

# Structure-based inhibitor optimization for the Nsp3 Macrodomain of SARS-CoV-2

Stefan Gahbauer<sup>1†</sup>, Galen J. Correy<sup>2†</sup>, Marion Schuller<sup>3</sup>, Matteo P. Ferla<sup>4,5</sup>, Yagmur Umay Doruk<sup>6</sup>, Moira Rachman<sup>1</sup>, Taiasean Wu<sup>7,8</sup>, Morgan Diolaiti<sup>6</sup>, Siyi Wang<sup>8</sup>, R. Jeffrey Neitz<sup>9</sup>, Daren Fearon<sup>10,11</sup>, Dmytro Radchenko<sup>12,13</sup>, Yuri Moroz<sup>13,14</sup>, John J. Irwin<sup>1</sup>, Adam R. Renslo<sup>6,9</sup>, Jenny C. Taylor<sup>4,5</sup>, Jason E. Gestwicki<sup>7,9</sup>, Frank von Delft<sup>10,11,15,16,17</sup>, Alan Ashworth<sup>6</sup>, Ivan Ahel<sup>3</sup>, Brian K. Shoichet<sup>1\*</sup> and James S. Fraser<sup>2\*</sup>

<sup>1</sup>Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA 94158, USA

<sup>2</sup>Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA 94158, USA

<sup>3</sup>Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford, OX1 3RE, UK

<sup>4</sup>Wellcome Centre for Human Genetics, University of Oxford, Old Road Campus, Oxford OX3 7BN, UK

<sup>5</sup>National Institute for Health Research Oxford Biomedical Research Centre, Oxford, OX4 2PG, UK

<sup>6</sup>Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA 94158, USA.

<sup>7</sup>Institute for Neurodegenerative Disease, University of California San Francisco, San Francisco, CA 94158, USA

<sup>8</sup>Chemistry and Chemical Biology Graduate Program, University of California San Francisco, San Francisco, CA 94158, USA

<sup>9</sup>Department of Pharmaceutical Chemistry and Small Molecule Discovery Center, University of California, San Francisco, California 94158, USA

<sup>10</sup>Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK.

<sup>11</sup>Research Complex at Harwell, Harwell Science and Innovation Campus, Didcot OX11 0FA, UK

<sup>12</sup>Enamine Ltd., Chervonotkatska Street 78, Kyiv 02094, Ukraine

<sup>13</sup>Taras Shevchenko National University of Kyiv, Volodymyrska Street 60, Kyiv, 01601, Ukraine

<sup>14</sup>Chemspace, Chervonotkatska Street 78, Kyiv, 02094, Ukraine

<sup>15</sup>Centre for Medicines Discovery, University of Oxford, South Parks Road, Headington, OX3 7DQ, UK.

<sup>16</sup>Structural Genomics Consortium, University of Oxford, Old Road Campus, Roosevelt Drive, Headington OX3 7DQ, UK.

<sup>17</sup>Department of Biochemistry, University of Johannesburg, Auckland Park 2006, South Africa.

\*Corresponding author. Email: [bshoichet@gmail.com](mailto:bshoichet@gmail.com) (B.K.S.); [jfraser@fraserlab.com](mailto:jfraser@fraserlab.com) (J.S.F.)

†These authors contributed equally to this work

## Summary

Computational fragment-linking and ultra-large library docking identifies potent inhibitors of the SARS-CoV-2 macrodomain.

## 1 Abstract

2 The nonstructural protein 3 (NSP3) of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)  
3 contains a conserved macrodomain enzyme (Mac1) that is critical for pathogenesis and lethality. While small  
4 molecule inhibitors of Mac1 have great therapeutic potential, few have been described. Here, we report the  
5 structure-based development of several chemical scaffolds exhibiting low- to sub-micromolar affinity for Mac1  
6 through iterations of computer-aided design, structural characterization by ultra-high resolution X-ray protein  
7 crystallography, and binding evaluation with in-solution assays. Potent scaffolds were designed with *in silico*  
8 linkage of previously obtained fragment hits and ultra-large library docking screens of more than 450 million  
9 molecules. In total, 160 hits comprising 119 different scaffolds were discovered and 152 Mac1-ligand  
10 complex crystal structures were determined, typically to 1 Å resolution or better. The structure-activity-  
11 relationships emerging from this study may template future drug development against Mac1.

## 12 Introduction

13 The macrodomain of SARS-CoV-2 NSP3 (Mac1) presents an intriguing target for drug discovery (1–5). Upon  
14 viral infection, host cells initiate an innate interferon-mediated immune response leading to the expression of  
15 poly-(ADP-ribose)-polymerases (PARPs), which catalyze the antiviral post-translational addition of ADP-  
16 ribose (ADPr) to a large range of target proteins (6). Mac1 enzymatically reverses this mono-ADP-  
17 ribosylation, thereby counteracting immune signaling (7). Promisingly, inactivation of Mac1 by single-point  
18 mutations in the ADPr-binding site demonstrated significant reduction in lethality and pathogenicity in mice  
19 after SARS-CoV infection (8). Small molecule inhibitors of SARS-CoV-2 Mac1 might therefore offer novel  
20 therapeutics to mitigate COVID-19 (9, 10).

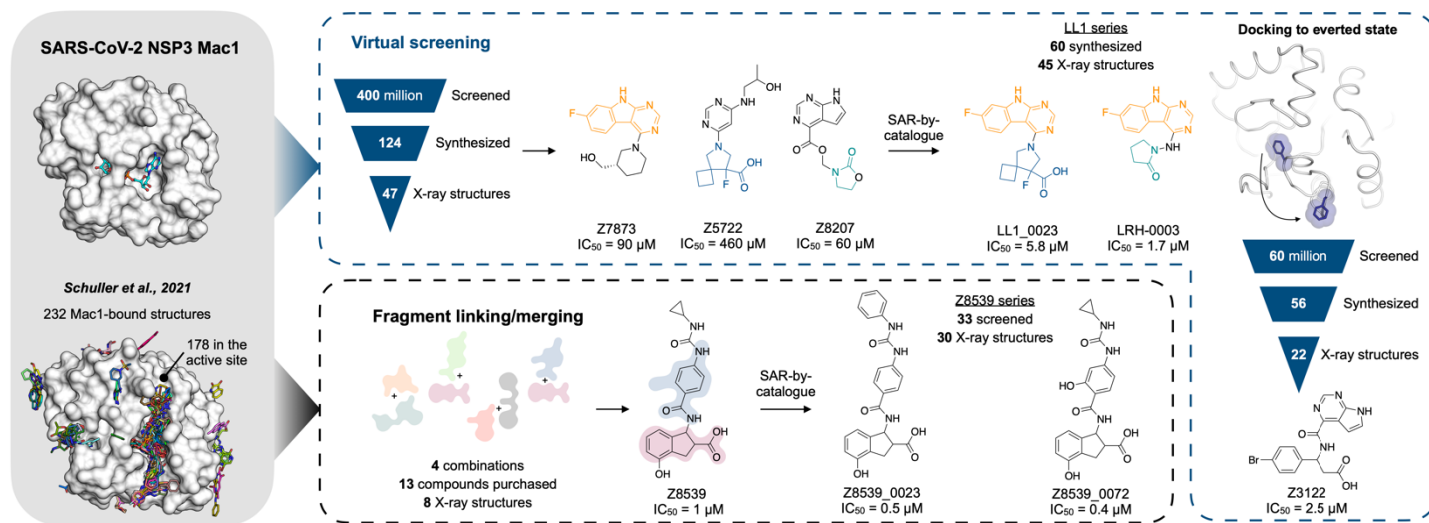
21 A challenge for the development of such inhibitors has been the lack of small molecule modulators of  
22 macrodomain activity, other than ADPr. Only recently have quantitative assays been developed (10, 11).  
23 Accordingly, to map the recognition determinants of Mac1, we adopted a biophysical approach, screening for  
24 fragment ligands using protein crystallography, molecular docking, isothermal titration calorimetry (ITC),  
25 differential scanning fluorimetry (DSF) and a novel binding assay based on homogeneous time-resolved  
26 fluorescence (HTRF) (12). Mac1 proved to be unusually amenable to structure-determination, enabling us to  
27 determine the structures of over 230 fragment complexes, typically to ultra-high resolution (often better than  
28 1.1 Å), affording us a detailed map of enzyme hot-spots with chemical matter of sufficient potency with which  
29 to optimize a quantitative assay (12, 13).

30 Nevertheless, our best fragments remained of modest potency, with none more potent than 180 μM. Here we  
31 describe efforts at optimisation using two strategies (Fig. 1). In the first, we sought to link and merge pairs of  
32 fragments to create larger molecules that exploited multiple hot-spots, so reaching higher affinities. This used  
33 a new fragment-linking method (12, 14), adapted to explore a virtual library of 22 billion readily-synthesizable  
34 molecules (15). In a second approach, we exploited the hot-spots revealed by the initial fragments to guide  
35 computational docking of ultra-large chemical libraries of lead-like molecules, potentially more potent than the  
36 fragments docked in our original study (12). These two approaches led to compounds with IC<sub>50</sub> values of 1–7  
37 μM for the merged fragments and 2.5–500 μM for the docking hits (Fig. 1). Subsequent structure-based  
38 optimization improved IC<sub>50</sub> values to 0.4 μM and 1.7 μM in the fragment-merging and docking campaigns,  
39 respectively. X-ray crystal structures were determined for initial fragment-linking and docking hits and for  
40 molecules optimized for affinity, providing templates for further drug development campaigns against this  
41 promising antiviral target.

## 42 Results

### 43 Hit discovery through fragment-merging

44 The large collection of Mac1-fragment crystal structures revealed interaction patterns between initial ligands  
45 and the Mac1 active site (12). The largest subset of fragments bound in the adenine-recognition subsite,  
46 hydrogen-bonding to Asp22 and Ile23, and stacking with Phe156. Another group of mainly acidic fragments  
47 occupied a sub-pocket formed by the backbone NH groups of Phe156 and Asp157, which we labeled the  
48 “oxyanion subsite”. Although ADPr itself does not directly interact with this oxyanion site, the most potent  
49 compound that emerged from the fragment screen (ZINC263392672, PDB 5RSG, IC<sub>50</sub> = 180 μM) placed a  
50 pyrrolo-pyrimidine group in the adenine subsite and carboxylate in the oxyanion subsite, suggesting that  
51 molecules able to bridge between both subsites hold potential for potent ligand design. An interactive dataset  
52 of the initial hits can be found at <https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mac1>.



**Figure 1.** Overview of the structure-based strategies used to discover ligands that bind to the NSP3 macrodomain of SARS-CoV-2 (Mac1).

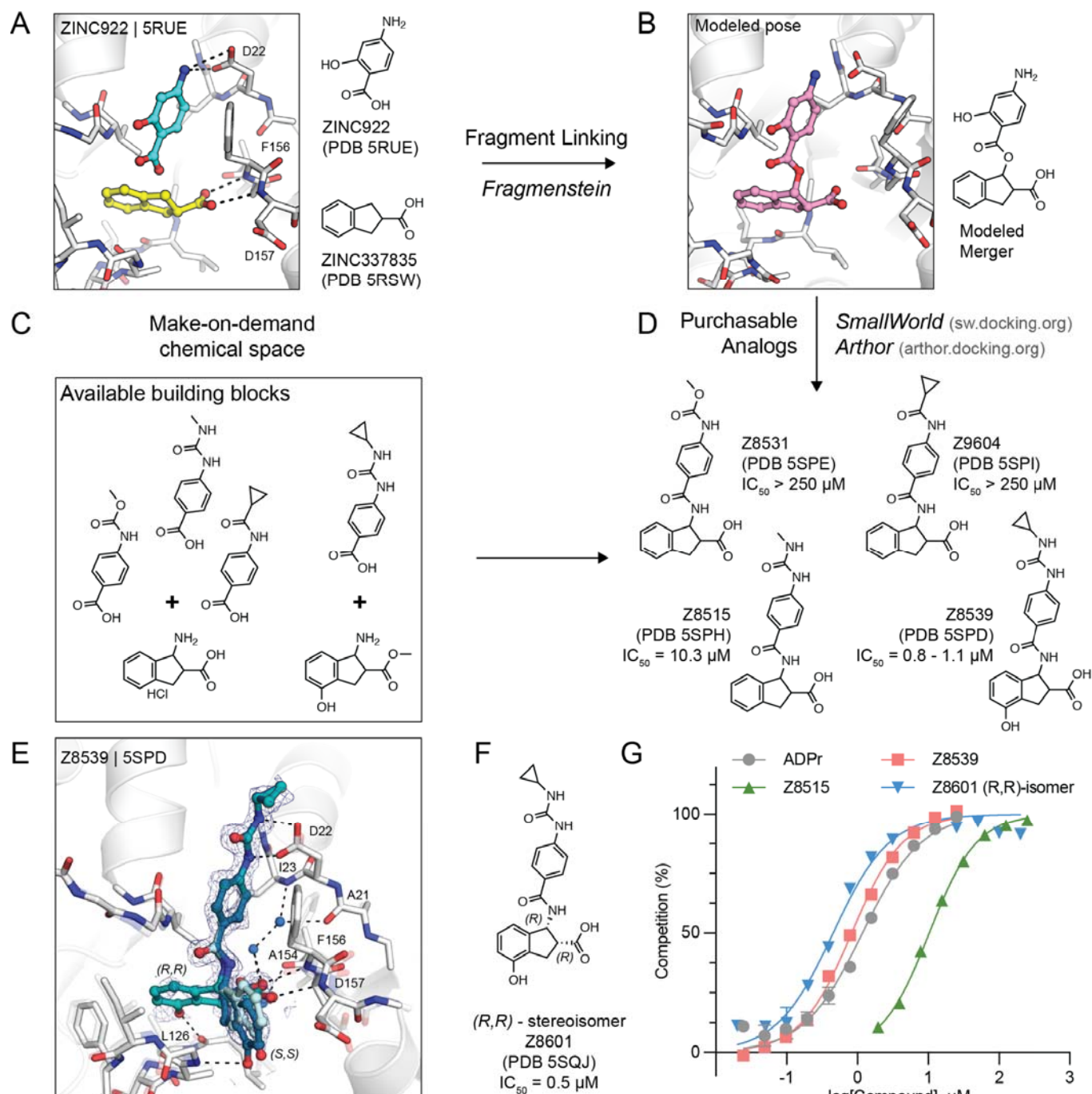
Consequently, we pursued an automated *in silico* fragment-linking or -merging approach, termed *Fragmenstein* (14). From their crystallographic binding poses, fragments were merged based on superposed atoms or linked via hydrocarbon ethers. These virtually merged scaffolds were automatically modeled into the protein binding pocket by ensuring faithful placement of corresponding molecular segments onto the position of the original fragments (Fig. 2A,B). These virtually merged molecules became templates to search the make-on-demand chemical library of the Enamine REAL database, using the 2D molecular similarity search engine SmallWorld (<http://sw.docking.org>) and the substructure browser Arthor (<http://arthor.docking.org>) (16). We pursued four combinations of fragment hits to explore linked or merged scaffolds. Specifically, ZINC337835 (PDB 5RSW) was linked with ZINC922 (PDB 5RUE) (Fig. 2 and Fig. S1) or ZINC98208711 (PDB 5RU5) (Fig. S2), ZINC26180281 (PDB 5RSF) was merged with ZINC89254160\_N3 (PDB 5RSJ) (Fig. S2), and Z44592329 (PDB 5S2F) was merged with ZINC13514509 (PDB 5RTN) (Fig. S2). A total of 16 purchasable analogs (four for each linked or merged scaffold) were prioritized, of which 13 were successfully synthesized by Enamine. In subsequent crystal soaking experiments using the pan-dataset density analysis (PanDDA) algorithm to identify hits (17), 8/13 (~60%) bound to Mac1 (Fig. 2, Fig. S1 and Fig. S2), seven had dose-responsive thermal up-shift of at least 0.5°C in DSF (data file S1), and two molecules had measurable binding to Mac1 in a HTRF-based ADPr-conjugated peptide displacement assay (Fig. 2).

### Identification of promising fragment merger

The linked scaffold combining the fragment hit ZINC922 (PDB 5RUE), occupying the adenine-recognizing subsite, with ZINC337835 (PDB 5RSW), placing a carboxylic acid at the oxyanion subsite, provided a promising template for a molecular scaffold bridging between both subsites (Fig. 2B). While the exact hypothetical merger was not readily available from the make-on-demand chemical space, we found four close analogs that were: Z4718398531 (Z8531), Z4574659604 (Z9604), Z4718398515 (Z8515), and Z4718398539 (Z8539) (data file S1). The main difference between these four accessible scaffolds and the initial merger model was the substitution of the fragment-linking ester by an amide, and the removal of the phenolic function of ZINC922 (Fig. 2D), both of which likely improve the *in vivo* stability of the molecules. The four analogs also differed in the substituents extending from the aniline amine, and Z8539 adds a hydroxyl-group to the indane of the initial fragment hit ZINC337835.

