Unexpected single ligand occupancy and negative cooperativity in the SARS-CoV-2 Main protease

Simone Albani[**1,2], Elisa Costanzi**[**3,4], Gia Linh Hoang**[**5], Maria Kuzikov[6,7], Marcus Frings[8], Narjes Ansari[9], Nicola Demitri[3], Toan T. Nguyen[10], Valerio Rizzi[11], Jörg B. Schulz[5,12], Carsten Bolm[8], Andrea Zaliani[6,7], Paolo Carloni[1,5,10], Paola Storici*[3], Giulia Rossetti*[1,5,12,13]**

**** shared first authorship**
**** to whom correspondence should be sent**

[1] Institute for Neuroscience and Medicine (INM-9), Forschungszentrum Jülich, Jülich, 52425, Germany
[2] Faculty of Mathematics, Computer Science and Natural Sciences, RWTH Aachen, Aachen, 52062, Germany
[3] Elettra - Sincrotrone Trieste S.C.p.A., SS 14 - km 163, 5 in AREA Science Park, Trieste, Basovizza, 34149, Italy
[4] present address: Center for Structural Studies, Heinrich-Heine Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany
[5] JARA-Brain Institute Molecular Neuroscience and Neuroimaging, Research Center Jülich GmbH and RWTH Aachen University, Aachen, Germany
[6] Fraunhofer Institute for Translational Medicine and Pharmacology (ITMP), Schnackenburgallee 114, 22525 Hamburg, Germany
[7] Fraunhofer Cluster of Excellence for Immune-Mediated Diseases (CIMD), Theodor Stern Kai 7, 60590 Frankfurt, Germany
[8] Institute of Organic Chemistry, RWTH Aachen University, Landoltweg 1, 52074 Aachen, Germany
[9] Atomistic Simulations, Italian Institute of Technology, Via Enrico Melen, 83, 16152, Genova, Italy
[10] Key Laboratory for Multiscale Simulation of Complex Systems, and Department of Theoretical Physics, Faculty of Physics, University of Science, Vietnam National University - Hanoi, 334 Nguyen Trai street, Thanh Xuan, Hanoi 11400, Vietnam
[11] School of Pharmaceutical Sciences, University of Geneva, Rue Michel Servet 1, 1206, Genève, Switzerland
[12] Department of Neurology, Medical Faculty, RWTH Aachen University, Aachen, Germany
[13] Jülich Supercomputing center (JSC), Forschungszentrum Jülich, Jülich, 52425, Germany

Abstract

Many homodimeric enzymes tune their function by exploiting either negative or positive cooperativity between subunits. In the SARS-CoV-2 Main protease (Mpro) homodimer, the latter has been suggested by symmetry in most of the 500 reported protease/ligand complex structures solved by macromolecular crystallography. Here we apply the latter to both covalent and non-covalent ligands in complex with Mpro. Strikingly, our experiments show that occupation of both active sites of the dimer originates from an excess of ligands. Indeed, co-crystals obtained using a 1:1 ligand/protomer stoichiometry leads to single occupation only. The empty binding site exhibits a catalytic-inactive geometry in solution, as suggested by molecular dynamics simulations. Thus, Mpro operates through negative cooperativity, with asymmetric activity of the catalytic sites. This allows it to function with a wide range of substrate concentrations, making it resistant to saturation and potentially difficult to shut down - all properties advantageous for the virus' adaptability and resistance.

Introduction

A significant fraction of enzymes are homodimers with one catalytic site in each subunit[1], active only in their dimeric states [2–7]. This hints to an allosteric communication between the two sites and hence to cooperativity [8], which can be exploited for enzymatic function. The substrate affinity of a subunit upon substrate binding in the other one may increase ("*positive* cooperativity", PC), thus increasing the enzymes' sensitivity: a small change in ligand concentration gives rise to a large change in the concentration of the bound state of the protein [9]. However, the allosteric interaction between subunits following the binding of the first ligand may also decrease the affinity for the second ligand into the other subunit ("*negative* cooperativity", NC), allowing to maintain enzymatic reactivity even in an excess of the substrate. This is a crucial feature for branching points in metabolic networks, that is the case where an intermediate species is chemically made or transformed by multiple enzymatic processes [9,10]. Besides providing fundamental insight on enzymatic function, understanding the nature of cooperativity helps develop strategies for drug design [11–14].

Several types of measurements have been used to investigate cooperativity in homodimeric enzymes: (i) *detection of occupancy status of ligands in the active sites:* the presence of both subunits in apo or holo form hints to PC, while presence of a ligand (substrate or inhibitor) only in one binding site suggests NC; (ii) *the ligand input-output response measure:* if a low ligand concentration leads to basically no output whilst a larger ligand content leads to almost maximal output, PC may be operative. However, if ligand depletion is considered, such response can also be characteristic of NC (especially when the ligand is appreciably depleted due to very high binding affinity) [15]. The situation is further complicated by the fact that NC cannot be distinguished from independent binding at multiple sites by equilibrium measurements [16]. These two situations are not identical over the complete time courses of the binding reaction; but so far, the proposed approaches in pre-equilibrium conditions to distinguish between a NC model and a model where independent binding to multiple binding sites occurs, can only evaluate how well the models fit the data, but not infer on the model itself [16]. (iii) *The value of the Hill Coefficient (HC)*, *detected by input/output curves' slopes*: HC greater than 1 suggests PC, whereas lower than 1 hints to NC [17,9,18]. However, this criterion has been criticized because (1) it assumes that ligands bind to enzyme simultaneously [18,19], although ligand binding can alter the subunit-dimer equilibrium becoming not simultaneous; (2) it does not consider the possibility that HC can be greater than 1 for covalent ligands, irrespectively of the nature of cooperativity[20]. (iv) *The symmetric nature of the homodimer structure:* fully symmetric subunits may be characteristic of PC while asymmetric ones (both in the apo form and in the doubly occupied form) may be specific for NC [9,21–24].

