Iterative computational design and crystallographic screening identifies potent inhibitors targeting the Nsp3 macrodomain of SARS-CoV-2

Stefan Gahbauer[a,1] 🄳, Galen J. Correy[b,1] 🄳, Marion Schuller[c], Matteo P. Ferla[d,e] 🄳, Yagmur Umay Doruk[f] 🄳, Moira Rachman[a] 🄳, Taiasean Wu[g,h], Morgan Diolaiti[f] 🄳, Siyi Wang[h], R. Jeffrey Neitz[i] 🄳, Daren Fearon[j,k] 🄳, Dmytro S. Radchenko[l,m] 🄳, Yurii S. Moroz[m,n] 🄳, John J. Irwin[a], Adam R. Renslo[f,i] 🄳, Jenny C. Taylor[d,e] 🄳, Jason E. Gestwicki[g,i], Frank von Delft[j,k,o,p,q] 🄳, Alan Ashworth[f], Ivan Ahel[c], Brian K. Shoichet[a,2] 🄳, and James S. Fraser[b,2] 🄳

Edited by Lila Gierasch, University of Massachusetts Amherst, Amherst, MA; received July 27, 2022; accepted November 28, 2022

The nonstructural protein 3 (NSP3) of the severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) contains a conserved macrodomain enzyme (Mac1) that is critical for pathogenesis and lethality. While small-molecule inhibitors of Mac1 have great therapeutic potential, at the outset of the COVID-19 pandemic, there were no well-validated inhibitors for this protein nor, indeed, the macrodomain enzyme family, making this target a pharmacological orphan. Here, we report the structure-based discovery and development of several different chemical scaffolds exhibiting low- to sub-micromolar affinity for Mac1 through iterations of computer-aided design, structural characterization by ultra-high-resolution protein crystallography, and binding evaluation. Potent scaffolds were designed with in silico fragment linkage and by ultra-large library docking of over 450 million molecules. Both techniques leverage the computational exploration of tangible chemical space and are applicable to other pharmacological orphans. Overall, 160 ligands in 119 different scaffolds were discovered, and 153 Mac1-ligand complex crystal structures were determined, typically to 1 Å resolution or better. Our analyses discovered selective and cell-permeable molecules, unexpected ligand-mediated conformational changes within the active site, and key inhibitor motifs that will template future drug development against Mac1.

coronavirus | macrodomain | virtual screening | fragment-based drug discovery

The macrodomain of severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) nonstructural protein 3 (NSP3) (Mac1) presents an intriguing target for drug discovery (1–5). Upon viral infection, host cells initiate an innate interferon-mediated immune response leading to the expression of poly-(ADP-ribose)-polymerases (PARPs), which catalyze the antiviral posttranslational addition of ADP-ribose (ADPr) to a large range of target proteins (6). Mac1 enzymatically reverses this mono-ADP-ribosylation, counteracting immune signaling (7). Promisingly, inactivation of Mac1 by single-point mutations in the ADPr-binding site significantly reduced lethality and pathogenicity in mice after SARS-CoV infection (8). Small-molecule inhibitors of SARS-CoV-2 Mac1 should therefore offer novel therapeutics to mitigate COVID-19 (9, 10).

A challenge for the development of such inhibitors has been the lack of small-molecule modulators of macrodomain activity, other than ADPr; indeed, only recently have quantitative assays been developed (10, 11). This is true not only for Mac1 from SARS-CoV-2, but also for the overall family of enzymes, which lack good chemical matter by which their activity can be probed, despite their importance in several areas of health and diseases. Accordingly, to map the recognition determinants of Mac1, we adopted a biophysical approach, screening for fragment ligands using protein crystallography, molecular docking, isothermal titration calorimetry (ITC), differential scanning fluorimetry (DSF), and a binding assay based on homogeneous time-resolved fluorescence (HTRF) (12). Mac1 proved to be unusually amenable to structure determination, enabling us to determine the structures of over 230 fragment complexes, typically to ultra-high resolution (often better than 1.1 Å), affording us a detailed map of enzyme hot spots with chemical matter of sufficient potency with which to optimize a quantitative assay (12, 13).

Nevertheless, our best fragments remained of modest potency, with none more potent than 180 μM. Here, we describe the discovery and optimization of potent macrodomain ligands using two strategies (Fig. 1). In the first, we sought to link and merge pairs of fragments to create larger molecules that exploited multiple hot spots, so reaching higher affinities. This used a fragment-linking method (12), adapted to explore a virtual library of 22 billion readily synthesizable molecules (14). In the second approach, we exploited the hot spots revealed by the initial fragments to guide computational docking of ultra-large chemical libraries of lead-like molecules, potentially more potent than the fragments docked in our original study (12). Both approaches ultimately led to compounds with IC$_{50}$ values as low as 0.4 μM for the merged fragments and as low as 1.7 μM for the docking hits (Fig. 1). These represent the

Significance

The nonstructural protein 3 (NSP3) of the severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) contains a conserved macrodomain enzyme (Mac1) that is critical for pathogenesis and lethality. There are currently no well-validated inhibitors for this protein. Here, we discovered and optimized several different classes of ligands that bind to Mac1 with low- to sub-micromolar affinity. Ligands were designed by linking together small-molecule fragments and by ultra-large library docking of 450 million molecules. Overall, we discovered 160 ligands in 119 different scaffolds, and 152 Mac1-ligand complex crystal structures were determined. Our analyses discovered selective and cell-permeable molecules, unexpected ligand-mediated protein dynamics within the active site, and key structural information that will guide future drug development for this important antiviral target.

[1]S.G. and G.J.C. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: bshoichet@gmail.com or jfraser@fraserlab.com.
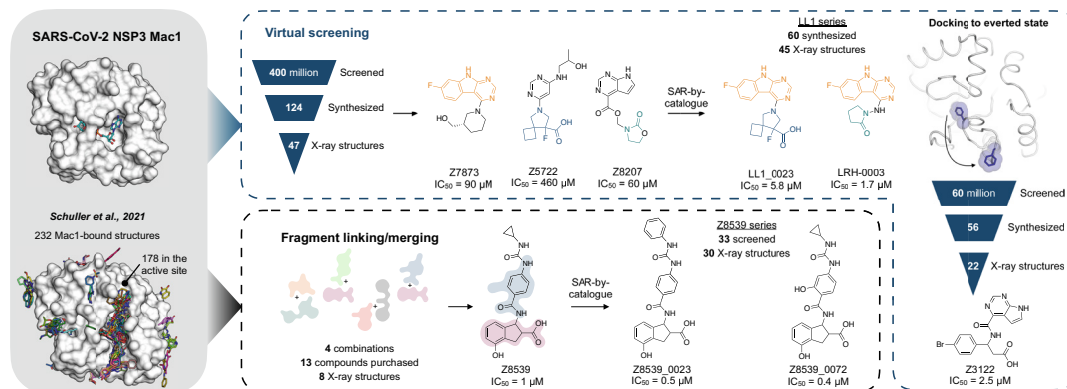
Fig. 1. Overview of the structure-based strategies used to discover ligands that bind to the NSP3 macrodomain of SARS-CoV-2 (Mac1).

most potent inhibitors reported for any member of the broad family of macrodomains. Furthermore, the many X-ray crystal structures determined here provide an extensive resource for drug development campaigns against this promising antiviral target.

Results

Hit Discovery through Fragment Merging. The large collection of Mac1-fragment crystal structures revealed interaction patterns between initial ligands and the Mac1 active site, thus providing various templates for fragment-based drug discovery (12). The largest subset of fragments bound in the adenine recognition subsite, hydrogen bonding to Asp22 and Ile23, and stacking with Phe156. Another group of mainly acidic fragments occupied a subpocket formed by the backbone NH groups of Phe156 and Asp157, which we labeled the "oxyanion subsite." Although ADPr itself does not directly interact with this oxyanion site, the most potent compound that emerged from the fragment screen (ZINC263392672, Protein Data Bank (PDB) deposition 5RSG, $IC_{50} = 180 \mu M$) placed a pyrrolopyrimidine group in the adenine subsite and carboxylate in the oxyanion subsite, suggesting that molecules able to bridge between both subsites hold potential for potent ligand design. An interactive dataset of the initial hits can be found at https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mac1.