Remarkably, all four analogs were confirmed to bind Mac1 in crystallographic soaking experiments, with high fidelity between the predicted binding pose and the crystallographic result (Fig. 2E and Fig. S1). In the HTRF-based binding assay (12), Z8531 and Z9604 had  $IC_{50}$  values above 250  $\mu M$ , while Z8515 and Z8539 had  $IC_{50}$  values of 7.9  $\mu M$  and 0.8–1.1  $\mu M$ , respectively. The more potent analogs both share a phenylurea group occupying the adenine subsite to stack with Phe156 and form bidentate hydrogen bonds between the urea and Asp22. Z8539 is among the most potent Mac1 compounds described with an affinity comparable to ADPr in the HTRF assay (0.9–1.3  $\mu M$ ) (Fig. 2G). The  $K_D$  of the ADPr-conjugated peptide used in the HTRF assay was determined to be 2.7  $\mu M$  by ITC (data file S1), therefore, the measured  $IC_{50}$  values of the molecules are similar to the binding affinities estimated using the Cheng-Prusoff equation (18).





**Figure 2. *In silico* fragment-linking targeting the adenosine site of Mac1.** **A)** Binding pose of two fragments identified in the previously reported fragment screen (12). Fragment-protein hydrogen bonds are shown with dashed black lines. **B)** Theoretical linked scaffold of ZINC922 and ZINC337835 generated using *Fragemstein* (14). **C)** Availability of corresponding chemical building blocks and reactions in the Enamine REAL database. **D)** Readily accessible analogs of the theoretical scaffold shown in (B). **E)** X-ray crystal structure of Mac1 bound to Z8539. Three conformations of Z8539 [(R,R) and two (S,S)] could be resolved in the PanDDA event map (blue mesh contoured at 2  $\sigma$ ). Water molecules that form bridging hydrogen bonds between Z8539 and the protein are shown as blue spheres. The apo state of Mac1 is shown with transparent white sticks. **F)** 2D structure of the most potent (R,R)-stereoisomer of Z8539 (Z8601). **G)** ADPr-peptide competition (%) of Z8539, Z8515 and Z8601 on Mac1 determined by an HTRF-based displacement assay. ADPr was used as reference. Data are presented as the mean  $\pm$  SEM of at least two technical replicates.

All four molecules possess two chiral centers in the acid-bearing indane group, and initially the compounds were synthesized as diastereomeric mixtures, with evidence for at least two of the four diastereomers observed in the PanDDA event map for Z8539 (Fig. 2A). Chiral separation and testing of Z8539 confirmed that the (R,R) stereoisomer (Z8601), most faithful to the initial fragment hits, had the highest affinity for Mac1 with an  $IC_{50}$  of 0.5  $\mu M$ , i.e. two-fold more potent than the diastereomeric mixture (Fig. 2F). In this

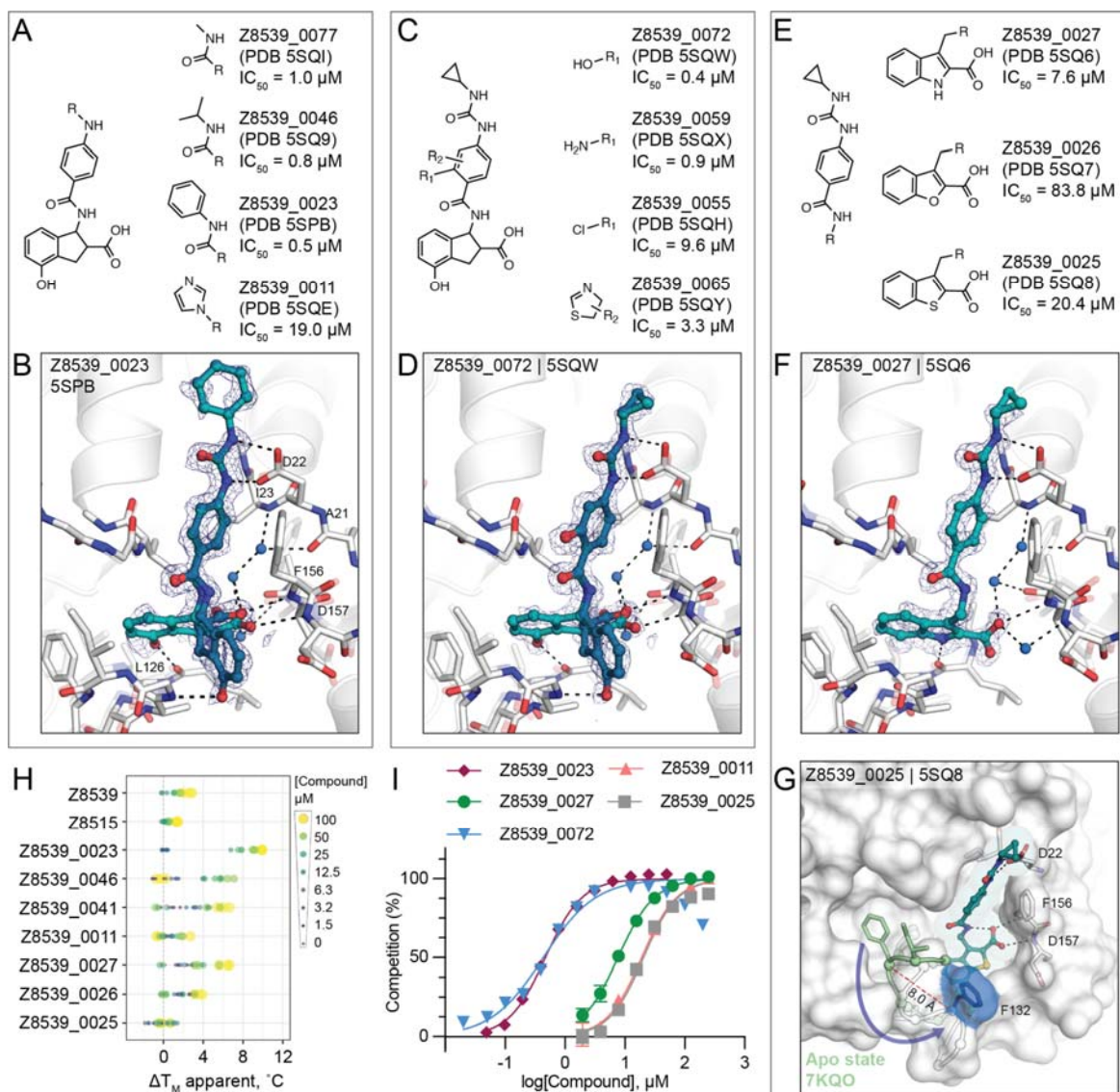
1 configuration, the indane group partially inserts into the phosphate binding domain and the terminal phenol  
2 hydrogen-bonds to the backbone oxygen of Leu126. In the binding pose of the (*S,S*) stereoisomer ( $IC_{50} = 2.9$   
3  $\mu\text{M}$ ), the phenol is mainly solvent exposed and the hydroxyl hydrogen-bonds with the backbone nitrogen of  
4 Gly130 (**Fig. 3A**). By contrast, the two *trans* diastereomers showed reduced affinities with  $IC_{50}$  values  
5 between 43 and 55  $\mu\text{M}$ . The X-ray crystal structure shows that the carboxylic acid of the (*R,S*) isomer only  
6 forms a single hydrogen bond to the oxyanion subsite (**Fig. S3**), while a structure of the (*S,R*) isomer was not  
7 obtained. The (*R,R*) stereoisomer (Z8601) was tested for off-target activity against two human  
8 macrodomains, MacroD2 and TARG1, using an adapted HTRF-based peptide displacement assay. The  
9 human proteins MacroD2 and TARG1 were chosen to test selectivity because they are the most similar  
10 human proteins to SARS-CoV-2 Mac1 (5). Z8601 showed no displacement of the ADPr-conjugated substrate  
11 at 50  $\mu\text{M}$  against either target and approximately 50% displacement at high concentrations of 1 mM (**Fig. S4**).  
12 The selectivity of this scaffold for the viral over the tested human macrodomains is likely related to sequence  
13 differences within the ADPr-binding pockets between all three proteins: while Ala52 in the viral Mac1 offers  
14 ample space to accommodate the compound's phenyl-urea functional group, MacroD2 and TARG1 carry  
15 considerably larger residues at the corresponding position, namely Leu50 and Cys104, respectively (**Fig. S4**).  
16

17 The 1.05 Å resolution crystal structure of Mac1 in complex with the (*R,R*) isomer of Z8539 (Z8601) reveals an  
18 extended water-mediated hydrogen bond network between the ligand's central amide, its carboxylic acid and  
19 Ile23, Ala21 as well as Ala154 (**Fig. 2, Fig. S3**). Interestingly, methylation of Z8539's central amide group  
20 (Z8539\_0056, **Fig. S3**) rendered the compound inactive, likely because of the interruption of this network. It is  
21 uncertain whether the initially generated ester-linked merger (**Fig. 2B**) can form this water network, and our  
22 preference for readily synthesized molecules may have conferred an unexpected advantage over the initial  
23 theoretical merger.  
24

### 25 **Structure-based optimization of the merged scaffold**

26 To further explore the Z8539 scaffold, we generated a structure-activity relationship (SAR) series (**Fig. 3A-E**).  
27 Here too, 2D-based similarity searches of the Enamine REAL database were used to find readily accessible  
28 and SAR-useful analogs, while analogs unavailable in the REAL database were also designed.  
29 Approximately 21,000 analogs (roughly 4,000 mono-anions) were identified via *SmallWorld* and subsequently  
30 docked against the Mac1-Z8539 crystal structure. Visual inspection of top-ranked (mostly) anionic  
31 compounds led to the selection of 19 readily accessible make-on-demand analogs, while nine compounds  
32 were manually designed; of these 28, 26 were successfully synthesized at Enamine. Of these 26 analogs, 23  
33 were confirmed to bind Mac1 by crystallography and 20 showed activity in the HTRF assay (**data file S1**,  
34 **data file S2, data file S3**).  
35

36 Most analogs bore modification of the cyclopropyl-phenylurea group of Z8539 (**Fig. 3A**). Removal of the  
37 cyclopropyl (Z8539\_0041, PDB 5SPA) or replacement by either methyl (Z8539\_0077, PDB 5SQI) or isobutyl  
38 (Z8539\_0046, PDB 5SQ9) did not substantially change binding affinity, however, phenyl replacement  
39 (Z8539\_0023, PDB 5SPB) improved the  $IC_{50}$  to 0.5  $\mu\text{M}$  and showed a significantly increased thermal up-shift  
40 of 9°C in DSF (for the stereoisomeric mixture) (**Fig. 3B,H**). The resulting diphenyl-urea superimposes well  
41 with known fragment hits, e.g. Z44592329 (PDB 5S2F) or Z321318226 (PDB 5S2G) (12) (**Fig. S2**).  
42 Compound Z8539\_0011 ( $IC_{50} = 19 \mu\text{M}$ ) contains an imidazole moiety that forms an additional hydrogen bond  
43 to Lys55 (**data file S3A.5**). Addition of hydrogen bond donors such as amine (Z8539\_0059, PDB 5SQX) or  
44 hydroxyl (Z8539\_0072, PDB 5SQW) at the amide-ortho-position of the central benzene yielded relatively  
45 potent analogs with affinities of 0.9  $\mu\text{M}$  and 0.4  $\mu\text{M}$ , respectively (**Fig. 3C,D**). The corresponding crystal  
46 structures do not reveal additional interactions between the newly introduced substituents and the protein,  
47 however, the binding poses of the ligands indicate the formation of an internal hydrogen bond with the  
48 molecules' central amides (**Fig. 3D**). Furthermore, the hydroxyl of Z8539\_0072 formed a hydrogen bond with  
49 the backbone nitrogen of Lys11 of a symmetry mate, which closely matches the lattice interaction seen in the  
50 initial fragment hit ZINC922 (**Fig. S5**). Z8539\_0072 did not show any off-target activity against either human  
51 TARG1 and MacroD2 at a concentration of 50  $\mu\text{M}$  or 1 mM (**Fig. S4**), indicating selectivity for the viral Mac1  
52 protein.



**Figure 3. Structure-based optimization of Z8539.** **A)** Modifications of the cyclopropyl-phenylurea group. **B)** X-ray crystal structure of Mac1 bound to Z8539\_0023. The PanDDA event map is shown around the ligand (blue mesh contoured at 2  $\sigma$ ). **C)** Modifications of the central benzene. **D)** X-ray crystal structure of Mac1 bound to Z8539\_0072. **E)** Modifications of the indane group. **F)** X-ray crystal structure of Mac1 bound to Z8539\_0027. **G)** X-ray crystal structure of Mac1 bound to Z8539\_0025. The Gly130-Phe132 loop is aligned to the apo-state conformation in green (PDB 7KQO). The Z8539\_0025-Mac1 structure is shown with a transparent white surface. **H)** DSF-derived temperature upshifts. Data are presented for three technical replicates. **I)** HTRF-based peptide displacement dose-response curves. Data are presented as the mean  $\pm$  SEM of at least two technical replicates.

Finally, we tested analogs modulating the acid-carrying indane group (**Fig. 3E**). Of particular interest were achiral analogs where the indane was replaced by benzothiophene (Z8539\_0025, PDB 5SQ8), benzofuran (Z8539\_0026, PDB 5SQ7) or indole (Z8539\_0027, PDB 5SQ6). The indole analog had low micromolar affinity ( $IC_{50} = 7.6 \mu M$ ) for Mac1 and the crystal structure revealed a hydrogen bond between the indole amine and Leu126 (**Fig. 3F**). The lower affinity of this indole versus the parent compound may reflect the sub-optimal placement of the carboxylate in the oxyanion subsite. Surprisingly, although the benzothiophene ( $IC_{50} = 20 \mu M$ ) and furan ( $IC_{50} = 84 \mu M$ ) analogs only differ in one atom compared to the indole analog, the crystal structures in complex with Mac1 indicate that they adopt different poses, with a substantial rearrangement of the protein (**Fig. 3G**). The compounds' cyclopropyl-phenylurea groups are shifted by 2.7 Å compared to the parent Z8539, while the benzothiophene or -furan groups are tilted by roughly 65° relative to the indole group in Z8539\_0027, leaving the phosphate binding region vacant but enabling intramolecular hydrogen bonding between the carboxylic acid and the central amide. The loop formed by residues Ala129 to Pro136 adopts an everted conformation in which Phe132 is displaced by 8 Å and becomes almost fully solvent exposed, indicating high conformational flexibility in the phosphate binding region. Intriguingly, the



1 displaced phenylalanine is reported to be crucial for catalytic function of macrodomains, e.g. mutation of  
2 Phe272 in human MacroD1 reduced enzymatic activity by approximately two-fold (19).