From the discussion above, it is apparent that establishing unambiguously the nature of cooperation (especially NC) may be highly non-trivial. This is the case of the SARS-CoV-

2 main protease (Mpro hereafter) [25,26], a fundamental target against the virus [27]. This enzyme is active only as a homodimer [28], with the N-finger of one monomer shaping the substrate-binding site of the other [26](Fig. 1). This suggests a cooperation between the binding sites [28]. However, the type of cooperativity has not been unambiguously demonstrated. From one hand, PC has been suggested by the following facts: (i) HC is greater than 1 [25,29,30], however this could be caused by the fact that most of its ligand are covalent binders, as well as by the fact that ligand binding might not be simultaneous [25]. (ii) Almost all of the ligand/protein complexes solved by macromolecular X-ray crystallography (MX) contain two ligands per dimer (as shown by an inspection of the 500 structures in the PDB DataBank (https://www.rcsb.org, Tab. 1 in Methods) [31], (iii) the apo-Mpro and almost all (99%) of the ligand/Mpro complex MX structures exhibit dimeric symmetry. However, these facts could be the consequence of the excess of ligands added in the crystallization procedure (saturating both active sites) that might in turn cause the protein to crystallize as a homodimer with only one monomer in the asymmetric unit (See Methods for details).

On the other hand, NC could be suggested by observing that: (i) some ligand/protein complex X-ray structures do exhibit asymmetry: namely, one subunit is not be obtained by a symmetry operator on the other one, and the crystallographic unit contains the whole functional dimer(s). However, the overall number of such structures is very small (1.8%). (ii) A symmetry-breaking process of apo-Mpro occurs once passing from the solid state to aqueous solution, as seen by long-time-scale molecular dynamics simulations [32,33]. Such symmetry-breaking has not been discussed for the doubly occupied enzyme. (iii) The enzymatic activity increases with the addition of catalytically-inactive monomers in solution for SARS-CoV Mpro, which shares 96% sequence identity [34]. However, one has to be careful in drawing conclusions from one protein to the other, as they exhibit significant catalytic differences[35].

To gain insight on the biophysics of this fascinating protein, here we attempt to establish the true nature of the enzyme's cooperativity by applying an arsenal of biophysical methods. First, we ask ourselves whether the doubly occupancy might arise by the fact that an excess of ligand is used. We address this by using MX and binding assays. Next, we investigate the impact of solvation, which leads to a loss of symmetry of the apo-protein on passing from the solid state to aqueous solution [32]. Anticipating our results, we show that in co-crystals obtained in conditions of 1:1 ligand/protomer stoichiometry, the protein features NC with only one ligand bound in one active site, possibly because binding of one ligand in site distorts the other one. This contrasts with what was found in the 500 MX ligand-bound structures solved so far, which might have been obtained in excess of ligands.

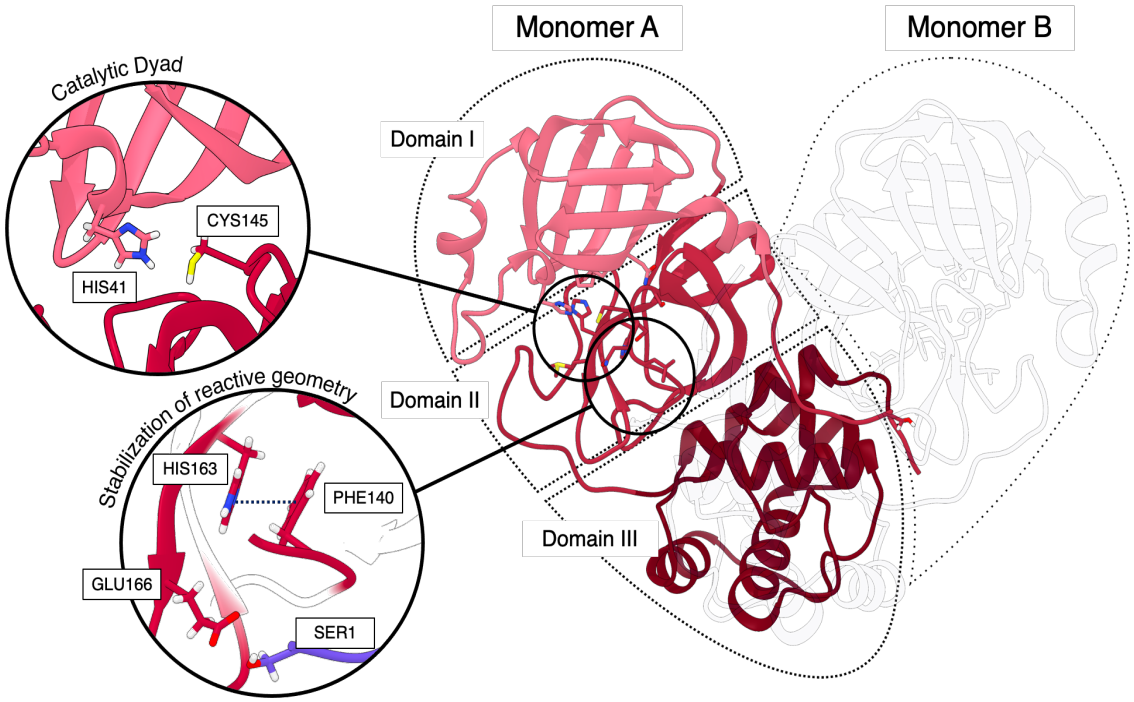


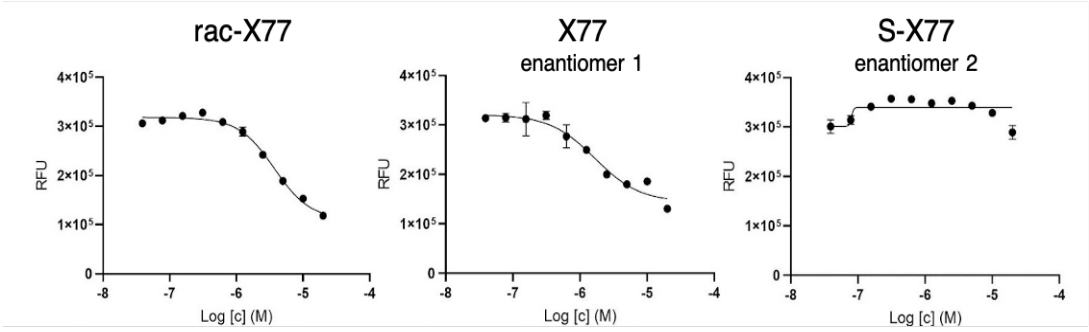
Figure 1. *Ribbon representation of Mpro's subunit "A", shown in shades of red, and "B", in the foreground, represented with gray low-opacity ribbons (PDBid 7PHZ). Each subunit consists of three domains. The first two are the chymotrypsin-like β-barrel domains I and II (residues 10 to 99 and 100 to 182, respectively) with six-stranded antiparallel β barrels that harbor the substrate-binding site between them. The catalytic center is a CYS–HIS dyad. The last domain (residues 198-303) is a globular cluster of five helices, involved in dimerization of the enzyme. The insets show details of the catalytic dyad and of the interactions that stabilize the reactive geometry and that were previously reported to be fundamental for site activation/deactivation* [32]*, namely the hydrophobic interactions between PHE140 and HIS163, and the proximity of GLU166 to the SER1 of the adjacent protomer, which allows for the formation of inter-protomer H-bonds. .*

