

## CORONAVIRUS

# 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. Correy<sup>1</sup>, Daniel W. Kneller<sup>2,3†</sup>, Gwyndalyn Phillips<sup>2,3</sup>, Swati Pant<sup>2,3</sup>, Silvia Russi<sup>4</sup>, Aina E. Cohen<sup>4</sup>, George Meigs<sup>5,6</sup>, James M. Holton<sup>4,5,6</sup>, Stefan Gahbauer<sup>7</sup>, Michael C. Thompson<sup>8</sup>, Alan Ashworth<sup>9</sup>, Leighton Coates<sup>3,10</sup>, Andrey Kovalevsky<sup>2,3\*</sup>, Flora Meilleur<sup>2,11\*</sup>, James S. Fraser<sup>1\*</sup>

The nonstructural protein 3 (NSP3) macrodomain of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Mac1) removes adenosine diphosphate (ADP) ribosylation posttranslational modifications, playing a key role in the immune evasion capabilities of the virus responsible for the coronavirus disease 2019 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 the 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 reevaluation of macrodomain catalytic mechanisms and will guide the optimization of Mac1 inhibitors.

## INTRODUCTION

Viral infection triggers the release of interferons, a group of secreted cytokines that are key components of the innate immune response (1, 2). Among the several hundred interferon-stimulated genes are poly(adenosine diphosphate-ribose) polymerases (PARPs), a family of enzymes that catalyze the transfer of ADP-ribose from nicotinamide adenine dinucleotide (NAD<sup>+</sup>) to proteins as a posttranslational modification. Both mono- and poly-ADP ribosylation are important for signal transduction of the interferon response (3, 4). In severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, ADP ribosylation is mediated by PARP9 and its binding partner DTX3L (5). To counteract this host defense mechanism, SARS-CoV-2 encodes a macrodomain, Mac1, with mono(ADP-ribosyl) hydrolase activity that reverses the protective effect of ADP ribosylation (Fig. 1A) (6). An active site mutation that inactivates Mac1 has been shown to render SARS-CoV nonlethal in a mouse model of viral infection (7), and interferon-stimulated cells can clear SARS-CoV virus bearing that active site mutation but not wild-type Mac1. Accordingly,

Mac1 is a potential therapeutic target for small-molecule inhibitors. The Mac1 active site is highly conserved among the SARS-CoV, SARS-CoV-2, and MERS (Middle Eastern Respiratory Syndrome) viruses, which raises the possibility that a Mac1 inhibitor could be used as a broad-spectrum antiviral against coronaviruses. Among human proteins, PARP9, PARP14, and hMacroD1/2 are most similar, with sequence identities between 22 and 24% to SARS-CoV-2 Mac1 (8). Although the key adenosine diphosphate ribose (ADPr) recognition motifs are conserved between viral and human enzymes, low sequence similarity in regions near to the ADPr recognition motifs suggests that cross-reactivity of inhibitors between viral and human proteins is unlikely (8).

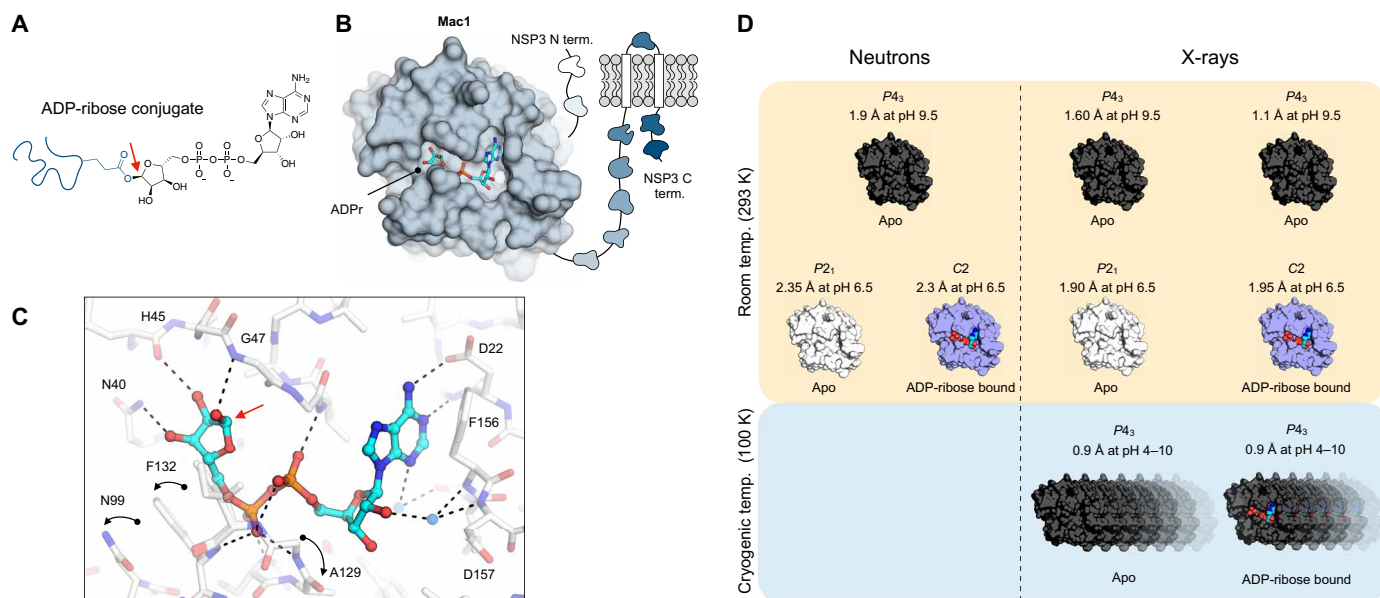
The SARS-CoV-2 macrodomain is encoded as a domain of non-structural protein 3 (NSP3), a large multidomain protein found in all coronaviruses (Fig. 1B) (9). Mac1 adopts an  $\alpha/\beta/\alpha$ -sandwich fold and binds ADPr through an extensive network of hydrogen bonds in a well-defined cleft (Fig. 1, B and C) (6). Several conformational changes are stabilized upon ADPr binding, including rotation of Phe<sup>132</sup> to accommodate the terminal ribose of ADPr and two peptide flips to bind the diphosphate portion of ADPr (10). Although the structural determinants of ADPr binding are well established, the catalytic mechanism of Mac1 is unknown. On the basis of structural homology to the human macrodomain hMacroD2 (11), a substrate-assisted mechanism has been proposed for viral macrodomains (8); however, this has not been confirmed experimentally. Determining the precise orientations of active site water molecules and the protonation states of key catalytic site residues can discriminate whether the proposed water nucleophile is appropriately orientated for the hydrolysis reaction in the ADPr-bound state or whether other catalytic mechanisms should be considered.

Although x-ray crystallography can provide high-resolution structural information, hydrogen atoms are not normally observed in x-ray diffraction experiments (12). This means that key information may be obscured, including residue protonation states,

<sup>1</sup>Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA 94158, USA. <sup>2</sup>Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. <sup>3</sup>National Virtual Biotechnology Laboratory, U.S. Department of Energy, Washington, DC 20585, USA. <sup>4</sup>Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA. <sup>5</sup>Department of Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. <sup>6</sup>Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA. <sup>7</sup>Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94158, USA. <sup>8</sup>Department of Chemistry and Biochemistry, University of California, Merced, Merced, CA 95343, USA. <sup>9</sup>Helen Diller Family Comprehensive Cancer, University of California, San Francisco, San Francisco, CA 94158, USA. <sup>10</sup>Second Target Station, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. <sup>11</sup>Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695, USA.

\*Corresponding author. Email: kovalevskyay@ornl.gov (A.K.); meilleurf@ornl.gov (F.M.); jfraser@fraserlab.com (J.S.F.)

†Present address: New England Biolabs Inc., Ipswich, MA 01938, USA.



**Fig. 1. The NSP3 macrodomain (Mac1) reverses mono-ADP ribosylation.** (A) Chemical structure of ADPr conjugated to a glutamate residue. The C1' covalent attachment point is shown with a red arrow. (B) Cartoon of the multidomain 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. The C1' covalent attachment point is shown with a red arrow. (D) Summary of the crystal structures reported in this work.

hydrogen bond networks, and water orientations. Neutron crystallography allows hydrogens to be visualized at modest resolutions ( $<2.5$  Å) (13, 14), which can reveal details missing from x-ray structures. In practice, hydrogen is typically exchanged with deuterium in protein samples for neutron crystallography, because protons have a small negative neutron scattering length (NSL), as well as a large incoherent cross section, whereas deuterium has a comparatively large positive NSL and small incoherent cross section (15). The incoherent cross section of hydrogen contributes noise to the diffraction experiment; replacing hydrogen with deuterium increases the signal-to-noise ratio of the data and improves the quality of the diffraction experiment. Furthermore, the low brightness of neutron sources means that the volume of crystals required is 100- to 1000-fold larger than crystals that are suitable for experiments at synchrotron light sources (16). Because protein crystals rarely grow to this size, relatively few drug targets have been structurally characterized by neutron crystallography (17, 18).

In addition to informing catalytic mechanisms, neutron crystallography can augment structure-based drug design, which depends on an accurate model of the shape, electrostatic potential, solvation, and flexibility of the target site (19, 20). We recently screened more than 2500 fragments against Mac1 using x-ray crystallography (10). Structures of 234 fragments bound to Mac1 were determined, including almost 200 in the active site. This screen has given the community an extensive set of chemically diverse starting points for fragment-based ligand discovery against Mac1. As expected, on the basis of their low molecular weight, the fragments bound Mac1 weakly (10). 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

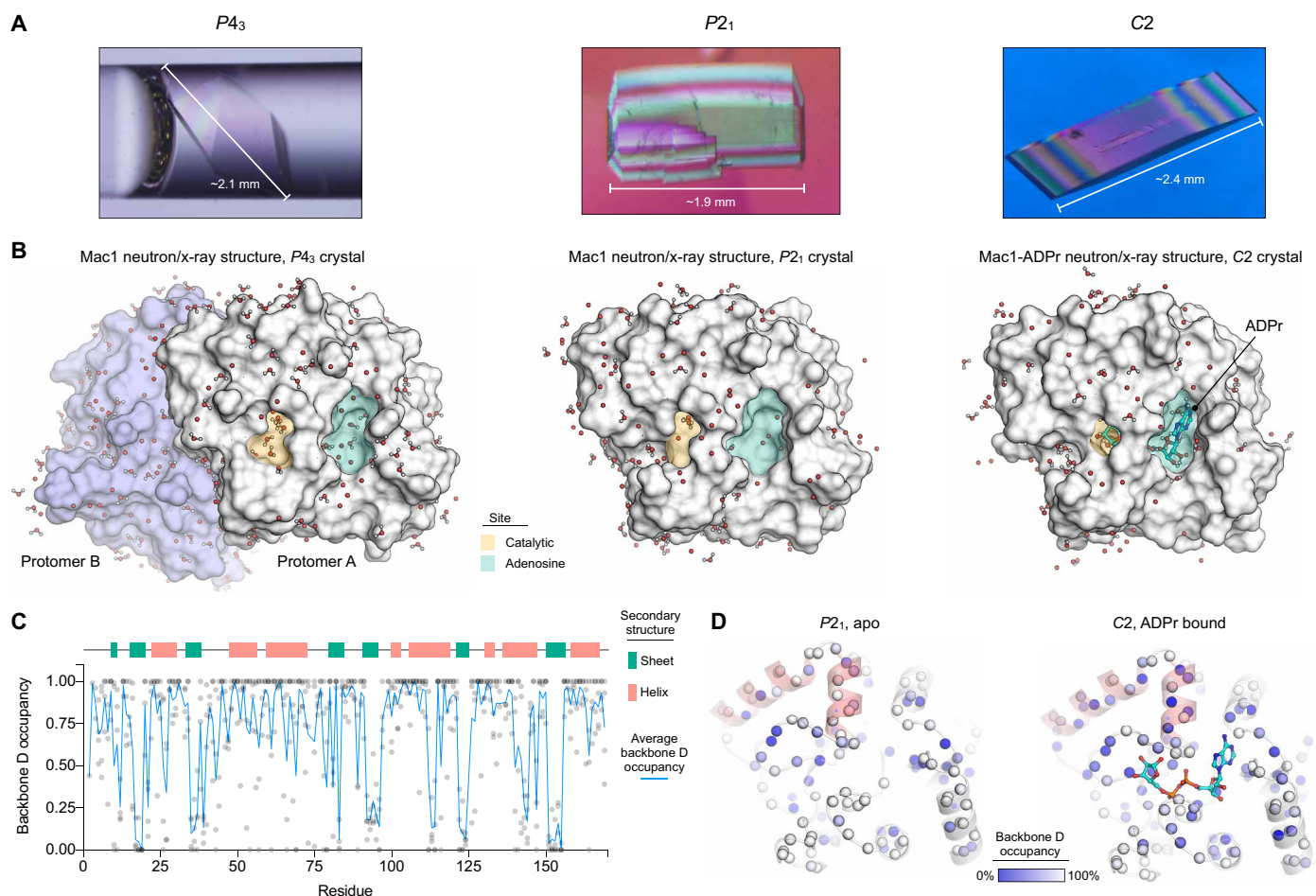
grown at pH 6.5 and 9.5, and a neutron structure of the ADPr-Mac1 complex was determined at pH 6.5. Protonation states and active site hydrogen bond networks were characterized in the three structures, revealing the molecular basis for functionally relevant flexibility in active site loops and proton 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, temperature, and pH. The water networks reorganize upon ADPr binding, with a network of tightly bound water molecules acting as protein-ligand bridges. Last, 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.