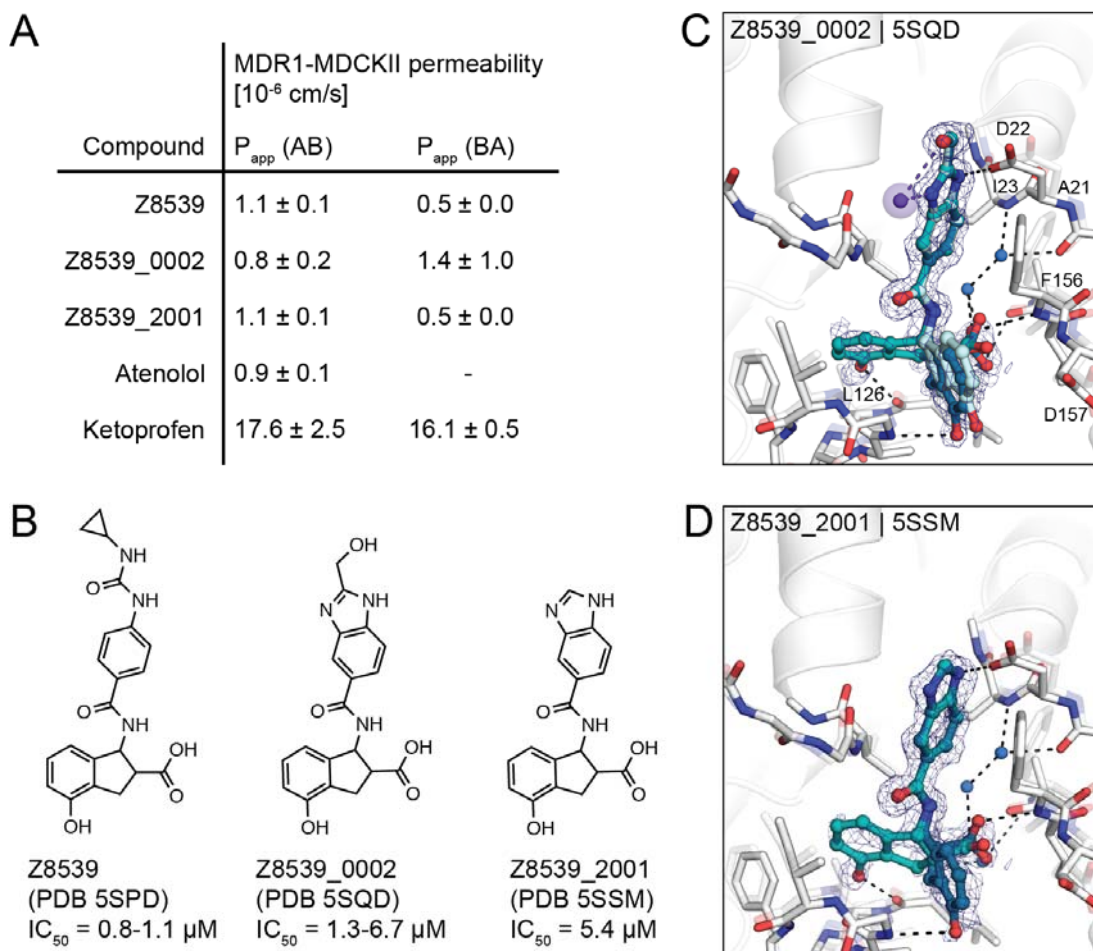
3  
4 Although this compound series led to potent molecules, the Z8539 scaffold had low cell permeability (11  
5 nm/s) in MDCK cells (**Fig. 4A**), which likely limits its potential antiviral activity. As carboxyl bioesters were  
6 not readily available for make-on-demand synthesis, we attempted to increase membrane permeability by  
7 replacing the cyclopropyl-phenylurea with a benzodiazol group, which only marginally reduced the IC<sub>50</sub> value  
8 versus the parent urea (**Fig. 4B**). Z8539\_0002 contains a methanol group that was designed to maintain the  
9 bidentate interaction with Asp22, however, the crystal structure instead indicated a hydrogen bond formed  
10 with the symmetry mate in the crystal lattice (**Fig. 4C**). Removing the alcohol group did not affect the binding  
11 affinity (Z8539\_2001, **Fig. 4D**) and both compound analogs were selective for viral Mac1 over both tested  
12 human macrodomains (**Fig. S4**). However, despite lacking the urea, these compounds had similar P<sub>app</sub>  
13 values compared to Z8539 (**Fig. 4A**), indicating that the carboxylate is most likely responsible for the  
14 observed low cell membrane permeability.

### 15 16 **Hit discovery with computational docking**

17 In addition to fragment-merging or -linking, molecules in lead-like chemical space were virtually screened  
18 against Mac1 using ultra-large library computational docking (20). Molecules were screened against two  
19 different protein models, either using an ADPr-bound structure (PDB 6W02 (21)) or subsequently using a  
20 structure bound to a first-round lead-like docking hit (see below). The first screen of approximately 350 million  
21 molecules of the ZINC15 database (22), belonging mainly to the make-on-demand chemical libraries of  
22 Enamine and WuXi AppTec, with molecular weight ranging from 250 to 350 amu and calculated (c)logP  
23 below 3.5, was performed against the same docking template that we previously used in the computational  
24 fragment screen (ADPr-bound Mac1, PDB 6W02) (12). Molecules were targeted to the adenosine binding  
25 pocket of Mac1; molecules that docked to form polar interactions with the adenine-recognizing residues  
26 Asp22, Ile23 and Phe156, or with residues within the phosphate binding region such as Val49 or Ile132, were  
27 prioritized for experimental testing. Overall, 78 highly ranked molecules were selected for experimental  
28 testing, of which 22 (28%) were confirmed to bind Mac1 in crystallographic soaking screens, 11 (14%)  
29 showed binding in the HTRF assay at concentrations below 1 mM, and 30 (38%) revealed statistically  
30 significant thermal up-shifts of  $\geq 0.5^\circ\text{C}$  in DSF (**data file S1**).

31  
32 In a second docking campaign, scoring parameters were optimized based on the results from the  
33 computational and crystallographic fragment screens as well as the first lead-like docking campaign (23).  
34 Here, the crystal structure of Mac1 in complex with Z6511 (PDB 5SOI, **Fig. 5L**) was used and the docking  
35 parameters were calibrated to ensure higher ranking of 172 previously confirmed fragment hits against a  
36 background of 2,384 molecules (mostly fragments) that did not bind to Mac1 in the crystal soaking  
37 experiments. Compared to the first docking model, this new screen better ranked acidic compounds  
38 interacting with the oxyanion subsite (Methods). Approximately 300 million compounds were docked,  
39 including ca. 250 million neutral and anionic compounds with molecular weights between 250 and 350 amu  
40 and clogP below 3.5 from the ZINC15 library (22), and 50 million compounds from in-house virtual anion  
41 libraries (with molecular weights between 250-400 amu) containing additional, mostly negatively charged  
42 molecules from the Enamine REAL database (15). From among the top-ranking molecules, 46 were obtained  
43 from Enamine, 25 (54%) of which were confirmed to bind Mac1 by X-ray crystallography, five (11%) showed  
44 activity in the HTRF binding assay at concentrations below 250  $\mu\text{M}$  and eight (18%) were classified as hits in  
45 the DSF experiment (Methods).

46  
47 In summary, 124 molecules were selected from virtually screening more than 400 million distinct molecules in  
48 lead-like chemical space, resulting in the identification of 50 Mac1 ligands (40% hit rate) (**Fig. 5**). Of these, 47  
49 were confirmed by crystallographic screening, and 13 showed measurable binding in the HTRF-based  
50 peptide displacement assay with IC<sub>50</sub> values ranging from 42 to 504  $\mu\text{M}$ . Only three molecules that showed  
51 ADPr-peptide competition in the HTRF assay were not confirmed by X-ray crystallography (F6831, Z2051,  
52 Z3271). The seemingly much higher hit-rate in the crystallographic soaking versus the HTRF-based peptide  
53 displacement experiments likely reflects the higher compound concentrations used in crystal soaking (10-20  
54 mM) compared to the highest tested concentration in the HTRF-based assay e.g. 1 mM in the first docking  
55 campaign and 250  $\mu\text{M}$  for the second campaign. Thirty eight compounds showed significant thermal upshifts  
56 of more than  $0.5^\circ\text{C}$  in DSF (**data file S1**), thereby compounds with activity in the HTRF assay often had  
57 upshifts of  $>1^\circ\text{C}$ . Ten compounds were confirmed by all three techniques.

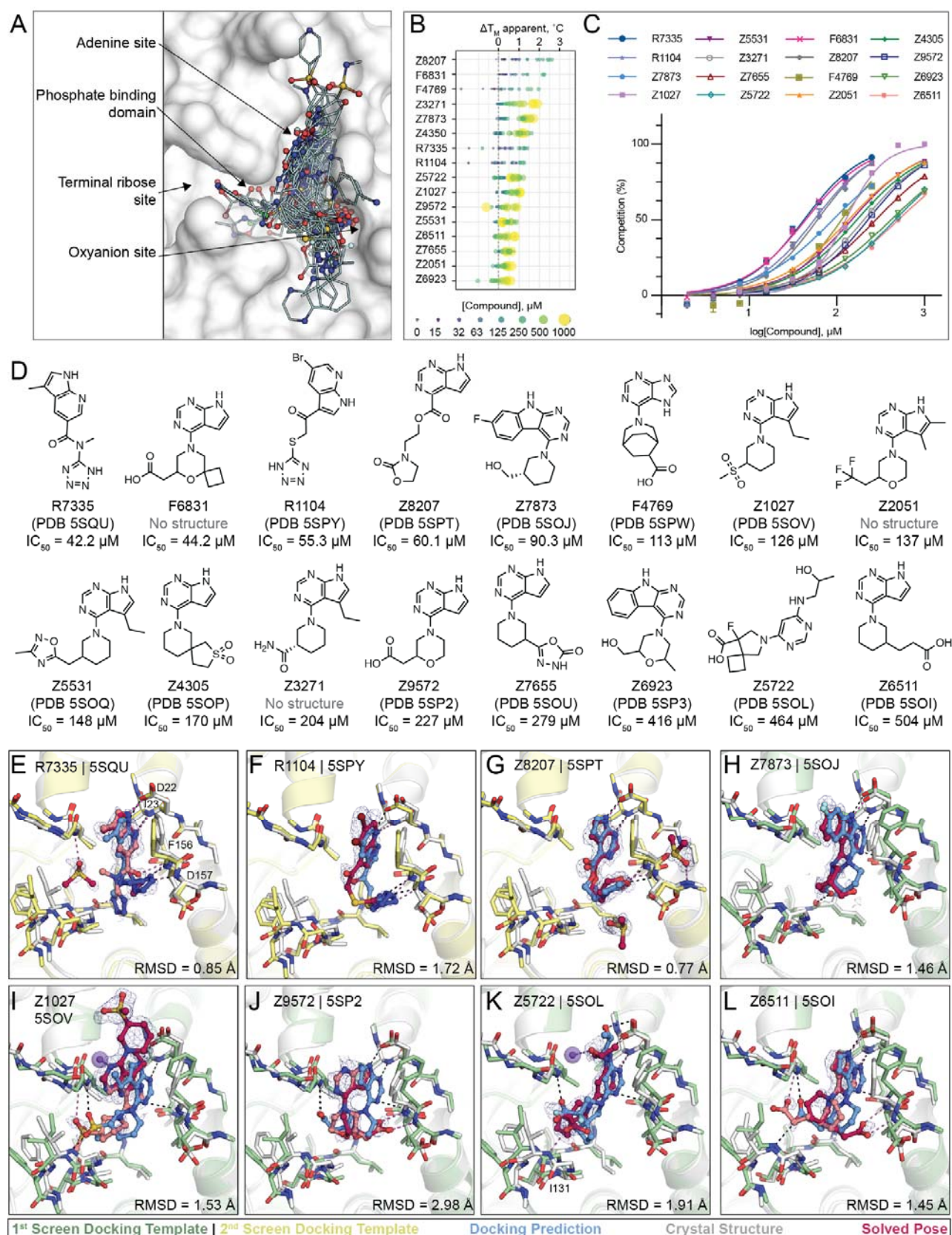


**Figure 4. Z8539 analog with enhanced cell membrane permeability.** **A**) Apparent permeability (P<sub>app</sub>) assayed with MDR1-MDCKII cells. Permeability was measured in apical (A)-to-basolateral (B) direction and vice versa. Atenolol and Ketoprofen were included as control compounds. **B**) 2D structures of Z8539, Z8539\_0002 and Z8539\_2001. **C**) X-ray crystal structure of Mac1 bound to Z8539\_0002. Hydrogen bonding interactions between ligand and the Lys11 backbone nitrogen of a symmetry mate are shown with purple dashes/spheres. PanDDA event maps are shown around the ligand (blue mesh contoured at 2 σ). **D**) Crystal structure of Mac1 bound to Z8539\_2001.

#### Docking hits explore the targeted adenosine binding pocket

Consistent with the docking predictions, almost all of the hits bound to the adenosine binding pocket in the Mac1 active site. A common structural motif among docking hits was a pyrimidine-containing headgroup that interacted with the adenine-recognizing residues of Mac1 (Asp22, Ile23, Ala154). Additional polar or even anionic moieties of docking hits typically bound in either the phosphate subsite or interacted with the oxyanion subsite (**Fig. 5A**). Two compounds, namely F9192 (PDB 5SPO) and Z4273 (PDB 5SPU) did not bind within the active site but occupied a shallow pocket near the terminal ribose binding site (**Fig. S6, data file S4B.29, B.39**). Although we previously identified several fragments binding in this site, they lack high quality interactions and are therefore unlikely to serve as starting points for ligands targeting this site. Good agreement between computationally predicted and crystallographically determined binding poses with Hungarian (symmetry corrected) root mean square deviations (RMSD) below 2 Å (24) was achieved for molecules with measurable binding affinity (e.g. R7335, R1104, Z8207, Z7873, **Fig. 5D-G, data file S1**), whereas larger deviations between docked and experimentally solved binding modes were observed for compounds with binding affinities outside of the tested range. For molecules predicted to place large, often cyclic moieties into the phosphate binding region, the corresponding crystal structures suggested binding modes extending from the adenine subsite to areas outside of the ADPr-binding active site, e.g. Z9710 (PDB 5SOK), Z8186 (PB 5SP1) or Z3280 (PDB 5SON) (see **data file S4**).





**Figure 5. Large scale docking targeting the adenosine site of Mac1.** **A)** Binding poses of 47 docking hits confirmed by X-ray crystallography. The ADPr-bound structure of Mac1 (PDB 6W02) is shown with a white surface. **B)** Thermal upshifts measured by DSF. Data are presented for three technical replicates. **C)** HTRF-based peptide displacement dose-response curves. Data are presented as the mean ± SEM of at least 2 repeat measurements. **D)** 2D structures of docking hits with activity in the HTRF assay. **E-L)** Crystal structures of Mac1 bound to R7335, R1104, Z8207, Z7873,

1  
2  
3  
4  
5  
6  
7

1 Z1027, Z9572, Z5722, Z6511, respectively. The protein structure used in the first docking screen is shown in green, the  
2 structure from the second screen is colored yellow. The predicted binding poses are shown in blue. Protein crystal  
3 structures are shown in gray and the solved binding poses are shown in red, with alternative ligand conformations  
4 colored salmon. Hydrogen bonding interactions between ligands and the Lys11 backbone nitrogen of a symmetry mate  
5 are shown with purple dashes/spheres. Hungarian Root Mean Square deviations (RMSD) between the docked and  
6 solved ligand poses were calculated with DOCK6. PanDDA event maps are shown for each ligand (contoured at  $2\sigma$ ).  
7

8 Although many different headgroups for the adenine subsite were explored among docking hits (see **data file**  
9 **S4**), molecules that were active in the peptide-displacement assay typically shared a pyrrolopyrimidine  
10 scaffold forming hydrogen bonds with Asp22, Ile23 and stacking with Phe156, e.g R7335 (PDB 5SQU),  
11 Z8207 (PDB 5SPT), Z6511 (PDB 5SOI) (see **Fig. 5C**, **Fig. 5D,F,K**). Two compounds, Z7837 (PDB 5SOJ)  
12 (**Fig. 5G**) and Z6923 (PDB 5SP3), extend the bicyclic purine headgroups into tricyclic pyrimidoindole  
13 scaffolds revealing moderate  $IC_{50}$  values of up to 90  $\mu$ M, indicating favorable shape complementarity of larger  
14 segments in the adenine subsite compared to the nucleobase of ADPr. Of note, similar to what we observed  
15 in the fragment screen, four adenine-containing compounds (Z1211, Z4827, Z0893, Z0078) were not  
16 correctly synthesized and showed alkyl derivatives from the N3 rather than the intended N9 nitrogen in their  
17 corresponding crystal structures (see **data file S4**, **data file S1**) (12).  
18

19 Among the most potent molecules were anions placing acidic functional groups such as a carboxylate  
20 (F6831, F4769, Z9572, **Fig. 5C**, **Fig. 5J**) or a tetrazole (R7335, R1104, **Fig. 5E,F**) in the oxyanion subsite.  
21 Interestingly, Z8207 (**Fig. 5G**) places oxazolidin-2-one, a polar but neutral functional group, in the oxyanion  
22 site, and has an  $IC_{50}$  of 60  $\mu$ M. Ketone groups at the oxyanion site offer neutral alternatives to acid functional  
23 groups characteristic of many of the Mac1 inhibitors found to date (below). Two docking hits with measurable  
24  $IC_{50}$  values inserted carboxylates into the phosphate binding region: Z5722 ( $IC_{50} = 464 \mu$ M, **Fig. 5K**) uses a  
25 rigid acid-carrying spiro-octane group to hydrogen bond with Ile131, while Z6511 ( $IC_{50} = 504 \mu$ M, **Fig. 5L**)  
26 projects a flexible butyrate side chain toward the oxyanion site.  
27