Consequently, we sought to improve the affinity of the individual fragments by fusing pairs together into a larger, more potent molecule. Such fragment linking has traditionally been considered technically difficult (15), as the linkage must minimally disturb the positioning of the two original fragments, and such a molecule must be synthetically accessible. Here, we tried to do so using an automated fragment-linking approach, *Fragmenstein*, that searches purchasable chemical space to find molecules that could meet the design. 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. 2 *A* and *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 *SI Appendix,* Fig. S1) or ZINC98208711 (PDB 5RU5) (*SI Appendix,* Fig. S2), ZINC26180281 (PDB 5RSF) was merged with ZINC89254160_N3 (PDB 5RSJ) (*SI Appendix,* Fig. S2), and Z44592329 (PDB 5S2F) was merged with ZINC13514509 (PDB

5RTN) (*SI Appendix,* 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 and *SI Appendix,* Figs. S1 and S2). Compounds were assessed for thermal upshift by DSF across seven concentrations in triplicate. Seven compounds induced thermal upshift (Dataset S1). A thermal upshift was defined as an increase in Tma of at least 0.5 °C (equivalent to three SDs), with statistically significant ($P < 0.05$) dose responsiveness by Spearman's correlation. 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) (Dataset 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 the 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 *SI Appendix,* Fig. S1). In the HTRF-based binding assay (12), Z8531 and Z9604 had IC_{50} values above 250 μM, while Z8515 and Z8539 had IC_{50} values of 7.9 μM and 0.8 to 1.1 μM, 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 to 1.3 μM) (Fig. 2G). The K_D of the ADPr-conjugated peptide used in the HTRF assay was determined to be 2.7 μM by ITC (Dataset S1); therefore, the measured IC_{50} values of the molecules are similar to the binding affinities estimated using the Cheng–Prusoff equation (18).

All the 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
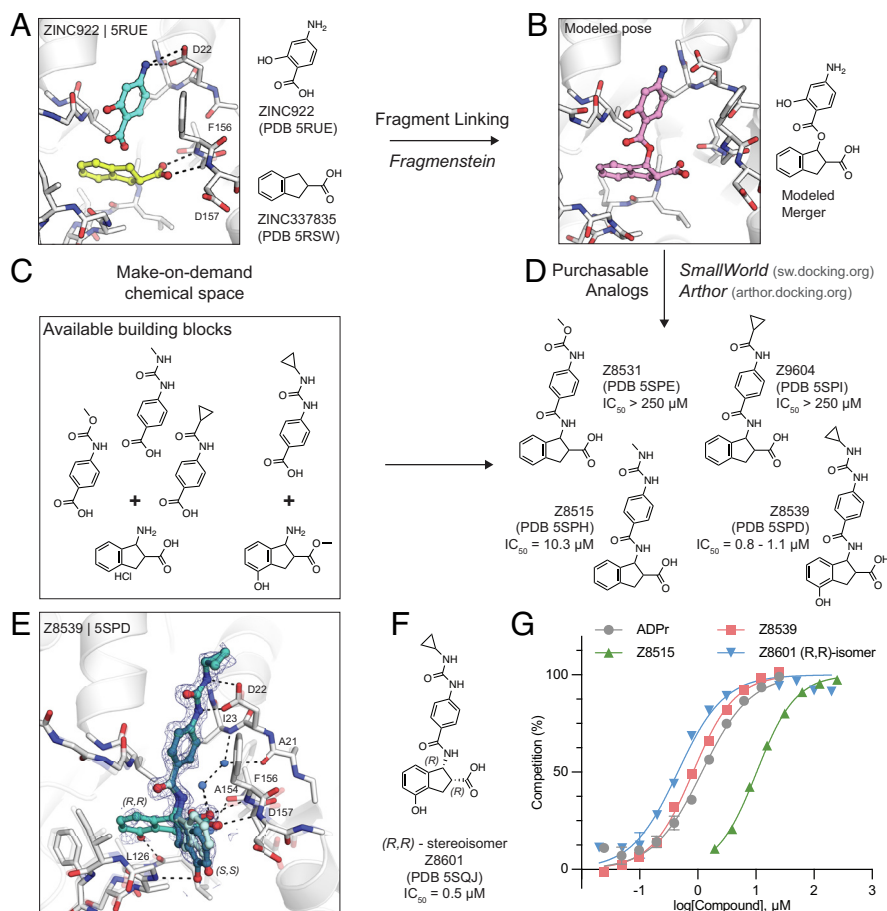
Fig. 2. In silico fragment linking targeting the adenosine site of Mac1. *A*) Binding pose of two fragments identified in the previously reported fragment screen (12). Fragment protein hydrogen bonds are shown with dashed black lines. *B*) Theoretical linked scaffold of ZINC922 and ZINC337835 generated using *Fragmenstein*. *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 σ). 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 ± SEM of at least two technical replicates.

(Fig. 2*A*). 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 μM, i.e., twofold more potent than the diastereomeric mixture (Fig. 2*F*). In this configuration, the indane group partially inserts into the phosphate-binding domain and the terminal phenol hydrogen bonds with the backbone oxygen of Leu126. In the binding pose of the (*S,S*) stereoisomer (IC_{50} = 2.9 μM), the phenol is mainly solvent exposed and the hydroxyl hydrogen bonds with the backbone nitrogen of Gly130 (Fig. 3*A*). By contrast, the two *trans* diastereomers showed reduced affinities with IC_{50} values between 43 and 55 μM. The X-ray crystal structure shows that the carboxylic acid of the (*R,S*) isomer only forms a single hydrogen bond to the oxyanion subsite (*SI Appendix*, Fig. S3), while a structure of the (*S,R*) isomer was not obtained. The (*R,R*) stereoisomer (Z8601) was tested for off-target activity against two human macrodomains, MacroD2 and TARG1, using an adapted HTRF-based peptide displacement assay. The human proteins MacroD2 and TARG1 were chosen to test selectivity because they are the most similar human proteins to SARS-CoV-2 Mac1 (5). Z8601 showed no displacement of the ADPr-conjugated substrate at 50 μM against either target and approximately 50% displacement at high concentrations of 1 mM (*SI Appendix*, Fig. S4). The selectivity of this scaffold for the viral over the tested human macrodomains is likely related to sequence differences within the ADPr-binding pockets between all the three proteins: while Ala52

in the viral Mac1 offers ample space to accommodate the compound's phenyl-urea functional group, MacroD2 and TARG1 carry considerably larger residues at the corresponding position, namely Leu50 and Cys104, respectively (*SI Appendix*, Fig. S4).

The 1.05 Å resolution crystal structure of Mac1 in complex with the (*R,R*) isomer of Z8539 (Z8601) reveals an extended water-mediated hydrogen bond network between the ligand's central amide, its carboxylic acid and Ile23, Ala21, as well as Ala154 (Fig. 2 and *SI Appendix*, Fig. S3). Methylation of Z8539's central amide group (Z8539_0056, *SI Appendix*, Fig. S3) rendered the compound inactive, likely because of the interruption of this network. Interestingly, the initially generated ester-linked merger (Fig. 2*B*) that could be custom synthesized by Enamine showed much weaker binding to Mac1 (Z8507, IC_{50} = 144 μM) than the make-on-demand analogs Z8539 and Z8515. The corresponding crystal structure of Mac1 in complex with the ester Z8507 (PDB 8ERS) shows a similar water network as in the Z8539-bound structure; however, the central ester oxygen atom of Z8507 may not form the same favorable interactions with this water network as compared to Z8539's central amine (*SI Appendix*, Fig. S1). Thus, our preference for readily synthesized molecules may have conferred an unexpected advantage over the initial theoretical merger.

Structure-Based Optimization of the Merged Scaffold. To further explore the Z8539 scaffold, we generated a structure–activity relationship (SAR) series (Fig. 3 *A–E*). Here too, 2D-based similarity
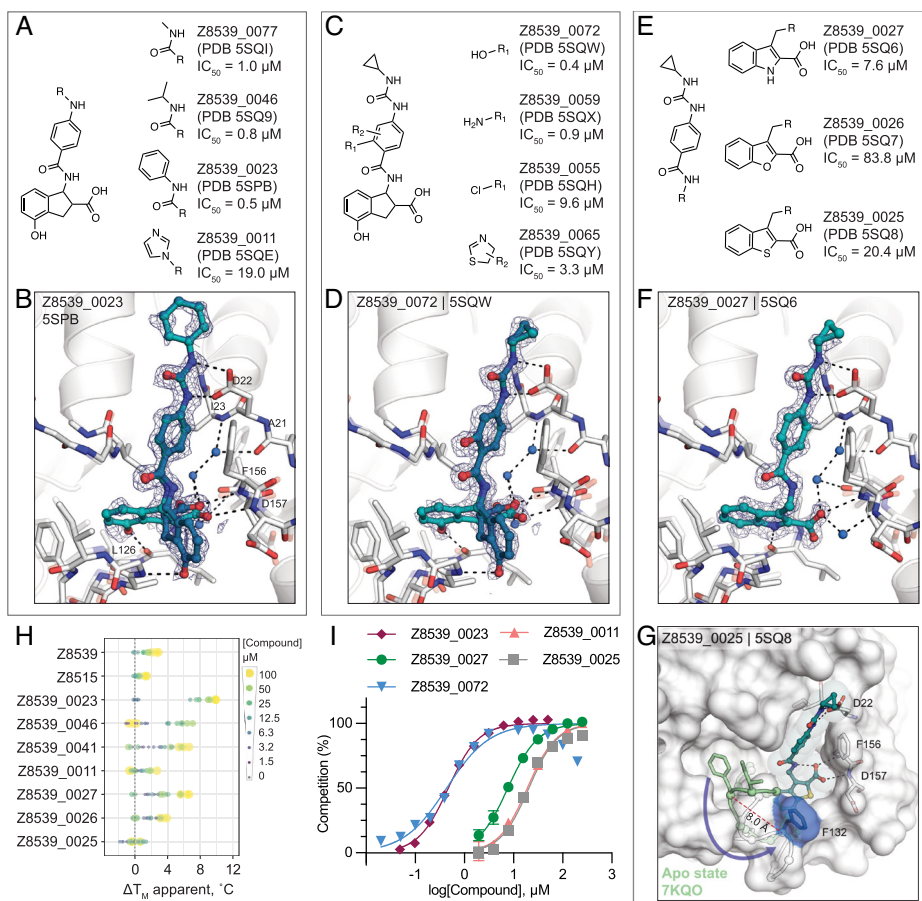
Fig. 3. Structure-based optimization of Z8539. *A*) Modifications of the cyclopropyl-phenylurea group. *B*) X-ray crystal structure of Mac1 bound to Z8539_0023. The PanDDA event map is shown around the ligand (blue mesh contoured at 2 σ). *C*) Modifications of the central benzene. *D*) X-ray crystal 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 ± SEM of at least two technical replicates.