## RESULTS

### Room temperature Mac1 structures from distinct crystal forms at different pHs

#### Mac1 neutron structure at pH 9.5 (1.9 Å, $P4_3$ crystal)

We grew neutron diffraction-quality crystals at pH 9.5 using a Mac1 construct that crystallizes in the  $P4_3$  space group (Fig. 2A) (10). Hydrogen/deuterium exchange (HDX) was achieved by growing crystals using solutions prepared with  $D_2O$ , followed by soaking the crystal in a mother liquor prepared with  $D_2O$ . Neutron diffraction data to 1.9 Å were collected at room temperature using the Macromolecular Neutron Diffractometer (MaNDi) at the Oak Ridge National Laboratory (ORNL; table S1) (21–23). To enable joint neutron/x-ray refinement in Phenix (24), 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



**Fig. 2. Crystal structures of Mac1 determined using neutron diffraction.** (A) Neutron-quality Mac1 crystals grown in the *P4<sub>3</sub>*, *P2<sub>1</sub>*, and *C2* space groups. The *P4<sub>3</sub>* crystal is shown in the quartz capillary used for data collection, while the *P2<sub>1</sub>* and *C2* crystals are shown before mounting. (B) Molecular surface showing Mac1 neutron structures. D<sub>2</sub>O 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 *P2<sub>1</sub>* and *C2* Mac1 structures. Backbone nitrogens are shown with blue spheres, colored by backbone D occupancy. The helices composed of residues 50 to 70 are shaded red. The average backbone D occupancy for these helices was 82% in the *P2<sub>1</sub>* structure and 52% in the *C2* structure.

sites, both hydrogen and deuterium were modeled at exchangeable positions, and deuterium occupancies were automatically refined (24). 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 D<sub>2</sub>O 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 Phe<sup>156</sup>/Asp<sup>157</sup> (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 (PDB) with the accession code 7TX3.

#### Mac1 x-ray structure at pH 9.5 (1.1 Å, *P4<sub>3</sub>* crystal)

To obtain more structural information about Mac1 at room temperature, we collected higher-resolution diffraction data using the *P4<sub>3</sub>* crystals and synchrotron radiation (table S1). The sensitivity of the *P4<sub>3</sub>* crystals to radiation damage was assessed by collecting four datasets from similarly sized crystals (~0.1 mm<sup>3</sup>), with the absorbed

x-ray dose varied from 73 to 539 kilograys (kGy). There was no evidence for general or specific radiation damage, as measured by per-image resolution,  $R_{\text{CP}}$ , 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* $\alpha$  root mean square fluctuation (RMSF) < 0.1 Å, Pearson correlation coefficient (*r*) for *C* $\alpha$  *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, and 7TWI. On the basis of having similarly high resolution but slightly lower *C* $\alpha$  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* $\alpha$  root mean square deviation (RMSD) < 0.05 Å, Pearson



$r > 0.93$  for Ca  $B$ -factors] (fig. S3). Although the crystals used for x-ray diffraction were grown under the same conditions as the neutron crystal [100 mM CHES (pH 9.5) and 32% polyethylene glycol (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 solvent molecules between the two structures shows that their positions are conserved (fig. S3). Of the 317 D<sub>2</sub>O molecules modeled in the 1.9/1.6-Å neutron/x-ray structure, 257 had a matching H<sub>2</sub>O within 0.5 Å in the 1.1-Å x-ray structure. The solvent molecules 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 crystal)**

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 P2<sub>1</sub> space group (Fig. 2, A and B) (25). The difference between the two constructs was slight; the P2<sub>1</sub> construct contained a Gly-Glu sequence immediately following the TEV cleavage site, while the P4<sub>3</sub> construct had Ser-Met at the equivalent position. HDX was performed in a similar manner to the P4<sub>3</sub> crystals, and diffraction data were collected to 2.35 Å at MaNDi (table S1). As with the P4<sub>3</sub> crystals, an x-ray dataset was collected with the same crystal (table S1), and joint neutron/x-ray 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<sub>3</sub> structure (Ca 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 (25), 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 cocrystallized ADPr with the P2<sub>1</sub> construct (Fig. 2, A and B). Unlike the P4<sub>3</sub> construct, where crystal packing prevents cocrystallization with ADPr (10), cocrystals with the P2<sub>1</sub> construct grew readily. Neutron diffraction data were collected to 2.3 Å using the IMAGINE beamline at the High Flux Isotope Reactor (HFIR) at ORNL (26), and an x-ray dataset was collected to 1.95 Å using the same crystal. The space group of the cocrystal 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  $mF_o - DF_c$  electron density for ADPr bound in the Mac1 active site (fig. S1), and the ADPr occupancy was refined to 93%. The terminal ribose of ADPr adopts the  $\beta$  epimer, with the C1'' hydroxyl hydrogen-bonded to the backbone nitrogen of Gly<sup>48</sup>. Alignment of the Mac1-ADPr neutron/x-ray structure with the three apo structures shows that the structures are highly similar (Ca RMSD < 0.3 Å) (fig. S3). Structural differences include the peptide flips of Gly<sup>48</sup> and Ala<sup>129</sup> that allow two new hydrogen bonds with the diphosphate portion of ADPr and a coupled conformational change in the Phe<sup>132</sup> and Asn<sup>99</sup> side chains to accommodate the terminal ribose of ADPr (fig. S1). These changes have been documented across several crystal forms (6, 25, 27, 28). The main difference with the ADPr-bound structure

obtained with the P4<sub>3</sub> crystal form is that crystal packing interactions prevent the flip of Gly<sup>48</sup>, forcing ADPr to adopt the  $\alpha$  epimer (10). 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.

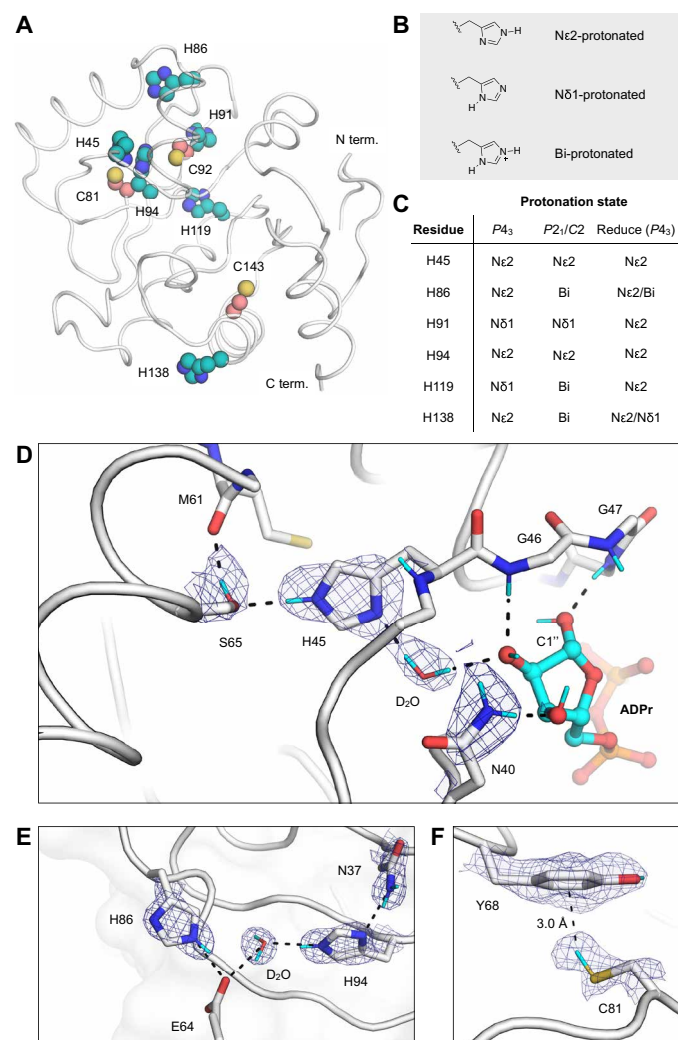
#### **Features of deuteration between the different crystal structures**

Next, we examined whether the refined deuterium occupancy in the neutron structures could reveal information about Mac1 structure and dynamics (29, 30). Although deuterium occupancy was refined at all exchangeable positions, we restricted our analysis to backbone amide deuteriums. For the P4<sub>3</sub> 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<sub>1</sub> and C2 crystals was 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  $\beta$  strands (Fig. 2C and fig. S4), which is consistent with the protected nature of backbone amides in  $\beta$  sheets (29). The protective effect was pronounced in the solvent-exposed  $\beta$  strand that includes residues 80 to 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<sub>3</sub> 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  $\alpha$  helices composed of residues 50 to 70 (Fig. 2D). This may reflect increased protection from HDX caused by stabilization of the Gly<sup>46-48</sup> 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  $\alpha$  helix composed of residues 138 to 145 has relatively low HDX in the P2<sub>1</sub> and C2 crystals, which may be connected to 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).

#### **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  $F_o - F_c$  NSL difference density maps (Fig. 3, B and C, and fig. S5). The histidine protonation states were the same across both crystal forms, except for His<sup>86</sup>, His<sup>119</sup>, and His<sup>138</sup>, which were singly protonated in the P4<sub>3</sub> structure but doubly protonated in the P2<sub>1</sub> structure (Fig. 3C). The importance of experimentally determining protonation states is highlighted by the mismatch between the protonation states predicted by the Reduce program (31) and the states assigned using neutron crystallography (Fig. 3C). Although the protonation states of the two buried histidines, His<sup>45</sup> and His<sup>94</sup>, were assigned correctly by Reduce, there was a discrepancy in the protonation states for the remaining



**Fig. 3. Protonation states of Mac1 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 on the basis of NSL density maps (maps are shown in fig. S5). The protonation states assigned to the high-resolution  $P_{43}$  x-ray structure (PDB code 7KQO) by the program Reduce are also shown. (D) NSL density maps reveal the hydrogen bond network connecting His<sup>45</sup> 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 a blue mesh (contoured at  $2.5 \sigma$ ). (E) An extensive hydrogen bond network connects the Asn<sup>37</sup> side chain with a surface histidine (His<sup>86</sup>). The  $P_{43}$  structure (protomer A, PDB code 7TX3) and the corresponding  $2mF_o - DF_c$  NSL density map are shown with a blue mesh (contoured at  $2.5 \sigma$ ). (F) An aromatic-thiol bond was observed between Tyr<sup>68</sup> and Cys<sup>81</sup> in protomer A of the  $P_{43}$  structure (PDB code 7TX3). The  $2mF_o - DF_c$  NSL density map is shown with a blue mesh (contoured at  $1 \sigma$ ).

histidines in at least one of the  $P_{43}$  protomers (Fig. 3C). To test whether histidine protonation states could be assigned from the previously published high-resolution x-ray structures of Mac1 (10), we calculated  $F_o - F_c$  electron density difference maps from the 0.86-Å x-ray structure determined using the  $P_{43}$  crystals (PDB code 7KQO) and the 0.77-Å structure determined using C2 crystals (PDB code 7KR0). Although the protonation state of His<sup>45</sup> and

His<sup>94</sup> could be assigned on the basis of 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 are available at ultrahigh resolution.