### 28 **Ligand-mediated stabilization of alternative protein conformations**

29 Surprisingly, in the crystal structures of three docking hits, namely Z4305 (PDB 5SOP,  $IC_{50} = 170 \mu$ M), F4769  
30 (PDB 5SPW,  $IC_{50} = 113 \mu$ M) and Z5531 (PDB 5SOQ,  $IC_{50} = 148 \mu$ M) the compounds appear to stabilize  
31 alternative, open states of the phosphate binding region, wherein the loop formed by residues Leu127 to  
32 Pro136 adopts an everted conformation relative to the apo structure (**Fig. 6A-C**). Compared to the previously  
33 described structures of Mac1 bound to Z8539\_0025 or Z8539\_0026 (**Fig. 3G**), the docking hits induced even  
34 larger rearrangements within the active site. All three compounds occupy the adenosine subpocket, forming  
35 hydrogen bonds between their pyrrolo-pyrimidine containing groups and Asp22 as well as Ile23. Z4305 and  
36 F4769 interact with the oxyanion subsite via sulfone or carboxylic acid, respectively (**Fig. 6A,B**). Both  
37 compounds stabilize the same loop rearrangement in which the  $C\alpha$  of Phe132 is displaced by 11 Å versus  
38 the canonical closed state, which does not seem able to accommodate the rigid and large non-aromatic cyclic  
39 moieties of the molecules, which would clash with Gly130. Z5331 stabilized a similar everted loop  
40 conformation (**Fig. 6C**). Whereas Z5331 does not interact with the oxyanion subsite, it inserts methyl-  
41 oxadiazole into the phosphate binding region, forming direct and water-mediated hydrogen bonds with  
42 Ser128 and Val49, respectively (**Fig. 6C**). As opposed to Z4305 and F4769, the central piperidine of Z5531  
43 does not clash with Gly130, however, its methyl-oxadiazole would clash with Phe132 in the apo form. A  
44 similar conformational change in the Phe132-containing loop was observed for the merged fragment Z8580  
45 (**Fig. S7**). The observed ligand-induced flexibility within the active site of Mac1 may hint to a catalytic  
46 mechanism requiring conformational flexibility to efficiently bind, cleave and release ADPr from different  
47 target proteins (13, 25).  
48

### 49 **Docking to everted protein conformation**

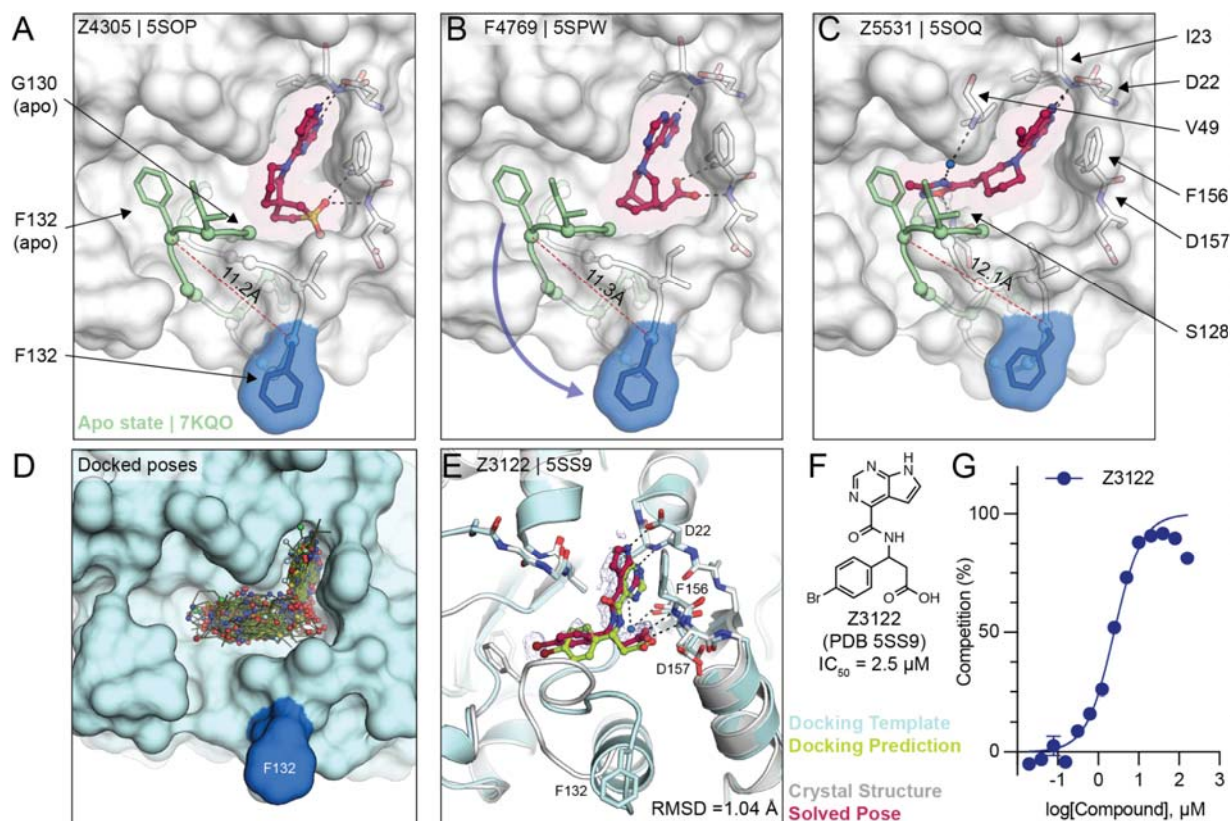
50 To investigate the potential ligandability of the everted Mac1 conformation, we virtually screened roughly 60  
51 million anionic compounds of the ZINC22 virtual library (<https://cartblanche22.docking.org>) against the open  
52 state structure discovered in complex with Z4305. Ligands of this open state are predicted to bind with similar  
53 headgroups in the adenine site as closed state ligands, including polar interactions with Asp22 and Ile23, and  
54 stacking with Phe156. In addition, compared to the closed state, Ser128 was more solvent-exposed and was  
55 therefore targeted by molecules selected from this docking screen. Interactions with these three anchor  
56 points (Asp22, Ile23 and Ser128) were used to select molecules for experimental testing leading to a final set  
57 of 56 molecules that were synthesized by Enamine. On testing, 22 of these (39%) bound to Mac1 in crystal



1 soaking experiments, of which five showed activity in the HTRF-based peptide displacement assay. While  
2 docking generated more favorable scores for the molecules against the open state than the closed state  
3 (**data file S1**), in the crystal structures all 22 hits bound to the closed state (see **data file S4**). Still, among the  
4 five in-solution hits, Z3122 (PDB 5SS9, see **Fig. 6F**) had an  $IC_{50}$  of 2.5  $\mu$ M against Mac1 and had no  
5 measurable activity against the human macrodomains TARG1 or MacroD2 at 160  $\mu$ M (**Fig. S8**), offering yet  
6 another promising, selective scaffold for future optimization.

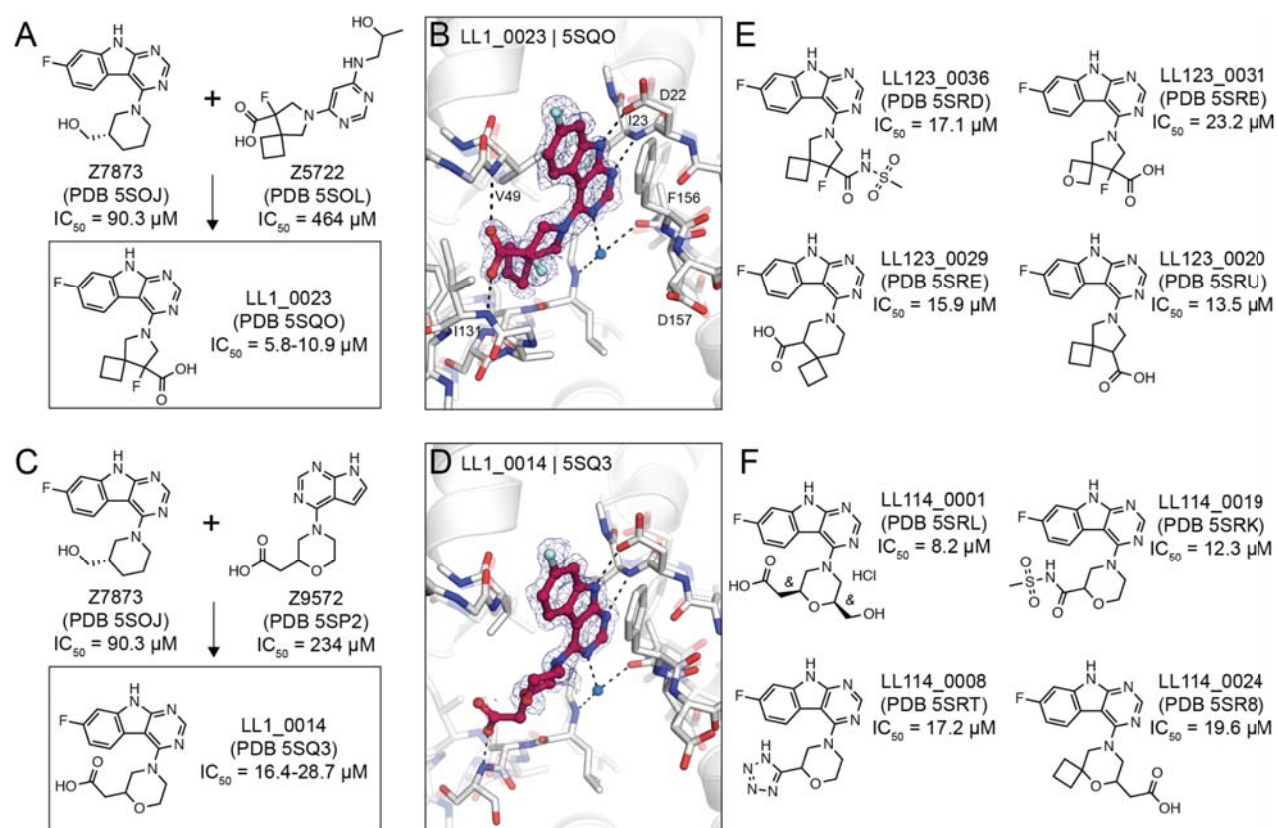
### 7 **Structure-based optimization of docking hits**

8 To improve the affinity of initial docking hits, we explored combinations of molecular substructures bound at  
9 different subsites, templated by their crystal structures. The fluoro-pyrimidindole of Z7873 (PDB 5SOJ),  
10 occupying the adenine-subsite, was introduced into docking hits with mainly bicyclic purine scaffolds (e.g.  
11 Z9572, Z6511, Z5531) or combined with the spiro-octane-carboxylic acid of Z5722 (**Fig. 7A**). Nine analogs  
12 designed with this strategy were accessible in the Enamine REAL database and were synthesized for testing  
13 against Mac1. Of these nine, seven were confirmed to bind crystallographically, five were active in the DSF  
14 assay (see **data file S1**), and four bound in the HTRF assay (**Fig. S8**). Low micromolar affinities were  
15 measured for LL1\_0023 (PDB 5SQO,  $IC_{50}$  = 6-10  $\mu$ M) and LL1\_0014 (PDB 5SQ3,  $IC_{50}$  = 16-29  $\mu$ M), both  
16 containing the pyrimidindole headgroup to occupy the adenine subsite and placing carboxylic acid in the  
17 phosphate binding region (**Fig. 7B,D**). Both compounds showed 50% displacement of the ADPr-conjugated  
18 peptide when tested against TARG1 and MacroD2 at 1 mM, whereas only LL1\_0023 was active against  
19 TARG1 at 50  $\mu$ M (**Fig. S6**). Compared to the Z8539 scaffold, LL1\_0023 was 3-fold more permeable in  
20 MDR1-MDCKII cells.  
21



23  
24  
25 **Figure 6. Stabilization of everted phosphate binding region by docking hits. A,B,C)** The ligand-bound Mac1 crystal  
26 structures are shown in gray with Phe132 highlighted in blue. The Gly130-Phe132 loop of the Mac1 apo structure is  
27 depicted in green. Experimentally determined ligand-binding poses are shown in red. **D)** Predicted binding poses of  
28 molecules docked against the Z4305-bound Mac1 structure (PDB 5SOP). **E)** Crystal structure of Z3122 (red) bound to  
29 Mac1 (gray) compared to the predicted complex (Mac1 in blue, Z3122 in green). The PanDDA event map is shown  
30 around the ligand (blue mesh contoured at 2  $\sigma$ ). The Hungarian RMSD between solved and docked binding poses was  
31 calculated with DOCK6. **F)** Chemical structure of Z3122. **G)** HTRF-derived ADPr-peptide competition curve of Z3122.  
32 Data are presented as the mean  $\pm$  SEM of three technical repeats.





**Figure 7. Structure-based optimization of docking hits. A)** Design of LL1\_0023. **B)** X-ray crystal structure of LL1\_0023. The PanDDA event map is shown around the ligand (contoured at  $2 \sigma$ ). Hydrogen bonds are shown with dashed black lines. **C, D)** Design and X-ray crystal structure of LL1\_0014, respectively. **E, F)** Selected analogs of LL1\_0023 and LL1\_0014, respectively.

Thirteen analogs of the LL1\_0023 scaffold were selected and synthesized from the Enamine chemical space to investigate structure-activity-relationship for this scaffold. Eleven of these bound in crystal soaking experiments (**data file S5**), while nine analogs had  $IC_{50}$  values below  $200 \mu M$  in the ADPr-peptide displacement assay (**Fig. S6**). No improvement of affinity was achieved by replacing the carboxylic acid of LL1\_0023 by sulfonamide (LL123\_0036, PDB 5SRD,  $IC_{50} = 17 \mu M$ , **Fig. 7E**), or replacing the cyclobutane with oxetane (LL123\_0031, PDB 5SRB,  $IC_{50} = 23 \mu M$ ). In addition, modifications of the compound's core spiro-octane e.g. replacement by spiro-nonane (LL123\_0029, PDB 5SRE,  $IC_{50} = 16 \mu M$ ) or removal of a fluoro group (LL123\_0020, PDB 5SRU,  $IC_{50} = 14 \mu M$ , **Fig. 7E**) did not change affinity notably. Correspondingly, removal or neutralization of the acidic functional group by methylation increased  $IC_{50}$  values to over  $200 \mu M$  (**data file S5**).

For the LL1\_0014 scaffold, 18 analogs were designed and synthesized by Enamine, 16 of which bound to Mac1 in the soaking or HTRF-based binding experiments. Here, addition of an ethanolic group to the central morpholino group, reflecting the initial docking hit Z7873 (**Fig. 7C**), improved the  $IC_{50}$  value to  $8 \mu M$  (LL114\_0001, PDB 5SRL, **Fig. 7F**). Similarly, the addition of cyclobutane to the morpholino group, which mimicked the docking hit F6831 (**Fig. 5C**), showed slight improvement of affinity (LL114\_0024, PDB 5SR8,  $IC_{50} = 20 \mu M$ ). Furthermore, exchanging the carboxylic acid by bioisosteres such as sulfonamide (LL114\_0019, PDB 5SRK,  $IC_{50} = 12 \mu M$ ) or tetrazole (LL114\_0008, PDB 5SRT,  $IC_{50} = 17 \mu M$ ) seemed to moderately improve the ligands' binding affinities (**Fig. 7F**). In subsequent screens against TARG1 and MacroD2, only the tetrazole-containing analog (LL114\_0008) showed measurable peptide displacement against TARG1 and MacroD2 at  $50 \mu M$  (**Fig. S8**). Additional analogs are shown in the Supporting Information (see **data file S5**).