Results and Discussion

Macromolecular crystallography and binding essays. Using non-saturating conditions, namely 1:1 and 1:2 ligand/monomer stoichiometries (LMS), we solved 4 new MX structures to be added to the ~500 already deposited Mpro/ligand complex structures, which were possibly all determined in excess of ligand and almost in their entirety exhibiting double occupancy of the ligand.

The first ligand is benzyl *N*-[(2*S*)-4-methyl-1-[[(2*S*)-4-methyl-1-[[(2*S*)-4-methyl-1-oxopentan-2-yl]amino]-1-oxopentan-2-yl]amino]-1-oxopentan-2-yl]carbamate, **MG-132** in **Chart 1**, which forms a covalent bond with CYS145, and its IC50 for MPro is 7.4 μM [36,37]. The MX structure bound to **MG-132** with double occupancy was solved previously by some of us at 1.94 Å in the C2 space group (PDBid 7NF5) and also, at 1.68 Å resolution in the P212121 space group (PDBid 7BE7), in condition of excess of ligands [36].

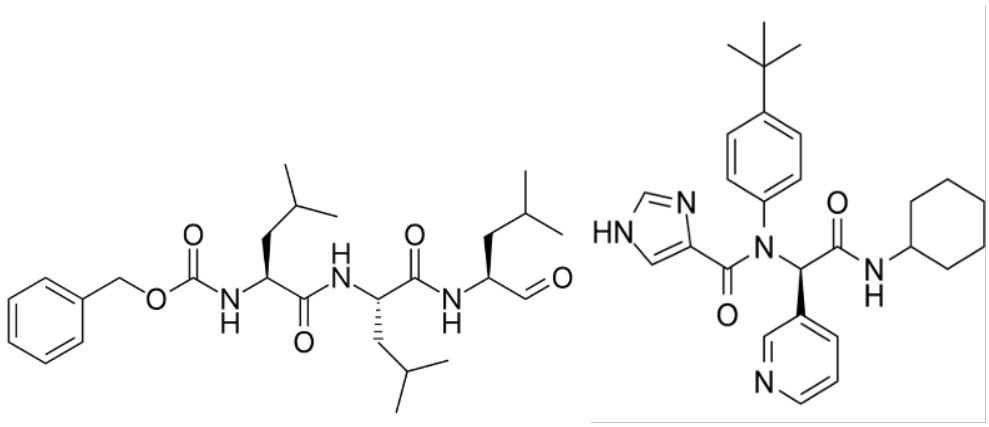
The second ligand is the R-enantiomer *N*-(4-tert-butylphenyl)-*N*-[(1*R*)-2-(cyclohexylamino)-2-oxo-1-(pyridin-3-yl)ethyl]-1*H*-imidazole-4-carboxamide, **X77** in **Chart 1**, which forms only non covalent interactions. The MX structure with double ligand occupation was reported at 2.1 Å in the C2 space group (PDBid: 6W63) and we reproduced it in our crystallization condition in spacegroup P212121 (PDBid: 7PHZ). Its inhibitory activity for Mpro, along with that of the S-enantiomer (S-**X77** hereafter) and that of the racemate (rac-**X77**), were not known when we started this study. They were measured here employing a Förster resonance energy transfer (FRET) with a dual-labeled substrate, DABCYL-KTSAVLQ↓SGFRKM-EDANS (Bachem #4045664) containing a protease specific cleavage site after the GLN. In the intact peptide, EDANS fluorescence is quenched by the DABCYL group. Its inhibitory activities are reported in Fig. 2 as dose-response curves. **X77** and S-**X77** were identified by a comparison with X-ray experiments, where the two enantiomers were separately co-crystallized with Mpro, solving 4 crystal structures, with the two enantiomers at two different concentrations (see below paragraph). The inhibitory activities are reported in Fig. 2 as dose-response curves. The racemate showed an IC50 of 3.7 μM, while that of **X77** is 1.7 μM. The S-**X77** curve could not allow IC50 calculation as no real dose-response could be measured: likely, this enantiomer could not properly bind to stop the reaction. Indeed, this was confirmed by solving the crystal structure with the S-**X77** enantiomer (see below paragraph).



	3CLpro DESY
Bottom	110451
Top	318161
LogEC50	-5.430
HillSlope	-1.620
EC50	3.719e-006

	3CLpro DESY
Bottom	143791
Top	320879
LogEC50	-5.777
HillSlope	-1.293
EC50	1.670e-006

Figure 2. *Dose Response curves for rac-**X77, X77 and** S-**X77** and in biochemical assay for Mpro. S-**X77** does not exert any inhibitory activity Mpro.*



MG-132 **X77**

Chart 1

***Complex with MG-132*.** The 150 µM protein solution was incubated with the **MG-132** inhibitor in non-saturating conditions, namely 1:1 and 1:2 LMS following our standard protocol to obtain crystals in space group P212121 with the entire dimer/a.u.. The two binding sites of our resulting crystal structures, solved at 1.85 Å and 1.60 Å resolution respectively (PDBid: 8P55, and 8P54), showed clear dissimilarities: The difference electron density map of one subunit showed a continuous positive electron density that well fit the **MG-132** moiety, while in the other subunit no residual electron density was present, suggesting an empty pocket (Fig. 3A). Even after refinement, no further density appeared in the second binding site (Fig. 3B). This establishes single occupancy of the ligand.

