LSCALE (92) was used to determine the wavelength-normalization curve using the intensities of symmetry-equivalent reflections measured at different wavelengths. No explicit absorption corrections were applied. These data were then merged in SCALA (93). The neutron data collection statistics are shown in **Table S1**. Following the neutron data collection, the room-temperature X-ray diffraction datasets were collected from the same crystals in the same manner as described previously for the P4₃ crystals. The X-ray data collection statistics are shown in **Table S1**.

Joint neutron/X-ray refinement

Joint neutron/X-ray refinement was performed in the same manner as the P4₃ neutron/X-ray experiment. Phases were obtained by molecular replacement using the same coordinates as the P4₃ experiments. For the ADPr co-crystal structure, binding of ADPr was confirmed by inspection of mF_O-DF_C electron density map. Restraints for the β isomer of ADPr were generated with phenix.elbow from the corresponding SMILES string. Coordinates and structure factor intensities were deposited in the PDB with accession code 7TX4 for the P2₁ apo structure and 7TX5 for the C2 ADPr-bound structure.

Structural analysis

Protein

RMSF values were calculated using a python script in PyMOL (version 2.5.1, Schrodinger). First, structures were aligned using C α atoms. Second, average coordinates for each C α were calculated across the structures. Third, the RMSF was calculated relative to the average C α coordinates. Normalized C α B-factors (B') were calculated using the following equation: $B' = (B - \mu_B) / \sigma_B$, where B is the C α B-factor, μ_B is the average B-factor across all C α atoms and σ_B is the C α B-factor standard deviation. For residues with alternative conformations, only the A conformation was included in the calculations of RMSF/average B-factor. To color structures by RMSF, the B-factor column in PDB files was replaced with the RMSF values. RMSD values were calculated using PyMOL with five cycles of outlier rejection.

The protonation states of histidine residues were assigned based on NSL density maps after refinement with either the doubly or deprotonated tautomers (maps are shown in **Fig. S5**). For the doubly protonated tautomers,

the occupancy of the Nε2/Nδ1 deuterium atoms was set to 100%, and automatic refinement of occupancy was disabled.

Water

RSCC values for D₂O molecules were calculated using phenix.real_space_correlation. The number of contacts made by water molecules was defined as the number of protein nitrogen or oxygen atoms within 3.5 Å of a water, with distances calculated using PyMOL. Normalized B-factors (B') for water molecules were calculated with the following equation: $B' = (B - \mu B) / \sigma B$, where B is the water B-factor, μB is the average B-factor across all water molecules and σB is the water B-factor standard deviation.

To assess the reproducibility of water orientation refinement, we ran 100 independent rounds of refinement starting from the three jointly refined Mac1 structures. Deuterium atoms were added to all D₂O oxygen atoms prior to refinement, and a random offset of 0.5 Å was applied to deuterium atoms using phenix.pdbtools. This procedure meant that each round of refinement started with a random distribution of D₂O orientations, but with the same oxygen position. The variability in D₂O orientations after refinement was quantified by calculating the RMSF value for each water deuterium atom across the 100 structures (**Fig. S8**). Agreement between the orientation of D₂O molecules in the active sites of protomer A and B of the P4₃ structure was assessed by calculating the distance between the average position of each D₂O deuterium from the 100 refined structures (**Fig. 5E**).

The conservation of water molecules between structures was calculated with a python script in PyMOL. First, symmetry mates (including waters) within 5 Å of protein heavy atoms were generated and appended to each coordinate file. Next, the structures were aligned by Cα atoms, and the closest water in the structure being compared was determined, as well as the B-factors for both waters and the minimum distance to a protein hydrogen bond donor/acceptor.

The density map for bridging water molecules in the 232 previously reported Mac1 fragment-bound structures (14) was calculated using the GROMAPS tool (56) in GROMACS (94). First, water molecules related by symmetry within 5 Å of protein heavy atoms were appended to each fragment structure. Next, bridging water molecules were identified by selecting waters within 3.5 Å of both a protein and a fragment hydrogen bond acceptor/donor. No angle requirements were applied to the selection of bridging water molecules. Next, the fragment coordinates were aligned to protomer A of the 0.85 Å P4₃ apo structure (PDB code 7KQO) and the coordinates of the bridging water molecules were saved and concatenated to form a single trajectory in PDB format. For GROMACS to accept the trajectory, it was necessary to add up to six waters to each structure so that the number of waters in each structure was the same. These waters were selected at least 25 Å from the active site. The grid resolution for calculating the density map was set to 0.5 Å.