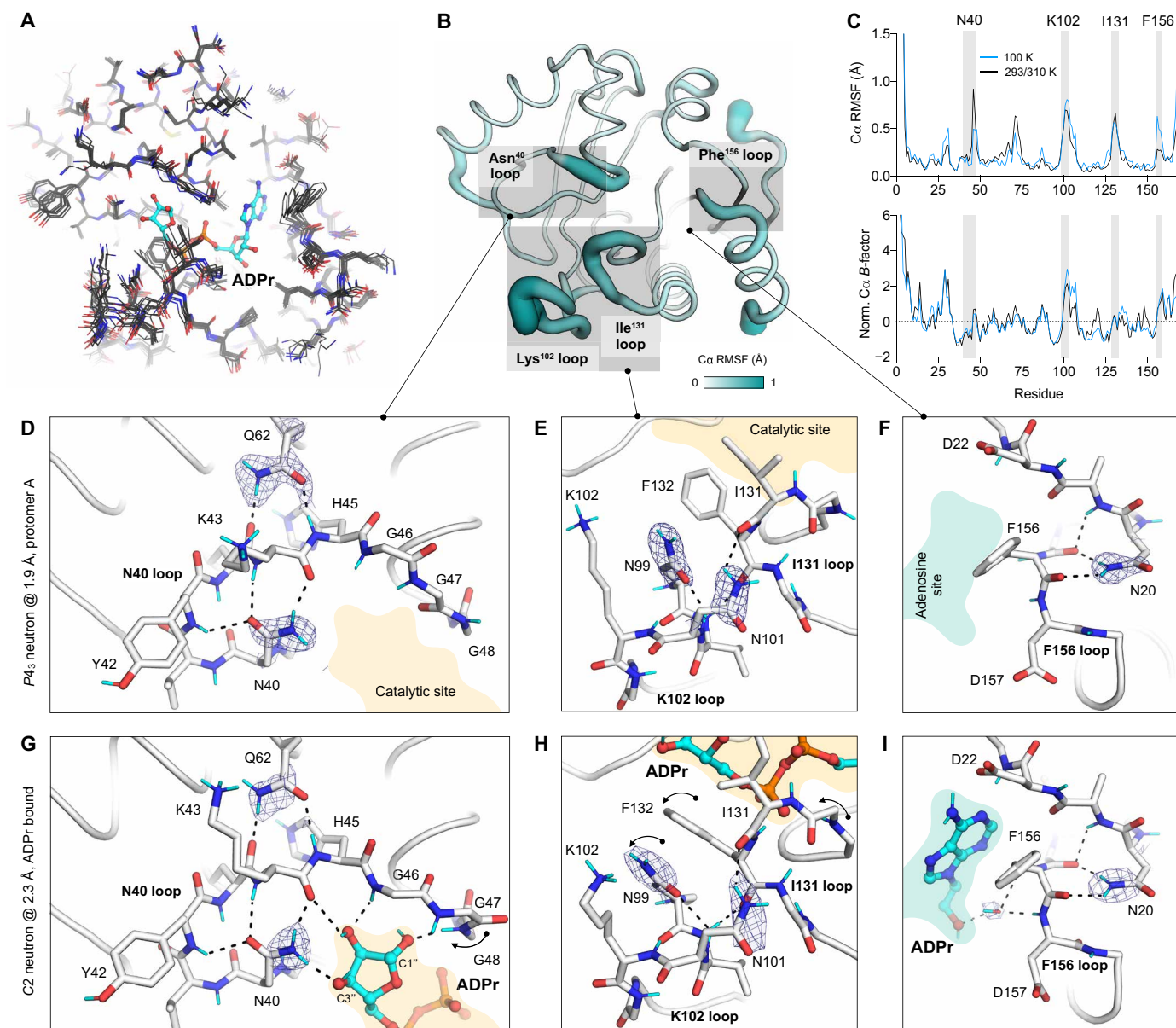
Next, we examined the hydrogen bonding networks around the buried histidines. His<sup>45</sup> is singly protonated on the Ne2 nitrogen and hydrogen bonds to the Ser<sup>65</sup> side chain and the backbone carbonyl of Gln<sup>62</sup> (Fig. 3D). The neutron structure of Mac1 with ADPr bound shows that this network extends to the active site via a D<sub>2</sub>O molecule, which forms a hydrogen bond with the C2'' hydroxyl of the terminal ribose (Fig. 3D). It has previously been shown that mutation of His<sup>45</sup> to an alanine substantially decreases Mac1 activity (8). This water-mediated network provides a potential link between His<sup>45</sup> and Mac1 activity; His<sup>45</sup> could participate either through stabilization of the terminal ribose in a catalytically productive conformation, as proposed previously (8), or through the action of His<sup>45</sup> as a general base in the hydrolysis reaction. The other buried histidine, His<sup>94</sup>, participates in an extensive hydrogen bond network connecting the core of the Mac1 to its surface (Fig. 3E). Asn<sup>37</sup> is hydrogen-bonded to the singly protonated His<sup>94</sup>, and the network extends to His<sup>86</sup> via a D<sub>2</sub>O molecule and the Glu<sup>64</sup> backbone (Fig. 3E). The residues comprising these two networks are conserved in the SARS-CoV and MERS macrodomains (8).

### Cysteines

The acid association constant ( $pK_a$ ) 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  $P_{43}$  and  $P_{21}$  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 Cys<sup>81</sup> thiol deuterium atom 3.1 Å from the Tyr<sup>68</sup> aromatic ring (Fig. 3F), which is consistent with a rare thiol-aromatic bond (32). However, this peak was weaker or absent in protomer B and the  $P_{21}/C_2$  structures (fig. S5). We had similar difficulty locating the Cys<sup>143</sup> thiol deuterium atom; a positive peak in the  $F_o - F_c$  NSL density map was present in protomer B of the  $P_{43}$  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.

### 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. 4, A and B). The resulting ensemble shows that structure variation is mainly restricted to two regions around the active site: the loop that positions Phe<sup>156</sup> in the adenosine-binding site (the Phe<sup>156</sup> loop) and the two loops that position Phe<sup>132</sup> in the catalytic site (the Ile<sup>131</sup> loop and the Lys<sup>102</sup> loop). Next, we compared structural variation across structures determined at cryogenic temperature and ambient temperature (Fig. 4C). The  $\alpha$  RMSF was highly correlated (Pearson  $r = 0.93$ ), indicating that structural variation was independent of data collection temperature. The



**Fig. 4. Protein flexibility and hydrogen bond networks in the Mac1 active site.** (A) Alignment of  $P_4$ ,  $P_{21}$ , and C2 Mac1 neutron and x-ray crystal structures determined at 100/293/310 K (PDB codes 7KQO, 7KQP, 7KR0, 7KR1, 7TWH, 7TX3, 7TX4, and 7TX5). (B)  $C_\alpha$  RMSF calculated from structures shown in (A) mapped onto the Mac1 structure. (C) Top: Plot showing  $C_\alpha$  RMSF of Mac1 structures determined at 100 K (blue line) and those at 293/310 K (black line). Bottom: Plot showing  $C_\alpha$  B-factors from the Mac1 structure. B-factors were normalized by Z score as described in Materials and Methods. (D to F) Active site hydrogen bond networks assigned on the basis of NSL density maps for the  $P_4$  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  $2\sigma$  around the asparagine/glutamine side chains). Hydrogen bonds ( $<3.5\text{ \AA}$ ) are shown with dashed black lines. (G to I) Same as (D) to (F) but showing the hydrogen bond networks in the C2 ADPr cocrystal structure (PDB code 7TX5). The NSL density maps are shown with a blue mesh [contoured at 3, 2, and  $2.5\sigma$  in (G) to (I)].

$C_\alpha$  B-factors were also highly correlated between the two temperatures (Pearson  $r = 0.92$ ) (Fig. 4C). The region with the highest  $C_\alpha$  B-factors at both temperatures was the Lys<sup>102</sup> loop, which is consistent with the structural variation seen across different crystal forms (Fig. 4B).

#### Asn<sup>40</sup> network

The need for flexibility to allow substrate binding is often balanced by the need to stabilize a substrate for catalysis (33, 34). In Mac1,

Asn<sup>40</sup> plays an important role in substrate stabilization, based on its absolute conservation across evolutionarily diverse macrodomains, and loss of macrodomain function when Asn<sup>40</sup> is mutated to aspartic acid (8). In the ADPr-bound Mac1 neutron structure, Asn<sup>40</sup> forms a hydrogen bond with the C3' hydroxyl of ADPr, with the side-chain orientation clearly defined in the NSL density map (Fig. 4G). Asn<sup>40</sup> is preorganized to make this interaction, based on the similarity in the side-chain conformation in the apo and ADPr-bound



structures (Fig. 4, D and G, and fig. S1). This preorganization is achieved by hydrogen bonds to the backbone of Tyr<sup>42</sup> and Lys<sup>44</sup>, residues that are, in turn, stabilized by hydrogen bonds to the side chain of Gln<sup>62</sup> (Fig. 4, D and G). This network is conserved at 100 and 293 K, across different Mac1 crystal forms and between the SARS-CoV and MERS macrodomains (8). This extensive network stabilizes Asn<sup>40</sup> while allowing the adjacent glycine-rich loop (Gly<sup>46-48</sup>) to be flexible during ADPr binding.

### **Ile<sup>131</sup>/Lys<sup>102</sup> network**

Next, we examined hydrogen bond networks in the Ile<sup>131</sup>/Lys<sup>102</sup> loops, which undergo a series of conformational changes upon ADPr binding (Fig. 1C). The orientation of Asn<sup>99</sup> is clearly defined in NSL density maps in the three neutron structures (Fig. 4, E and H). The side-chain nitrogen is oriented toward the catalytic site, while the side-chain oxygen is hydrogen-bonded to the backbone nitrogen of Asn<sup>101</sup>. In the *P*<sub>43</sub> structure, the nearby Asn<sup>101</sup> forms a hydrogen bond with the backbone of Ile<sup>131</sup>, connecting the two loops. This network is disrupted in the *P*<sub>21</sub> crystal form, where Asn<sup>101</sup> forms a hydrogen bond with a symmetry mate. Overall, this network is conserved between the 100 K and 293 K *P*<sub>43</sub> x-ray structures, with only a slight shift in the Lys<sup>102</sup> 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 Lys<sup>102</sup> and Ile<sup>131</sup> loop movement is correlated, with interloop hydrogen bonds maintained across the ensemble. In the SARS-CoV macrodomain, an arginine and aspartic acid are present at the position equivalent to Asn<sup>99</sup> and Asn<sup>101</sup>. Although this removes a hydrogen bond between the Lys<sup>102</sup> loop and the Ile<sup>131</sup> loop, it introduces a new salt bridge to constrain the Asp<sup>98</sup> side chain (equivalent to Asn<sup>99</sup> in Mac1) (fig. S6). Together, analysis of the Ile<sup>131</sup>/Lys<sup>102</sup> loops shows how a key flexible region of Mac1 is supported by interloop hydrogen bond networks. This flexibility may be important for accommodating a wide range of ADP-ribosylated substrates.

### **Phe<sup>156</sup> network**

The adenosine portion of ADPr is recognized by hydrogen bonds to Asp<sup>22</sup> and Ile<sup>23</sup> and a water-mediated contact with the Phe<sup>156</sup>/Asp<sup>157</sup> backbone (Fig. 1C). The Phe<sup>156</sup> loop is stabilized by two hydrogen bonds that connect the Asn<sup>20</sup> side chain to the Val<sup>155</sup>/Phe<sup>156</sup> backbone (Fig. 4, F and I). The orientation of the Asn<sup>20</sup> side chain is clearly defined in the NSL density maps obtained from both the *P*<sub>43</sub> and C2 crystals (Fig. 4, F and I). Although multiple conformations of the Phe<sup>156</sup> side chain were seen in the cryogenic *P*<sub>43</sub> 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 Phe<sup>156</sup> side chain is structurally diverse, matching the diversity seen for fragment-bound structures (10). Both SARS-CoV and MERS have an asparagine at the position equivalent to Phe<sup>156</sup> and a cysteine at the position equivalent to Asn<sup>20</sup> (fig. S6). This indicates a degree of tolerance of the residues at this position, given that all three enzymes have similar ADPr-binding properties (6). One consequence of this amino acid change is the potential for  $\pi$ - $\pi$  stacking interactions at this site. Although the adenine portion of ADPr is too far to form  $\pi$ - $\pi$  stacking interactions, we previously observed several fragments making close contacts with the Phe<sup>156</sup> side chain (10). An inhibitor that relies too heavily on stacking interactions with Phe<sup>156</sup> may therefore have limited use as a broad-spectrum antiviral against coronaviruses.