searches of the Enamine REAL database were used to find readily accessible and SAR-useful analogs, while analogs unavailable in the REAL database were also designed. Approximately 21,000 analogs (roughly 4,000 mono-anions) were identified via *SmallWorld* and subsequently docked against the Mac1-Z8539 crystal structure. Visual inspection of top-ranked (mostly) anionic compounds led to the selection of 19 readily accessible make-on-demand analogs, while nine compounds were manually designed; of these 28 and 26 were successfully synthesized at Enamine. Of these 26 analogs, 23 were confirmed to bind Mac1 by crystallography and 20 showed activity in the HTRF assay (Datasets S1–S3).

Most analogs bore modification of the cyclopropyl-phenylurea group of Z8539 (Fig. 3*A*). Removal of the cyclopropyl (Z8539_0041, PDB 5SPA) or replacement by either methyl (Z8539_0077, PDB 5SQI) or isobutyl (Z8539_0046, PDB 5SQ9) did not substantially change binding affinity; however, phenyl replacement (Z8539_0023, PDB 5SPB) improved the IC$_{50}$ to 0.5 μM and showed a significantly increased thermal upshift of 9 °C in DSF (for the stereoisomeric mixture) (Fig. 3 *B–H*). The resulting diphenyl-urea superimposes well with known fragment hits, e.g., Z44592329 (PDB 5S2F) or Z321318226 (PDB 5S2G) (12) (*SI Appendix*, Fig. S2). Compound Z8539_0011 (IC$_{50}$ = 19 μM) contains an imidazole moiety that forms an additional hydrogen bond to Lys55 (Dataset S3A.5). Addition of hydrogen bond donors such as amine (Z8539_0059, PDB 5SQX) and hydroxyl (Z8539_0072, PDB 5SQW) at the amide-ortho-position of the central benzene yielded relatively potent analogs with affinities of 0.9 μM and 0.4 μM, respectively (Fig. 3 *C* and *D*). The corresponding crystal structures do not reveal additional interactions between the newly

introduced substituents and the protein; however, the binding poses of the ligands indicate the formation of an internal hydrogen bond with the molecules' central amides (Fig. 3*D*). Furthermore, the hydroxyl of Z8539_0072 formed a hydrogen bond with the backbone nitrogen of Lys11 of a symmetry mate, which closely matches the lattice interaction seen in the initial fragment hit ZINC922 (*SI Appendix*, Fig. S5). Z8539_0072 did not show any off-target activity against either human TARG1 or MacroD2 at a concentration of 50 μM or 1 mM (*SI Appendix*, Fig. S4), indicating selectivity for the viral Mac1 protein.

Finally, we tested analogs modulating the acid-carrying indane group (Fig. 3*E*). 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 μM) for Mac1 and the crystal structure revealed a hydrogen bond between the indole amine and Leu126 (Fig. 3*F*). The lower affinity of this indole versus the parent compound may reflect the suboptimal placement of the carboxylate in the oxyanion subsite. Surprisingly, although the benzothiophene (IC$_{50}$ = 20 μM) and furan (IC$_{50}$ = 84 μM) 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. 3*G*). 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
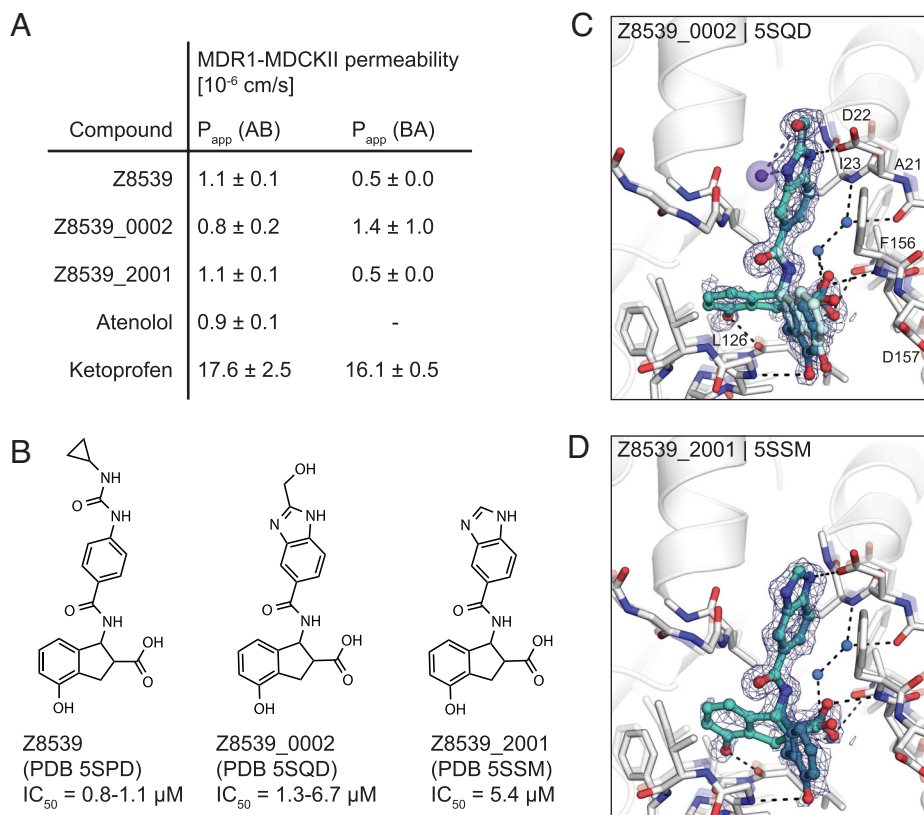
Fig. 4. Z8539 analog with enhanced cell membrane permeability. *A*) Apparent permeability (P_{app}) assayed with MDR1-MDCKII cells. Permeability was measured in apical (A)-to-basolateral (*B*) direction and vice versa. Atenolol and ketoprofen were included as control compounds. *B*) 2D structures of 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.

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 displaced phenylalanine is reported to be crucial for catalytic function of macrodomains, e.g., mutation of Phe272 in human MacroD1 reduced enzymatic activity by approximately twofold (19). This truly atomic structure–activity relationship offers an unprecedented insight into the complex nature of protein–ligand interactions.

Although this compound series led to potent molecules, the Z8539 scaffold had low cell permeability (11 nm/s) in MDCK cells (Fig. 4*A*), which likely limits its potential antiviral activity. As carboxyl bioisosteres were not readily available for make-on-demand synthesis, we attempted to increase membrane permeability by replacing the cyclopropyl-phenylurea with a benzodiazol group, which only marginally reduced the IC$_{50}$ value versus the parent urea (Fig. 4*B*). Z8539_0002 contains a methanol group that was designed to maintain the bidentate interaction with Asp22; however, the crystal structure instead indicated a hydrogen bond formed with the symmetry mate in the crystal lattice (Fig. 4*C*). Removing the alcohol group did not affect the binding affinity (Z8539_2001, Fig. 4*D*) and both compound analogs were selective for viral Mac1 over both tested human macrodomains (*SI Appendix*, Fig. S4). However, despite lacking the urea, these compounds had similar P_{app} values compared to Z8539 (Fig. 4*A*), indicating that the carboxylate is most likely responsible for the observed low cell membrane permeability. Competitive substitutions of carboxylates for Mac1 inhibitors are presented at the end of the manuscript.

In summary, the in silico fragment-linking approach employed here led to a promising and potent inhibitor scaffold based on

only two fragments out of the roughly 200 fragment hits in the active site; many others remain to be considered. This method to explore the recent huge expansion of purchasable chemical space (20) may now allow the discovery of compounds that merge and minimally displace the key interactions of the parent fragments, which has previously limited fragment merging approaches. The combination of fragment-linking and large chemical library exploration might offer a pragmatic and relatively rapid strategy to generate active chemical matter for a vast group of protein targets with little to no known chemical matter.