The binding pose of the ligand is the same as that observed in the doubly occupied enzyme previously solved [36](adduct root-mean-square deviation of 0.75 and 0.32 Å with 7NF5 chain "A" and 7BE7 chain "A", respectively). The b-factors of chain "B" (not containing the ligand) are larger than those of "A" (Fig. 4A and S2).
A fully consistent picture is obtained by letting the double occupied crystal for two months in their growing solution: co-crystals of Mpro obtained in excess of **MG-132** as described in ref. [36], after two months turned showing a positive Fo-Fc difference map corresponding to the covalently bound ligand only in one chain while the other resulted empty (Fig. S3). The single occupied site crystals diffracts to resolutions similar to those of fresh crystals, around 1.85 Å. This shows that only one binding site remains occupied if the enzyme is allowed to have enough time to let one ligand break its covalent bond and diffuse. The results strikingly differ from freshly obtained crystals prepared with the same protocol, that clearly showed to have both sites occupied [36].

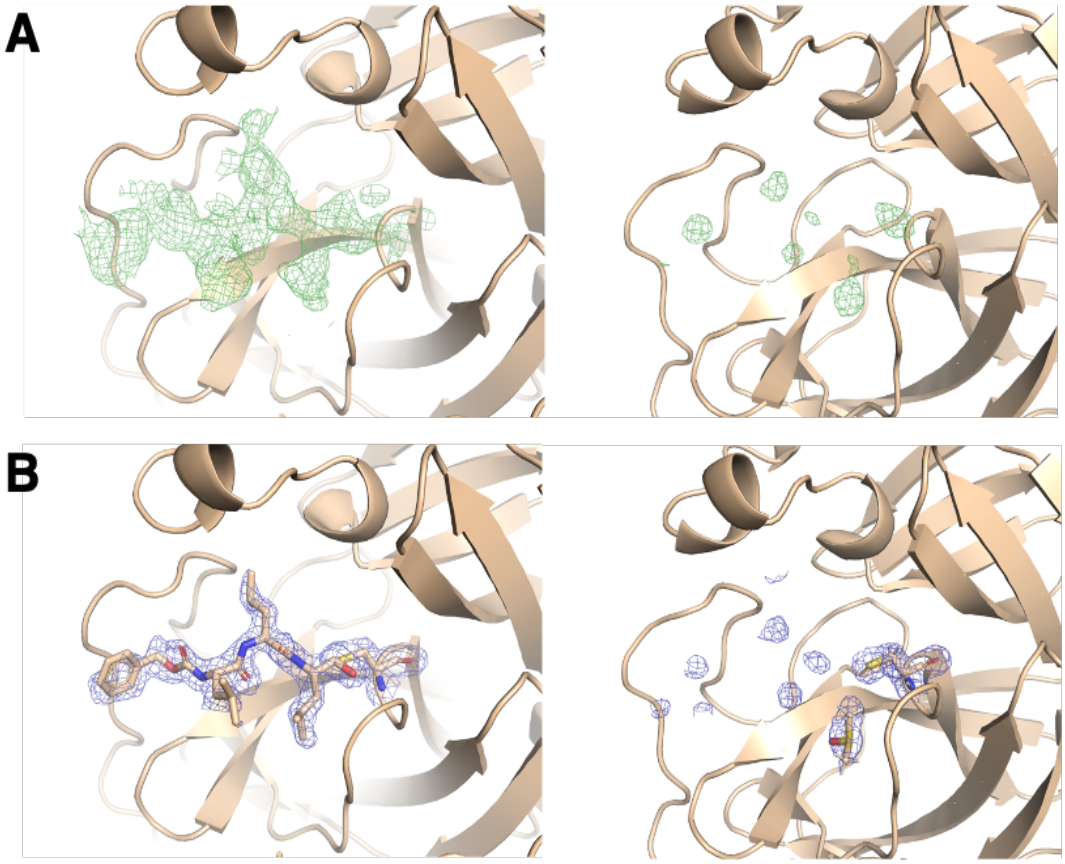


Figure 3. ***MG-132*** *complex: (**A**) Initial Fo-Fc maps contoured at 3 sigma for chain "A" (right) and chain "B" (left) of the complex obtained with ligand/protein ratio of 1:1 (PDBid 8P55). (**B**) final 2Fo-Fc maps contoured at 1 sigma for chain A (right) and chain B (left) of 8P55 (i.e. 75 μM of **MG-132**). Polder omit maps of the ligand placed in both chains were generated and confirm the results observed in the initial Fo-Fc difference maps (Fig. S1). **MG-132** is covalently bound to the sulfur atom of the catalytic CYS145. The nitrogen atoms of the backbone of this peptidic ligand act as hydrogen bond donor towards the residues HIS164 backbone and GLN189 sidechain. The last carboxyl and amide group in the ligand's backbone form two additional hydrogen bonds with the backbone of GLU 166. The terminal benzyl group is stabilized by hydrophobic contacts with the C atoms in the side chains of LEU167, PRO168, and GLN192 (Fig. 4B).*

Complex with X77**.** The MX structure was solved both in non saturating conditions (again 1:1 and 1:2 LMS, PDBid: 8P56 and 8P57), at a resolution ranging from 1.85 to 1.60 Å (Tab. S1 in the SI), and in excess of ligand (PDBid: 7PHZ). Please notice that all the structure crystallizes in the same e P212121 space group (with the entire dimer/a.u.).

As observed for **MG-132**, the ligand occupies only one active site in non-saturating conditions (Fig. 4C). The presence of the empty cavity is evident by Fourier difference map Fo-Fc, with reduced mobility in the ligand-bound subunit, again emerging by the values of the b-factors (Fig. 4C). The ligand occupies both sites when in excess (Fig. 4D) as it does in the reported X-ray structure (PDBid 6W63). However, also in this case, the b-factors of chain B are higher (Fig. 4D and Fig. S2).

As in the above case, the pose is the same as that of the structures in excess of the ligand by others (PDBid 6W63) or here (PDBid 7PHZ) (Fig. S4). In detail, **X77** carboxyl moieties accept hydrogen bonds from the backbone of the protein through the residues GLU166 and GLY143. The former residue can establish a hydrogen bond with imidazole's N_{ε} atom of **X77**. Also the pyridyl ring is stabilized by a hydrogen bond, in this case with the sidechain of HIS163. Additionally, water-mediated hydrogen bonds further contribute to the stability of the molecule (e.g. interaction between imidazole N_{γ} and HIS41 backbone) (Fig. 4B). Notably, as in the **MG-132** case, when the crystals obtained in excess of the ligand are left for two months in their crystallization solution before being flash-frozen for the diffraction experiments, the latter showed unambiguously only one occupied cavity, demonstrating that **X77** remained bound at one site while diffused from the other one (PDBid: 8P87) (Fig. S3).