## **Water networks in Mac1**

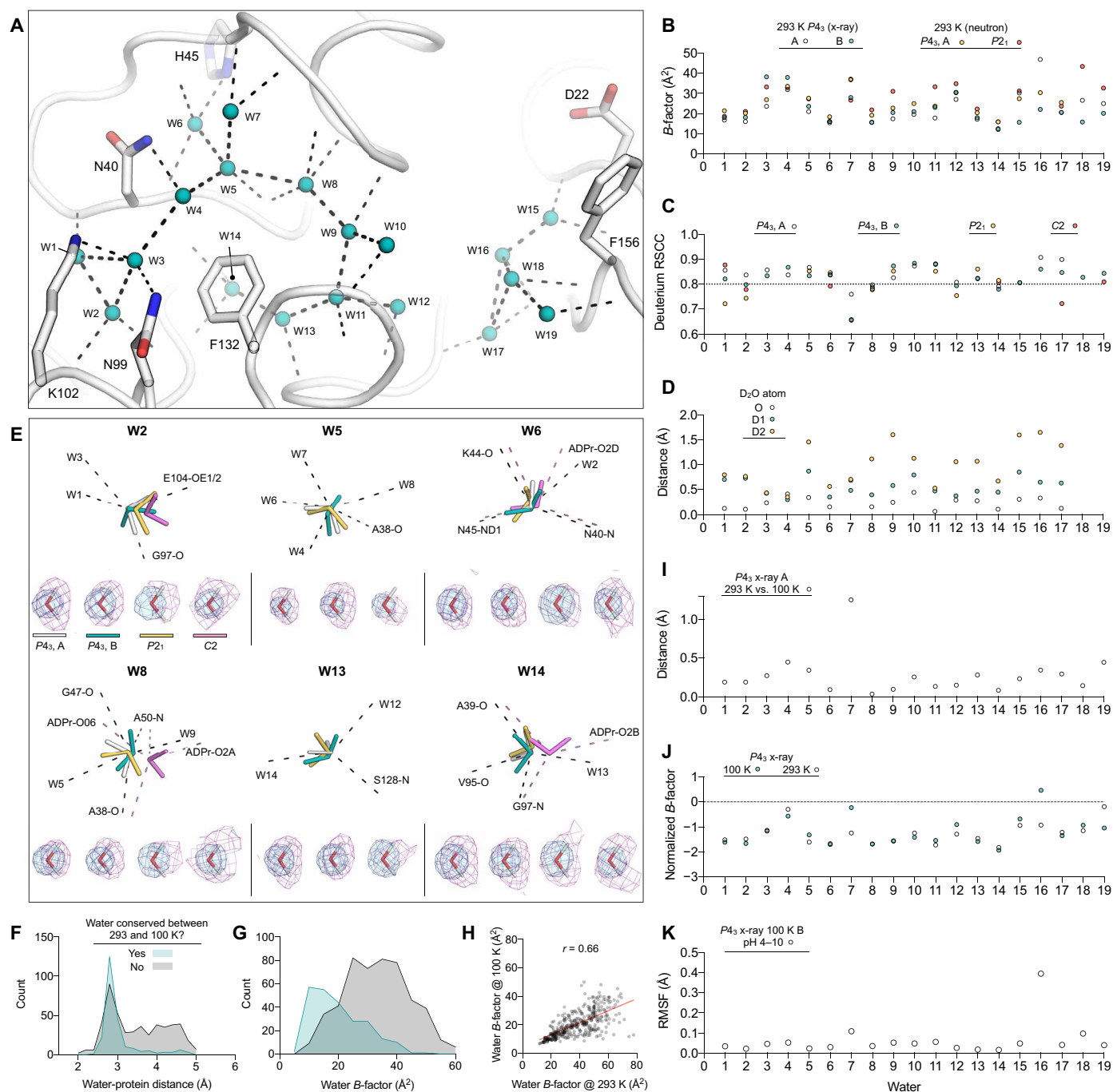
Water networks play an important role in protein-ligand recognition (35, 36); however, their contribution to ligand binding is often difficult to quantify. Here, the combination of ultrahigh-resolution cryogenic x-ray structures (0.77 and 0.85 Å), a high-resolution room temperature structure (1.1 Å), structures determined using different crystal forms (*P*<sub>43</sub>, *P*<sub>21</sub>, and C2), high-resolution neutron structures (1.9 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 *P*<sub>43</sub> crystals (Fig. 5A). There were 19 active site water molecules in total: 14 in the catalytic site and 5 in the adenosine site. The catalytic site network encircles the Phe<sup>132</sup> 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 *P*<sub>43</sub> ASU (Fig. 5B). 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 *P*<sub>21</sub> crystal. The water network was similar (RMSD = 0.45 Å), with only W10 and W16 not resolved in the *P*<sub>21</sub> structure. Together, these results indicate that the Mac1 active site water network is conserved across different crystal packing environments.

### **D<sub>2</sub>O 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 D<sub>2</sub>O molecules. To model D<sub>2</sub>O molecules in the neutron structures, deuterium atoms were added to D<sub>2</sub>O oxygens, and their orientations were refined using the 2F<sub>O</sub> - F<sub>C</sub> NSL density maps (24). Deuterium atoms were removed from D<sub>2</sub>O molecules that lacked density. Although all D<sub>2</sub>O molecules in the active site of the *P*<sub>43</sub> structure were modeled with deuterium, the lower resolution of the *P*<sub>21</sub> structure meant that six D<sub>2</sub>O molecules were modeled without deuterium (fig. S7). The quality of D<sub>2</sub>O modeling was assessed by calculating the real-space correlation coefficient (RSCC) (Fig. 5C) (37). The RSCC values were consistently high, indicating that there was good agreement between D<sub>2</sub>O orientations and NSL maps. Next, we assessed the reproducibility of D<sub>2</sub>O orientation in refinement by running 100 independent rounds of refinement starting from coordinates where the D<sub>2</sub>O deuterium atoms were randomly shifted by 0.5 Å (fig. S8). Deuterium atoms were included for all D<sub>2</sub>O molecules in this procedure. The RMSF of deuterium atoms calculated across the 100 structures provides a measure of the reproducibility of D<sub>2</sub>O orientations. For D<sub>2</sub>O molecules in the *P*<sub>43</sub> structure, RMSF values were low, consistent with highly reproducible D<sub>2</sub>O orientations (fig. S8). RMSF values for W3, W4, W9, W15, and W18 were comparatively high in the *P*<sub>21</sub> structure, supporting their omission from the final model. To measure the agreement between equivalent D<sub>2</sub>O molecules across the two *P*<sub>43</sub> protomers, we calculated the distance between the average deuterium coordinates from the 100 structures (Fig. 5D). The D1/D2 positions for W3, W4, W6, W11, and W14



**Fig. 5. Mac1 active site water networks are robust to changes in crystal packing, temperature, and pH.** (A) Active site water networks in the 293 K x-ray structure (PDB code 7TWH). Waters are shown as teal spheres, and hydrogen bonds are shown with dashed black lines. (B) B-factors for active site H<sub>2</sub>O/D<sub>2</sub>O molecules, numbered according to (A). (C) Average deuterium RSCC of active site D<sub>2</sub>O molecules calculated using 2mF<sub>O</sub> – DF<sub>C</sub> NSL density maps. An RSCC value of 0.8 has previously been used to assess whether a D<sub>2</sub>O is correctly oriented (37). (D) Variation in D<sub>2</sub>O orientation across 100 rounds of refinement. The distance between the average deuterium position across the two P<sub>4</sub><sub>3</sub> protomers is plotted along with the distance between D<sub>2</sub>O oxygen atoms. (E) Selected D<sub>2</sub>O molecules showing the 2mF<sub>O</sub> – DF<sub>C</sub> NSL density map (purple mesh contoured at 2  $\sigma$ ) and the 2mF<sub>O</sub> – DF<sub>C</sub> electron density map (blue mesh, 2  $\sigma$ ). Pink dashed lines show ADPr-specific hydrogen bonds. (F) Histogram showing water-protein distances in the 100 K P<sub>4</sub><sub>3</sub> x-ray structure (PDB code 7KQO). Waters were classified as conserved if a matching water molecule was found within 0.5 Å in the 293 K P<sub>4</sub><sub>3</sub> structure (PDB code 7TWH) (histogram bin width = 0.2 Å). (G) Same as (F) but showing water B-factors (histogram bin width = 5 Å<sup>2</sup>). (H) Scatterplot showing correlation between water B-factors at 100 and 293 K. (I) Distances between H<sub>2</sub>O oxygen atoms in protomer A of the P<sub>4</sub><sub>3</sub> x-ray structures determined at 293 and 100 K (PDB codes 7TWH/7KQO). (J) Normalized B-factors of active site H<sub>2</sub>O molecules in the P<sub>4</sub><sub>3</sub> x-ray structures determined at 293 and 100 K. (K) RMSF values of water molecules calculated across the seven structures determined from pH 4 to 10.



were all within 0.6 Å on average, consistent with conserved D<sub>2</sub>O orientations between the P<sub>43</sub> protomers. On the basis of *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. 5E).

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

It is well established that crystal cryocooling causes crystal lattice contraction and remodeling of water networks at crystal packing interfaces (38–40). In the P<sub>43</sub> crystal form of Mac1, lattice contraction upon cryocooling was modest, with a 2% decrease in unit cell volume from 293 to 100 K (PDB codes 7TWH versus 7KQO). To investigate the effect of cryocooling on Mac1 solvation, we compared the overall water structure in the 1.1-Å P<sub>43</sub> structure with the previously published 0.85-Å structure determined at 100 K (PDB code 7KQO). 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. 5, F to H). 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 tended to have higher *B*-factors as well. The resolution gap between the two structures (0.25 Å) is not large enough to explain the difference in the number of waters modeled (41). In line with previous observations (40), it seems likely that this difference is due to increased solvent disorder at room temperature. However, unlike previous observations (40), 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 and fig. S3). Next, we compared active site water networks at 293 and 100 K (Fig. 5, I and J). Water positions and *B*-factors were highly conserved (RMSD = 0.38, Pearson *r* for *B*-factors = 0.62). As expected, the positions of waters with lower *B*-factors tended to be more similar.

#### **Mac1 active site water networks are conserved from pH 4 to 10**

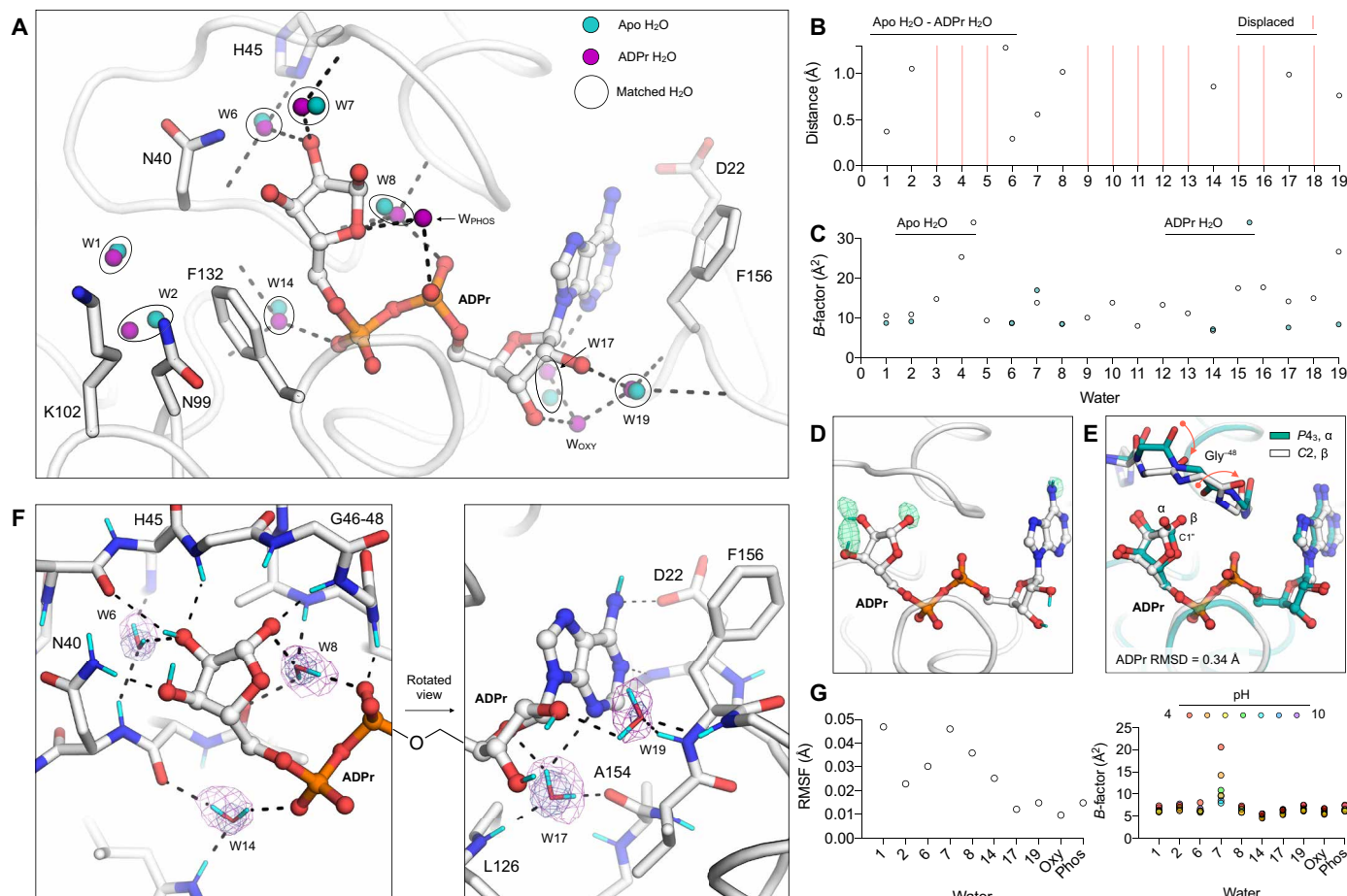
On the basis of the similarity of the Mac1 structures determined using crystals grown at pH 6.5 and 9.5, Mac1 tolerates a broad pH range. To extend this analysis to a wider pH range and higher resolution, while eliminating the impact of different crystal packing, we performed a pH shift experiment with the P<sub>43</sub> crystals. Briefly, crystals grown at pH 9.5 were soaked for 4 hours in a buffer composed of citric acid, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Hepes) and CHES, with pH values from 3 to 10 (42). 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). On the basis of 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 His<sup>91</sup> and His<sup>119</sup> (fig. S10). These residues are located at the interface between the two protomers in the P<sub>43</sub> ASU, and there was a pH-dependent shift in the side-chain  $\chi_1$  and  $\chi_2$  dihedral angles (fig. S10). Unexpectedly, we observed multiple buffer components binding to the oxyanion subsite (fig. S1). Binding was pH dependent, with density for citrate between pH 4 and 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 (43). Binding of the buffer components in the oxyanion subsite led to substantial conformational changes in the Phe<sup>156</sup> loop (fig. S9), mirroring the changes seen across different crystal forms (Fig. 4A) and upon fragment binding (10). There was an increase in backbone disorder from pH 4 to 10 for the  $\alpha$  helices composed of residues 20 to 31 and 156 to 169, as well as the Lys<sup>102</sup> loop (fig. S9). Unexpectedly, we observed a low-occupancy alternative conformation of the Ile<sup>131</sup> 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 Ala<sup>129</sup>. Overall, the active site water network was highly similar across the entire pH range (Fig. 5K and fig. S9).