Acknowledgements

The SARS-CoV-2 Mac1 expression plasmid (pMCSG53) was provided by Dr. Andrzej Joachimiak (Argonne National Laboratory) with support from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract HHSN272201700060C. We thank Dr. Hugh M. O'Neill from ORNL for assistance during expression of the P2₁/C2 constructs. Structural biology applications used in this project at UCSF were compiled and configured by SBGrid (95). This research was supported by the DOE Office of Science through the National Virtual Biotechnology Laboratory (NVBL), a consortium of DOE national laboratories focused on response to COVID-19, with funding provided by the

1 Coronavirus CARES Act. This research used resources at the Spallation Neutron Source and the High Flux
2 Isotope Reactor, which are DOE Office of Science User Facilities operated by the Oak Ridge National
3 Laboratory. The Office of Biological and Environmental Research supported research at ORNL's Center for
4 Structural Molecular Biology (CSMB), a DOE Office of Science User Facility. This research used resources of
5 the Spallation Neutron Source Second Target Station Project at Oak Ridge National Laboratory (ORNL). ORNL
6 is managed by UT-Battelle LLC for DOE's Office of Science, the single largest supporter of basic research in the
7 physical sciences in the United States. The synchrotron X-ray diffraction data used to determine Mac1 structures
8 from pH 4 to 10 were collected at beamline 8.3.1 of the Advanced Light Source. The ALS, a U.S. DOE Office of
9 Science User Facility under Contract No. DE-AC02-05CH11231, is supported in part by the ALS-ENABLE
10 program funded by the National Institutes of Health, National Institute of General Medical Sciences, grant P30
1 GM124169-01. The synchrotron X-ray diffraction data used to determine Mac1 structures at room were collected
2 at beamline 12-1 of the Stanford Synchrotron Radiation Lightsource (SSRL). Use of the SSRL, SLAC National
3 Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy
4 Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is
5 supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health,
6 National Institute of General Medical Sciences (P30GM133894). L.C. acknowledges support by the NIH (R01-
7 GM071939). This work was supported by NIH GM123159, NSF Rapid 2031205, and a TMC Award from the
8 UCSF Program for Breakthrough Biomedical Research, funded in part by the Sandler Foundation (to J.S.F.).

9 Author contributions

0 G. Correy expressed, purified and crystallized the P4₃ construct, reduced X-ray diffraction data for P4₃ crystals,
1 refined and analyzed the neutron and X-ray structures, prepared the figures and wrote the manuscript. D. Kneller
2 crystallized the P2₁/C2 construct, collected and reduced neutron/X-ray data from the P2₁/C2 construct. G.
3 Phillips expressed and purified the P2₁/C2 construct. S. Pant expressed and purified the P2₁/C2 construct. S.
4 Russi provided user support at SSRL, and assisted with the data collection strategy at SSRL. A. Cohen provided
5 user support at SSRL, and assisted with the data collection strategy at SSRL. G. Meigs provided user support
6 at ALS, and assisted with the data collection strategy at ALS. J. Holton provided user support at ALS, and
7 assisted with the data collection strategy at ALS. S. Gahbauer edited the manuscript. M. Thompson assisted
8 with the room temperature X-ray data collection strategy. A. Ashworth supervised work and arranged funding. L.
9 Coates collected and reduced neutron/X-ray data for P2₁/C2 construct. A. Kovalevsky collected and reduced X-
0 ray and neutron data for the P2₁/C2 construct and wrote the manuscript. F. Meilleur collected and reduced
1 neutron/X-ray data for the P4₃ construct and wrote the manuscript. J. Fraser supervised work, arranged for
2 funding and wrote the manuscript.