Inhibitors by Molecular Docking. Seeking even newer chemotypes, we docked ultra-large libraries of lead-like "tangible" (make-on-demand) molecules against Mac1 (20), leveraging the hot spots revealed by the initial fragment-binding experiment (12). Molecules were screened against two different protein models, either using an ADPr-bound structure [PDB 6W02 (21)] or subsequently using a structure bound to a first-round lead-like docking hit (see below). The first screen of approximately 350 million molecules of the ZINC15 database (22), belonging mainly to libraries from Enamine and WuXi AppTec, with molecular weight ranging from 250 to 350 amu and calculated (c)logP below 3.5, was performed against the same docking template that we previously used in the computational fragment screen (ADPr-bound Mac1, PDB 6W02) (12). Molecules were targeted to the adenosine-binding pocket of Mac1; molecules that docked to form polar interactions with the adenine-recognizing residues Asp22, Ile23, and Phe156, or with residues within the phosphate-binding region such as Val49 or Ile132, were prioritized for experimental testing. Overall, 78 highly ranked molecules were selected for experimental testing, of which 22 (28%) were confirmed to bind Mac1 in crystallographic soaking screens, 11 (14%) showed binding in the HTRF assay at
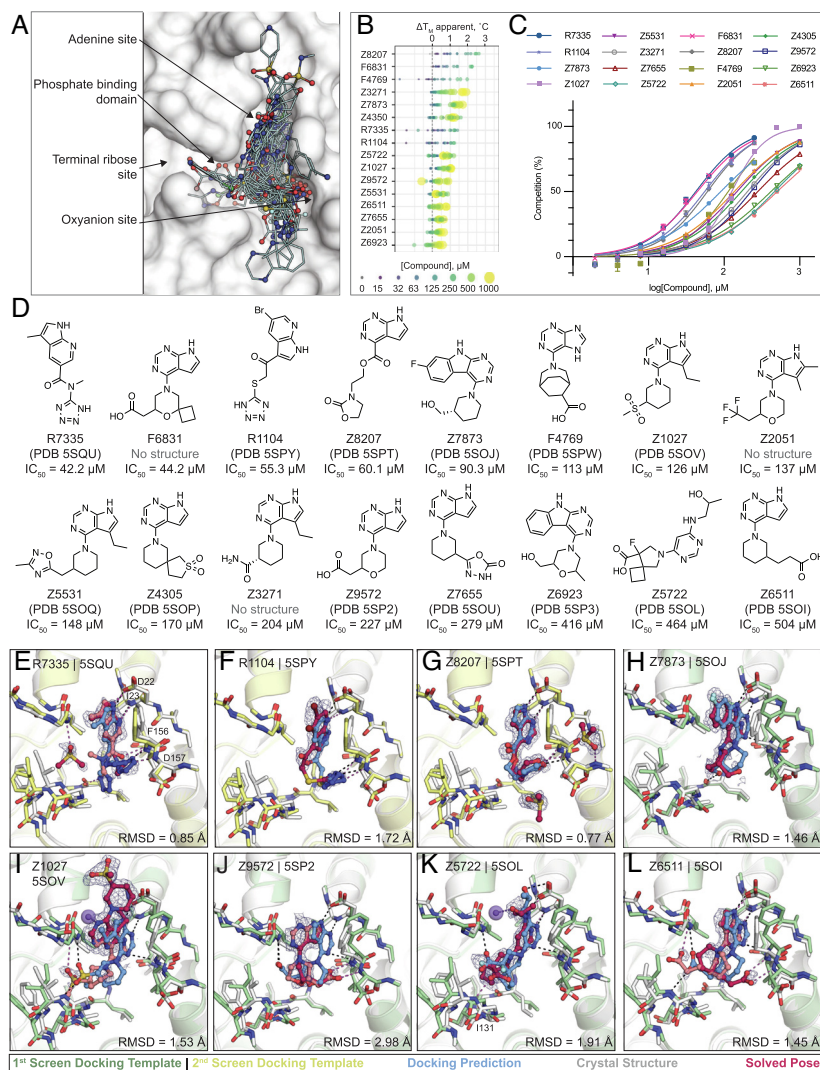
Fig. 5. Large scale docking targeting the adenosine site of Mac1. *A*) Binding poses of 47 docking hits confirmed by X-ray crystallography. The ADPr-bound structure of Mac1 (PDB 6W02) is shown with a white surface. *B*) Thermal upshifts measured by DSF. Data are presented for three technical replicates. *C*) HTRF-based peptide displacement dose–response curves. Data are presented as the mean ± SEM of at least two 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, Z1027, Z9572, Z5722, and Z6511, respectively. The protein structure used in the first docking screen is shown in green, the structure from the second screen is colored yellow. The predicted binding poses are shown in blue. Protein crystal structures are shown in gray and the solved binding poses are shown in red, with alternative ligand conformations colored salmon. Hydrogen bonding interactions between ligands and the Lys11 backbone nitrogen of a symmetry mate are shown with purple dashes/spheres. Hungarian RMSD between the docked and solved ligand poses were calculated with DOCK6. PanDDA event maps are shown for each ligand (contoured at 2 σ).

concentrations below 1 mM, and 30 (38%) revealed statistically significant thermal upshifts of ≥0.5 °C in DSF (Dataset S1).

In a second docking campaign, scoring parameters were optimized based on the results from the computational and crystallographic fragment screens as well as the first lead-like docking campaign (23). Here, the crystal structure of Mac1 in complex with Z6511 (PDB 5SOI, Fig. 5*L*) was used and the docking parameters were calibrated to ensure higher ranking of 172 previously confirmed fragment hits against a background of 2,384 molecules (mostly fragments) that did not bind to Mac1 in the crystal soaking experiments. Compared to the first docking model, this screen better ranked acidic compounds interacting with the oxyanion subsite (*Methods*). Approximately 300 million compounds were docked, including ca. 250 million neutral and anionic compounds with molecular weights between 250 and 350 amu and clogP below 3.5 from the ZINC15 library (22), and 50 million compounds from in-house virtual anion libraries (with molecular weights between 250 and 400 amu) containing additional, mostly negatively charged molecules from the Enamine REAL database (14). From among the top-ranking molecules, 46 were obtained

from Enamine, 25 (54%) of which were confirmed to bind Mac1 by X-ray crystallography, five (11%) showed activity in the HTRF-binding assay at concentrations below 250 μM, and eight (18%) were classified as hits in the DSF experiment (Methods).

In summary, 124 molecules were selected from virtually screening more than 400 million distinct molecules in lead-like chemical space, finding 50 Mac1 ligands (40% hit rate) (Fig. 5). Of these, 47 were confirmed by crystallographic screening and 13 showed measurable binding in the HTRF-based peptide displacement assay with IC$_{50}$ values ranging from 42 to 504 μM. Only three molecules that showed ADPr-peptide competition in the HTRF assay were not confirmed by X-ray crystallography (F6831, Z2051, and Z3271). The seemingly much higher hit-rate in the crystallographic soaking versus the HTRF-based peptide displacement experiments likely reflects the higher compound concentrations used in crystal soaking (10 to 20 mM) compared to the highest tested concentration in the HTRF-based assay, e.g., 1 mM in the first docking campaign and 250 μM for the second campaign. The failure of crystallography to detect three of the compounds that showed peptide displacement in the HTRF assay could be due to the crystal lattice biasing protein

conformations that disfavor compound binding (24). Thirty-eight compounds showed significant thermal upshifts of more than 0.5 °C in DSF (Dataset S1), thereby compounds with activity in the HTRF assay often had upshifts of >1 °C. Ten compounds were confirmed by all the three techniques.

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 (*SI Appendix*, Fig. S6 and Dataset 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) RMSD below 2 Å (25) was achieved for molecules with measurable binding affinity (e.g., R7335, R1104, Z8207, Z7873, Fig. 5 D–G and Dataset 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) (Dataset S4).

Although many different headgroups for the adenine subsite were explored among docking hits (Dataset S4), molecules that were active in the peptide displacement assay typically shared a pyrrolopyrimidine scaffold forming hydrogen bonds with Asp22

and Ile23 and stacking with Phe156, e.g., R7335 (PDB 5SQU), Z8207 (PDB 5SPT), Z6511 (PDB 5SOI) (Fig. 5 C, D, F–K). Two compounds, Z7837 (PDB 5SOJ) (Fig. 5G) and Z6923 (PDB 5SP3), extend the bicyclic purine headgroups into tricyclic pyrimidoindole scaffolds revealing moderate IC$_{50}$ values of up to 90 μM, indicating favorable shape complementarity of larger segments in the adenine subsite compared to the nucleobase of ADPr. Of note, similar to what we observed in the fragment screen, four adenine-containing compounds (Z1211, Z4827, Z0893, and Z0078) were not correctly synthesized and showed alkyl derivatives from the N3 rather than the intended N9 nitrogen in their corresponding crystal structures (Datasets S1 and S4) (12).