Complex with S-X77. We obtained crystal structures in presence of the two enantiomers respectively at resolution 1.55 Å for enantiomer 1, and at resolution 1.47 Å for enantiomer 2. As shown in Fig. S5A, we could prove that enantiomer 1 had the R configuration by the unambiguous electron density reproducing the result obtained with the racemic mixture. In the crystal structure obtained in presence of enantiomer 2, instead we saw small blobs of electron density that could be modeled with a DMSO and water molecules (Fig. S5B). With the refinement of the structure obtained in presence of enantiomer S, small positive blobs of not modelable Fo-Fc were left. We repeated the crystallization experiments of both enantiomers using the highest reachable concentration, taking into account the DMSO tolerance of the protein. Crystallization trials were set up in presence of 5 mM inhibitor and the crystals diffracted at resolutions of 1.66 Å for enantiomer 1/R and 1.51 Å for enantiomer 2/S. For enantiomer 1/R the results reproduced the same results as for lower concentrations (Fig. S5C). Interestingly, for enantiomer 2/S, we obtained a positive Fo-Fc density that allowed the modeling of the enantiomer, as shown in Fig. S5D. Comparing the crystal structures of the R and S enantiomers, it was evident how the only functional group occupying the same position is the pyridine ring located in the S1 pocket (Fig. S5EFG). The S enantiomer is mainly anchored there to the binding site, moreover the 2Fo-Fc density is less clear for this enantiomer, and its refined B-factors are higher, overall confirming the biochemical data obtained.

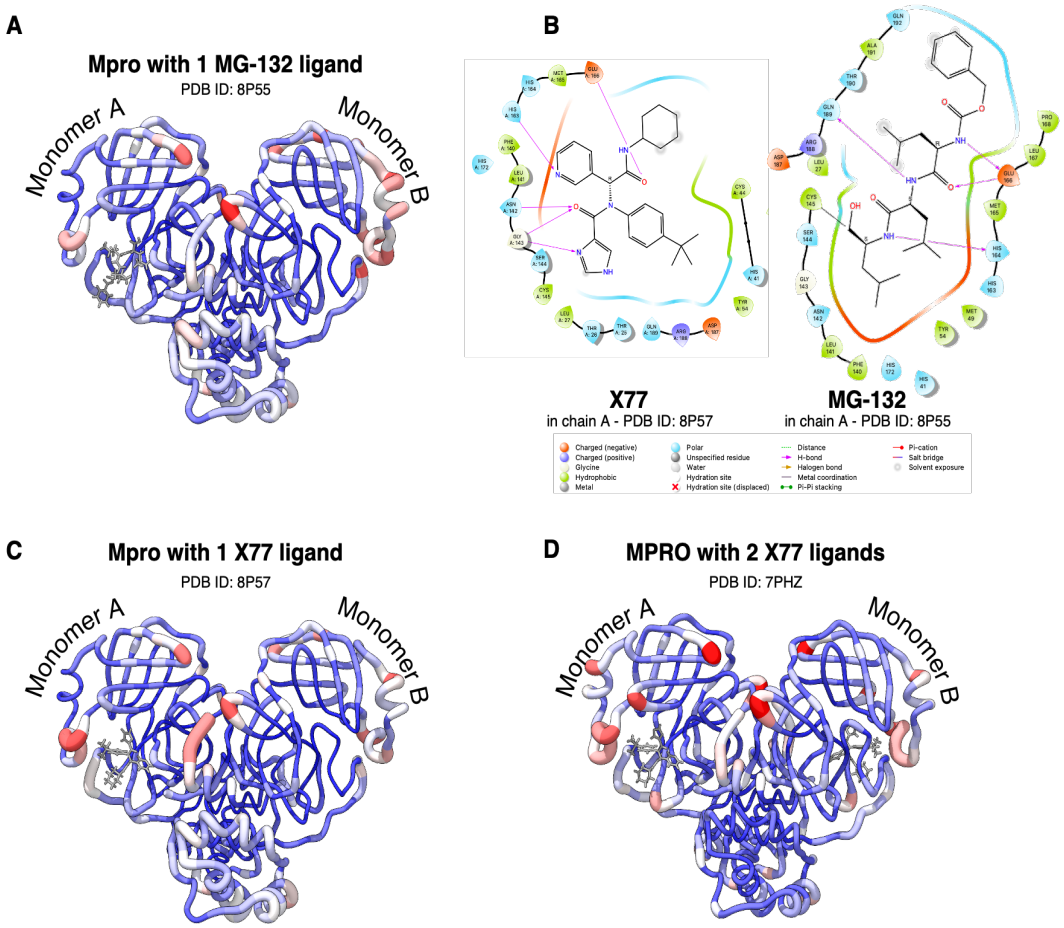


Figure 4. B-factors and binding modes in MX. (**A**) *B-factor 'cartoon putty' representation of 8P55 :(in each dimer, left = chain A, right = chain B); Pink to red colors and a wider tube indicate regions with higher B-factors, whereas shades of blue and a narrower tube indicate regions with lower B-factors. (**B**) 2D schematic representation of the interactions between Mpro and the ligands **X77** and **MG-132**, as observed in the PDB structures with ID 8P57 and 8P55. 'Cartoon putty' representation of the B-factors of structures 8P57 (**C**) and 7PHZ (**D**).*

X77/Mpro in aqueous solution. Here we use MD to investigate the structural changes of three X77/Mpro complex structures (PDBids 8P57, 7PHZ, 6W63) solved in different saturation conditions and space groups, on passing from the solid state to the aqueous solution. Specifically, we perform 500ns-long AMBER-based molecular dynamics simulations in explicit solvent of these systems. The Mpro structure and ligand pose remains stable during 500 ns of unbiased simulations for all the three simulated systems (see Fig. S6). The number of contacts between the two subunits is conserved for the systems with both cavities occupied (7PHZ, 6W63, Fig. S7), independently of the space group, while, for the single-cavity occupied system, this number increases, tightening up the subunit-to-subunit interaction (Fig. S7).