3 Competing interests

4 A. Ashworth is a co-founder of Tango Therapeutics, Azkarra Therapeutics, Ovibio Corporation; a consultant for
5 SPARC, Bluestar, ProLynx, Earli, Cura, GenVivo and GSK; a member of the SAB of Genentech, GLAdiator,
6 Circle and Cambridge Science Corporation; receives grant/research support from SPARC and AstraZeneca;
7 holds patents on the use of PARP inhibitors held jointly with AstraZeneca which he has benefitted financially
8 (and may do so in the future). J. Fraser is a consultant for, has equity in, and receives research support from
9 Relay Therapeutics.

1 Data and materials availability

2 All data generated or analyzed during this study are included in this article and its Supplementary Information.
3 Crystallographic coordinates and structure factors for all structures have been deposited in the Protein Data
4 Bank with the following accessing codes: 7TWF, 7TWG, 7TWH, 7TWI, 7TWJ, 7TWN, 7TWO, 7TWP, 7TWQ,
5 7TWR, 7TWS, 7TWT, 7TWV, 7TWW, 7TWX, 7TWY, 7TX0, 7TX1, 7TX3, 7TX4, 7TX5.

6 References

- 7 1. M. G. Katze, Y. He, M. Gale Jr, Viruses and interferon: a fight for supremacy. *Nat. Rev. Immunol.* **2**, 675–687 (2002).
- 8 2. L. B. Ivashkiv, L. T. Donlin, Regulation of type I interferon responses. *Nat. Rev. Immunol.* **14**, 36–49 (2014).
- 9 3. J. W. Schoggins, Interferon-Stimulated Genes: What Do They All Do? *Annu Rev Virol.* **6**, 567–584 (2019).
- 0 4. W. M. Schneider, M. D. Chevillotte, C. M. Rice, Interferon-stimulated genes: a complex web of host defenses. *Annu.*
1 *Rev. Immunol.* **32**, 513–545 (2014).
- 2 5. N. C. Hoch, L. M. Polo, ADP-ribosylation: from molecular mechanisms to human disease. *Genet. Mol. Biol.* **43**,
3 e20190075 (2019).
- 4 6. N. C. Hoch, Host ADP-ribosylation and the SARS-CoV-2 macrodomain. *Biochem. Soc. Trans.* **49**, 1711–1721
5 (2021).
- 6 7. J. G. M. Rack, L. Palazzo, I. Ahel, (ADP-ribosyl)hydrolases: structure, function, and biology. *Genes Dev.* **34**, 263–
7 284 (2020).
- 8 8. L. C. Russo, R. Tomasin, I. A. Matos, A. C. Manucci, S. T. Sowa, K. Dale, K. W. Caldecott, L. Lehtiö, D.
9 Schechtman, F. C. Meotti, A. Bruni-Cardoso, N. C. Hoch, The SARS-CoV-2 Nsp3 macrodomain reverses
0 PARP9/DTX3L-dependent ADP-ribosylation induced by interferon signaling. *J. Biol. Chem.* **297**, 101041 (2021).
- 1 9. Y. M. O. Alhammad, M. M. Kashipathy, A. Roy, J.-P. Gagné, P. McDonald, P. Gao, L. Nonfoux, K. P. Battaile, D. K.
2 Johnson, E. D. Holmstrom, G. G. Poirier, S. Lovell, A. R. Fehr, The SARS-CoV-2 Conserved Macrodomain Is a
3 Mono-ADP-Ribosylhydrolase. *J. Virol.* **95** (2021), doi:10.1128/JVI.01969-20.
- 4 10. A. R. Fehr, R. Channappanavar, G. Jankevicius, C. Fett, J. Zhao, J. Athmer, D. K. Meyerholz, I. Ahel, S. Perlman,
5 The Conserved Coronavirus Macrodomain Promotes Virulence and Suppresses the Innate Immune Response
6 during Severe Acute Respiratory Syndrome Coronavirus Infection. *MBio.* **7** (2016), doi:10.1128/mBio.01721-16.
- 7 11. H. Yang, Z. Rao, Structural biology of SARS-CoV-2 and implications for therapeutic development. *Nat. Rev.*
8 *Microbiol.* **19**, 685–700 (2021).
- 9 12. J. Lei, Y. Kusov, R. Hilgenfeld, Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein.
0 *Antiviral Res.* **149**, 58–74 (2018).
- 1 13. M.-H. Lin, S.-C. Chang, Y.-C. Chiu, B.-C. Jiang, T.-H. Wu, C.-H. Hsu, Structural, Biophysical, and Biochemical
2 Elucidation of the SARS-CoV-2 Nonstructural Protein 3 Macro Domain. *ACS Infect Dis.* **6**, 2970–2978 (2020).
- 3 14. M. Schuller, G. J. Correy, S. Gahbauer, D. Fearon, T. Wu, R. E. Díaz, I. D. Young, L. Carvalho Martins, D. H. Smith,
4 U. Schulze-Gahmen, T. W. Owens, I. Deshpande, G. E. Merz, A. C. Thwin, J. T. Biel, J. K. Peters, M. Moritz, N.
5 Herrera, H. T. Kratochvil, QCRG Structural Biology Consortium, A. Aimon, J. M. Bennett, J. Brandao Neto, A. E.
6 Cohen, A. Dias, A. Douangamath, L. Dunnett, O. Fedorov, M. P. Ferla, M. R. Fuchs, T. J. Gorrie-Stone, J. M. Holton,
7 M. G. Johnson, T. Krojer, G. Meigs, A. J. Powell, J. G. M. Rack, V. L. Rangel, S. Russi, R. E. Skyner, C. A. Smith, A.
8 S. Soares, J. L. Wierman, K. Zhu, P. O'Brien, N. Jura, A. Ashworth, J. J. Irwin, M. C. Thompson, J. E. Gestwicki, F.
9 von Delft, B. K. Shoichet, J. S. Fraser, I. Ahel, Fragment binding to the Nsp3 macrodomain of SARS-CoV-2 identified
0 through crystallographic screening and computational docking. *Sci Adv.* **7** (2021), doi:10.1126/sciadv.abf8711.