### Towards potent neutral Mac1 inhibitors

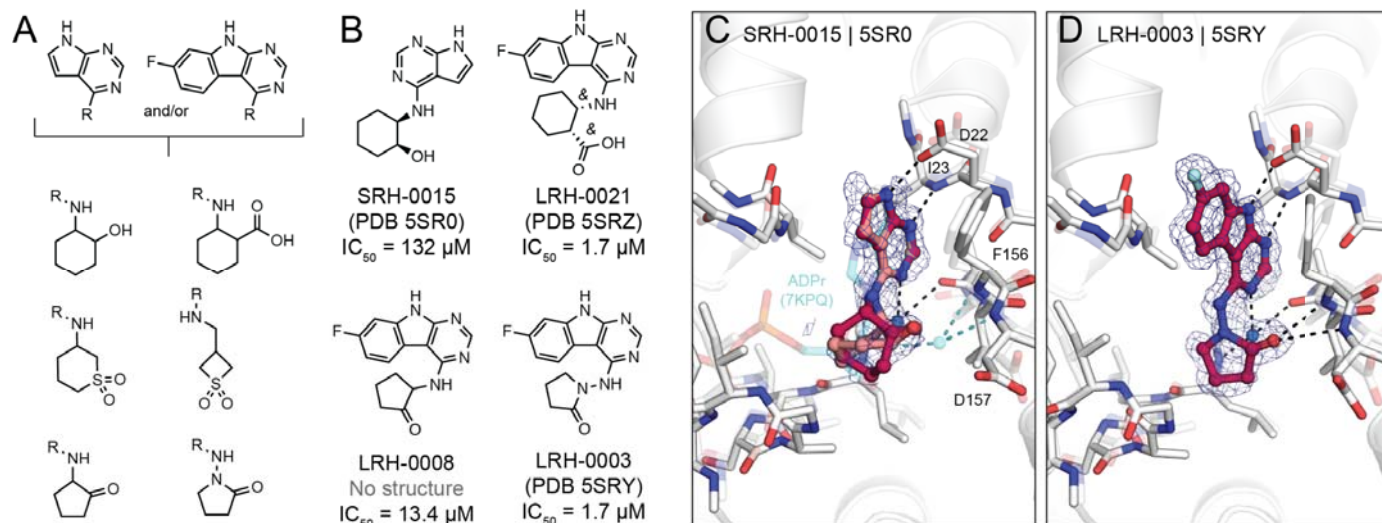
Although our initial SAR for Mac1 ligands showed the benefit of carboxylate binding to the oxyanion subsite, ADPr instead interacts with this subsite via a water-mediated hydrogen bond to a neutral ribose hydroxyl. The development of non-anionic inhibitors might hold several advantages for antiviral drug discovery, especially

1 considering drugs will need to cross cell membranes to engage viral targets residing within infected host  
2 cells. To identify neutral alternatives to carboxylate and other anions at this site, we designed a small set of  
3 analogs by linking the previously identified pyrrolo-pyrimidine or pyrimidindole to small moieties bearing  
4 hydrogen bond donor or acceptor functionality (e.g., sulfones, hydroxyls, pyridines, or ketones) (see **Fig. 8A**,  
5 **data file S6**). A total of 124 molecules (290 enantiomers) were generated in 3D conformer libraries for  
6 computational docking (see Methods). We selected 21 compounds based on the predicted docking poses of  
7 which 20 were synthesized by Enamine. Fourteen of these 20 molecules (70%) were confirmed to bind to  
8 Mac1 by X-ray crystallography and four (20%) showed binding in the HTRF-based assay.

9  
10 Promisingly, SRH-0015 (PDB 5SR0,  $IC_{50} = 132 \mu M$ , MW = 232 amu), notably active for its small size, placed  
11 a hydroxyl group towards the oxyanion subsite, mimicking the placement of a ribose-hydroxyl group of ADPr  
12 (see **Fig. 8B,C**). While the crystal structure of Mac1 in complex with ADPr revealed a water-mediated  
13 hydrogen bond between the corresponding ADPr-hydroxyl and the oxyanion site, the structure of the Mac1-  
14 SRH-0015 complex does not suggest direct or water-mediated hydrogen bonding (**Fig. 8C**). The most  
15 promising analogs from this series were LRH-0008 ( $IC_{50} = 13.4 \mu M$ ) and LRH-0003 (PDB 5SRY,  $IC_{50} = 1.7 \mu M$ )  
16 (see **Fig. 8B,D**). These compounds contain fluoro-pyrimidindole headgroups joined to 2-  
17 aminocyclopentan-1-one or 1-aminopyrrolidin-2-one rings, respectively. The crystal structure of LRH-0003  
18 bound to Mac1 revealed favorable placement of its hydrazide carbonyl function at the oxyanion site, enabling  
19 simultaneous hydrogen bonding to both NH groups of Phe156 as well as Asp157 (**Fig. 8D**). The enhanced  
20 potency of LRH-0003 versus LRH-0008 is consistent with a stronger hydrogen bonding interaction in the  
21 former, given the greater basicity of the hydrazide carbonyl present in LRH-0003 as compared to the ketone  
22 in LRH-0008. Notably, the similar anionic analog LRH-0021 (PDB 5SRZ, **data file S6**) was equipotent to  
23 LRH-0003 indicating that neutral compounds can indeed offer competitive alternatives to anionic Mac1  
24 ligands. Although the anionic compound LRH-0021 showed binding to TARG1 at 160  $\mu M$ , neither neutral  
25 compounds LRH-0003 and LRH-0008 had measurable binding to TARG1 and MacroD2 (**Fig. S8**), suggesting  
26 higher selectivity for the neutral isosteres.

## 27 Discussion

28  
29 Like many targets to emerge from SARS-CoV-2, Mac1 is both highly attractive and challenging. While animal  
30 studies in SARS have highlighted its crucial role in viral pathogenesis, there were no reliable chemical tools,  
31 or really inhibitors of any kind, for the enzyme. Fortunately, Mac1 crystallized readily and diffracted to ultra-  
32 high resolution (often better than 1 Å), supporting fragment-based exploration of its recognition determinants,  
33 both empirically and computationally (12). Capitalizing on this, over 230 fragment structures were  
34 determined. The binding poses of the ligands tiled the active site of the enzyme, but despite often favorable  
35 ligand efficiencies, none of the fragments had affinities more potent than 180  $\mu M$ . Here, we built on the  
36 molecular determinants revealed by the fragment structures to discover more potent molecules, making  
37 progress towards chemical probes and leads for drug development.



39  
40  
41 **Figure 8. Probing neutral functional groups in the Mac1 oxyanion subsite. A)** Design strategy of analog set. **B)**  
42 **Chemical structures of most potent hits. C)** Crystal structure of Mac1 bound to SRH-0015. ADPr and the water-mediated

1 hydrogen bond to the oxyanion subsite are shown for reference (PDB 7KQP, transparent cyan sticks/spheres). Both of  
2 the *trans* stereoisomers were modeled: the (*S,R*) is colored dark red and the (*R,S*) isomer is colored salmon. PanDDA  
3 event maps are shown around the ligand (blue mesh contoured at 2  $\sigma$ ). **D**) Crystal structure of the Mac1-LRH-0003  
4 complex.

5 Three key points emerge from this effort. First, an automated fragment merging and linking strategy, allied  
6 with searches of ultra-large libraries, identified molecules that combined key groups of pairs of fragments *and*  
7 were readily available from make-on-demand synthesis. These efforts led to the rapid discovery of molecules  
8 with low  $\mu\text{M}$  affinity that were subsequently optimized to affinities as low as 430 nM (compound Z8539\_0072),  
9 an overall improvement of >400-fold compared to the best starting fragment. Second, templated again by the  
10 ligand-recognition patterns revealed by the fragments, molecular docking screens found compounds with  
11 affinities down to 2.5  $\mu\text{M}$ , with several in the mid- $\mu\text{M}$  range that were also optimizable to the low  $\mu\text{M}$ . The best  
12 of these had ligand efficiencies that were measurably better than even the merged fragments. Third, while  
13 most of these molecules were anionic with high polar surface areas that reduced cell permeability, structure-  
14 based optimization found analogs with fewer hydrogen-donating groups like ureas, alcohols and phenols, and  
15 enabled the replacement of anionic warheads with neutral ones. This suggests that it may be possible to  
16 improve cell membrane permeability for several of the scaffold classes here.

17  
18 We used X-ray crystallography both as a primary screening tool to identify macrodomain-binding compounds  
19 from computational design, and to provide structural information to guide compound optimization. The  
20 success of this approach was partly due to the high-quality nature of the Mac1 crystals in the  $P4_3$  space  
21 group; they grew readily, withstood high concentrations of DMSO and diffracted consistently to <1 Å. The  
22 high resolution diffraction, coupled with analysis of electron density with PanDDA (17), allowed us to identify  
23 fragments with occupancies below 20% in the initial fragment screen (12). Low occupancy fragments  
24 included ZINC337835 and ZINC922, which were linked together in the present work to generate Z8539, a  
25 potent binder of Mac1 (**Fig. 2**), testifying to the potential of this approach. Although the initial fragments were  
26 soaked at high concentrations (10 mM), only hints of fragment binding were visible in  $F_o - F_c$  difference maps,  
27 and the fragment binding signal was largely obscured by ground-state solvent (**Fig. S9**). However, both  
28 ZINC337835 and ZINC922 could only be modeled unambiguously into PanDDA event maps (**Fig. S9**). This  
29 contradicts recent arguments that no useful conclusions can be derived from ligands modeled at the low  
30 occupancies detected by PanDDA (26). Our work, and that of others (27, 28), shows how low-occupancy  
31 ligands can inspire the design of more potent analogs. In addition to identifying the fragments that led to  
32 Z8539, PanDDA helped to identify the most potent stereoisomer of Z8539. We initially obtained this  
33 compound as a mixture of diastereomers, and although the density indicated that the major isomer was (*S,S*),  
34 inspection of the PanDDA event map at low contour level hinted that the (*R,R*) isomer might be present (**Fig.**  
35 **S9**). This prompted us to test the four diastereomers separately, which revealed that the (*R,R*) isomer was  
36 the most potent in solution, with good agreement between the fragments modeled using PanDDA and the  
37 theoretical model (**Fig. S1**).

38  
39 One notable complication to using X-ray crystallography to screen ligands is the influence of crystal lattice  
40 interactions on ligand binding (29). Our initial fragment screen revealed that the  $P4_3$  crystal form had a  
41 substantially higher hit rate compared to the  $C2$  crystal form (24% versus 6%) (12). We partly attributed the  
42 difference in hit rates to fortuitous crystal packing in the  $P4_3$  form: the backbone nitrogen of Lys11 on a  
43 symmetry mate is ideally positioned to interact with compounds binding in the adenine subsite. Indeed, 66 of  
44 the 123 fragments identified in or near the adenine subsite formed hydrogen bonds with Lys11 (12). Similarly,  
45 in the present work, several of the compounds that were identified by virtual screening, and subsequent  
46 optimization, adopted alternative conformations that were stabilized by hydrogen bonds with Lys11 (e.g.  
47 Z1027, Z9020, LL123\_0006 and LL123\_0016.). Although one might be tempted to discard these  
48 conformations as artifacts, our current work indicates that they can be useful. One of the two fragments that  
49 were linked to create Z8539 contained a hydroxyl that formed a hydrogen bond with Lys11 (**Fig. S5**). The  
50 compound lacking the hydroxyl (4-aminobenzoic acid, ZINC920) did not bind to Mac1 in the fragment screen  
51 (12). Crystal lattice interaction may explain the large difference between predicted and observed binding mode  
52 for several of the hits from virtual screening (e.g. F9046, F0346, R3575, Z6744, Z6684, Z5740, Z6689,  
53 Z6567).

54  
55 We were surprised to find several ligands that induced large scale re-arrangement of the active site loop  
56 consisting of residues 127-136 (**Fig. 6, Fig. S7**). Conformational changes involving Ala129, Phe132 and  
57 Asn99 have been characterized in this loop in the ADPr-bound state (12) and in the ligand-free enzyme at



1 low pH (13), but these are relatively minor compared to the 7-12 Å shifts in Phe132 seen here. Everted loop  
2 conformations have also been observed for other macrodomains, including human MacroD1 (PDB 2X47) (30)  
3 and PARP14 (PDB 5O2D) (31) (**Fig. S7**). Despite the apparent flexibility of this region, our initial virtual  
4 screening campaign did not identify any compounds that stabilized the flipped conformation of Ala129 that is  
5 present in the ADPr-bound state, despite using this state as a template for docking (PDB 6W02) (**Fig. 5**).  
6 However, during compound optimization, several structures were determined with Ala129 in the flipped state.  
7 These included LL114\_0041, which places a carboxylic acid in the phosphate binding subsite, and  
8 LL123\_0020, which stabilizes a water molecule in a similar position (**Fig. S6**). A similar rearrangement in  
9 water networks was seen for the docking hit Z0828, although the shift in Ala129 was smaller (**Fig. S6**). These  
10 ligands offer new opportunities for structure-guided design efforts targeting the phosphate binding subsite of  
11 Mac1.

12  
13 Certain caveats merit discussion. The anti-viral or immunomodulating effect of the developed compounds has  
14 not been shown. This partly reflects limitations of the molecules themselves—e.g., their current low cell  
15 permeability—but it also reflects the lack of suitable cell-based assays to monitor the effect of Mac1 inhibition  
16 on interferon signaling. The development of such assays is an urgent need in the field; currently, our only way  
17 to measure the efficacy of Mac1 inhibitors, outside of the enzyme itself, is *in vivo*. On a technical level, while  
18 hit rates of computational docking were high in the X-ray soaking assay, only a few truly potent compounds  
19 were identified in the HTRF-based binding assay. Furthermore, while many docking predicted poses  
20 corresponded well to the crystallographically determined poses, compared to the previous fragment docking  
21 screen, larger deviations between docked and crystallographic poses were sometimes observed, especially  
22 among molecules that were predicted to insert deep into the phosphate-binding pocket. Also, ligand-induced  
23 stabilization of alternate conformations of the mobile active site loop was not predicted. While docking against  
24 a Mac1 structure with the everted Phe132 loop conformation (PDB 5SOP) led to a potent 2.5 μM inhibitor  
25 (Z3122), the Mac1-Z3122 crystal structure showed binding in the closed state (**Fig. 6E**). In addition to  
26 shortcomings of computational docking, our fragment-linking strategy relied on the access to chemicals  
27 mimicking theoretically linked scaffolds. In our case, the purchasable analogs offered promising templates,  
28 however, some differed noticeably from the initial model e.g. they replaced a central hydrogen bond acceptor  
29 (ester group) with a donor (amide group). Although this exchange seemed actually beneficial in our Z8539-  
30 series, similar changes might lead to loss of activity in other cases.

31  
32 These caveats do not affect the central observations of this study. From an initial mapping of the Mac1  
33 binding site with >230 fragment crystal structures (12), fragment-linking and -merging led to compounds that  
34 bound >400-fold better than the best fragment. The same mapping identified hot spots that supported ultra-  
35 large library docking that identified mid- and low μM binders falling into still newer families. Overall, the  
36 determination of 150 new Mac1-ligand crystal structures supported the discovery and optimization of 19 low-  
37 and sub-μM compounds falling into eight different scaffolds and chemotypes, while another 28 compounds in  
38 eleven scaffolds were discovered in the 10 to 50 μM range. While these compounds retain permeability  
39 liabilities, structure-based optimization suggests routes to improving their physical properties, including by  
40 reducing hydrogen-bond donors and swapping anionic for neutral warheads, without substantial loss of  
41 affinity for the enzyme. From a technical standpoint, the rich of structure-activity-relationships combined with  
42 X-ray crystal structures for most compounds described here creates a dataset for benchmarking and  
43 improving computational techniques for drug discovery, such as free energy perturbation (32, 33). From a  
44 therapeutic perspective, the compounds and structures described in this study will support progress towards  
45 first-in-class antiviral therapeutics targeting the NSP3 macrodomain of SARS-CoV-2.

46  
47  
48

## 1 **Methods**

### 2 **Fragment merging/linking**

3 Fragment mergers and linkers were generated using *Fragmenstein* (14). Specifically, spatially superposed  
4 atoms or rings are combined, while attempting to maintain bonding, and separate fragments are linked,  
5 depending on distance, via a bond, oxygen bridge or hydrocarbon ether bridge. The resulting compounds are  
6 corrected for any defects, such as impossible valence, and minimized under strong constraints using  
7 PyRosetta. The merging and the search for purchasable similar compounds was performed similarly to the  
8 example Colab notebook for *Fragmenstein* (14). The structure PDB 6WOJ (7) was chosen as a template  
9 structure and was energy minimized with 15 cycles of FastRelax in PyRosetta restrained against the electron  
10 density map and with ADPr parameterised. The initial fragments were processed and merged pairwise. The  
11 mergers that were predicted with a combined RMSD less than 1 Å were sorted by Rosetta-predicted binding  
12 Gibbs free energy and the top mergers were manually inspected. The SmallWorld server was queried for  
13 purchasable compounds similar to the top merged compounds (16), which were then placed restrained to the  
14 initial fragments.