To understand how solvation can impact on the catalytic site, we next define an 'active' geometry: this features the PHE140/HIS163 intra-subunit hydrophobic contact and the inter-subunit interactions between m-shaped loop and the N-finger of the adjacent subunit (Fig. 5B) [35,38,39]. Such hydrophobic contacts of PHE140/HIS163 are analyzed in terms of the distance between the center of these two rings as a function of time (d_{CM}). The empty binding cavity (subunit B of 8P57) becomes inactive after a short simulation time (Fig. 5C): d_{CM} passes from 0.46 (SD=0.13) nm to 0.84 (SD=0.08) nm. This is not the case for all the other occupied cavities, where dCM is 0.42 nm (SD ≤ 0.03) for the overall 500ns of MD (Fig. 5D). This suggests that, in the singly occupied protein, the presence of one ligand in one subunit might induce a non active geometry in the empty cavity of the adjacent subunit. Water plays a key role for this distortion: while basically absent in the occupied cavity (total number 0.07 (SD=0.26)), as many as (total number 4.29 (SD=1.19)) are present in the empty one (Fig. 6D). As a result, HIS163 and PHE140 pi-pi stacking is broken, leading to the inactive state (Fig. 5CD). A Principal Component Analysis (PCA) on each subunit further shows that the largest scale motion of subunit A is anticorrelated to that of subunit B: the former causes the closing, the latter the opening of the binding cavity (Fig. 6ABC). Interestingly, the trend of water occupation is also observed in the fully occupied enzymes. In subunit A they have 0.12 (SD=0.34) and 0.08 (SD=0.26) number of water molecules, respectively, and in subunit B, 0.28 (SD=0.49) and 0.91 (SD=0.83), respectively. Notably, toward the end of the simulations, both subunits A are without water molecules, while both subunits B are with two water molecules on average (Fig. 6D). This is more clear cut in the asymmetric space group crystals (PDBid 7PHZ).

Next, we considered that the m-shaped loop of one subunit and the N-finger of the adjacent subunit interact via hydrogen bonds: i.e. GLU166 and PHE140 of one subunit, and SER1 of the other. Such hydrogen bonds are only preserved in the fully symmetric double occupied enzyme (symmetric space group) while it breaks for both subunits in the asymmetric space groups (P212121), either single or double occupied. These results suggest that symmetry might impact the stability of such interaction and, in turn, of the 'active' geometry (Fig. S8). However, such observation should be taken with care, since the highly flexible structure of the N-term, plus the presence of artificial capping (see Methods) might impact significantly on its dynamic behavior.

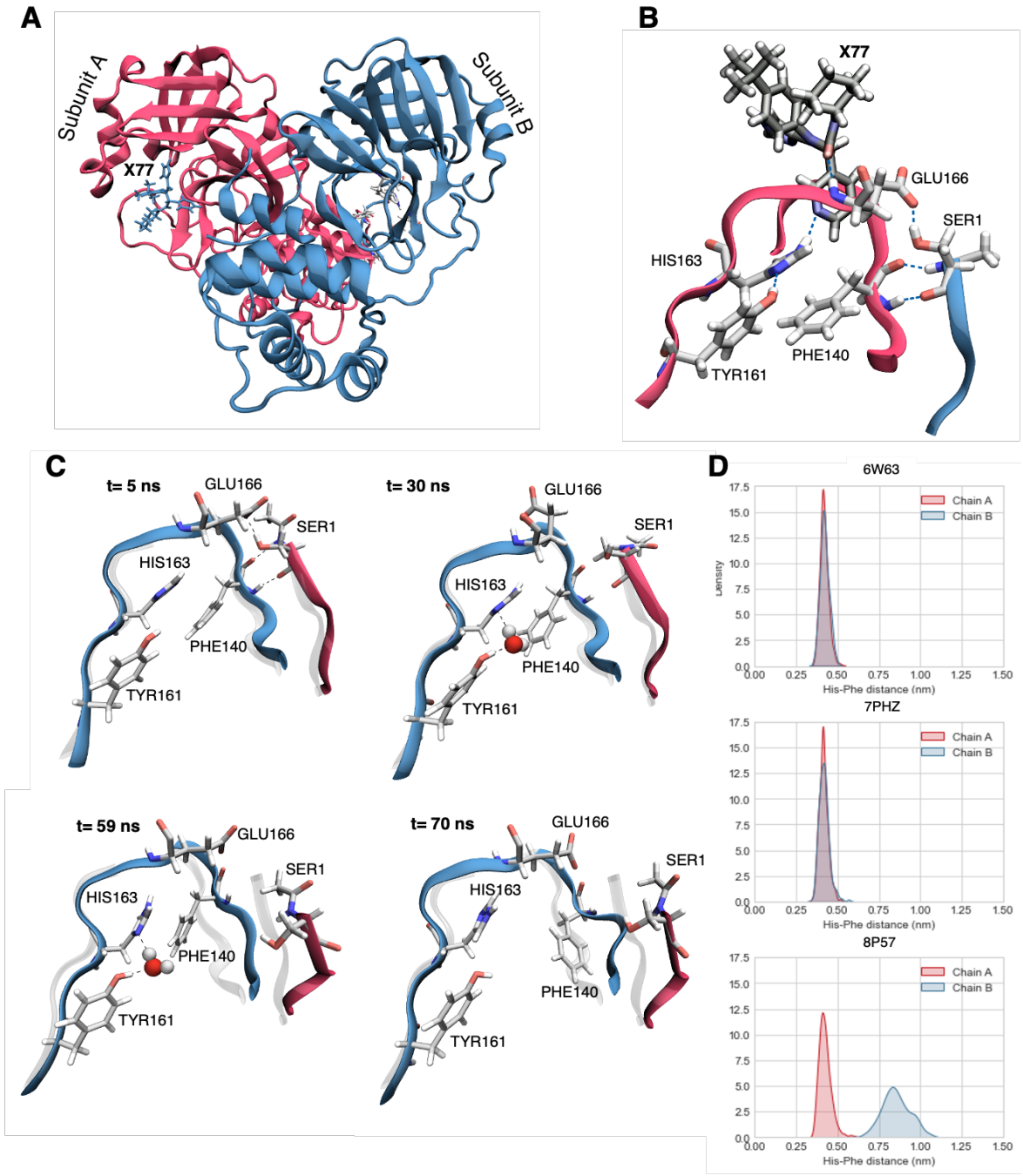


Figure 5. MD of X77/protein complexes in water solutions. (A) *Cartoon representation of Mpro structure (in red, subunit A; in blue, subunit B; the ligand **X77** is represented in blue sticks). (**B**) Hydrogen bond network in the binding site of subunit A in 8P57 after 5 ns of simulation. **(C)** Symmetry breaking happening at the level of the active site of subunit B in the unbiased simulation of 8P57. Residues that are relevant to the process are represented with gray sticks. One water molecule enters the binding site forming a bridge between HIS163 and TYR161. When the molecule exits the binding site, the hydrophobic contact between HIS163 and PHE140 is broken and the binding site inactivated. (**D**) Distribution of distance between HIS163 and PHE140 rings during the last 400 ns of simulation of 6W63, 7PHZ, 8P57.*