15. 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).
16. 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).
17. A. Vrielink, N. Sampson, Sub-Angstrom resolution enzyme X-ray structures: is seeing believing? *Curr. Opin. Struct. Biol.* **13**, 709–715 (2003).
18. M. P. Blakeley, P. Langan, N. Niimura, A. Podjarny, Neutron crystallography: opportunities, challenges, and limitations. *Curr. Opin. Struct. Biol.* **18**, 593–600 (2008).
19. D. W. Kneller, G. Phillips, K. L. Weiss, S. Pant, Q. Zhang, H. M. O'Neill, L. Coates, A. Kovalevsky, Unusual zwitterionic catalytic site of SARS-CoV-2 main protease revealed by neutron crystallography. *J. Biol. Chem.* **295**, 17365–17373 (2020).
20. T. Chatake, A. Ostermann, K. Kurihara, F. G. Parak, N. Niimura, Hydration in proteins observed by high-resolution neutron crystallography. *Proteins.* **50**, 516–523 (2003).
21. J. Schiebel, R. Gaspari, T. Wulsdorf, K. Ngo, C. Sohn, T. E. Schrader, A. Cavalli, A. Ostermann, A. Heine, G. Klebe, Intriguing role of water in protein-ligand binding studied by neutron crystallography on trypsin complexes. *Nat. Commun.* **9**, 3559 (2018).
22. D. A. A. Myles, Neutron protein crystallography: current status and a brighter future. *Curr. Opin. Struct. Biol.* **16**, 630–637 (2006).
23. J. D. Ng, J. K. Baird, L. Coates, J. M. Garcia-Ruiz, T. A. Hodge, S. Huang, Large-volume protein crystal growth for neutron macromolecular crystallography. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **71**, 358–370 (2015).
24. M. P. Blakeley, Neutron crystallography aids in drug design. *IUCrJ.* **3**, 296–297 (2016).
25. M. Adachi, T. Ohhara, K. Kurihara, T. Tamada, E. Honjo, N. Okazaki, S. Arai, Y. Shoyama, K. Kimura, H. Matsumura, S. Sugiyama, H. Adachi, K. Takano, Y. Mori, K. Hidaka, T. Kimura, Y. Hayashi, Y. Kiso, R. Kuroki, Structure of HIV-1 protease in complex with potent inhibitor KNI-272 determined by high-resolution X-ray and neutron crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4641–4646 (2009).
26. R. Knihtila, G. Holzapfel, K. Weiss, F. Meilleur, C. Mattos, Neutron Crystal Structure of RAS GTPase Puts in Question the Protonation State of the GTP γ -Phosphate. *J. Biol. Chem.* **290**, 31025–31036 (2015).
27. P. J. Whittle, T. L. Blundell, Protein structure--based drug design. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 349–375 (1994).
28. A. C. Anderson, The process of structure-based drug design. *Chem. Biol.* **10**, 787–797 (2003).
29. M. Congreve, C. W. Murray, T. L. Blundell, Structural biology and drug discovery. *Drug Discov. Today.* **10**, 895–907 (2005).
30. R. S. Viridi, R. V. Bavisotto, N. C. Hopper, N. Vuksanovic, T. R. Melkonian, N. R. Silvaggi, D. N. Frick, Discovery of Drug-Like Ligands for the Mac1 Domain of SARS-CoV-2 Nsp3. *SLAS Discov.* **25**, 1162–1170 (2020).
31. M. Dasovich, J. Zhuo, J. A. Goodman, A. Thomas, R. L. McPherson, A. K. Jayabalan, V. F. Busa, S.-J. Cheng, B. A. Murphy, K. R. Redinger, Y. M. O. Alhammad, A. R. Fehr, T. Tsukamoto, B. S. Slusher, J. Bosch, H. Wei, A. K. L. Leung, High-Throughput Activity Assay for Screening Inhibitors of the SARS-CoV-2 Mac1 Macrodomain. *ACS Chem. Biol.* (2021), doi:10.1021/acscchembio.1c00721.
32. L. Coates, B. Sullivan, The macromolecular neutron diffractometer at the spallation neutron source. *Methods Enzymol.* **634**, 87–99 (2020).
33. L. Coates, M. J. Cuneo, M. J. Frost, J. He, K. L. Weiss, S. J. Tomanicek, H. McFeeters, V. G. Vandavasi, P. Langan, E. B. Iverson, The Macromolecular Neutron Diffractometer MaNDi at the Spallation Neutron Source. *J. Appl. Crystallogr.* **48**, 1302–1306 (2015).