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

Ligand-Mediated Stabilization of Alternative Protein Conformations. Surprisingly, in the crystal structures of three docking hits, namely Z4305 (PDB 5SOP, IC$_{50}$ = 170 μM), F4769 (PDB 5SPW, IC$_{50}$ = 113 μM), and Z5531 (PDB 5SOQ, IC$_{50}$ = 148 μM), the compounds appear to stabilize alternative, open states of the phosphate-binding region, wherein the loop formed by residues Leu127 to Pro136 adopts an everted conformation relative to the apo structure (Fig. 6 A–C). Compared to the previously described structures of Mac1 bound to Z8539_0025 or Z8539_0026 (Fig. 3G), the docking hits induced even larger rearrangements within the active site. The magnitude of the
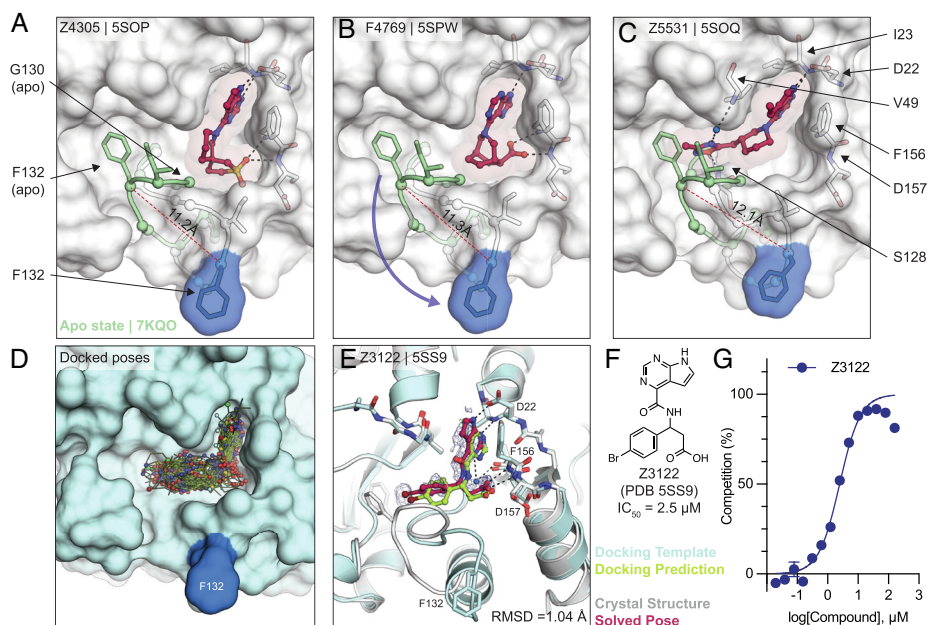

Fig. 6. Stabilization of everted phosphate-binding region by docking hits. *A–C*) The ligand-bound Mac1 crystal structures are shown in gray with Phe132 highlighted in blue. The Gly130-Phe132 loop of the Mac1 apo structure is depicted in green. Experimentally determined ligand-binding poses are shown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

loop rearrangement is surprising given the tightly packed Mac1 crystal lattice, formed prior to ligand soaking. All the three compounds occupy the adenosine subpocket, forming hydrogen bonds between their pyrrolopyrimidine-containing groups and Asp22 as well as Ile23. Z4305 and F4769 interact with the oxyanion subsite via sulfone and carboxylic acid, respectively (Fig. 6 A and B). Both compounds stabilize the same loop rearrangement in which the Cα of Phe132 is displaced by 11 Å versus the canonical closed state, which does not seem to accommodate the rigid and large nonaromatic cyclic moieties of the molecules, which would clash with Gly130. Z5331 stabilized a similar everted loop conformation (Fig. 6C). Whereas Z5331 does not interact with the oxyanion subsite, it inserts methyl-oxadiazole into the phosphate-binding region, forming direct and water-mediated hydrogen bonds with Ser128 and Val49, respectively (Fig. 6C). As opposed to Z4305 and F4769, the central piperidine of Z5531 does not clash with Gly130; however, its methyl-oxadiazole would clash with Phe132 in the apo form. A similar conformational change in the Phe132-containing loop was observed for the merged fragment Z8580 (*SI Appendix,* Fig. S7). The observed ligand-induced flexibility within the active site of Mac1 may hint to a catalytic mechanism requiring conformational flexibility to efficiently bind, cleave, and release ADPr from different target proteins (13, 26).

Docking to Everted Protein Conformation. To investigate the potential ligandability of the everted Mac1 conformation, we virtually screened roughly 60 million anionic compounds of the ZINC22 virtual library [https://cartblanche22.docking.org (27)] against the open state structure discovered in complex with Z4305. Ligands of this open state are predicted to bind with similar headgroups in the adenine site as closed state ligands, including polar interactions with Asp22 and Ile23, and stacking with Phe156. In addition, compared to the closed state, Ser128 was more solvent exposed and was therefore targeted by molecules selected from this docking screen. Interactions with these three anchor points (Asp22, Ile23, and Ser128) were used to select molecules for experimental testing, leading to a final set of 56 molecules that were synthesized by Enamine. On testing, 22 of these (39%) bound to Mac1 in crystal soaking experiments, of which five showed activity in the HTRF-based peptide displacement assay. While docking generated more favorable scores for the molecules against the open state than the closed state (Dataset S1), in the crystal structures, all 22 hits bound to the closed state (Dataset S4). Still, among the five in-solution hits, Z3122 (PDB 5SS9, see Fig. 6F) had an IC$_{50}$ of 2.5 μM against Mac1 and had no measurable activity against the human macrodomains TARG1 or MacroD2 at 160 μM (*SI Appendix,* Fig. S8), offering yet another promising, selective scaffold for future optimization.

Structure-Based Optimization of Docking Hits. To improve the affinity of initial docking hits, we explored combinations of molecular substructures bound at different subsites, templated by their crystal structures. The fluoro-pyrimidoindole of Z7873 (PDB 5SOJ), occupying the adenine subsite, was introduced into docking hits with mainly bicyclic purine scaffolds (e.g. Z9572, Z6511, Z5531) or combined with the spiro-octane-carboxylic acid of Z5722 (Fig. 7A). Nine analogs designed with this strategy were accessible in the Enamine REAL database and were synthesized for testing against Mac1. Of these nine, seven were confirmed to bind crystallographically, five were active in the DSF assay (Dataset S1), and four bound in the HTRF assay (*SI Appendix,* Fig. S8). Low micromolar affinities were measured for LL1_0023 (PDB 5SQO, IC$_{50}$ = 6-10 μM) and LL1_0014 (PDB 5SQ3, IC$_{50}$ = 16-29 μM), both containing the pyrimidoindole headgroup to occupy the adenine subsite and placing carboxylic acid in the phosphate-binding region (Fig. 7 B–D). Both compounds showed 50% displacement of the ADPr-conjugated peptide when tested against TARG1 and MacroD2 at 1 mM, whereas only LL1_0023 was active against TARG1 at 50 μM (*SI Appendix,* Fig. S6). Compared to the Z8539 scaffold, LL1_0023 was 3-fold more permeable in MDR1-MDCKII cells.

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 (Dataset S5), while nine analogs had IC$_{50}$ values below 200 μM in the ADPr-peptide displacement assay (*SI Appendix,* 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 μM, Fig. 7E), or replacing the cyclobutane with oxetane (LL123_0031, PDB 5SRB, IC$_{50}$ = 23 μM). In addition,
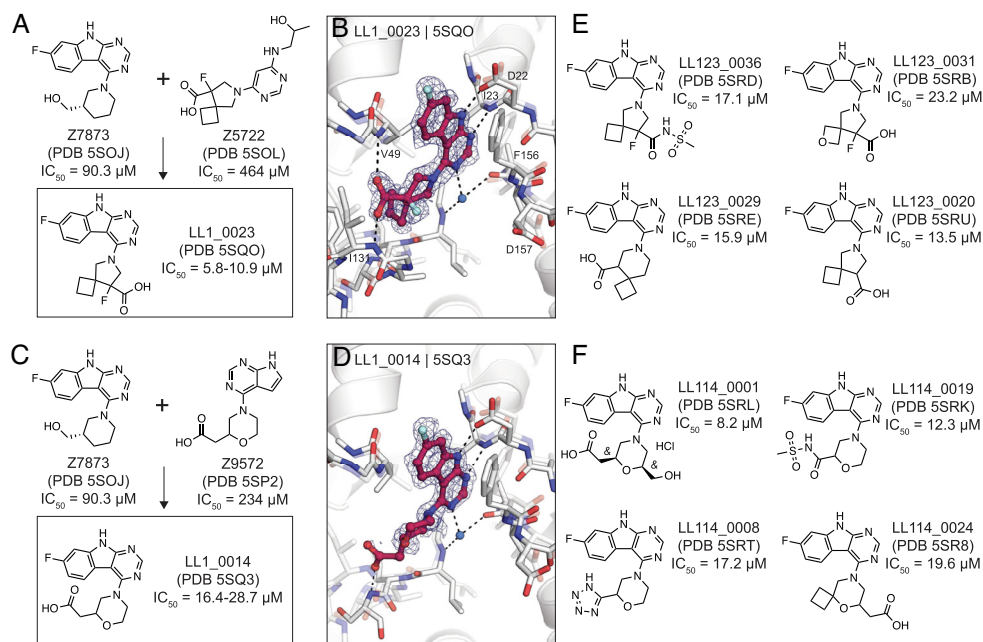