#### **Reorganization 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 P<sub>43</sub> crystals (PDB codes 7KQO and 7KQP) (Fig. 6, A and B). In total, 11 waters are displaced by ADPr binding, either through direct displacement by ADPr (W4-5, W9-13, W15-16, and W18) or through displacement by a shift in protein conformation (W3, displaced by Phe<sup>132</sup>). Of the remaining waters, six form bridging hydrogen bonds between ADPr and Mac1 (W6, W7, W8, W14, W17, and W19). These waters tend to have more protein contacts and lower *B*-factors compared to the displaced waters (Fig. 5A). A notable exception is W19, which connects the oxyanion subsite to the C2' hydroxyl and is stabilized in the ADPr-bound structure (Fig. 6, A and C). There were minor adjustments in the bridging waters that bind ADPr; W6, W7, W8, W14, and W17 all shift between 0.3 and 1 Å (Fig. 6B). In addition to the water displacement and reorganization, two new waters are found in the ADPr-bound structure. W<sub>OXY</sub> extends the hydrogen bond network near the oxyanion subsite, and W<sub>PHOS</sub> is hydrogen-bonded to the terminal ribose and adjacent phosphate of ADPr (Fig. 6A). In total, the reorganization of water networks upon ADPr binding fits a model where weakly bound waters (as measured by the 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 structure determined using the C2 cocrystal system and neutron diffraction. Although the ADPr used for cocrystallization was hydrogenated, HDX occurred when the crystal was soaked in the deuterated buffer. This is supported by positive peaks in the F<sub>0</sub> – F<sub>C</sub> NSL density map calculated before modeling ADPr deuterium atoms (Fig. 6D). In the neutron structure, the terminal ribose adopts the  $\beta$  epimer, with rearrangement of the Gly<sup>46-48</sup> loop accommodating the C1' hydroxyl (Fig. 6E). This is in contrast to the P<sub>43</sub> structure, where crystal packing interactions preclude rearrangement of the Gly<sup>46-48</sup> loop, which forces the terminal ribose to adopt the  $\alpha$  epimer (Fig. 6E) (10). Despite the different epimers adopted by the terminal ribose and rearrangement of the Gly<sup>46-48</sup> loop, the ADPr-binding pose was very similar between the C2 and P<sub>43</sub> structures (RMSD = 0.34 Å, excluding the C1' hydroxyl). Likewise, the positions of bridging D<sub>2</sub>O molecules were highly conserved (RMSD for W6, W8, W14, and 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<sub>OXY</sub> and W<sub>PHOS</sub>), only W<sub>OXY</sub> was identified in the



**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 (P4<sub>3</sub> 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 HDX in ADPr cocrystallized with Mac1 (C2, PDB code 7TX5). The mF<sub>o</sub> - DF<sub>c</sub> NSL density map calculated after joint neutron/x-ray refinement but before adding deuterium atoms to ADPr is shown with a 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 P4<sub>3</sub> and C2 crystals. The two epimers of the terminal ribose are marked (α and β), and the conformational change required to bind the α epimer 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 2mF<sub>o</sub> - DF<sub>c</sub> 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 2mF<sub>o</sub> - DF<sub>c</sub> 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.

C2 structure. The absence of W<sub>PHOS</sub> is due to crystal packing interactions; a close contact between Ala<sup>69</sup> of a symmetry mate forces the Ile<sup>131</sup> side chain to rotate and occupy the W<sub>PHOS</sub> position (fig. S1).

Refinement of the positions of ADPr deuteriums and the orientations of bridging D<sub>2</sub>O 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 Ala<sup>38</sup>, the backbone nitrogen of Ala<sup>50</sup>, and one of the ADPr phosphate oxygens. This highly coordinated D<sub>2</sub>O is appropriately oriented to act as a nucleophile in the putative substrate-assisted hydrolysis mechanism (8), with the D<sub>2</sub>O oxygen oriented toward the C1' carbon. In addition, in the catalytic site, W6 forms a bridging interaction between the deprotonated N81 of His<sup>45</sup> and the C2' hydroxyl of ADPr. The nearby Asn<sup>40</sup>, which plays a key role in Mac1 function (8), forms a 3-Å hydrogen bond

with the C3' hydroxyl. In the adenosine site, W19 acts as a bridge between the oxanion subsite and the C2' hydroxyl (Fig. 6F). The second bridging D<sub>2</sub>O in this site, W17, forms hydrogen bonds to the backbone oxygen of Ala<sup>154</sup> and the backbone nitrogen of Leu<sup>126</sup> 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 to 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 (10). 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 to 10 were excellent, with reflections recorded beyond 0.8 Å at all pHs. Data were truncated to 0.9 Å to achieve 100% completeness, except for 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). Similar to 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 the stabilization of protomer A by ADPr binding. At pH 4, there was weak density for an alternative conformation of His<sup>45</sup> in protomer A (fig. S10). The conformational change involved an 8-Å shift in the His<sup>45</sup> side chain, with an adjustment in the backbone torsion angles of Gly<sup>46/47</sup> and rearrangement of water molecules to occupy the cavity left by His<sup>45</sup> (fig. S10). The change in His<sup>45</sup> conformation is likely induced by protonation at pH 4, indicating that the doubly protonated imidazole group 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 *P*<sub>43</sub> 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 (10). 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 (10). 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 (Fig. 7A) (44). 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 52 of them. The remaining 12 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 to 3.5 Å and over a wide range of angles (Fig. 7, E to G). The flexibility in bridging interactions with W17 may explain the large number of carboxylic acid-containing fragments that bound to the oxyanion subsite (10). Of the 64 fragments interacting with W17, 47 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).

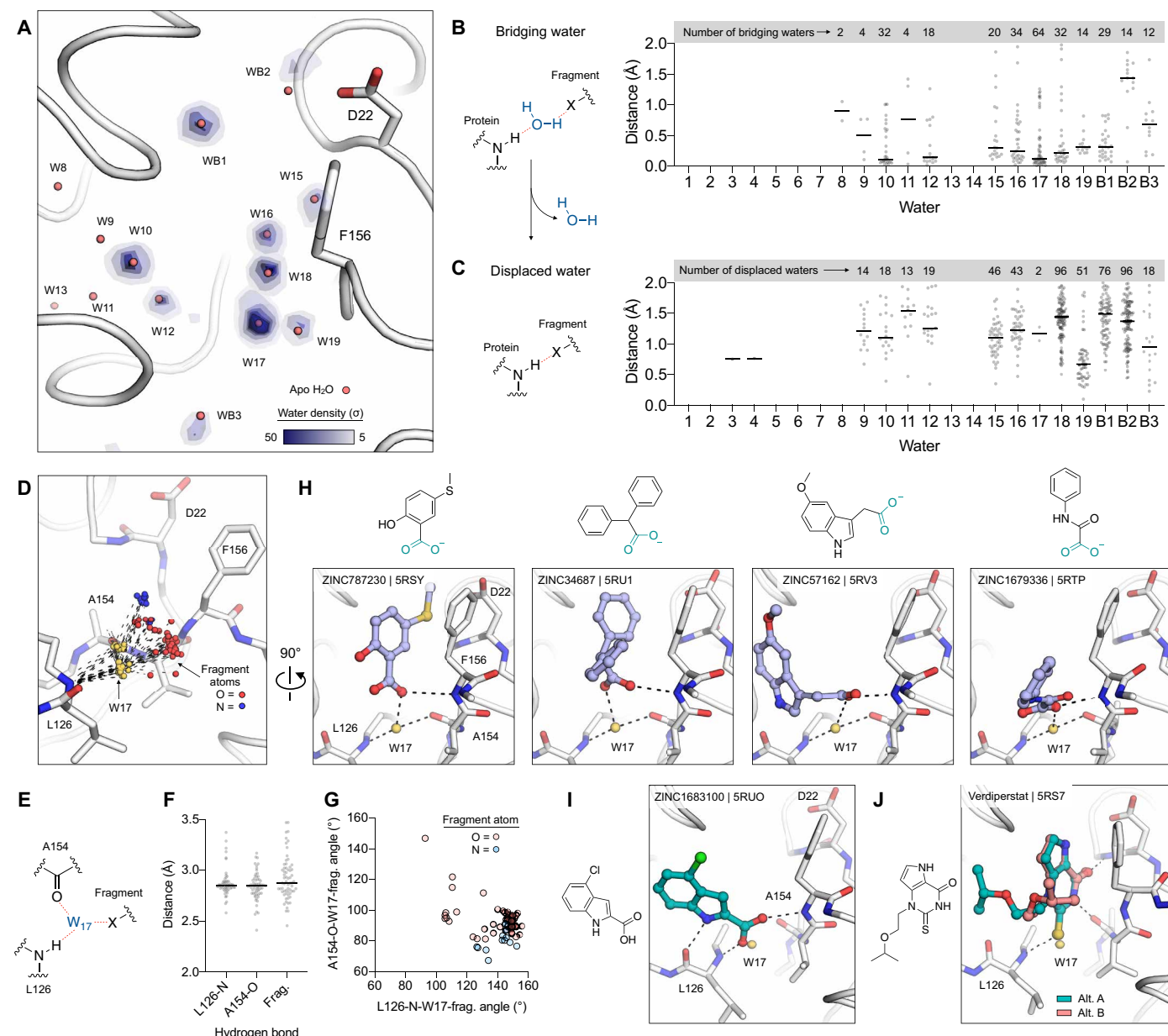
Next, we examined the displacement of waters from the adenosine network by fragment binding. A water was classified as displaced if a fragment atom was within 2 Å of the water in the apo structure (Fig. 7C). On average, each fragment binding in the Mac1 active site displaced 1.6 waters from the active 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 (10). Waters that were highly ordered in the apo structure were displaced infrequently or not at all. Notably, W17 was only displaced by 2 fragments of 178 binding in the adenosine site (Fig. 7C). In both cases, the fragments form extensive hydrogen bond networks to neighboring residues (Fig. 7, I and J). The indole-containing ZINC1683100 forms hydrogen bonds to the backbone nitrogen/oxygen of Leu<sup>129</sup> and to the oxyanion subsite (Fig. 7I), while the thioketone of verdiperstat forms hydrogen bonds with the Leu<sup>126</sup> 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 most of the waters in the adenosine site network are available for bridging interactions or displacement (Fig. 7, B and 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. 4, D to 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 (Figs. 5 and 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 shows that the water networks in the Mac1 active site are unexpectedly 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 2 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 are consistent with previous large-scale surveys of ligand-water-protein interactions (45, 46). Although ADPr recognition involves a series of changes in Mac1 conformation that either introduce new polar contacts with ADPr (e.g., the Gly<sup>48</sup>/Ala<sup>129</sup> peptide flips) or remove