34. L. Coates, A. D. Stoica, C. Hoffmann, J. Richards, R. Cooper, The macromolecular neutron diffractometer (MaNDi) at the Spallation Neutron Source, Oak Ridge: enhanced optics design, high-resolution neutron detectors and simulated diffraction. *J. Appl. Crystallogr.* **43**, 570–577 (2010).
35. P. V. Afonine, M. Mustyakimov, R. W. Grosse-Kunstleve, N. W. Moriarty, P. Langan, P. D. Adams, Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 1153–1163 (2010).
36. 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).
37. F. Meilleur, A. Kovalevsky, D. A. A. Myles, IMAGINE: The neutron protein crystallography beamline at the high flux isotope reactor. *Methods Enzymol.* **634**, 69–85 (2020).
38. D. N. Frick, R. S. Virdi, N. Vuksanovic, N. Dahal, N. R. Silvaggi, Molecular Basis for ADP-Ribose Binding to the Mac1 Domain of SARS-CoV-2 nsp3. *Biochemistry.* **59**, 2608–2615 (2020).
39. A. A. Kossiakoff, Protein dynamics investigated by the neutron diffraction-hydrogen exchange technique. *Nature.* **296**, 713–721 (1982).
40. J. C.-H. Chen, B. L. Hanson, S. Z. Fisher, P. Langan, A. Y. Kovalevsky, Direct observation of hydrogen atom dynamics and interactions by ultrahigh resolution neutron protein crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 15301–15306 (2012).
41. J. J. Skinner, W. K. Lim, S. Bédard, B. E. Black, S. W. Englander, Protein dynamics viewed by hydrogen exchange. *Protein Sci.* **21**, 996–1005 (2012).
42. J. M. Word, S. C. Lovell, J. S. Richardson, D. C. Richardson, Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **285**, 1735–1747 (1999).
43. G. Duan, V. H. Smith, D. F. Weaver, Characterization of aromatic-thiol π -type hydrogen bonding and phenylalanine-cysteine side chain interactions through ab initio calculations and protein database analyses. *Mol. Phys.* **99**, 1689–1699 (2001).
44. R. Wolfenden, Enzyme catalysis: conflicting requirements of substrate access and transition state affinity. *Mol. Cell. Biochem.* **3**, 207–211 (1974).
45. D. Herschlag, A. Natarajan, Fundamental challenges in mechanistic enzymology: progress toward understanding the rate enhancements of enzymes. *Biochemistry.* **52**, 2050–2067 (2013).
46. A. R. Fehr, G. Jankevicius, I. Ahel, S. Perlman, Viral Macrod domains: Unique Mediators of Viral Replication and Pathogenesis. *Trends Microbiol.* **26**, 598–610 (2018).
47. J. F. Darby, A. P. Hopkins, S. Shimizu, S. M. Roberts, J. A. Brannigan, J. P. Turkenburg, G. H. Thomas, R. E. Hubbard, M. Fischer, Water Networks Can Determine the Affinity of Ligand Binding to Proteins. *J. Am. Chem. Soc.* **141**, 15818–15826 (2019).
48. M. L. Samways, R. D. Taylor, H. E. Bruce Macdonald, J. W. Essex, Water molecules at protein-drug interfaces: computational prediction and analysis methods. *Chem. Soc. Rev.* **50**, 9104–9120 (2021).
49. A. Eriksson, O. Caldararu, U. Ryde, E. Oksanen, Automated orientation of water molecules in neutron crystallographic structures of proteins. *Acta Crystallogr D Struct Biol.* **76**, 1025–1032 (2020).
50. D. H. Juers, B. W. Matthews, Reversible lattice repacking illustrates the temperature dependence of macromolecular interactions. *J. Mol. Biol.* **311**, 851–862 (2001).
51. 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).
52. D. A. Keedy, H. van den Bedem, D. A. Sivak, G. A. Petsko, D. Ringe, M. A. Wilson, J. S. Fraser, Crystal cryocooling distorts conformational heterogeneity in a model Michaelis complex of DHFR. *Structure.* **22**, 899–910 (2014).