Fig. 7. Structure-based optimization of docking hits. A) Design of LL1_0023. B) X-ray crystal structure of LL1_0023. The PanDDA event map is shown around the ligand (contoured at 2 σ). Hydrogen bonds are shown with dashed black lines. C and D) Design and X-ray crystal structure of LL1_0014, respectively. E and F) Selected analogs of LL1_0023 and LL1_0014, respectively.

modifications of the compound's core spiro-octane, e.g., replacement by spiro-nonane (LL123_0029, PDB 5SRE, IC$_{50}$ = 16 μM) or removal of a fluoro group (LL123_0020, PDB 5SRU, IC$_{50}$ = 14 μM, 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 μM (Dataset 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 μM (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 μM). Furthermore, exchanging the carboxylic acid by bioisosteres such as sulfonamide (LL114_0019, PDB 5SRK, IC$_{50}$ = 12 μM) or tetrazole (LL114_0008, PDB 5SRT, IC$_{50}$ = 17 μM) 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 μM (SI Appendix, Fig. S8). Additional analogs are shown in the Supporting Information (Dataset S5).

In summary, large library docking and subsequent structure-based optimization revealed several potent inhibitors of Mac1, structurally unrelated to those obtained from fragment linking. This expanded the number of low μM scaffolds, each topologically unrelated to the others, to at least five families of molecules inhibiting a key viral enzyme for which none had been previously known.

Toward 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 nonanionic inhibitors might hold several advantages for antiviral drug discovery, especially considering drugs will need to cross cell membranes to engage viral targets residing within infected host cells. To identify neutral alternatives to carboxylate and other anions at this site, we designed a small set of analogs by linking the previously identified pyrrolopyrimidine or pyrimidoindole to small moieties bearing hydrogen bond donor or acceptor functionality (e.g., sulfones, hydroxyls, pyridines, or ketones) (Fig. 8A and Dataset S6). A total of 124 molecules (290 enantiomers) were generated in 3D conformer libraries for computational docking (see Methods). We selected 21 compounds based on the predicted docking poses of which 20 were synthesized by Enamine. Fourteen of these 20 molecules

(70%) were confirmed to bind to Mac1 by X-ray crystallography and four (20%) showed binding in the HTRF-based assay.

Promisingly, SRH-0015 (PDB 5SR0, IC$_{50}$ = 132 μM, MW = 232 amu), notably active for its small size, placed a hydroxyl group toward the oxyanion subsite, mimicking the placement of a ribose-hydroxyl group of ADPr (Fig. 8 B and C). While the crystal structure of Mac1 in complex with ADPr revealed a water-mediated hydrogen bond between the corresponding ADPr-hydroxyl and the oxyanion site, the structure of the Mac1-SRH-0015 complex does not suggest direct or water-mediated hydrogen bonding (Fig. 8C). The most promising analogs from this series were LRH-0008 (IC$_{50}$ = 13.4 μM) and LRH-0003 (PDB 5SRY, IC$_{50}$ = 1.7 μM) (Fig. 8 B–D). These compounds contain fluoro-pyrimidoindole headgroups joined to 2-aminocyclopentan-1-one or 1-aminopyrrolidin-2-one rings, respectively. The crystal structure of LRH-0003 bound to Mac1 revealed favorable placement of its hydrazide carbonyl function at the oxyanion site, enabling simultaneous hydrogen bonding to both NH groups of Phe156 and Asp157 (Fig. 8D). The enhanced potency of LRH-0003 versus LRH-0008 is consistent with a stronger hydrogen bonding interaction in the former, given the greater basicity of the hydrazide carbonyl present in LRH-0003 as compared to the ketone in LRH-0008. Notably, the similar anionic analog LRH-0021 (PDB 5SRZ, Dataset S6) was equipotent to LRH-0003, indicating that neutral compounds can indeed offer competitive alternatives to anionic Mac1 ligands. Encouragingly, both LRH-0003 and LRH-0021 obtained high permeability values in MDR1-MDCKII cell-based assays of 138 and 120 nm/s in apical to basal and 243 and 91 nm/s in basal to apical direction, respectively. Thereby, the carboxylate of LRH-0021 may form an internal hydrogen bond to the compound's central amine group, leading to improved membrane permeability. Although the anionic compound LRH-0021 showed binding to TARG1 at 160 μM, neither neutral compounds LRH-0003 and LRH-0008 had measurable binding to TARG1 and MacroD2 (SI Appendix, Fig. S8), suggesting higher selectivity for the neutral isosteres.

Discussion

Like many antiviral targets to emerge from SARS-CoV-2, Mac1 is both highly attractive and challenging. While animal studies in SARS have highlighted its crucial role in viral pathogenesis, there were no reliable chemical tools, or really inhibitors of any kind, for the enzyme. Fortunately, Mac1 crystallized readily and diffracted to ultra-high resolution (often better than 1 Å), supporting fragment-based exploration of its recognition determinants, both empirically and computationally (12). Capitalizing on this, over 230 fragment structures were determined. The binding poses of the ligands tiled the active site of the enzyme, but despite often
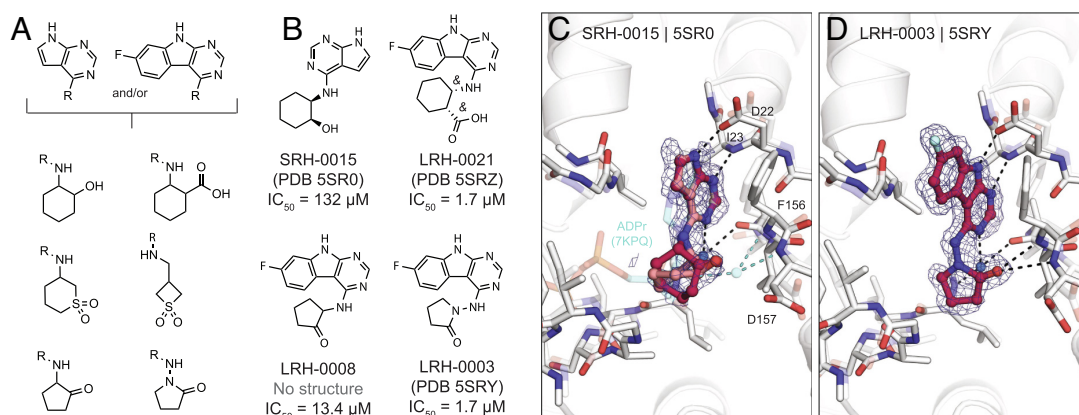

Fig. 8. Probing neutral functional groups in the Mac1 oxyanion subsite. A) Design strategy of analog set. B) Chemical structures of most potent hits. C) Crystal structure of Mac1 bound to SRH-0015. ADPr and the water-mediated hydrogen bond to the oxyanion subsite are shown for reference (PDB 7KQP, transparent cyan sticks/spheres). Both of the trans stereoisomers were modeled: the (S,R) is colored dark red and the (R,S) isomer is colored salmon. PanDDA event maps are shown around the ligand (blue mesh contoured at 2 σ). D) Crystal structure of the Mac1-LRH-0003 complex.

favorable ligand efficiencies, none of the fragments had affinities more potent than 180 µM. Here, we built on the molecular determinants revealed by the fragment structures to discover potent, selective, and cell-permeable molecules, making progress toward chemical probes and leads for drug development.

Four key points emerge from this effort. First, an automated fragment merging and linking strategy, allied with searches of ultra-large libraries, identified molecules that combined key groups of pairs of fragments and were readily available from make-on-demand synthesis. This led to the rapid discovery of molecules with low µM affinity that were subsequently optimized to affinities as low as 430 nM (compound Z8539_0072), an overall improvement of >400-fold compared to the best starting fragment. Second, templated again by the ligand recognition patterns revealed by the fragments, molecular docking screens found compounds with affinities down to 2.5 µM, with several in the mid-µM range that were also optimizable to the low µM. The best of these had ligand efficiencies that were measurably better than even the merged fragments. Third, while most of these molecules were anionic with high polar surface areas that reduced cell permeability, structure-based optimization found analogs with fewer hydrogen-donating groups such as ureas, alcohols, and phenols and enabled the replacement of anionic warheads with neutral ones. This suggests that it may be possible to improve cell membrane permeability for several of the scaffold classes here. Finally, these efforts occurred against an understudied target from an enzyme family without validated chemical probes, hinting at the potential of structure-based approaches to advance chemical matter against other understudied proteins.