**Fig. 7. Water networks mediate fragment binding in the adenosine site of Mac1.** (A) Structure of Mac1 showing the density of fragment-protein bridging water molecules, calculated using the 232 previously reported fragment structures (10). The P4<sub>3</sub> 100 K structure (PDB code 7KQO) is shown with a white stick/cartoon representation. Water density was calculated with the GROMaps tool (44) and is contoured from 5 to 50  $\sigma$ . (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 Leu<sup>126</sup> and Ala<sup>154</sup> are shown with black dashed lines, and the fragment atoms are shown with spheres colored by atom type. (E to G) The W17-Leu<sup>126</sup>/Ala<sup>154</sup> 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 and J) W17 is displaced by 2 of the 178 fragments binding in the adenosine site. Two conformations of veridiperstat were observed.

steric blocks (e.g., the coupled Phe<sup>132</sup>/Asn<sup>99</sup> 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 lead 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 not only acts as a bridging water in the ADPr complex but 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 (Figs. 6A and 7A). This correlation could be leveraged in ligand design to minimize 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 (47). Second, the water network analysis identified buried waters that can be targeted for displacement. This strategy is frequently invoked in ligand optimization efforts (36, 48, 49), but it is unclear how broadly applicable this strategy is for optimizing potency. The large number of bridging waters identified here provides an unmatched opportunity to systematically examine the thermodynamics of water displacement (for example, the 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 (Figs. 5E and 6F). The orientation of water molecules included in receptor models can have a major impact on the success of virtual screening (50). One caveat to using the neutron-determined orientations is that only orientations of highly ordered water molecules could be determined with high confidence (Fig. 5, C and D). 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 (51).

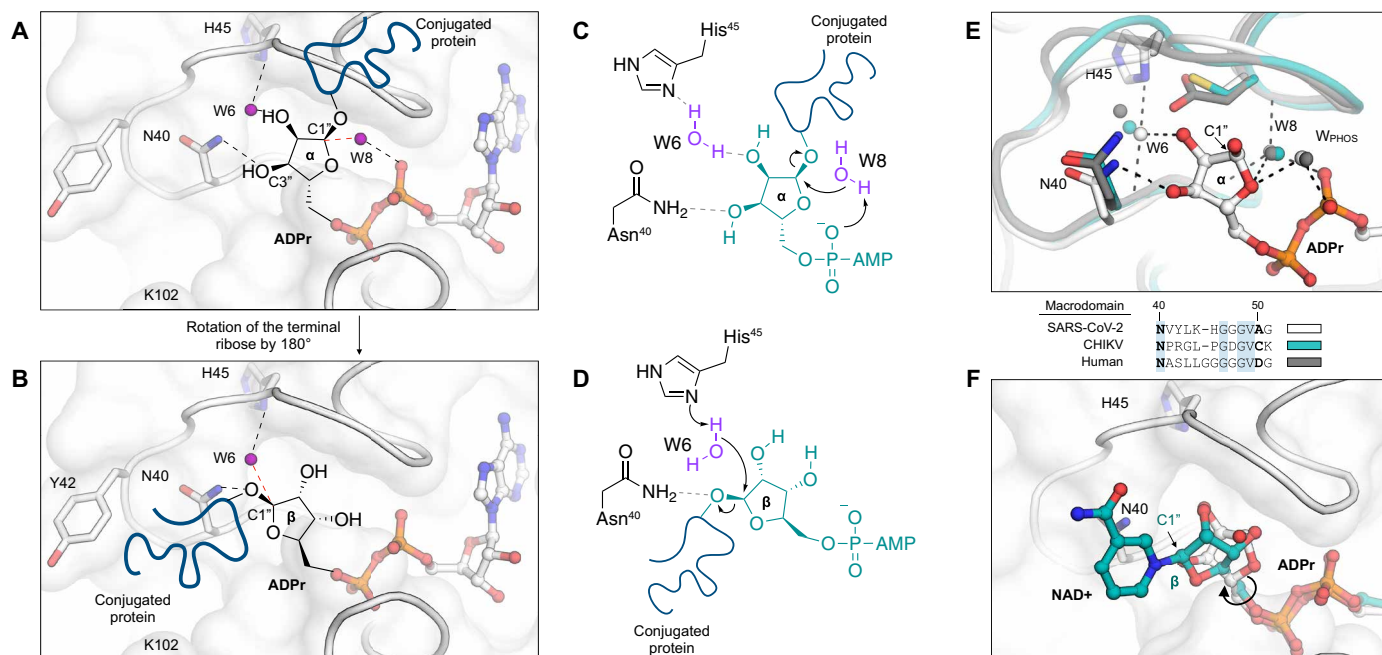
We found that the conformational ensemble of Mac1 was unusually robust to temperature and pH perturbation in the crystal. Room temperature crystallography can reveal low-occupancy, functionally relevant conformational states of proteins (39, 52–54). The 1.1-Å room temperature structure of Mac1 reported here, in combination with the previously reported 0.85-Å structure determined using the same crystal form at 100 K (10), 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 (54–56), 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 (57), the pH shift experiments revealed that the Mac1 active site residues and water networks are remarkably robust (Fig. 5K, fig. S9), and this is matched by the invariance in ADPr binding across the entire pH range (fig. S9). To the extent of our knowledge, the pH dependence of enzyme activity has not been determined for viral or human macrodomains. On the basis of the structural rearrangement of His<sup>45</sup> and the adjacent glycine-rich loop seen at pH 4 (fig. S10), we predict a corresponding loss of activity at low pH. The structural data reported here will assist the interpretation of future pH activity data. 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. 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 found conformational diversity by comparing Mac1 structures across different crystal forms (Fig. 4C). This ensemble of structures captures most of the structural diversity seen upon ADPr and fragment binding (fig. S6). One notable exception is the Phe<sup>132</sup> 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). Unexpectedly, the relatively constant position of Phe<sup>132</sup> across the apo structures occurs despite substantial variation in the Asn<sup>99</sup>/Lys<sup>102</sup> 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 (10); fragments are unable to overcome the energetic penalty required to stabilize the Phe<sup>132</sup> 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 focused 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. The correlation between the structural diversity observed across different space groups and the diversity seen in fragment/ADPr-bound structures indicates that different crystal packing interactions can reveal functionally important structural differences, instead of being merely structural artifacts.

Although the physiologically relevant substrates of Mac1 are unknown, macrodomains are thought to primarily hydrolyze ADPr conjugated to glutamate or aspartate side chains at the C1' position (Fig. 1A) (58). On the basis of the x-ray and neutron crystal structures of the SARS-CoV-2 macrodomain with ADPr bound, two substrate-binding modes are plausible, distinguishable based on the conformation of the terminal ribose (Fig. 8, A and B) and leading to different potential catalytic mechanisms. In the first mode, which matches the conformation seen in the P<sub>4</sub> crystal structure, the terminal ribose adopts the  $\alpha$  epimer, with the conjugated peptide exiting the active site through the gap between the Gly<sup>46–48</sup> loop and the Ile<sup>131</sup> 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) (10). In this binding mode, the proximity of the C1' carbon to the phosphate groups and the absence of another suitable general base implicate the previously proposed substrate-assisted hydrolysis mechanism with either W<sub>PHOS</sub> or W8 acting as a nucleophile (Fig. 8C) (8, 11). If the hydrolysis reaction involves an S<sub>N</sub>2-type nucleophilic substitution at the C1' carbon, then we favor W8 as the water nucleophile, because it can approach the sp<sup>3</sup> carbon with a collinear trajectory relative to the peptide leaving group (Fig. 8C), and the neutron crystal structure with ADPr bound shows that the D<sub>2</sub>O oxygen is oriented toward the C1' carbon (Fig. 6G). This is in contrast to the water nucleophile proposed by Jankevicius *et al.* (11) in hMacroD2, which is equivalent to W<sub>PHOS</sub> in Mac1 (Fig. 8E).

In the second possible ADPr-binding mode, the ribose adopts the  $\beta$  epimer, and the terminal ribose is rotated 180° about the C4'–C5' bond relative to the epimer observed in the P<sub>4</sub> crystals (Fig. 8B). In



**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 conformation that is compatible with a substrate-assisted mechanism. The terminal ribose adopts the  $\alpha$  epimer, 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  $\beta$  epimer of the terminal ribose that is compatible with His<sup>45</sup> acting as a general base to activate W6 as a nucleophile. (C and 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 to 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<sup>+</sup> (PDB code 6MEB). The terminal ribose is rotated  $\sim 180^\circ$  relative to ADPr. This configuration matches the model proposed in (B) and (D).

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 His<sup>45</sup> as a general base in the hydrolysis mechanism for coronaviral macrodomains (Fig. 8D). First, the neutron structures show that His<sup>45</sup> has the appropriate orientation and protonation state to act as a general base, with a deprotonated N $\delta$ 1 nitrogen and a D<sub>2</sub>O molecule hydrogen-bonded between His<sup>45</sup> and ADPr (Fig. 3D). Second, His<sup>45</sup> is conserved across all known coronavirus macrodomains, with mutation to aspartic acid eliminating ADPr-hydrolysis activity (8). Note that His<sup>45</sup> is absent in alphavirus macrodomains and the human macrodomain hMacroD2, although both macrodomains contain a nearby nucleophilic residue (Asp<sup>102</sup> in hMacroD2 and Cys<sup>34</sup> 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<sup>+</sup> bound (Fig. 8F) (59). The nicotinamide group is oriented toward a groove formed between Tyr<sup>42</sup> and Lys<sup>102</sup>. We previously identified several fragments binding in this site (10) 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 (59). Last, this binding mode hints at an alternative role for Asn<sup>40</sup> in ADPr hydrolysis. Asn<sup>40</sup> is strictly conserved across

macrodomains, and mutation to alanine eliminates activity (7). In the substrate-assisted mechanism, this functional importance has been attributed to the hydrogen bond formed between Asn<sup>40</sup> and the C3'' hydroxyl of ADPr. In the alternative mechanism proposed here (Fig. 8, B and D), Asn<sup>40</sup> 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 nonhydrolyzable ADPr-ribosylated protein or peptide). As has been suggested previously (8), 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 site 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.



## MATERIALS AND METHODS

### P4<sub>3</sub> crystals

#### Protein expression and purification

The gene encoding SARS-CoV-2 NSP3 Mac1 (residues 3 to 169) was cloned into a pET-22b(+) expression plasmid with a TEV-cleavable N-terminal 6-His tag (Genscript). The protein was expressed and purified as described previously (10).

#### Crystallization for neutron diffraction

Neutron diffraction-quality crystals were grown using sitting drop vapor diffusion. Crystallization drops were set up on 22-mm silicized glass coverslips (Hampton Research, HR3-231) with 65  $\mu$ l of protein, 55  $\mu$ l of reservoir [32% PEG 3000 and 100 mM CHES (pH<sub>apparent</sub> 9.1; pD 9.5)], and 10  $\mu$ l of seed stock. This crystallization condition was previously identified using commercial crystallization screens (10) and was chosen on the basis of the high-resolution diffraction data obtained from these crystals ( $<0.9$  Å). Seed stocks were prepared as described previously (10) and were stored at  $-80^{\circ}\text{C}$ . The sitting drops were equilibrated against a 4-ml reservoir solution stored in a small plastic beaker, using two 10-cm plastic petri dishes taped together to create a sealed environment. Crystals reached their maximum size of approximately 2 mm by 2 mm by 0.5 mm after 1 week. To increase the deuterium content, crystals were transferred from crystallization drops into an artificial reservoir solution containing 32% PEG 3000 and 100 mM CHES (pH 9.5), prepared using D<sub>2</sub>O. Because crystals were tightly bound to the coverslips, it was necessary to use a fine needle to dislodge them. Crystals were allowed to soak in the artificial reservoir solution for 24 hours, before being transferred into a 2-mm thin-walled quartz capillary (Hampton Research, HR6-150). To keep crystals hydrated, 40  $\mu$ l of the D<sub>2</sub>O-containing reservoir solution was pipetted into one end, and the capillaries were sealed with wax at both ends (Hampton Research, HR4-328).