53. O. Carugo, D. Bordo, How many water molecules can be detected by protein crystallography? *Acta Crystallogr. D Biol. Crystallogr.* **55**, 479–483 (1999).
54. J. Newman, Novel buffer systems for macromolecular crystallization. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 610–612 (2004).
55. L. Abas, M. Guppy, Acetate: a contaminant in Hepes buffer. *Anal. Biochem.* **229**, 139–140 (1995).
56. R. Briones, C. Blau, C. Kutzner, B. L. de Groot, C. Aponte-Santamaría, GROmaps: A GROMACS-Based Toolset to Analyze Density Maps Derived from Molecular Dynamics Simulations. *Biophys. J.* **116**, 4–11 (2019).
57. C. S. Poornima, P. M. Dean, Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions. *J. Comput. Aided Mol. Des.* **9**, 500–512 (1995).
58. C. Barillari, J. Taylor, R. Viner, J. W. Essex, Classification of water molecules in protein binding sites. *J. Am. Chem. Soc.* **129**, 2577–2587 (2007).
59. M. L. Raymer, P. C. Sanschagrin, W. F. Punch, S. Venkataraman, E. D. Goodman, L. A. Kuhn, Predicting conserved water-mediated and polar ligand interactions in proteins using a K-nearest-neighbors genetic algorithm. *J. Mol. Biol.* **265**, 445–464 (1997).
60. B. C. Roberts, R. L. Mancera, Ligand-protein docking with water molecules. *J. Chem. Inf. Model.* **48**, 397–408 (2008).
61. J. Michel, J. Tirado-Rives, W. L. Jorgensen, Energetics of displacing water molecules from protein binding sites: consequences for ligand optimization. *J. Am. Chem. Soc.* **131**, 15403–15411 (2009).
62. B. Breiten, M. R. Lockett, W. Sherman, S. Fujita, M. Al-Sayah, H. Lange, C. M. Bowers, A. Heroux, G. Krilov, G. M. Whitesides, Water networks contribute to enthalpy/entropy compensation in protein-ligand binding. *J. Am. Chem. Soc.* **135**, 15579–15584 (2013).
63. J. L. Thomaston, N. F. Polizzi, A. Konstantinidi, J. Wang, A. Kolocouris, W. F. DeGrado, Inhibitors of the M2 Proton Channel Engage and Disrupt Transmembrane Networks of Hydrogen-Bonded Waters. *J. Am. Chem. Soc.* **140**, 15219–15226 (2018).
64. E. B. Lenselink, T. Beuming, W. Sherman, H. W. T. van Vlijmen, A. P. IJzerman, Selecting an optimal number of binding site waters to improve virtual screening enrichments against the adenosine A2A receptor. *J. Chem. Inf. Model.* **54**, 1737–1746 (2014).
65. M. E. Wall, G. Calabró, C. I. Bayly, D. L. Mobley, G. L. Warren, Biomolecular Solvation Structure Revealed by Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **141**, 4711–4720 (2019).
66. J. S. Fraser, M. W. Clarkson, S. C. Degnan, R. Erion, D. Kern, T. Alber, Hidden alternative structures of proline isomerase essential for catalysis. *Nature.* **462**, 669–673 (2009).
67. 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.
68. A. Ebrahim, B. T. Riley, D. Kumaran, B. Andi, M. R. Fuchs, S. McSweeney, D. A. Keedy, The temperature-dependent conformational ensemble of SARS-CoV-2 main protease (M pro). *bioRxiv* (2021), doi:10.1101/2021.05.03.437411.
69. M. Fischer, Macromolecular room temperature crystallography. *Q. Rev. Biophys.* **54**, e1 (2021).
70. A. Broom, R. V. Rakotoharisoa, M. C. Thompson, N. Zarifi, E. Nguyen, N. Mukhametzhanov, L. Liu, J. S. Fraser, R. A. Chica, Ensemble-based enzyme design can recapitulate the effects of laboratory directed evolution in silico. *Nat. Commun.* **11**, 4808 (2020).
71. C. M. Driggers, R. B. Cooley, B. Sankaran, L. L. Hirschberger, M. H. Stipanuk, P. A. Karplus, Cysteine dioxygenase structures from pH4 to 9: consistent cys-persulfenate formation at intermediate pH and a Cys-bound enzyme at