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

One notable complication to using X-ray crystallography to screen ligands is the influence of crystal lattice interactions on ligand binding (31). Our initial fragment screen revealed that the $P4_3$ crystal form had a substantially higher hit rate compared to the C2 crystal form (24% versus 6%) (12). We partly attributed the difference in hit rates to fortuitous crystal packing in the $P4_3$ form: the backbone nitrogen of Lys11 on a symmetry mate is ideally positioned to interact with compounds binding in the adenine subsite. Indeed, 66 of the 123 fragments identified in or near the adenine subsite formed hydrogen bonds with Lys11 (12). Similarly, in the present work, several of the compounds that were identified by virtual screening, and subsequent optimization, adopted alternative conformations that were stabilized by hydrogen bonds with Lys11 (e.g., Z1027, Z9020, LL123_0006, and LL123_0016). Although one might be tempted to discard these conformations as artifacts, our current work indicates that they can be useful. One of the two fragments that were linked to create Z8539 contained a hydroxyl that formed a hydrogen bond with Lys11 (*SI Appendix,* Fig. S5). The compound lacking the hydroxyl (4-aminobenzoic acid, ZINC920) did not bind to Mac1 in the fragment screen (12). Crystal lattice interaction may explain the large difference been predicted and observed binding mode for several of the hits from virtual screening (e.g., F9046, F0346, R3575, Z6744, Z6684, Z5740, Z6689, Z6567).

We were surprised to find several ligands that induced large scale rearrangement of the active site loop consisting of residues 127 to 136 (Fig. 6 and *SI Appendix,* Fig. S7). Conformational changes involving Ala129, Phe132, and Asn99 have been characterized in this loop in the ADPr-bound state (12) and in the ligand-free enzyme at low pH (13), but these are relatively minor compared to the 7-12 Å shifts in Phe132 seen here. Everted loop conformations have also been observed for other macrodomains, including human MacroD1 (PDB 2X47) (32) and PARP14 (PDB 5O2D) (33) (*SI Appendix,* Fig. S7). Despite the apparent flexibility of this region, our initial virtual screening campaign did not identify any compounds that stabilized the flipped conformation of Ala129 that is present in the ADPr-bound state, despite using this state as a template for docking (PDB 6W02) (Fig. 5). However, during compound optimization, several structures were determined with Ala129 in the flipped state. These included LL114_0041, which places a carboxylic acid in the phosphate-binding subsite, and LL123_0020, which stabilizes a water molecule in a similar position (*SI Appendix,* Fig. S6). A similar rearrangement in water networks was seen for the docking hit Z0828, although the shift in Ala129 was smaller (*SI Appendix,* Fig. S6). These ligands offer opportunities for structure-guided design efforts targeting the phosphate-binding subsite of Mac1.

Certain caveats merit discussion. The antiviral or immunomodulating effect of the developed compounds has not been shown. This partly reflects limitations of the molecules themselves—e.g., their current low cell permeability—but it also reflects the lack of suitable cell-based assays to monitor the effect of Mac1 inhibition on interferon signaling. The development of such assays is an urgent need in the field; currently, our only way to measure the efficacy of Mac1 inhibitors, outside of the enzyme itself, would be in vivo. Cellular target engagement experiments by means of cellular thermal shift assays (CETSA) (34) will inform whether the discovered, more permeable inhibitors bind Mac1 in a cellular environment and will be a priority as the next generation of analogs are designed. In addition, while we tested several Mac1 inhibitors for off-target activity against two human macrodomains (TARG1, MacroD2), advanced lead compounds should be tested against the other human macrodomains (e.g. MacroD1, PARG, or ARH3) to assess possible adverse side effects.

On a technical level, while hit rates of computational docking were high in the X-ray soaking assay, only a few truly potent compounds were identified in the HTRF-based binding assay. Furthermore, while many docking predicted poses corresponded well to the crystallographically determined poses, compared to the previous fragment docking screen, larger deviations between docked and crystallographic poses were sometimes observed, especially among molecules that were predicted to insert deep into the phosphate-binding pocket. Also, ligand-induced stabilization of alternative conformations of the mobile active site loop was not predicted. While the alternative conformation of Mac1 reported here is a promising target for ligand discovery, identifying alternative states prospectively and effectively

exploiting them remain as major challenges for the field. Computational approaches, including molecular dynamics simulations, can be used to identify alternative protein states (35); however, the challenge of deciding which alternative states should be targeted, and how to use them to drive new ligand discovery, remains. This is highlighted by the similarity of docking scores of two everted-state stabilizers (Z4305 and Z5531) when docked to both the closed and everted Mac1 structures (Dataset S1), which suggests that our docking approach would have difficulty identifying the correct protein conformation. Moreover, while docking against a Mac1 structure with the everted Phe132 loop conformation (PDB 5SOP) led to a potent 2.5 μM inhibitor (Z3122), the Mac1-Z3122 crystal structure showed binding in the closed state (Fig. 6E). In addition to shortcomings of computational docking, our fragment-linking strategy relied on the access to chemicals mimicking theoretically linked scaffolds. In our case, the purchasable analogs offered promising templates; however, some differed noticeably from the initial model, e.g., they replaced a central hydrogen bond acceptor (ester group) with a donor (amide group). Although this exchange seemed actually beneficial in our Z8539 series, similar changes might lead to loss of activity in other cases.

These caveats do not diminish the central observations of this study. From an initial mapping of the Mac1 binding site with >230 fragment crystal structures (12), fragment-linking and -merging led to compounds that bound >400-fold better than the best fragment. The same mapping identified hot spots that supported ultra-large library docking that identified mid- and low-μM binders falling into still newer families. Overall, the determination of 150 Mac1-ligand crystal structures supported the discovery and optimization of 19 low- and sub-μM compounds falling into eight different scaffolds and chemotypes, while another 28 compounds in 11 scaffolds were discovered in the 10 to 50 μM range. While these compounds retain permeability liabilities, structure-based optimization suggests routes to improving their physical properties, including by reducing hydrogen bond donors and swapping anionic for neutral warheads, without substantial loss of affinity for the enzyme. From a technical standpoint, the rich structure–activity relationships combined with X-ray crystal structures for most compounds described here creates a dataset for benchmarking and improving computational techniques for drug discovery, such as free energy perturbation (36, 37). From a therapeutic perspective, the compounds and structures described in this study will support progress toward antiviral therapeutics targeting the NSP3 macrodomain of SARS-CoV-2.

Materials and Methods

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

Computational Docking. Docking calculations were performed with DOCK3.7 (23, 38) using precomputed scoring grids for rapid evaluation of docked molecules. Scoring grids for van der Waals interactions were generated with CHEMGRID and electrostatic potentials within the targeted binding pocket were calculated by numerical solution of the Poisson-Boltzmann equation with QNIFFT (39). Therefore,

AMBER united-atom charges (40) were assigned to the protein and selected structural water molecules. Ligand desolvation scoring grids were computed using Solvmap (41). Virtual compound screening libraries were selected from the ZINC15, ZINC20, and ZINC22 databases (16, 22, 27). A detailed description of the docking procedure is given in the Supplementary Materials and Methods.

Crystallographic Ligand Screening. Crystals of SARS-CoV-2 NSP3 Mac1 were grown using an expression construct that crystallized in the $P4_3$ space group, as described previously (12) (Dataset S2). Details of crystallization, ligand soaking, X-ray diffraction data collection, ligand modeling, and structure refinement are given in the *SI Appendix, Materials and Methods*.

HTRF Assay. Binding of the compounds to macrodomain proteins was assessed by the displacement of an ADPr-conjugated biotin peptide from His$_6$-tagged protein using a HTRF-technology-based screening assay which was performed as previously described (12). The expression sequences used for SARS-CoV-2 Mac1, and the human macrodomains TARG1 and MacroD2, are listed in Dataset S2. All proteins were expressed and purified as described previously for SARS-CoV-2 Mac1 (12). Further details of the HTRF protocol are given in the *SI Appendix, Materials and Methods*.

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

DSF. DSF and associated compound handling was performed as described (12), with 5 μM dye "Fluorescent Yellow" (Jacquard iDye Cat #JID1405) used in place of SYPRO Orange. Compounds were tested in triplicate, at seven concentrations in two-fold serial dilutions, at a top concentration of either 1,000 or 100 μM. To minimize experimental noise arising from plate location-based effects on Tma, all concentrations and replicates for each compound were tested within a single region of the microtiter plates, including a DMSO-only control from which ΔTmas were calculated for that compound. Data were analyzed using DSFworld (42) by fitting raw RFU values from 25 to 85 °C to the second DSFworld model (single transition with initial decay). For each compound, to determine statistical significance of the dose responsive thermal shift, the Spearman's coefficient was calculated between compound concentration and ΔTma. A "DSF positive" was defined as any compound which exhibited both i) thermal upshift of ≥0.5 °C from the DMSO control at any tested concentration, and ii) dose responsiveness, defined as a positive Spearman estimate and P value ≤ 0.05, calculated between compound concentration and ΔTma. All data used to determine temperature shifts by DSF are included in Dataset S7, and the mean Tmas and SDs are included in Dataset S8.