#### Neutron/x-ray diffraction data collection, reduction, and refinement

Neutron time-of-flight diffraction data were collected at room temperature on the MaNDi instrument at the Spallation Neutron Source (SNS) (21, 60). An incident neutron wavelength band-pass of 2 to 4 Å was used with a  $\Delta\phi$  value of  $10^{\circ}$  between frames. A total of 10 diffraction images were collected with an average exposure of 16 hours per frame. Exposure was set to 80 C of charge on the SNS target. At 1.4-MW operation, charge accumulates at a rate of approximately 5 C/hour. Fluctuation in exposure time occurs because the accelerator does not operate 100% of the time. The neutron dataset was reduced using the Mantid package (61) and integrated using three-dimensional profile fitting (62). The data were wavelength normalized using LAUENORM from the LAUEGEN suite (63–65).

Following neutron diffraction data collection, an x-ray dataset was collected on the same crystal at room temperature using a microfocus rotating anode x-ray diffractometer (Rigaku HighFlux HomeLab equipped with a MicroMax-007 HF x-ray generator, Osmic VariMax optics, and an Eiger R 4M hybrid photon counting detector). A total of 360 diffraction images were collected with a  $\Delta\phi$  value of  $0.5^{\circ}$  and an exposure time of 10 s per frame. X-ray diffraction data were integrated using the CrysAlis Pro software suite (Rigaku Inc.). Diffraction data were then reduced and scaled using the Aimless (66) program from the CCP4 suite (67). The x-ray data collection statistics are shown in table S1.

The Mac1 structure was initially refined against the x-ray data only. Phases were obtained by molecular replacement with Phaser

(68) using the coordinates of the apo Mac1 as the search model (monomer A of PDB code 7KQO, stripped of alternative conformations and waters). To debias refinement from the coordinates used for molecular replacement, the *B*-factors of 7KQO were reset to  $20\text{ Å}^2$ , and the coordinates were randomly offset by  $0.3\text{ Å}$  using *phenix.pdbtools*. The structure was refined with *phenix.refine* (69) using default parameters and five macrocycles. Manual model building was performed with Coot (70). In the early stages of refinement, waters were added automatically in *phenix.refine* to peaks in the  $mF_o - DF_c$  map  $3.5\sigma$  or higher. CHES was modeled into the adenosine site with restraints generated using *phenix.elbow* (71). 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 (31) run through *phenix.ready\_set*. Hydrogens were refined at their riding positions. Cycles of refinement and model building continued until  $R_{\text{free}}$  no longer decreased. The coordinates were prepared for joint neutron/x-ray refinement by adding deuterium atoms to all exchangeable positions and D<sub>2</sub>O 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* (24) with hydrogens/deuteriums refined individually. Hydrogen/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 were 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). D<sub>2</sub>O orientations were automatically refined with *phenix.refine*. The fit of D<sub>2</sub>O molecules into  $2mF_o - DF_c$  NSL density maps was assessed by visual inspection, and the orientations of D<sub>2</sub>O molecules were adjusted when necessary. Deuteriums were deleted from solvent 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 Research, HR3-171). Crystallization drops contained 12  $\mu$ l of Mac1 at 40 mg/ml, 2  $\mu$ l of seeds, and 10  $\mu$ l of reservoir consisting of 32% PEG 3000 and 100 mM CHES (pH 9.5) with 500  $\mu$ l of reservoir in each well. Crystals reached a maximum size of approximately 1 mm by 1 mm by 0.25 mm after 4 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  $\mu$ 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 were 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 (72) using the parameters listed in table S1. X-ray diffraction data were indexed, integrated and scaled using XDS (73), and merged with Aimless (66). 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 of protein, 100 nl of reservoir solution [28% PEG 3000 and 100 mM CHES (pH 9.5)], and 100 nl of seed stock, with 30  $\mu$ l in the reservoirs. Crystals grew to their maximum size in 2 days with dimensions of approximately 0.4 mm by 0.4 mm by 0.1 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) (42) and 28% PEG 3000, with the pH adjusted from 3 to 10. Crystals were soaked for 4 hours and 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-Aldrich, A0752) (from a 500 mM stock prepared in H<sub>2</sub>O). Diffraction data were collected at beamline 8.3.1 of the Advanced Light Source (ALS) for three to six 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 was observed in the apo crystal structures. Restraints were 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), and 7TX1 (pH 10; ADPr). Data reduction and refinement statistics are presented in table S1, with the statistics produced with *phenix.table\_one*.

#### P2<sub>1</sub>/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 *Escherichia coli* (BL21-DE3). A detailed procedure for Mac1 expression and purification has been published elsewhere (25).

##### Crystallization for neutron diffraction

Purified Mac1 was stored at 4°C and filtered through a 0.2- $\mu$ m centrifugal filter before crystallization trials. Initial crystals grown in a 20- $\mu$ l drop of Mac1 (~17.5 mg/ml) mixed with 28% PEG 4000 and 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- $\mu$ l drops of Mac1 (~19 mg/ml) mixed at a 1:1 ratio with 25% PEG 4000 and 0.1 M MES (pH 6.5) seeded with 0.2  $\mu$ l of microseeds (1:3000 dilution). This crystallization condition was previously identified using commercial crystallization screens (25) and was chosen on the basis of the high-resolution diffraction data obtained from these crystals (~1.1 Å). 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% D<sub>2</sub>O. 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 min. 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 cocrystal previously grown in a 24-well plate using a 24% PEG 4000 and 0.1 M MES (pH 6.5) well solution, and then the 9-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% D<sub>2</sub>O. 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 HFIR at ORNL (26, 74, 75). The crystal was held stationary at room temperature, and diffraction data for each image were collected for 23 hours using neutrons between 3.3 and 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 (64, 76). The program LSCALE (77) 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 (78). 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  $P4_3$  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  $P4_3$  neutron/x-ray experiment. Phases were obtained by molecular replacement using the same coordinates as the  $P4_3$  experiments. For the ADPr cocrystal structure, binding of ADPr was confirmed by inspection of the  $mF_o - DF_c$  electron density map. Restraints for the  $\beta$  epimer 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  $P2_1$  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 Ca atoms. Second, average coordinates for each Ca were calculated across the structures. Third, the RMSF was calculated relative to the average Ca coordinates. Normalized Ca  $B$ -factors ( $B'$ ) were calculated using the following equation:  $B' = (B - \mu B)/\sigma B$ , where  $B$  is the Ca  $B$ -factor,  $\mu B$  is the average  $B$ -factor across Ca atoms, and  $\sigma B$  is the Ca  $B$ -factor SD. Because of their high  $B$ -factor, residues 1 to 10 and 165 to 170 were excluded from the normalization procedure. 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 on the basis of NSL density maps after refinement with either the doubly or deprotonated side chain (maps are shown in fig. S5). For the doubly protonated side chains, the occupancy of the  $N\epsilon_2/N\delta_1$  deuterium atoms was set to 100%, and automatic refinement of occupancy was disabled.

### Water

RSCC values for  $D_2O$  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,  $\mu B$  is the average  $B$ -factor across all water molecules, and  $\sigma B$  is the water  $B$ -factor SD.

To assess the reproducibility of  $D_2O$  orientation refinement, we ran 100 independent rounds of refinement starting from the three jointly refined Mac1 structures. Deuterium atoms were added to all  $D_2O$  oxygen atoms before 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  $D_2O$  orientations but with the same oxygen position. The variability in  $D_2O$  orientations after refinement was quantified by calculating the RMSF value for each  $D_2O$  deuterium atom across the 100 structures (fig. S8). Agreement between the orientation of  $D_2O$  molecules in the active sites of protomer A and B of the  $P4_3$

structure was assessed by calculating the distance between the average position of each  $D_2O$  deuterium from the 100 refined structures (Fig. 5D).

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 Ca 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 (10) was calculated using the GROMaps tool in GROMACS (44). 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  $P4_3$  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.

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <https://science.org/doi/10.1126/sciadv.abo5083>

[View/request a protocol for this paper from Bio-protocol.](#)

## REFERENCES AND NOTES

1. L. B. Ivashkiv, L. T. Donlin, Regulation of type I interferon responses. *Nat. Rev. Immunol.* **14**, 36–49 (2014).
2. J. W. Schoggins, Interferon-stimulated genes: What do they all do? *Annu Rev Virol.* **6**, 567–584 (2019).
3. N. C. Hoch, L. M. Polo, ADP-ribosylation: From molecular mechanisms to human disease. *Genet. Mol. Biol.* **43**, e20190075 (2019).
4. J. G. M. Rack, L. Palazzo, I. Ahel, (ADP-ribosyl)hydrolases: Structure, function, and biology. *Genes Dev.* **34**, 263–284 (2020).
5. L. C. Russo, R. Tomasini, I. A. Matos, A. C. Manucci, S. T. Sowa, K. Dale, K. W. Caldecott, L. Lehtiö, D. Schechtman, F. C. Meotti, A. Bruni-Cardoso, N. C. Hoch, The SARS-CoV-2 Nsp3 macrodomain reverses PARP9/DTX3L-dependent ADP-ribosylation induced by interferon signaling. *J. Biol. Chem.* **297**, 101041 (2021).
6. Y. M. O. Alhammad, M. M. Kashipathy, A. Roy, J.-P. Gagné, P. McDonald, P. Gao, L. Nonfoux, K. P. Bataille, D. K. Johnson, E. D. Holmstrom, G. G. Poirier, S. Lovell, A. R. Fehr, The SARS-CoV-2 conserved macrodomain is a mono-ADP-ribosylhydrolase. *J. Virol.* **95**, e01969-20 (2021).
7. A. R. Fehr, R. Channappanavar, G. Jankevicius, C. Fett, J. Zhao, J. Athmer, D. K. Meyerholz, I. Ahel, S. Perlman, The conserved coronavirus macrodomain promotes virulence and suppresses the innate immune response during severe acute respiratory syndrome coronavirus infection. *MBio* **7**, 01721-16 (2016).
8. 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).
9. J. Lei, Y. Kusov, R. Hilgenfeld, Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral Res.* **149**, 58–74 (2018).
10. M. Schuller, G. J. Correy, S. Gahbauer, D. Fearon, T. Wu, R. E. Díaz, I. D. Young, L. C. Martins, D. H. Smith, U. Schulze-Gahmen, T. W. Owens, I. Deshpande, G. E. Merz, A. C. Thwin, J. T. Biel, J. K. Peters, M. Moritz, N. Herrera, H. T. Kratochvil; QCRG Structural Biology Consortium, A. Aimon, J. M. Bennett, J. B. Neto, A. E. Cohen, A. Dias, A. Douangamath, L. Dunnett, O. Fedorov, M. P. Ferla, M. R. Fuchs, T. J. Gorrie-Stone, J. M. Holton,