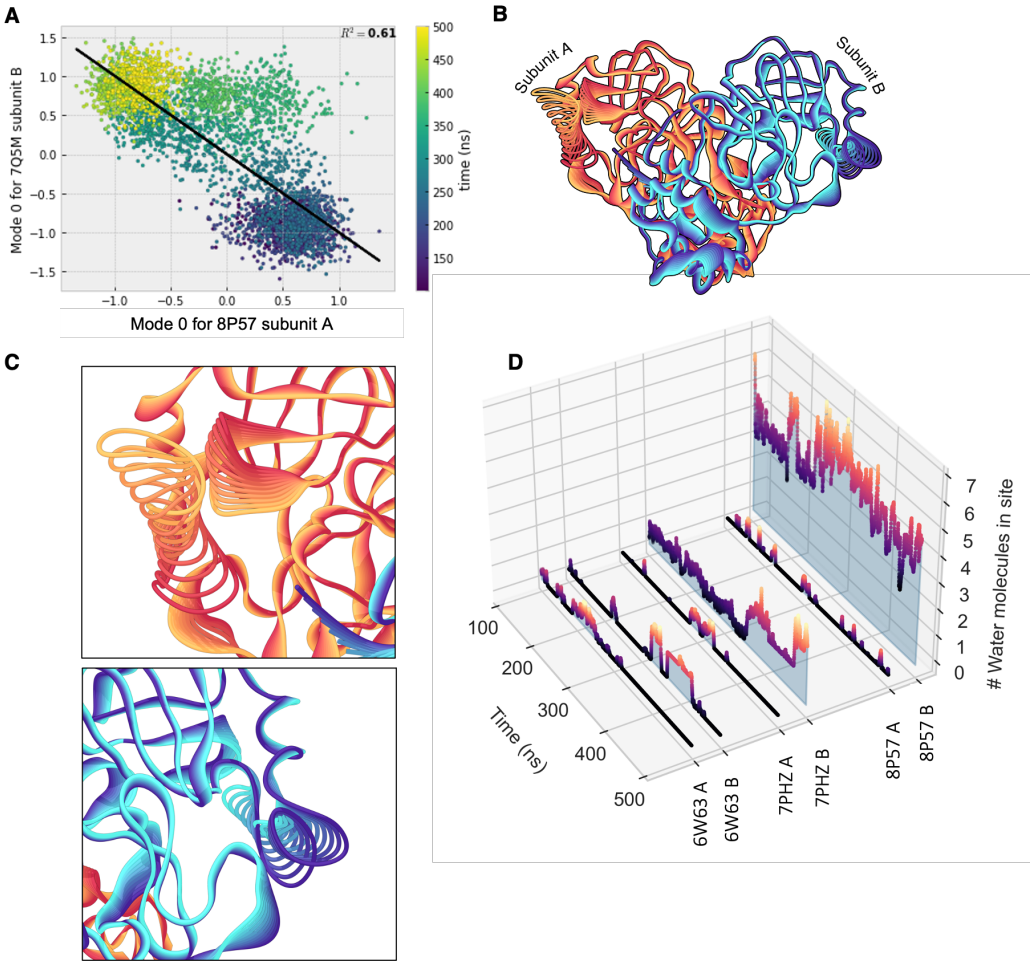


Figure 6 (A-C): PCA results for the singly occupied protein complex (8P57): ***(A)*** *Values of the trajectories of subunits A and B projected on the eigenvector of the first principal component and their correlation; the analysis was performed also for the double occupied enzymes, but no clear correlation was found (see Figure S9) (**B**) Structure of the protein deformed along the first eigenvector of the first principal component and (**C**) details of the binding sites. Subunits A and B are colored with a gradient from yellow to dark red, and from dark blue to cyan, respectively. The gradient from a light to a dark color is inverted in subunit B to reflect the anti-correlation shown in panel A. (**D**) Number of water molecules in the binding sites in subunits A and B of* 6W63, 7PHZ, 8P57 *during our 500-ns MD simulation.*

In conclusion, we observe a concerted opening of one site, while closing off the other one in the single occupied protein. In addition, the cavity which is not occupied (in subunit B) is highly hydrated in contrast to the other one. This latter trend is observed (albeit to a lesser extent) also for the doubly occupied enzymes.

A comment on saturating conditions. S-**X77** (Chart S1) does not exert any inhibitory activity in concentration 20 μM or lower (Fig. 2). Strikingly, however, in excess concentration, it does bind the enzyme. The MX structure of the adduct has been solved here and it shows double occupancy (See SI). This may be caused by the well-known

high flexibility of the active site cavities [33,40], which allows the distorted second binding site to eventually accommodate the ligand in saturating condition.
We conclude that Mpro can bind inhibitors if added in excess, forming doubly occupied adducts, even if the ligands exhibit no inhibitory activity.