MDR1-MDCK II Cell Permeability. Permeability of compounds was assessed using canine MDR1 knockout and human MDR1 knockin MDCKII cells (MDR1-MDCKII) (Sigma-Aldrich, MTOX1303) in confluent monolayers expressing P-glycoprotein (P-gp) at Enamine biological services Bienta LTD (Kyiv, Ukraine). Additional information is provided in the *SI Appendix, Materials and Methods*.

Data, Materials, and Software Availability. All study data are included in the article and/or SI Appendix.

Author affiliations: [a]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, CA 94158; [b]Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA 94158; [c]Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK; [d]Wellcome Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK; [e]National Institute for Health Research Oxford Biomedical Research Centre, Oxford OX4 2PG, UK; [f]Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA 94158; [g]Institute for Neurodegenerative Disease, University of California San Francisco, San Francisco, CA 94158; [h]Chemistry and Chemical Biology Graduate Program, University of California San Francisco, San Francisco, CA 94158; [i]Department of Pharmaceutical Chemistry and Small Molecule Discovery Center, University of California, San Francisco, CA 94158; [j]Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot OX11 0DE, UK; [k]Research Complex at Harwell Harwell Science and Innovation Campus, Didcot OX11 0FA, UK; [l]Enamine Ltd., Kyiv 02094, Ukraine; [m]Taras Shevchenko National University of Kyiv, Kyiv 01601, Ukraine; [n]Chemspace, Kyiv 02094, Ukraine; [o]Centre for Medicines Discovery, University of Oxford, Headington OX3 7DQ, UK; [p]Structural Genomics Consortium, University of Oxford, Headington OX3 7DQ, UK; and [q]Department of Biochemistry, University of Johannesburg, Auckland Park 2006, South Africa

1. W. Yan, Y. Zheng, X. Zeng, B. He, W. Cheng, Structural biology of SARS-CoV-2: Open the door for novel therapies. *Signal Transduct. Target Ther.* **7**, 26 (2022).
2. A. K. L. Leung, D. E. Griffin, J. Bosch, A. R. Fehr, The conserved macrodomain is a potential therapeutic target for coronaviruses and alphaviruses. *Pathogens* **11**, 94 (2022).
3. H. Yang, Z. Rao, Structural biology of SARS-CoV-2 and implications for therapeutic development. *Nat. Rev. Microbiol.* **19**, 685–700 (2021).
4. W. Fu et al., The search for inhibitors of macrodomains for targeting the readers and erasers of mono-ADP-ribosylation. *Drug. Discov. Today* **26**, 2547–2558 (2021).
5. J. G. M. Rack et al., Viral macrodomains: A structural and evolutionary assessment of the pharmacological potential. *Open Biol.* **10**, 200237 (2020).
6. L. C. Russo et al., The SARS-CoV-2 Nsp3 macrodomain reverses PARP9/DTX3L-dependent ADP-ribosylation induced by interferon signaling. *J. Biol. Chem.* **297**, 101041 (2021).
7. Y. M. O. Alhammad et al., The SARS-CoV-2 conserved macrodomain is a Mono-ADP-Ribosylhydrolase. *J. Virol.* **95**, e01969-20 (2021).
8. A. R. Fehr et al., The conserved coronavirus macrodomain promotes virulence and suppresses the innate immune response during severe acute respiratory syndrome coronavirus infection. *MBio* **7**, e01721 (2016).
9. L. M. Sherrill et al., Design, synthesis and evaluation of inhibitors of the SARS-CoV-2 nsp3 macrodomain. *Bioorg. Med. Chem.* **67**, 116788 (2022).
10. A. Roy et al., Discovery of compounds that inhibit SARS-CoV-2 Mac1-ADP-ribose binding by high-throughput screening. *Antiviral Res.* **203**, 105344 (2022).
11. S. T. Sowa et al., A molecular toolbox for ADP-ribosyl binding proteins. *Cell Rep. Methods* **1**, 100121 (2021).
12. M. Schuller et al., Fragment binding to the Nsp3 macrodomain of SARS-CoV-2 identified through crystallographic screening and computational docking. *Sci. Adv.* **7**, eabf8711 (2021).
13. G. J. Correy et al., The mechanisms of catalysis and ligand binding for the SARS-CoV-2 NSP3 macrodomain from neutron and x-ray diffraction at room temperature. *Sci. Adv.* **8**, eabo5083 (2022).
14. O. O. Grygorenko et al., Erratum: Generating multibillion chemical space of readily accessible screening compounds. *iScience* **23**, 101873 (2020).
15. S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531–1534 (1996).
16. J. J. Irwin et al., ZINC20–A free ultralarge-scale chemical database for ligand discovery. *J. Chem. Inf. Model.* **60**, 6065–6073 (2020).
17. N. M. Pearce et al., A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat. Commun.* **8**, 15123 (2017).
18. C. Yung-Chi, W. H. Prusoff, Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol.* **22**, 3099–3108 (1973).
19. X. Yang et al., Molecular basis for the MacroD1-mediated hydrolysis of ADP-ribosylation. *DNA Repair* **94**, 102899 (2020).
20. J. Lyu et al., Ultra-large library docking for discovering new chemotypes. *Nature* **566**, 224–229 (2019).
21. K. Michalska et al., Crystal structures of SARS-CoV-2 ADP-ribose phosphatase: From the apo form to ligand complexes. *IUCrJ* **7**, 814–824 (2020).
22. T. Sterling, J. J. Irwin, ZINC 15–Ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).
23. B. J. Bender et al., A practical guide to large-scale docking. *Nat. Protoc.* **16**, 4799–4832 (2021).
24. P. M. Collins et al., Achieving a good crystal system for crystallographic X-ray fragment screening. *Methods Enzymol.* **610**, 251–264 (2018).
25. W. J. Allen, R. C. Rizzo, Implementation of the Hungarian algorithm to account for ligand symmetry and similarity in structure-based design. *J. Chem. Inf. Model.* **54**, 518–529 (2014).
26. Q. Zhao et al., Enhanced sampling approach to the induced-fit docking problem in protein-ligand binding: The case of mono-ADP-ribosylation hydrolase inhibitors. *J. Chem. Theory Comput.* **17**, 7899–7911 (2021).
27. B. Tingle, ZINC-22-A free multi-billion-scale database of tangible compounds for ligand discovery. ChemRxiv (2022). https://doi.org/10.26434/chemrxiv-2022-82czl (Accessed 31 October 2022).
28. M. Jaskolski, A. Wlodawer, Z. Dauter, W. Minor, B. Rupp, Group depositions to the Protein Data Bank need adequate presentation and different archiving protocol. *Protein. Sci.* **31**, 784–786 (2022).
29. W. Jahnke et al., Fragment-to-lead medicinal chemistry publications in 2019. *J. Med. Chem.* **63**, 15494–15507 (2020).
30. I. J. P. de Esch, D. A. Erlanson, W. Jahnke, C. N. Johnson, L. Walsh, Fragment-to-lead medicinal chemistry publications in 2020. *J. Med. Chem.* **65**, 84–99 (2022).
31. C. R. Søndergaard, A. E. Garrett, T. Carstensen, G. Pollastri, J. E. Nielsen, Structural artifacts in protein-ligand X-ray structures: Implications for the development of docking scoring functions. *J. Med. Chem.* **52**, 5673–5684 (2009).
32. D. Chen et al., Identification of macrodomain proteins as novel O-acetyl-ADP-ribose deacetylases. *J. Biol. Chem.* **286**, 13261–13271 (2011).
33. M. Schuller et al., Discovery of a selective allosteric inhibitor targeting macrodomain 2 of polyadenosine-diphosphate-ribose polymerase 14. *ACS Chem. Biol.* **12**, 2866–2874 (2017).
34. R. Jafari et al., The cellular thermal shift assay for evaluating drug target interactions in cells. *Nat. Protoc.* **9**, 2100–2122 (2014).
35. A. Kuzmanic, G. R. Bowman, J. Juarez-Jimenez, J. Michel, F. L. Gervasio, Investigating cryptic binding sites by molecular dynamics simulations. *Acc. Chem. Res.* **53**, 654–661 (2020).
36. L. Wang et al., Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J. Am. Chem. Soc.* **137**, 2695–2703 (2015).
37. V. Gapsys et al., Accurate absolute free energies for ligand–protein binding based on non-equilibrium approaches. *Commun. Chem.* **4**, 1–13 (2021).
38. R. G. Coleman, M. Carchia, T. Sterling, J. J. Irwin, B. K. Shoichet, Ligand pose and orientational sampling in molecular docking. *PLoS ONE* **8**, e75992 (2013).
39. K. Gallagher, K. Sharp, Electrostatic contributions to heat capacity changes of DNA-ligand binding. *Biophys. J.* **75**, 769–776 (1998).
40. S. J. Weiner et al., A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).
41. M. M. Mysinger, B. K. Shoichet, Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* **50**, 1561–1573 (2010).
42. T. Wu et al., Three essential resources to improve differential scanning fluorimetry (DSF) experiments. bioRxiv (2020). https://doi.org/10.1101/2020.03.22.002543 (Accessed 25 March 2020).