- 1 higher pH. *J. Mol. Biol.* **425**, 3121–3136 (2013).
- 2 72. B. C. Gibb, Hitting the buffers. *Nat. Chem.* **13**, 1023–1024 (2021).
- 3 73. J. G. M. Rack, D. Perina, I. Ahel, Macrod domains: Structure, Function, Evolution, and Catalytic Activities. *Annu. Rev.*
4 *Biochem.* **85**, 431–454 (2016).
- 5 74. R. G. Hammond, N. Schormann, R. L. McPherson, A. K. L. Leung, C. C. S. Deivanayagam, M. A. Johnson, ADP-
6 ribose and analogues bound to the deMARylating macrodomain from the bat coronavirus HKU4. *Proc. Natl. Acad.*
7 *Sci. U. S. A.* **118** (2021), doi:10.1073/pnas.2004500118.
- 8 75. F. Meilleur, L. Coates, M. J. Cuneo, A. Kovalevsky, D. A. A. Myles, The Neutron Macromolecular Crystallography
9 Instruments at Oak Ridge National Laboratory: Advances, Challenges, and Opportunities. *Crystals*. **8**, 388 (2018).
- 0 76. O. Arnold, J. C. Bilheux, J. M. Borreguero, A. Buts, S. I. Campbell, L. Chapon, M. Doucet, N. Draper, R. Ferraz Leal,
1 M. A. Gigg, V. E. Lynch, A. Markvardsen, D. J. Mikkelsen, R. L. Mikkelsen, R. Miller, K. Palmen, P. Parker, G.
2 Passos, T. G. Perring, P. F. Peterson, S. Ren, M. A. Reuter, A. T. Savici, J. W. Taylor, R. J. Taylor, R. Tolchenov, W.
3 Zhou, J. Zikovsky, Mantid—Data analysis and visualization package for neutron scattering and μ SR experiments.
4 *Nucl. Instrum. Methods Phys. Res. A*. **764**, 156–166 (2014).
- 5 77. B. Sullivan, R. Archibald, P. S. Langan, H. Dobbek, M. Bommer, R. L. McFeeters, L. Coates, X. Wang, F. Gallmeier,
6 J. M. Carpenter, V. Lynch, P. Langan, Improving the accuracy and resolution of neutron crystallographic data by
7 three-dimensional profile fitting of Bragg peaks in reciprocal space. *Acta Crystallogr D Struct Biol.* **74**, 1085–1095
8 (2018).
- 9 78. J. R. Helliwell, J. Habash, D. W. J. Cruickshank, M. M. Harding, T. J. Greenhough, J. W. Campbell, I. J. Clifton, M.
0 Elder, P. A. Machin, M. Z. Papiz, S. Zurek, The recording and analysis of synchrotron X-radiation Laue diffraction
1 photographs. *J. Appl. Crystallogr.* **22**, 483–497 (1989).
- 2 79. J. W. Campbell, Q. Hao, M. M. Harding, N. D. Nguti, C. Wilkinson, LAUEGEN version 6.0 and INTLDM. *J. Appl.*
3 *Crystallogr.* **31**, 496–502 (1998).
- 4 80. G. P. Bourenkov, A. N. Popov, H. D. Bartunik, Bayesian approach to the analysis of time-resolved protein Laue
5 diffraction data. *Acta Crystallogr. A*. **52**, 50–50 (1996).
- 6 81. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. D Biol.*
7 *Crystallogr.* **69**, 1204–1214 (2013).
- 8 82. 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.
9 W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A.
0 Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**,
1 235–242 (2011).
- 2 83. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic
3 software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- 4 84. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A.
5 Urzhumtsev, P. H. Zwart, P. D. Adams, Towards automated crystallographic structure refinement with phenix.refine.
6 *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
- 7 85. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol.*
8 *Crystallogr.* **66**, 486–501 (2010).
- 9 86. N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, electronic Ligand Builder and Optimization Workbench
0 (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1074–1080
1 (2009).
- 2 87. C. S. Bury, J. C. Brooks-Bartlett, S. P. Walsh, E. F. Garman, Estimate your dose: RADDose-3D. *Protein Sci.* **27**,
3 217–228 (2018).