### 15 **Computational docking**

16 Docking calculations were performed with DOCK3.7 (23, 34) using precomputed scoring grids for rapid  
17 evaluation of docked molecules. Scoring grids for van der Waals interactions were generated with  
18 CHEMGRID and electrostatic potentials within the targeted binding pocket were calculated by numerical  
19 solution of the Poisson-Boltzmann equation with QNIFFT (35). Therefore, AMBER united-atom charges (36)  
20 were assigned to the protein and selected structural water molecules. Ligand desolvation scoring grids were  
21 computed using Solvmap (37).

22  
23  
24 In the first docking screen, the crystal structure of SARS-CoV-2 NSP3 Mac1 bound to ADPr (PDB 6W02 (21))  
25 was used as a template for docking. All water molecules except for HOH324, HOH344, HOH383 and  
26 HOH406 as well as chain B were removed. Next, the Mac1-ADPr complex with selected water molecules was  
27 prepared for docking following the protein prepwizard protocol of Maestro (Schrödinger v. 2019-3) (38).  
28 Accordingly, Epik was used to add protons and protonation states were optimized with PROPKA at pH 7 (39).  
29 The complex was energetically minimized using the OPLS3e force field. Thereby, the maximum heavy-atom  
30 root-mean-square deviation from the initial crystal structure was 0.3 Å. The atomic coordinates of the  
31 adenosine substructure within the co-crystallized ADPr molecule were used to generate 45 matching spheres  
32 for placement of ligand atoms by the docking program (23). For the calculation of the binding pocket  
33 electrostatic potential, the dielectric boundary between the low dielectric protein environment and high  
34 dielectric solvent was moved outwards from the protein surface by 1.9 Å using spheres generated by  
35 Sphgen. In addition, partial atomic charges of backbone amide hydrogen atoms of residues Ile23 and Phe156  
36 were increased by 0.2 elementary charge units (e) while partial charges of the corresponding backbone  
37 carbonyl oxygen atoms were reduced by the same amount, hence, retaining the residues' net charges.  
38 Furthermore, the dielectric boundary was extended by 0.4 Å from the protein surface for the generation of  
39 ligand desolvation scoring grids (23). At the time we launched the first lead-like docking screen against Mac1,  
40 ADPr was the only known ligand of the enzyme. Consequently, we calibrated the docking parameters  
41 according to their ability to place and score adenosine, adenine and ribose within the adenosine-binding site  
42 against a background of 250 property-matched decoys generated with the DUDE-Z approach (40). In  
43 addition, an Extrema set was screened to ensure prioritization of mono-anions and neutral molecules (23).

44  
45 A total of 330,324,265 molecules with molecular weights ranging from 250 to 350 amu and calculated (c)logP  
46 below 3.5 from the ZINC15 lead-like library were screened (22). In total, 316,505,043 compounds were  
47 successfully scored, each exploring on average 3,111 orientations and 405 conformations leading to the  
48 evaluation of roughly 175 trillion complexes in 65,794 core hours or roughly 66 hours on a 1000-core cluster.  
49 The predicted poses of the top-scored 500,000 molecules were filtered for internal molecular strain (total  
50 strain <6.5 TEU; maximum single torsion strain <1.8 TEU (41)) and their ability to form hydrogen bonds to  
51 residues Asp22, Ile23, Gly48, Val49, Gly130 or Phe156. Molecules with unsatisfied hydrogen bond donors or  
52 more than three unsatisfied acceptors were deprioritized. Finally, 90 molecules were purchased from  
53 Enamine, of which 78 (87%) were successfully synthesized.

54  
55 For the second docking campaign, the crystal structure of Mac1 in complex with the first-round docking hit  
56 ZINC000078036511 (Z6511, PDB 5SOI) was used as the structural template. Chain B and all water residues  
57 were removed and the Z6511-Mac1 complex (using conformation B of the ligand) was prepared according to

1 the protein prepwizdard protocol using Maestro (see above) (38). During scoring grid preparation, the low  
2 dielectric protein environment was extended by 1.8 Å outwards from the protein surface. In addition, the  
3 partial atomic charge of the backbone amide hydrogen atom of Ile23 was increased by 0.4 e whereas the  
4 partial charges of the backbone amide hydrogen atoms of Phe156 and Asp157 were increased by 0.2 e  
5 without modulating the net charge of the residues. Forty five matching spheres for ligand placement by  
6 docking were generated based on atomic coordinates obtained from various first-round lead-like docking hits  
7 as well as previously described fragments: ZINC000078036511 (PDB 5SOI), ZINC000292637864 (PDB  
8 5SOT), ZINC901381520 (PDB 5S6W), ZINC57162 (PDB 5RV3), ZINC26180281 (PDB 5RSF) and  
9 ZINC336438345 (PDB 5RSE) (12). The described docking parameters were evaluated by control calculations  
10 ensuring the enrichment of 142 previously identified fragment ligands and 24 first round lead-like docking hits  
11 against a background of 2,384 experimentally determined non-binders (2,333 fragments, 51 lead-like  
12 molecules).

13  
14 Using the ZINC15 database, 246,246,485 neutral and monoanionic molecules from the lead-like set were  
15 docked against this Mac1 model, resulting in the scoring of 156 trillion complexes where each scored  
16 molecule was on average sampled in 3,431 orientations and 428 conformations within 63 hours on a 1000-  
17 core computer cluster. In addition, an in-house anion library containing (mostly) negatively charged molecules  
18 with molecular weight between 250 and 400 amu from the 22B Enamine REAL database was screened. In  
19 total, 39 million anions were identified by performing SMART pattern searches in RDKit ([www.rdkit.org](http://www.rdkit.org)) of  
20 carboxylic acid and 33 bioisosteres. In the docking screen, 37,556,136 molecules were scored, each sampled  
21 in 4,134 orientations and 343 conformations on average resulting in the evaluation of 19.5 trillion complexes  
22 in approximately 20 hours on a 1000-core computer cluster. A final set of ca. 16 million mostly anionic  
23 molecules from the February-2020 release of Enamine REAL was docked against Mac1. Within 10,703 core  
24 hours, 15,957,174 molecules were scored by evaluating a total of 12 trillion complexes where each molecule  
25 sampled on average 5,142 orientations and 495 conformations.

26  
27 The top 1 million scored compounds from each screen were investigated for intramolecular strain (total strain  
28 <7.5 TEU, maximum single torsion strain <2.5 TEU (41)) and hydrogen bonding with Asp22, Ile23, Gly48,  
29 Val49, Phe156 and Asp157. Molecules with unsatisfied hydrogen bond donors or more than three unsatisfied  
30 acceptors were not considered for experimental evaluation. The second docking campaign led to 54  
31 molecules that we selected for synthesis at Enamine of which 46 (85%) were obtained. The small analog set  
32 designed to probe neutral alternatives of negatively moieties binding in the oxyanion subsite were docked  
33 using the parameters from the second large-scale docking campaign. Molecules were protonated using  
34 ChemAxon Jchem 2019.15 (<https://chemaxon.com/>) at pH 7.4, rendered into 3D with Corina (v.3.6.0026,  
35 Molecular Networks GmbH, <https://mn-am.com/products/corina/>) and conformational libraries were generated  
36 with Omega (v.2.5.1.4, OpenEye Scientific Software; <https://www.eyesopen.com/omega>).

37  
38 A third docking screen was performed against the Z4305-stabilized, everted conformation of Mac1 (PDB  
39 5SOP). Before docking to the open structure, MDMix (42) (that utilizes AMBER18 (43)) was performed to  
40 assess binding hotspots in this less explored state. For this, the protein was solvated in pre-equilibrated  
41 mixtures of 20% ethanol and water, as well as 20% methanol and water. Three replicates of 50 ns  
42 simulations (six simulations total) were performed. Settings for minimization, equilibration and the production  
43 phase were set to default (42). After the simulation, all trajectories in the three independent simulations for  
44 each solvent mixture were aligned, after which the observed density was converted to binding free energies  
45 using the inverse Boltzmann relationship. Low energy regions were visualized and inferred to be probable  
46 binding hotspots.

47  
48 The crystal structure of Mac1 in complex with Z4305 was prepared for docking following the same steps as  
49 above, i.e. protonation, minimization and grid preparation. The dielectric boundary between the low dielectric  
50 protein environment and high dielectric solvent was moved outwards from the protein surface by 1.9 Å. Forty  
51 five matching spheres were generated based on 26 atomic coordinates of Z4305 and Z5531 as well as 19  
52 randomly placed spheres covering the oxyanion subsite and the surface near Ser128. Partial atomic charges  
53 of backbone amide hydrogen atoms of residues Ile23 were increased by 0.4 elementary charge units, while  
54 backbone amide hydrogen atoms of residues Phe156, Asp157 and Ser128 were increased by 0.2 elementary  
55 charge units. Partial charges of the corresponding backbone carbonyl oxygen atoms were reduced by the  
56 same amount. The described docking parameters were evaluated by control calculations the same way as  
57 described above for the second docking screen.



1  
2 Using a new virtual library, ZINC22 (<https://cartblanche22.docking.org>), a collection of 60,732,663  
3 monoanionic lead-like compounds (heavy atom count 17 to 25) were screened. Within 40 hours on a 1000-  
4 core computer cluster, roughly 57 million compounds were scored, each sampled in approx. 5,336  
5 orientations and 359 conformations resulting in more than 54 trillion complexes. The molecules that reached  
6 a total score threshold of -35 kcal/mol (comprising 4.3 M molecules) were filtered for internal molecular strain  
7 (<6.5 TEU; maximum single torsion strain <1.8 TEU) after which 1.7 M molecules remained. Next, molecules  
8 with more than one unsatisfied hydrogen bond donor or more than three unsatisfied acceptors were removed.  
9 Four independent sets were clustered by similarity for visual inspection, namely compounds able to interact  
10 with i) Asp22, Asp157 and Ser128, ii) Asp22, Phe156 and Ser128, iii) Ile23, Phe156 and Ser128, and iv)  
11 Asp22, Ile23 and Ser128, which led to 2, 249, 1761 and 2,249 compounds, respectively, ultimately leading to  
12 70 being purchased from Enamine, of which 56 (80%) could be synthesized.

### 13 14 **Crystallization and ligand soaking**

15 Crystals of SARS-CoV-2 NSP3 Mac1 were grown using an expression construct that crystallized in the P4<sub>3</sub>  
16 space group, as described previously (12) (**data file S2**). This construct crystallizes with two molecules in the  
17 asymmetric unit: the active site of protomer A is accessible to ligands while the active site of protomer B is  
18 obstructed by a crystal lattice interaction (12). The P4<sub>3</sub> crystal system was chosen because the crystals grow  
19 readily, diffract to atomic resolution and tolerate soaks in 10% DMSO for at least 6 hours (12). Briefly, crystals  
20 were grown by microseeding in 96-well sitting drop plates (SWISSCI, 3W96T-UVP), using 30 µl of 28% PEG  
21 3000 and 100 mM CHES pH 9.5 in the reservoir and crystallization drops containing 100 nl seeds, 100 nl  
22 reservoir and 200 nl protein (40 mg/ml in 150 mM NaCl, 20 mM Tris pH 8.5, 5% glycerol and 2 mM DTT).  
23 Crystals grew to maximum size in ~24 hours at 19°C. Compounds were prepared in DMSO to 100 mM, or to  
24 the maximum concentration allowed by solubility (see **data file S2** for compound concentrations).  
25 Compounds in DMSO were added to crystallization drops using acoustic dispensing with an Echo 650 liquid  
26 handler (Labcyte) (44). Soaks were performed with either 40 or 80 nl of compound per crystallization drop,  
27 giving a nominal concentration of 10 or 20 mM (see **data file S2**). After incubating for 2-4.5 hours at room  
28 temperature, crystals were vitrified in liquid nitrogen using a Nanuq cryocooling device (Mitegen). No  
29 additional cryoprotectant was added prior to vitrification. Although there was no observed decrease in  
30 diffraction quality with increased soak time (**Fig. S9**), certain compounds (namely Z8601 and LRH-0003)  
31 induced substantial disintegration of crystals after two hours, possibly linked to disruption of the crystal lattice  
32 by binding of compounds to the protomer B active site (12). Despite the crystal disintegration, reflections  
33 were recorded to <1 Å for crystals soaked with both compounds (**data file S2**).

### 34 35 **X-ray diffraction data collection and data reduction,**

36 Diffraction datasets were collected at beamline 8.3.1 at the Advanced Light Source, beamlines 12-1 and 12-2  
37 at the Stanford Synchrotron Radiation Lightsource, or beamline 17-ID-2 at the National Synchrotron Light  
38 Source II. The data collection parameters used at each beamline are listed in **data file S2**. X-ray diffraction  
39 images were indexed, integrated and scaled with XDS (45), using a reference P4<sub>3</sub> dataset to ensure  
40 consistent indexing. The high resolution limit for each dataset was chosen based on a CC<sub>1/2</sub> value of ~0.3 in  
41 the highest resolution shell (46). The diffraction resolution of crystals frequently exceeded the maximum  
42 resolution achievable with the experimental set-up; for these datasets, the high resolution limit was set to  
43 achieve ~95% completeness in the highest resolution shell. Data were merged with Aimless (47), and free R  
44 flags were copied from a reference P4<sub>3</sub> dataset. Structure factors intensities for all datasets have been  
45 uploaded to Zenodo in MTZ format (DOI: 10.5281/zenodo.6688239). For some compounds, datasets were  
46 collected from multiple crystals. Data collection and reduction statistics for all datasets summarized are in  
47 **data file S2**.

### 48 49 **Ligand identification, modeling and refinement**

50 All datasets were initially refined with the Dimple pipeline (48) run through CCP4 (49) using a starting model  
51 refined from a crystal soaked only in DMSO (dataset UCSF-P0110 in **data file S2**). Ligands were identified  
52 using PanDDA version 0.2.14 (17), with a ground-state map calculated using 34 datasets collected from  
53 crystals soaked only in DMSO. PanDDA was run an additional two times with ground-state maps calculated  
54 using 35 or 62 datasets from the ligand-soaked crystals where no ligands were detected. This procedure led  
55 to the identification of an additional 19 binding events, four of which were not identified in the first PanDDA  
56 run. Datasets used for ground-state map calculation for each of the PanDDA runs are annotated in **data file**  
57 **S2**. For ligands with multiple crystals/datasets, only the highest occupancy event was modeled. Ligands were

1 modeled into PanDDA event maps using COOT version 0.8.9.2 (50) with ligand restraints generated using  
2 *phenix.elbow* (51) or ACEDRG (52) from a SMILES strings, or from coordinates generated using LigPrep  
3 version 2022-1 (53). Based on the background density correction (BDC) values, ligand occupancies ranged  
4 from ~10-90% (**data file S2**). Many of the ligands had multiple conformations and/or isomers present. The  
5 isomers modeled, and the estimated ratios based on PanDDA event maps, are listed in **data file S2**.  
6 Datasets were collected from soaks performed with two batches of Z8539\_0002; one of the datasets was  
7 modeled with the (*R,R*) and (*S,S*) isomers (PDB 5SQD), while the other was only modeled with the (*R,S*)  
8 isomer (PDB 5SSN). Two compounds, LRH-0022 (PDB 5SRH) and LRH-0031 (PDB 5SRI), were only  
9 modeled with their pyrimido-indole core.

10  
11 For all ligands, we modeled changes in protein structure and water in the ligand binding sites into PanDDA  
12 event maps. Alternative conformations were included for residues where the heavy-atom RMSD value of the  
13 ligand-bound model to the ground-state model was greater than 0.15 Å. This cut-off was chosen with  
14 reference to the RMSD values for the 34 ground-state structures, where 99.7% of residues had RMSD values  
15 <0.15 Å (**Fig. S9**). In these multi-conformer models, the ground-state model was assigned the alternative  
16 occupancy identifier (altloc) A and the ligand-bound state was assigned altloc B (and C/D when overlapping  
17 conformations/isomers were present). Water molecules modeled into PanDDA event maps were assigned  
18 altloc B, and ground-state water molecules were included within 2.5 Å of ligand-bound state ligands or water  
19 (assigned altloc A).