- 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. S. Soares, J. L. Wierman, K. Zhu, P. O'Brien, N. Jura, A. Ashworth, J. J. Irwin, M. C. Thompson, J. E. Gestwicki, F. von Delft, B. K. Shoichet, J. S. Fraser, I. Ahel, Fragment binding to the Nsp3 macrodomain of SARS-CoV-2 identified through crystallographic screening and computational docking. *Sci Adv.* **7**, eabf8711 (2021).
11. 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).
12. A. Vrielink, N. Sampson, Sub-angstrom resolution enzyme x-ray structures: Is seeing believing? *Curr. Opin. Struct. Biol.* **13**, 709–715 (2003).
13. M. P. Blakeley, P. Langan, N. Niimura, A. Podjarny, Neutron crystallography: Opportunities, challenges, and limitations. *Curr. Opin. Struct. Biol.* **18**, 593–600 (2008).
14. 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).
15. D. A. A. Myles, Neutron protein crystallography: Current status and a brighter future. *Curr. Opin. Struct. Biol.* **16**, 630–637 (2006).
16. 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).
17. M. P. Blakeley, Neutron crystallography aids in drug design. *IUCrJ.* **3**, 296–297 (2016).
18. 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  $\gamma$ -phosphate. *J. Biol. Chem.* **290**, 31025–31036 (2015).
19. A. C. Anderson, The process of structure-based drug design. *Chem. Biol.* **10**, 787–797 (2003).
20. M. Congreve, C. W. Murray, T. L. Blundell, Structural biology and drug discovery. *Drug Discov. Today* **10**, 895–907 (2005).
21. L. Coates, B. Sullivan, The macromolecular neutron diffractometer at the spallation neutron source. *Methods Enzymol.* **634**, 87–99 (2020).
22. 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. Cryst.* **48**, 1302–1306 (2015).
23. 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. Cryst.* **43**, 570–577 (2010).
24. 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).
25. K. Michalska, Y. Kim, R. Jedrejczak, 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).
26. 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).
27. M.-H. Lin, S.-C. Chang, Y.-C. Chiu, B.-C. Jiang, T.-H. Wu, C.-H. Hsu, Structural, biophysical, and biochemical elucidation of the SARS-CoV-2 nonstructural protein 3 macro domain. *ACS Infect. Dis.* **6**, 2970–2978 (2020).
28. 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).
29. A. A. Kossiakoff, Protein dynamics investigated by the neutron diffraction-hydrogen exchange technique. *Nature* **296**, 713–721 (1982).
30. 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).
31. 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).
32. G. Duan, V. H. Smith, D. F. Weaver, Characterization of aromatic-thiol  $\pi$ -type hydrogen bonding and phenylalanine-cysteine side chain interactions through ab initio calculations and protein database analyses. *Mol. Phys.* **99**, 1689–1699 (2001).
33. R. Wolfenden, Enzyme catalysis: Conflicting requirements of substrate access and transition state affinity. *Mol. Cell. Biochem.* **3**, 207–211 (1974).
34. D. Herschlag, A. Natarajan, Fundamental challenges in mechanistic enzymology: Progress toward understanding the rate enhancements of enzymes. *Biochemistry* **52**, 2050–2067 (2013).
35. 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).
36. 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).
37. 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).
38. D. H. Juers, B. W. Matthews, Reversible lattice repacking illustrates the temperature dependence of macromolecular interactions. *J. Mol. Biol.* **311**, 851–862 (2001).
39. 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).
40. 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).
41. O. Carugo, D. Bordo, How many water molecules can be detected by protein crystallography? *Acta Crystallogr. D Biol. Crystallogr.* **55**, 479–483 (1999).
42. J. Newman, Novel buffer systems for macromolecular crystallization. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 610–612 (2004).
43. L. Abas, M. Guppy, Acetate: A contaminant in Hepes buffer. *Anal. Biochem.* **229**, 139–140 (1995).
44. R. Briones, C. Blau, C. Kutzner, B. L. de Groot, C. Aponte-Santamaria, GROMaps: A GROMACS-based toolset to analyze density maps derived from molecular dynamics simulations. *Biophys. J.* **116**, 4–11 (2019).
45. 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).
46. 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).
47. B. C. Roberts, R. L. Mancera, Ligand-protein docking with water molecules. *J. Chem. Inf. Model.* **48**, 397–408 (2008).
48. 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).
49. 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).
50. 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).
51. 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).
52. 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).
53. 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**, e36307 (2018).
54. 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* 10.1101/2021.05.03.437411 (2021).
55. M. Fischer, Macromolecular room temperature crystallography. *Q. Rev. Biophys.* **54**, e1 (2021).
56. A. Broom, R. V. Rakotoharisoa, M. C. Thompson, N. Zarifi, E. Nguyen, N. Mukhametzhano, 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).
57. 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 higher pH. *J. Mol. Biol.* **425**, 3121–3136 (2013).
58. J. G. M. Rack, D. Perina, I. Ahel, Macrod domains: Structure, function, evolution, and catalytic activities. *Annu. Rev. Biochem.* **85**, 431–454 (2016).
59. R. G. Hammond, N. Schormann, R. L. McPherson, A. K. L. Leung, C. C. S. Deivanayagam, M. A. Johnson, ADP-ribose and analogues bound to the deMARIylating macrodomain from the bat coronavirus HKU4. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2004500118 (2021).
60. F. Meilleur, L. Coates, M. J. Cuneo, A. Kovalevsky, D. A. A. Myles, The neutron macromolecular crystallography instruments at Oak Ridge National Laboratory: Advances, challenges, and opportunities. *Crystals* **8**, 388 (2018).
61. O. Arnold, J. C. Bilheux, J. M. Borreguero, A. Buts, S. I. Campbell, L. Chapon, M. Doucet, N. Draper, R. F. Leal, M. A. Gigg, V. E. Lynch, A. Markvardsen, D. J. Mikkelsen, R. L. Mikkelsen, R. Miller, K. Palmen, P. Parker, G. Passos, T. G. Perring, P. F. Peterson, S. Ren, M. A. Reuter, A. T. Savici, J. W. Taylor, R. J. Taylor, R. Tolchenov, W. Zhou, J. Zikovsky, Mantid—Data analysis and visualization package for neutron scattering and  $\mu$ SR experiments. *Nucl. Instrum. Methods Phys. Res. A* **764**, 156–166 (2014).

62. B. Sullivan, R. Archibald, P. S. Langan, H. Dobbek, M. Bommer, R. L. McFeeters, L. Coates, X. Wang, F. Gallmeier, J. M. Carpenter, V. Lynch, P. Langan, Improving the accuracy and resolution of neutron crystallographic data by three-dimensional profile fitting of Bragg peaks in reciprocal space. *Acta Crystallogr D Struct Biol.* **74**, 1085–1095 (2018).
  63. J. R. Helliwell, J. Habash, D. W. J. Cruickshank, M. M. Harding, T. J. Greenhough, J. W. Campbell, I. J. Clifton, M. Elder, P. A. Machin, M. Z. Papiz, S. Zurek, The recording and analysis of synchrotron X-radiation Laue diffraction photographs. *J. Appl. Cryst.* **22**, 483–497 (1989).
  64. J. W. Campbell, Q. Hao, M. M. Harding, N. D. Nguti, C. Wilkinson, LAUEGEN version 6.0 and INTLDM. *J. Appl. Crystallogr.* **31**, 496–502 (1998).
  65. G. P. Bourenkov, A. N. Popov, H. D. Bartunik, Bayesian approach to the analysis of time-resolved protein Laue diffraction data. *Acta Crystallogr. A* **52**, 50 (1996).
  66. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
  67. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
  68. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
  69. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
  70. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
  71. N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, Electronic Ligand Builder and Optimization Workbench (eLBOW): A tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1074–1080 (2009).
  72. C. S. Bury, J. C. Brooks-Bartlett, S. P. Walsh, E. F. Garman, Estimate your dose: RADDOSE-3D. *Protein Sci.* **27**, 217–228 (2018).
  73. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
  74. G. C. Schröder, W. B. O'Dell, D. A. A. Myles, A. Kovalevsky, F. Meilleur, IMAGINE: Neutrons reveal enzyme chemistry. *Acta Crystallogr D Struct. Biol.* **74**, 778–786 (2018).
  75. F. Meilleur, P. Munshi, L. Robertson, A. D. Stoica, L. Crow, A. Kovalevsky, T. Koritsanszky, B. C. Chakoumakos, R. Blessing, D. A. A. Myles, The IMAGINE instrument: First neutron protein structure and new capabilities for neutron macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 2157–2160 (2013).
  76. J. W. Campbell, LAUEGEN, an X-windows-based program for the processing of Laue diffraction data. *J. Appl. Cryst.* **28**, 228–236 (1995).
  77. S. Arzt, J. W. Campbell, M. M. Harding, Q. Hao, J. R. Helliwell, LSCALE—The new normalization, scaling and absorption correction program in the Daresbury Laue software suite. *J. Appl. Cryst.* **32**, 554–562 (1999).
  78. M. S. Weiss, Global indicators of X-ray data quality. *J. Appl. Cryst.* **34**, 130–135 (2001).
  79. G. Winter, R. J. Gildea, N. G. Paterson, J. Beale, M. Gerstel, D. Axford, M. Vollmar, K. E. McAuley, R. L. Owen, R. Flaig, A. W. Ashton, D. R. Hall, How best to use photons. *Acta Crystallogr. D Struct. Biol.* **75**, 242–261 (2019).
  80. P. A. Karplus, K. Diederichs, Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
  81. R. P. Joosten, T. A. H. te Beek, E. Krieger, M. L. Hekkelman, R. W. W. Hooft, R. Schneider, C. Sander, G. Vriend, A series of PDB related databases for everyday needs. *Nucleic Acids Res.* **39**, D411–D419 (2011).
- Human Services, under contract HHSN272201700060C. We thank H. M. O'Neill from ORNL for assistance during expression of the  $P2_1/C2$  constructs. 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 Coronavirus CARES Act. This research used resources at the Spallation Neutron Source and the High Flux Isotope Reactor, which are DOE Office of Science User Facilities operated by the Oak Ridge National Laboratory. The Office of Biological and Environmental Research supported research at ORNL's Center for Structural Molecular Biology (CSMB), a DOE Office of Science User Facility. This research used resources of the Spallation Neutron Source Second Target Station Project at Oak Ridge National Laboratory (ORNL). ORNL is managed by UT-Battelle LLC for DOE's Office of Science, the single largest supporter of basic research in the physical sciences in the United States. The synchrotron x-ray diffraction data used to determine Mac1 structures 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 Science User Facility under contract no. DE-AC02-05CH11231, is supported in part by the ALS-ENABLE program funded by the NIH, National Institute of General Medical Sciences, grant P30 GM124169-01. The synchrotron x-ray diffraction data used to determine Mac1 structures at room temperature were collected at beamline 12-1 of the Stanford Synchrotron Radiation Lightsource (SSRL). Use of the SSRL, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under contract no. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by the NIH, National Institute of General Medical Sciences (P30GM133894). **Funding:** This work was supported by NIH grant R01-GM071939 (to L.C.), NIH grant GM123159 (to J.S.F.), NIH grant GM124149 (to J.M.H.), NSF Rapid grant 2031205 (to J.S.F.), and TMC Award from the UCSF Program for Breakthrough Biomedical Research, funded in part by the Sandler Foundation (to J.S.F.). **Author contributions:** Expressed, purified, and crystallized the  $P4_3$  construct; reduced x-ray diffraction data for  $P4_3$  crystals; refined and analyzed the neutron and x-ray structures; prepared the figures; and wrote the manuscript: G.J.C. Crystallized the  $P2_1/C2$  construct and collected and reduced neutron/x-ray data from the  $P2_1/C2$  construct: D.W.K. Expressed and purified the  $P2_1/C2$  construct: G.P. and S.P. User support and assistance with the data collection strategy at SSRL: S.R. and A.E.C. User support and assistance with the data collection strategy at ALS: G.M. and J.M.H. Edited the manuscript: S.G. Assistance with the room temperature x-ray data collection strategy: M.C.T. Supervised work and arranged for funding: A.A. Collected and reduced neutron/x-ray data for  $P2_1/C2$  construct: L.C. Collected and reduced neutron/x-ray data for the  $P2_1/C2$  construct and wrote the manuscript: A.K. Collected and reduced neutron/x-ray data for the  $P4_3$  construct and wrote the manuscript: F.M. Supervised work, arranged for funding, and wrote the manuscript: J.S.F. **Competing interests:** A.A. is a cofounder of Tango Therapeutics, Azkarra Therapeutics, Ovbio Corporation, and Kyttaro; a consultant for SPARC, Bluestar, ProLynx, Earli, Cura, GenVivo, and GSK; a member of the SAB of Genentech, GLAdiator, and Circle; a member of the board of Cambridge Science Corporation and Cytomx; receives grant/research support from SPARC and AstraZeneca; and holds patents on the use of PARP inhibitors held jointly with AstraZeneca, which he has benefited financially (and may do so in the future). J.S.F. is a consultant for, has equity in, and receives research support from Relay Therapeutics. All other authors declare that they have no competing interests. **Data and materials availability:** All data generated or analyzed during this study are included in the paper and/or the Supplementary Materials. Crystallographic coordinates and structure factors for all structures have been deposited in the PDB with the following accession codes: 7TWF, 7TWG, 7TWH, 7TWI, 7TWN, 7TWO, 7TWP, 7TWQ, 7TWR, 7TWS, 7TWT, 7TWV, 7TWW, 7TWX, 7TWY, 7TX0, 7TX1, 7TX3, 7TX4, and 7TX5.

Submitted 8 February 2022

Accepted 11 April 2022

Published 27 May 2022

10.1126/sciadv.abo5083

**Acknowledgments:** The SARS-CoV-2 Mac1 expression plasmid for the  $P2_1/C2$  crystal form (pMCSG53) was provided by A. Joachimiak (Argonne National Laboratory) with support from the National Institute of Allergy and Infectious Diseases, NIH, Department of Health and

## 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. CorreyDaniel W. KnellerGwyndalyn PhillipsSwati PantSilvia RussiAina E. CohenGeorge MeigsJames M. HoltonStefan GahbauerMichael C. ThompsonAlan AshworthLeighton CoatesAndrey KovalevskyFlora MeilleurJames S. Fraser

*Sci. Adv.*, 8 (21), eabo5083. • DOI: 10.1126/sciadv.abo5083

### View the article online

<https://www.science.org/doi/10.1126/sciadv.abo5083>

### Permissions

<https://www.science.org/help/reprints-and-permissions>