- 1 88. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
- 2 89. G. C. Schröder, W. B. O'Dell, D. A. A. Myles, A. Kovalevsky, F. Meilleur, IMAGINE: neutrons reveal enzyme
3 chemistry. *Acta Crystallogr D Struct Biol.* **74**, 778–786 (2018).
- 4 90. F. Meilleur, P. Munshi, L. Robertson, A. D. Stoica, L. Crow, A. Kovalevsky, T. Koritsanszky, B. C. Chakoumakos, R.
5 Blessing, D. A. A. Myles, The IMAGINE instrument: first neutron protein structure and new capabilities for neutron
6 macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 2157–2160 (2013).
- 7 91. J. W. Campbell, LAUEGEN, an X-windows-based program for the processing of Laue diffraction data. *J. Appl.*
8 *Crystallogr.* **28**, 228–236 (1995).
- 9 92. S. Arzt, J. W. Campbell, M. M. Harding, Q. Hao, J. R. Helliwell, LSCALE – the new normalization, scaling and
0 absorption correction program in the Daresbury Laue software suite. *J. Appl. Crystallogr.* **32**, 554–562 (1999).
- 1 93. M. S. Weiss, Global indicators of X-ray data quality. *J. Appl. Crystallogr.* **34**, 130–135 (2001).
- 2 94. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance
3 molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX.* **1-2**, 19–25 (2015).
- 4 95. A. Morin, B. Eisenbraun, J. Key, P. C. Sanschagrin, M. A. Timony, M. Ottaviano, P. Sliz, Collaboration gets the most
5 out of software. *Elife.* **2**, e01456 (2013).
- 6 96. G. Winter, R. J. Gildea, N. G. Paterson, J. Beale, M. Gerstel, D. Axford, M. Vollmar, K. E. McAuley, R. L. Owen, R.
7 Flaig, A. W. Ashton, D. R. Hall, How best to use photons. *Acta Crystallographica Section D: Structural Biology.* **75**,
8 242–261 (2019).
- 9 97. P. A. Karplus, K. Diederichs, Linking crystallographic model and data quality. *Science.* **336**, 1030–1033 (2012).
- 0 98. R. P. Joosten, T. A. H. te Beek, E. Krieger, M. L. Hekkelman, R. W. W. Hooft, R. Schneider, C. Sander, G. Vriend, A
1 series of PDB related databases for everyday needs. *Nucleic Acids Res.* **39**, D411–9 (2011).