20  
21 Refinement of the ligand-bound multi-conformer models was performed with *phenix.refine* using five  
22 refinement macrocycles (54). Coordinates and atomic displacement parameters (ADPs) were refined for all  
23 protein heavy atoms, and hydrogens were refined using a riding model. Based on previous observations (17),  
24 the occupancy of the ligand-bound and ground-states were set to 2\*(1-BDC) and 1-2\*(1-BDC) respectively,  
25 and occupancy refinement was switched off. Water molecules were automatically added to peaks in the mF<sub>O</sub>-  
26 DF<sub>C</sub> difference density map >3.5 σ using *phenix.refine*. To prevent the multi-conformer water molecules being  
27 removed by the automatic solvent picking, the ligand- and ground-state waters were renamed from HOH to  
28 WWW. After one round of refinement, maps and coordinates were inspected, and additional water molecules  
29 were placed manually using COOT into peaks in the mF<sub>O</sub>-DF<sub>C</sub> difference map. Based on positive/negative  
30 peaks in the mF<sub>O</sub>-DF<sub>C</sub> difference maps after refinement, the occupancies for some ligands were adjusted  
31 (initial and adjusted occupancies are listed in **data file S2**). Next, a second round of refinement was  
32 performed with ADPs refined anisotropically for non-hydrogen atoms, with automatic water picking, and the  
33 refinement of water coordinates, switched off. Data refinement statistics are summarized in **data file S2**.  
34 Coordinates, structure factor intensities and PanDDA event maps for all datasets have been deposited in the  
35 Protein Data Bank under the group deposition IDs G\_1002236, G\_1002238 and G\_1002239. Additionally, the  
36 PanDDA input and output files have been uploaded to Zenodo (DOI: 10.5281/zenodo.6688239).

### 37 38 **Homogeneous Time Resolved Fluorescence assay**

39 Binding of the compounds to macrodomain proteins was assessed by the displacement of an ADPr  
40 conjugated biotin peptide from His<sub>6</sub>-tagged protein using a HTRF-technology based screening assay which  
41 was performed as previously described (12). The expression sequences used for SARS-CoV-2 Mac1, and  
42 the human macrodomains TARG1 and MacroD2, are listed in **data file S2**. All proteins were expressed and  
43 purified as described previously for SARS-CoV-2 Mac1 (12). Compounds were dispensed into ProxiPlate-384  
44 Plus (PerkinElmer) assay plates using an Echo 525 liquid handler (Labcyte). Binding assays were conducted  
45 in a final volume of 16 μl with 12.5 nM NSP3 Mac1 protein, 400 nM peptide ARTK(Bio)QTARK(Aoa-RADP)S  
46 (Cambridge Peptides), 1:20000 Anti-His6-Eu3+ cryptate (HTRF donor, PerkinElmer) and 1:125 Streptavidin-  
47 XL665 (HTRF acceptor, PerkinElmer) in assay buffer (25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic  
48 acid (HEPES) pH 7.0, 20 mM NaCl, 0.05% bovine serum albumin and 0.05% Tween-20). TARG1 and  
49 MacroD2 binding were measured at 100 nM and 12.5 nM, respectively. Assay reagents were dispensed  
50 manually into plates using a multichannel pipette while macrodomain protein and peptide were first dispensed  
51 and incubated for 30 min at room temperature. This was followed by addition of the HTRF reagents and  
52 incubation at room temperature for 1 h. Fluorescence was measured using a PHERAstar microplate reader  
53 (BMG) using the HTRF module with dual emission protocol (A = excitation of 320 nm, emission of 665 nm,  
54 and B = excitation of 320 nm, emission of 620 nm) or a Synergy H1 (Biotek) using the HTRF filter set (A =  
55 excitation 330/80 nm, emission of 620/10 nm, and B = excitation of 330/80 nm and emission of 665/8 nm).  
56 Raw data were processed to give an HTRF ratio (channel A/B × 10,000), which was used to generate IC<sub>50</sub>

1 curves. The IC<sub>50</sub> values were determined by nonlinear regression using GraphPad Prism v.8 (GraphPad  
2 Software, CA, USA).

### 5 Isothermal Titration Calorimetry and estimation of K<sub>i</sub> values

6 To determine K<sub>i</sub> values from the obtained HTRF IC<sub>50</sub>s, binding experiments were carried out on a VP-ITC  
7 microcalorimeter (MicroCal) to determine the dissociation constant, K<sub>D</sub>, of Mac1 for the ADPr-peptide used in  
8 the HTRF assay. The protein was dialysed overnight at 4°C in ITC buffer (25 mM HEPES pH 7.0 and 20 mM  
9 NaCl) using D-tube Dialysis Midi MWCO 3.5 kDa (Novagen) dialysis tubes before the experiment. Titration  
10 experiments were then performed at 22°C, a reference power of 12 μCal s<sup>-1</sup> and a stirring speed of 307 rpm  
11 with an initial injection of 2 μl followed by 27 identical injections of 10 μl (duration of 4 s per injection and  
12 spacing of 240 s between injections). Data were analyzed using the MicroCal PEAQ-ITC analysis software  
13 (Malvern). K<sub>i</sub> values were calculated using the Cheng-Prusoff equation (18).

### 15 Differential Scanning Fluorimetry

16 DSF and associated compound handling was performed as described (12), with 5 μM dye “Fluorescent  
17 Yellow” (Jacquard iDye Cat #JID1405) used in place of SYPRO Orange. Compounds were tested in triplicate,  
18 at seven concentrations in two-fold serial dilutions, at a top concentration of either 1000 or 100 μM. Data  
19 were analyzed using DSFworld (55) by fitting raw RFU values from 25 to 85°C to the second DSFworld  
20 model (single transition with initial decay). For each compound, the Spearman coefficient was calculated  
21 between compound concentration and ΔT<sub>m</sub>. A “DSF positive” compound was defined as any compound  
22 which met all three criteria: a positive mean thermal shift ≥0.5°C at any tested concentration, positive  
23 Spearman estimate, and Spearman p value ≤0.05. All data used to determine temperature shifts by DSF are  
24 included in **data file S7**.

### 26 MDR1-MDCK II cell permeability

27 Permeability of compounds was assessed using canine MDR1 knockout, human MDR1 knockin MDCKII cells  
28 (MDR1-MDCKII) (Sigma-Aldrich, MTOX1303) in confluent monolayers expressing P-glycoprotein (P-gp) at  
29 Enamine biological services Bienta LTD (Kyiv, Ukraine). Cell suspension (400 μl) was added to each well of  
30 high throughput screening multiwell insert system plates. Test compounds were prepared as 20 mM DMSO  
31 stocks. The test compound (300 μl) was dissolved in transport buffer (9.5 g/l Hanks’ balanced salt solution  
32 and 0.35 g/l NaHCO<sub>3</sub> with 0.81 mM MgSO<sub>4</sub>, 1.26 mM CaCl<sub>2</sub>, 25 mM HEPES, pH adjusted to 7.4) and added  
33 into filter wells whereas 1000 μl of transport buffer was added to transport analysis plate wells in order to  
34 determine apical (A) to basolateral (B) transport. Basolateral to apical transport was measured by adding  
35 1000 μl of the test compound solution into the transport analysis plate wells whereas 300 μl of buffer was  
36 used to fill the filter plate wells. Final concentrations of test compounds were 10 μM. Plates were incubated  
37 for 90 min at 37°C under continuous shaking (100 rpm), 75 μl aliquots were taken from the donor and  
38 receiver compartments for LC-MS/MS analysis. Samples were mixed with acetonitrile followed by protein  
39 sedimentation by centrifugation at 1000 rpm for 10min. HPLC coupled with tandem mass spectroscopy was  
40 performed using the Shimadzu Prominence HPLC system coupled with the API 5000 (PE Sciex)  
41 spectrometer. Both the positive and negative ion modes of the TurbolonSpray ion source were used. The  
42 apparent permeability (P<sub>app</sub>) was computed using the equation 1, where V<sub>A</sub> is the volume of transport buffer  
43 in acceptor well, Area is the surface area of the insert, Time is the assay time, [drug]<sub>acc</sub> is the peak area of  
44 test compound in acceptor well, and [drug]<sub>initial,d</sub> is the initial amount of the test compound in a donor well.

$$45 \quad P_{app} = \frac{V_A}{Area \times Time} \times \frac{[compound]_{acc}}{[compound]_{initial,d}} \quad (1)$$



## 1 Acknowledgments

2 The X-ray crystal structures reported in this work were determined using diffraction data collected at the ALS,  
3 the SSRL and the NSLS-II. The ALS, a U.S. DOE Office of Science User Facility under contract no. DE-  
4 AC02-05CH11231, is supported in part by the ALS-ENABLE program funded by the NIH, National Institute of  
5 General Medical Sciences, grant P30 GM124169-01. Use of the SSRL, SLAC National Accelerator  
6 Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy  
7 Sciences under contract no. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is  
8 supported by the DOE Office of Biological and Environmental Research and by the NIH, National Institute of  
9 General Medical Sciences (P30GM133894). This research used beamline 17-ID-2 of the NSLS-II, the DOE  
10 Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory  
11 under contract no. DE-SC0012704. The Center for BioMolecular Structure (CBMS) is primarily supported by  
12 the NIH, NIGMS through a Center Core P30 Grant (P30GM133893), and by the DOE Office of Biological and  
13 Environmental Research (KP1605010). Structural biology applications used for the analysis and modeling of  
14 crystallographic data were compiled and configured by SBGrid (56). We acknowledge the use of the Wynton  
15 high-performance compute cluster at UCSF.

16  
17 **Funding:** This work was supported by the NIH and NIAID Antiviral Drug Discovery (AViDD) grant  
18 (U19AI171110) (to J.S.F., B.K.S., A.R., A.A.), by NSF Rapid 2031205, and a TMC Award from the UCSF  
19 Program for Breakthrough Biomedical Research, funded in part by the Sandler Foundation (to J.S.F. and  
20 J.E.G.), by NIH R35GM122481 and DARPA HR0011-19-2-0020 (to B.K.S.); GM141299 (to J.E.G.); and by  
21 GM133836 and GM071896 (to J.J.I.); by the Wellcome Trust (210634 and 223107), Oxford University  
22 Challenge Seed Fund (USCF 456), Biotechnology and Biological Sciences Research Council  
23 (BB/R007195/1), Ovarian Cancer Research Alliance (813369) and Cancer Research United Kingdom  
24 (C35050/A22284) (to I.A.) as well as NIHR Oxford Biomedical Research Centre (to J.C.T.).

25  
26 **Author contributions:** S.G. evaluated modeled merged and linked scaffolds, performed similarity searches  
27 for purchasable analogs, selected merger analogs for experimental validation, performed computational  
28 docking screens and structure-based compound optimization, prepared figures and co-wrote the manuscript.  
29 G.J.C. crystallized, performed compound soaking, vitrified crystals, collected X-ray diffraction data; modeled,  
30 refined, and analyzed compound-protein complex structures; prepared figures and co-wrote the manuscript.  
31 M.S. performed and analyzed the HTRF functional assay at Oxford University and performed ITC  
32 experiments with the ADPr-conjugated substrate. M.P.F. developed Fragemstein, performed fragment  
33 linking and merging calculations, and edited the manuscript. Y.U.D. set up, performed, and analyzed the  
34 HTRF functional assay for Mac1, TARG1 and MacroD2 at UCSF. M.R. performed computational docking  
35 against the everted state structure. T.W. performed and analyzed DSF experiments. M.D. supervised and  
36 assisted in the HTRF experiments at UCSF. S.W. assisted in computational docking. R.J.N. assisted in  
37 structure-based optimization. D.F. prepared X-ray structures for upload to Fragalysis. D.S.R. and Y.M.  
38 supervised compound synthesis at Enamine and Chempspace. J.J.I. edited the manuscript and developed  
39 docking libraries. A.R. designed the neutral probe chemical library and edited the manuscript. J.C.T.  
40 supervised the development of Fragemstein. J.E.G. supervised the DSF experiments. F.v.D. supervised the  
41 work and edited the manuscript. A.A. supervised the work, edited the manuscript and arranged funding. I.A.  
42 supervised the work, edited the manuscript and arranged funding. B.K.S. supervised the work, evaluated  
43 computational docking results, co-wrote the manuscript and arranged funding. J.S.F. supervised the work,  
44 co-wrote the manuscript and arranged funding.

45  
46 **Competing interests:** A.A. is a co-founder of Tango Therapeutics, Azkarra Therapeutics, and Ovibio  
47 Corporation; a consultant for SPARC, Bluestar, ProLynx, Earli, Cura, GenVivo, and GSK; a member of the  
48 SAB of Genentech, GLAdiator, Circle, and Cambridge Science Corporation; receives grant/research support  
49 from SPARC and AstraZeneca; and holds patents on the use of PARP inhibitors held jointly with  
50 AstraZeneca, which he has benefited financially (and may do so in the future). J.E.G. is consultant for Protego  
51 BioPharma and Contour Therapeutics as well as founder of Kaizen Therapeutics. B.K.S. is co-founder of  
52 BlueDolphin, LLC, a molecular docking contract research organization, Epiodyne, and Deep Apple  
53 Therapeutics, Inc., both drug discovery companies, has recently consulted for Umbra, Abbvie, and Dice  
54 Therapeutics, and is on the SAB of Schrodinger. J.J.I. co-founded Deep Apple Therapeutics, Inc., and  
55 BlueDolphin, LLC. J.S.F. is a consultant for, has equity in, and receives research support from Relay  
56 Therapeutics. I.A. is a consultant for Dark Blue Therapeutics.

1 **Data and materials availability:** All data generated or analyzed during this study are included in the paper  
2 and/or the Supplementary Materials. Crystallographic coordinates and structure factors for all structures have  
3 been deposited in the PDB with the following accession codes: 5SOI, 5SOJ, 5SOK, 5SOL, 5SOM, 5SON,  
4 5SOO, 5SOP, 5SOQ, 5SOR, 5SOS, 5SOT, 5SOU, 5SOV, 5SOW, 5SOX, 5SOY, 5SOZ, 5SP0, 5SP1, 5SP2,  
5 5SP3, 5SP4, 5SP6, 5SP7, 5SP8, 5SP9, 5SPA, 5SPB, 5SPC, 5SPD, 5SPE, 5SPF, 5SPG, 5SPH, 5SPI,  
6 5SPJ, 5SPK, 5SPL, 5SPM, 5SPN, 5SPO, 5SPP, 5SPQ, 5SPR, 5SPS, 5SPT, 5SPU, 5SPV, 5SPW, 5SPX,  
7 5SPY, 5SPZ, 5SQ0, 5SQ1, 5SQ2, 5SQ3, 5SQ4, 5SQ5, 5SQ6, 5SQ7, 5SQ8, 5SQ9, 5SQA, 5SQB, 5SQC,  
8 5SQD, 5SQE, 5SQF, 5SQG, 5SQH, 5SQI, 5SQJ, 5SQK, 5SQL, 5SQM, 5SQN, 5SQO, 5SQP, 5SQQ, 5SQR,  
9 5SQS, 5SQT, 5SQU, 5SQV, 5SQW, 5SQX, 5SQY, 5SQZ, 5SR0, 5SR1, 5SR2, 5SR3, 5SR4, 5SR5, 5SR6,  
10 5SR7, 5SR8, 5SR9, 5SRA, 5SRB, 5SRC, 5SRD, 5SRE, 5SRF, 5SRG, 5SRH, 5SRI, 5SRJ, 5SRK, 5SRL,  
11 5SRM, 5SRN, 5SRO, 5SRP, 5SRQ, 5SRR, 5SRS, 5SRT, 5SRU, 5SRV, 5SRW, 5SRX, 5SRY, 5SRZ, 5SS0,  
12 5SS1, 5SS2, 5SS3, 5SS4, 5SS5, 5SS6, 5SS7, 5SS8, 5SS9, 5SSA, 5SSB, 5SSC, 5SSD, 5SSE, 5SSF,  
13 5SSG, 5SSH, 5SSI, 5SSJ, 5SSK, 5SSL, 5SSM, 5SSN, 5SSO, 5SSP, 5SSQ, 5SSR. The scaled and merged  
14 structure factor intensities for all datasets, and the output of Dimple and PanDDA, have been uploaded to  
15 Zenodo (DOI: 10.5281/zenodo.6688239).

16  
17  
18

## 1 References

- 2 1. W. Yan, Y. Zheng, X. Zeng, B. He, W. Cheng, Structural biology of SARS-CoV-2: open the door for novel  
3 therapies. *Signal Transduct Target Ther.* **7**, 26 (2022).
- 4 2. A. K. L. Leung, D. E. Griffin, J. Bosch, A. R. Fehr, The Conserved Macrodomain Is a Potential Therapeutic Target  
5 for Coronaviruses and Alphaviruses. *Pathogens.* **11** (2022), p. 94.
- 6 3. H. Yang, Z. Rao, Structural biology of SARS-CoV-2 and implications for therapeutic development. *Nat. Rev.*  
7 *Microbiol.* **19**, 685–700 (2021).
- 8 4. W. Fu, H. Yao, M. Bütepage, Q. Zhao, B. Lüscher, J. Li, The search for inhibitors of macrodomains for targeting the  
9 readers and erasers of mono-ADP-ribosylation. *Drug Discov. Today.* **26**, 2547–2558 (2021).
- 10 5. J. G. M. Rack, V. Zorzini, Z. Zhu, M. Schuller, D. Ahel, I. Ahel, Viral macrodomains: a structural and evolutionary  
11 assessment of the pharmacological potential. *Open Biol.* **10**, 200237 (2020).
- 12 6. L. C. Russo, R. Tomasin, I. A. Matos, A. C. Manucci, S. T. Sowa, K. Dale, K. W. Caldecott, L. Lehtiö, D.  
13 Schechtman, F. C. Meotti, A. Bruni-Cardoso, N. C. Hoch, The SARS-CoV-2 Nsp3 macrodomain reverses  
14 PARP9/DTX3L-dependent ADP-ribosylation induced by interferon signaling. *J. Biol. Chem.* **297**, 101041 (2021).
- 15 7. Y. M. O. Alhammad, M. M. Kashipathy, A. Roy, J.-P. Gagné, P. McDonald, P. Gao, L. Nonfoux, K. P. Battaile, D. K.  
16 Johnson, E. D. Holmstrom, G. G. Poirier, S. Lovell, A. R. Fehr, The SARS-CoV-2 Conserved Macrodomain Is a  
17 Mono-ADP-Ribosylhydrolase. *J. Virol.* **95** (2021), doi:10.1128/JVI.01969-20.
- 18 8. A. R. Fehr, R. Channappanavar, G. Jankevicius, C. Fett, J. Zhao, J. Athmer, D. K. Meyerholz, I. Ahel, S. Perlman,  
19 The Conserved Coronavirus Macrodomain Promotes Virulence and Suppresses the Innate Immune Response  
20 during Severe Acute Respiratory Syndrome Coronavirus Infection. *MBio.* **7** (2016), doi:10.1128/mBio.01721-16.
- 21 9. L. M. Sherrill, E. E. Joya, A. Walker, A. Roy, Y. M. Alhammad, M. Atobatele, S. Wazir, G. Abbas, P. Keane, J.  
22 Zhuo, A. K. L. Leung, D. K. Johnson, L. Lehtiö, A. R. Fehr, D. Ferraris, Design, synthesis and evaluation of  
23 inhibitors of the SARS-CoV-2 nsp3 macrodomain. *Bioorg. Med. Chem.* **67**, 116788 (2022).
- 24 10. A. Roy, Y. M. Alhammad, P. McDonald, D. K. Johnson, J. Zhuo, S. Wazir, D. Ferraris, L. Lehtiö, A. K. L. Leung, A.  
25 R. Fehr, Discovery of compounds that inhibit SARS-CoV-2 Mac1-ADP-ribose binding by high-throughput screening.  
26 *Antiviral Res.* **203**, 105344 (2022).
- 27 11. S. T. Sowa, A. Galera-Prat, S. Wazir, H. I. Alanen, M. M. Maksimainen, L. Lehtiö, A molecular toolbox for ADP-  
28 ribosyl binding proteins. *Cell Rep Methods.* **1**, 100121 (2021).
- 29 12. M. Schuller, G. J. Correy, S. Gahbauer, D. Fearon, T. Wu, R. E. Díaz, I. D. Young, L. Carvalho Martins, D. H.  
30 Smith, U. Schulze-Gahmen, T. W. Owens, I. Deshpande, G. E. Merz, A. C. Thwin, J. T. Biel, J. K. Peters, M.  
31 Moritz, N. Herrera, H. T. Kratochvil, QCRG Structural Biology Consortium, A. Aimon, J. M. Bennett, J. Brandao  
32 Neto, A. E. Cohen, A. Dias, A. Douangamath, L. Dunnett, O. Fedorov, M. P. Ferla, M. R. Fuchs, T. J. Gorrie-Stone,  
33 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,  
34 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.  
35 E. Gestwicki, F. von Delft, B. K. Shoichet, J. S. Fraser, I. Ahel, Fragment binding to the Nsp3 macrodomain of  
36 SARS-CoV-2 identified through crystallographic screening and computational docking. *Sci Adv.* **7** (2021),  
37 doi:10.1126/sciadv.abf8711.
- 38 13. G. J. Correy, D. W. Kneller, G. Phillips, S. Pant, S. Russi, A. E. Cohen, G. Meigs, J. M. Holton, S. Gahbauer, M. C.  
39 Thompson, A. Ashworth, L. Coates, A. Kovalevsky, F. Meilleur, J. S. Fraser, The mechanisms of catalysis and  
40 ligand binding for the SARS-CoV-2 NSP3 macrodomain from neutron and x-ray diffraction at room temperature. *Sci*  
41 *Adv.* **8**, eabo5083 (2022).
- 42 14. M. Ferla, Scaffold hopping between bound compounds by stitching them together like a reanimated corpse.  
43 *Fragmenstein*, (available at <https://github.com/matteoferla/Fragmenstein>).
- 44 15. O. O. Grygorenko, D. S. Radchenko, I. Dziuba, A. Chuprina, K. E. Gubina, Y. S. Moroz, Erratum: Generating  
45 Multibillion Chemical Space of Readily Accessible Screening Compounds. *iScience.* **23**, 101873 (2020).
- 46 16. J. J. Irwin, K. G. Tang, J. Young, C. Dandarchuluun, B. R. Wong, M. Khurelbaatar, Y. S. Moroz, J. Mayfield, R. A.  
47 Sayle, ZINC20—A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *J. Chem. Inf. Model.* **60**, 6065–



- 1 6073 (2020).
- 2 17. N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M.  
3 Deane, F. von Delft, A multi-crystal method for extracting obscured crystallographic states from conventionally  
4 uninterpretable electron density. *Nat. Commun.* **8**, 15123 (2017).
- 5 18. C. Yung-Chi, W. H. Prusoff, Relationship between the inhibition constant (KI) and the concentration of inhibitor  
6 which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
- 7 19. X. Yang, Y. Ma, Y. Li, Y. Dong, L. L. Yu, H. Wang, L. Guo, C. Wu, X. Yu, X. Liu, Molecular basis for the MacroD1-  
8 mediated hydrolysis of ADP-ribosylation. *DNA Repair* . **94**, 102899 (2020).
- 9 20. J. Lyu, S. Wang, T. E. Balius, I. Singh, A. Levit, Y. S. Moroz, M. J. O’Meara, T. Che, E. Alga, K. Tolmacheva, A. A.  
10 Tolmachev, B. K. Shoichet, B. L. Roth, J. J. Irwin, Ultra-large library docking for discovering new chemotypes.  
11 *Nature*. **566**, 224–229 (2019).
- 12 21. K. Michalska, Y. Kim, R. Jedrzejczak, N. I. Maltseva, L. Stols, M. Endres, A. Joachimiak, Crystal structures of  
13 SARS-CoV-2 ADP-ribose phosphatase: from the apo form to ligand complexes. *IUCrJ*. **7**, 814–824 (2020).
- 14 22. T. Sterling, J. J. Irwin, ZINC 15 – Ligand Discovery for Everyone. *Journal of Chemical Information and Modeling*. **55**  
15 (2015), pp. 2324–2337.
- 16 23. B. J. Bender, S. Gahbauer, A. Lutten, J. Lyu, C. M. Webb, R. M. Stein, E. A. Fink, T. E. Balius, J. Carlsson, J. J.  
17 Irwin, B. K. Shoichet, A practical guide to large-scale docking. *Nat. Protoc.* **16**, 4799–4832 (2021).
- 18 24. W. J. Allen, R. C. Rizzo, Implementation of the Hungarian algorithm to account for ligand symmetry and similarity in  
19 structure-based design. *J. Chem. Inf. Model.* **54**, 518–529 (2014).
- 20 25. Q. Zhao, R. Capelli, P. Carloni, B. Lüscher, J. Li, G. Rossetti, Enhanced Sampling Approach to the Induced-Fit  
21 Docking Problem in Protein-Ligand Binding: The Case of Mono-ADP-Ribosylation Hydrolase Inhibitors. *J. Chem.*  
22 *Theory Comput.* **17**, 7899–7911 (2021).
- 23 26. M. Jaskolski, A. Wlodawer, Z. Dauter, W. Minor, B. Rupp, Group depositions to the Protein Data Bank need  
24 adequate presentation and different archiving protocol. *Protein Sci.* **31**, 784–786 (2022).
- 25 27. W. Jahnke, D. A. Erlanson, I. J. P. de Esch, C. N. Johnson, P. N. Mortenson, Y. Ochi, T. Urushima, Fragment-to-  
26 Lead Medicinal Chemistry Publications in 2019. *J. Med. Chem.* **63**, 15494–15507 (2020).
- 27 28. I. J. P. de Esch, D. A. Erlanson, W. Jahnke, C. N. Johnson, L. Walsh, Fragment-to-lead medicinal chemistry  
28 publications in 2020. *J. Med. Chem.* **65**, 84–99 (2022).
- 29 29. C. R. Søndergaard, A. E. Garrett, T. Carstensen, G. Pollastri, J. E. Nielsen, Structural artifacts in protein-ligand X-  
30 ray structures: implications for the development of docking scoring functions. *J. Med. Chem.* **52**, 5673–5684 (2009).
- 31 30. D. Chen, M. Vollmar, M. N. Rossi, C. Phillips, R. Kraehenbuehl, D. Slade, P. V. Mehrotra, F. von Delft, S. K.  
32 Crosthwaite, O. Gileadi, J. M. Denu, I. Ahel, Identification of macrodomain proteins as novel O-acetyl-ADP-ribose  
33 deacetylases. *J. Biol. Chem.* **286**, 13261–13271 (2011).
- 34 31. M. Schuller, K. Riedel, I. Gibbs-Seymour, K. Uth, C. Sieg, A. P. Gehring, I. Ahel, F. Bracher, B. M. Kessler, J. M.  
35 Elkins, S. Knapp, Discovery of a Selective Allosteric Inhibitor Targeting Macrodomain 2 of Polyadenosine-  
36 Diphosphate-Ribose Polymerase 14. *ACS Chem. Biol.* **12**, 2866–2874 (2017).
- 37 32. L. Wang, Y. Wu, Y. Deng, B. Kim, L. Pierce, G. Krilov, D. Lupyan, S. Robinson, M. K. Dahlgren, J. Greenwood, D.  
38 L. Romero, C. Masse, J. L. Knight, T. Steinbrecher, T. Beuming, W. Damm, E. Harder, W. Sherman, M. Brewer, R.  
39 Wester, M. Murcko, L. Frye, R. Farid, T. Lin, D. L. Mobley, W. L. Jorgensen, B. J. Berne, R. A. Friesner, R. Abel,  
40 Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a  
41 modern free-energy calculation protocol and force field. *J. Am. Chem. Soc.* **137**, 2695–2703 (2015).
- 42 33. V. Gapsys, A. Yildirim, M. Aldeghi, Y. Khalak, D. van der Spoel, B. L. de Groot, Accurate absolute free energies for  
43 ligand–protein binding based on non-equilibrium approaches. *Communications Chemistry*. **4**, 1–13 (2021).
- 44 34. R. G. Coleman, M. Carchia, T. Sterling, J. J. Irwin, B. K. Shoichet, Ligand Pose and Orientational Sampling in  
45 Molecular Docking. *PLoS ONE*. **8** (2013), p. e75992.

- 1 35. K. Gallagher, K. Sharp, Electrostatic Contributions to Heat Capacity Changes of DNA-Ligand Binding. *Biophysical*  
2 *Journal*. **75** (1998), pp. 769–776.
- 3 36. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, P. Weiner, A new force field  
4 for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).
- 5 37. M. M. Mysinger, B. K. Shoichet, Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf.*  
6 *Model.* **50**, 1561–1573 (2010).
- 7 38. G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, Protein and ligand preparation:  
8 parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234  
9 (2013).
- 10 39. M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, J. H. Jensen, PROPKA3: Consistent Treatment of Internal  
11 and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **7**, 525–537 (2011).
- 12 40. R. M. Stein, Y. Yang, T. E. Balius, M. J. O'Meara, J. Lyu, J. Young, K. Tang, B. K. Shoichet, J. J. Irwin, Property-  
13 Unmatched Decoys in Docking Benchmarks. *J. Chem. Inf. Model.* **61**, 699–714 (2021).
- 14 41. S. Gu, M. S. Smith, Y. Yang, J. J. Irwin, B. K. Shoichet, Ligand Strain Energy in Large Library Docking. *J. Chem.*  
15 *Inf. Model.* **61**, 4331–4341 (2021).
- 16 42. D. Alvarez-Garcia, X. Barril, Molecular simulations with solvent competition quantify water displaceability and  
17 provide accurate interaction maps of protein binding sites. *J. Med. Chem.* **57**, 8530–8539 (2014).
- 18 43. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force  
19 field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
- 20 44. P. M. Collins, J. T. Ng, R. Talon, K. Nekrosiute, T. Krojer, A. Douangamath, J. Brandao-Neto, N. Wright, N. M.  
21 Pearce, F. von Delft, Gentle, fast and effective crystal soaking by acoustic dispensing. *Acta Crystallogr D Struct*  
22 *Biol.* **73**, 246–255 (2017).
- 23 45. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
- 24 46. P. A. Karplus, K. Diederichs, Linking crystallographic model and data quality. *Science*. **336**, 1030–1033 (2012).
- 25 47. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. D Biol.*  
26 *Crystallogr.* **69**, 1204–1214 (2013).
- 27 48. R. Keegan, M. Wojdyr, G. Winter, A. Ashton, DIMPLE: a difference map pipeline for the rapid screening of crystals  
28 on the beamline. *Acta Crystallogr. A Found. Adv.* **71**, s18–s18 (2015).
- 29 49. 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.  
30 W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A.  
31 Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.*  
32 **67**, 235–242 (2011).
- 33 50. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol.*  
34 *Crystallogr.* **66**, 486–501 (2010).
- 35 51. N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, electronic Ligand Builder and Optimization Workbench  
36 (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1074–1080  
37 (2009).
- 38 52. F. Long, R. A. Nicholls, P. Emsley, S. Graëulis, A. Merkys, A. Vaitkus, G. N. Murshudov, AceDRG: a  
39 stereochemical description generator for ligands. *Acta Crystallogr D Struct Biol.* **73**, 112–122 (2017).
- 40 53. L. L. C. Schrödinger, *LigPrep*.
- 41 54. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger,  
42 A. Urzhumtsev, P. H. Zwart, P. D. Adams, Towards automated crystallographic structure refinement with  
43 phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).

- 1 55. T. Wu, J. Yu, Z. Gale-Day, A. Woo, A. Suresh, M. Hornsby, J. E. Gestwicki, Three Essential Resources to Improve  
2 Differential Scanning Fluorimetry (DSF) Experiments. *bioRxiv* (2020), p. 2020.03.22.002543.
- 3 56. A. Morin, B. Eisenbraun, J. Key, P. C. Sanschagrin, M. A. Timony, M. Ottaviano, P. Sliz, Collaboration gets the  
4 most out of software. *Elife*. **2**, e01456 (2013).
- 5 57. R. Sharifi, R. Morra, C. D. Appel, M. Tallis, B. Chioza, G. Jankevicius, M. A. Simpson, I. Matic, E. Ozkan, B. Golia,  
6 M. J. Schellenberg, R. Weston, J. G. Williams, M. N. Rossi, H. Galehdari, J. Krahn, A. Wan, R. C. Trembath, A. H.  
7 Crosby, D. Ahel, R. Hay, A. G. Ladurner, G. Timinszky, R. S. Williams, I. Ahel, Deficiency of terminal ADP-ribose  
8 protein glycohydrolase TARG1/C6orf130 in neurodegenerative disease. *EMBO J.* **32**, 1225–1237 (2013).
- 9 58. G. Jankevicius, M. Hassler, B. Golia, V. Rybin, M. Zacharias, G. Timinszky, A. G. Ladurner, A family of  
10 macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nat. Struct. Mol. Biol.* **20**, 508–514 (2013).
- 11 59. A. H. Forst, T. Karlberg, N. Herzog, A.-G. Thorsell, A. Gross, K. L. H. Feijs, P. Verheugd, P. Kursula, B. Nijmeijer,  
12 E. Kremmer, H. Kleine, A. G. Ladurner, H. Schüler, B. Lüscher, Recognition of mono-ADP-ribosylated ARTD10  
13 substrates by ARTD8 macrodomains. *Structure*. **21**, 462–475 (2013).