Conclusions

Here, we have shown that X-ray structures at almost equimolar quantities of non-covalent and covalent ligands such as **X77** and **MG-132** (Chart 1) show only one active site occupied. The same asymmetry can be observed by leaving crystals of the doubly occupied enzyme in the drops for at least 2 months. Our MD simulations suggest that the single occupied protein undergoes a further breaking of symmetry [32] on passing from the solid state to solution. Water molecules enter the cavity, destabilize the PHE140-HIS163 contact (Fig. 5) and, consequently, the catalytically active conformation of the HIS41/CYS145 dyad. This water-occupancy trend is also observed in the doubly occupied enzymes, although to a lesser extent. A similar, water-triggered breaking of symmetry in solution had been observed also for the apo-protein [32,33] . The here observed destabilization is associated with anti-correlated motions of the two subunits that close up the occupied binding cavity, while opening up the empty one (Fig. 6ABC). This impacts on the catalytic activity of the empty cavity, as the occupation of the binding cavity of one Mpro subunit by **X77** causes the loss of the catalytically active conformation in the other one. This observation, along with the MX results, suggests that the unoccupied chain in the formed dimer has a reduced affinity for a second ligand.
Taken together, our results strongly suggest that NC[1] is operative for this enzyme. This would lead to two advantages in the life-cycle of the virus. First, it favors the ability to respond to a very wide range of ligand concentrations[41], making it very adaptable to the highly diverse local environments encountered by the enzyme during viral infection. Second, it allows the enzyme not to stop in saturation conditions [9]. Such features may contribute to the ability of viral enzymes to function in different hosts' conditions, and in turn for virus survival and quickly adaptability to the host's immune response and drug treatment. Our findings have significant implications for identifying effective inhibitors targeting Mpro, as well as other viral enzymes of the same family. Furthermore, the finding that non-inhibiting molecules can still bind to the enzyme's active sites emphasizes the importance of selecting appropriate reference compounds for ligand-based screening. Considering the structural asymmetry between the enzyme's binding sites is also crucial for precise drug design using structure-based methods. It's noteworthy that the virtual screening efficiency may vary between the two binding sites, necessitating careful consideration of their unique characteristics [33]. Furthermore, our study highlights the untapped potential of targeting the enzyme's dimerization interface, an area with limited exploration for this enzyme class[42]. Exploring this avenue allows for

[1] Further molecular simulations studies on the Mpro adduct with X77 and several other ligands confirm NC, including the loss of catalytic water molecule in one binding site (Toan T Nguyen et al, manuscript in preparation).

a broader range of ligands and holds great promise for advancing drug development strategies[11].

It is also worth to notice that our MX results contrast with what found so far in the 500 MX structures, which exhibit double occupation and, in most cases (>98%), cyclic symmetry. This suggests that these studies probably were conducted with 2:1 ligand/monomer stoichiometry or more. The ligands, if in excess, may not be an inhibitor of Mpro even if they form doubly occupied adducts. Indeed, while **X77** - an R-enantiomer structure - inhibits Mpro in the μM-high nM range (Fig. 2), the correspondent S-enantiomer (S-**X77)** does not exert any inhibitory activity in concentration 20 μM or lower (Fig. 2). In excess concentration, however, it does bind the enzyme, as shown by the X-ray structure of the S-enantiomer/Mpro adduct (See SI). This may be caused by the well-known high flexibility of the active site cavities [33,40], which allows the distorted second binding site to eventually accommodate the substrate in saturating condition.

Methods

1. Ligands.

MG-132 was purchased. **X77** was synthetized by us as follows. Rac-**X77** was prepared in two separate steps (Scheme S1). Although **X77** can be formed in only one step by the four-component Ugi reaction (Patent US9975885B2) we observed slightly higher yields when a pre-formed aldimine was utilized. Hence, reaction of 3-pyridine-carbox-aldehyde with 4-(*tert*-butyl)aniline in methanol at room temperature gave (*E*)-*N*-[4-(*tert*-butyl)-phenyl]-1-(pyridin-3-yl)-methanimine in quantitative yield. In the subsequent step this aldimine was then treated with 4-imidazole carboxylic acid and cyclohexyl iso-cyanide at 40 °C in methanol to furnish rac-**X77** in 45% yield after work-up and purification (Fig. S10 to S18). Finally, **X77** and S-**X77** were successfully separated by preparative HPLC with a chiral stationary phase.

2. Biochemical analyses of **X77,** S-**X77,** Rac-**X77**

The SARS-CoV-2 Mpro was synthesized using the ORF1ab polyprotein residues 3264–3569, (GenBank code: MN908947.3). Gene synthesis, protein production, and purification were as reported by Zhang et al., [26] where eluted fractions containing the target protein were pooled and subjected to buffer exchange in 20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, and 1 mM DTT, pH 7.8. The detection of enzymatic activity of the Mpro was performed under conditions reported by Kuzikov et al. [37].

Enzymatic activity was measured by a Förster resonance energy transfer (FRET), using the dual-labeled substrate, DABCYL-KTSAVLQ↓SGFRKM-EDANS (Bachem #4045664) containing a protease specific cleavage site after the GLN. In the intact peptide, EDANS fluorescence is quenched by the DABCYL group. Following enzymatic cleavage, generation of the fluorescent product was monitored (Ex/Em = 340/460 nm) (EnVision, PerkinElmer). The assay buffer contained 20 mM Tris (pH 7.3), 100 mM NaCl, and 1 mM EDTA. The assay was established in an automated screening format (384 well black

microplates, Corning, #3820) and optimized with respect to assay volume (10 µL), enzyme concentration (60 nM), substrate concentration (15 µM), incubation time (60 min with compounds, 15 min with substrate) and temperature (37 °C for incubation with compounds, 25 °C for incubation with substrate), DMSO tolerance (up to 5 v/v%), response to inhibition with known compounds as zinc pyrithione and the effects of reducing agents (DTT). **X77,** S-**X77** and Rac-**X77** were then profiled in triplicate in 11 points concentration responses, starting from 20 µM top concentration with 1:2 dilution steps.

3. Xray Crystallography

Crystallization. Crystallization of Mpro in complex with compounds, was carried out as previously described [36]. Briefly, Mpro, stored in 20 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 7.8, 1 mM DTT was incubated at 5 mg/ml (150 µM) with compounds (**X77/MG-132**) at either 75 μM or 150 μM final concentrations. For **X77** also a 500 μM concentration was used. Crystallization experiments were set up after 1h of incubation at RT, by seeding in sitting drops using the Morpheus® kit (Molecular Dimensions) with a Mosquito robot (STPlabtech Ltd., Melbourn Hertfordshire, UK). Crystals appeared within a couple of days, and were flash frozen in liquid nitrogen after a few days of growth. For S-**X77** a 5 mM concentration of molecule was needed. For the "old" crystals, crystallization was carried out as described, with 5 mM **MG-132** or **X77** respectively and crystals were flash-frozen in liquid nitrogen after at least 2 months from their first appearance.

The best diffracting crystals appeared in the following conditions:

- Mpro:**X77** 500 μM, condition F10: 0.1 M Tris/BICINE pH 8.5; 0.12M D-Glucose; 0.12M D-Mannose; 0.12M D-Galactose; 0.12M L-Fucose; 0.12M D-Xylose; 0.12M N-Acetyl-D-Glucosamine; 20% v/v Ethylene glycol; 10 % w/v PEG 8000.
- Mpro:**X77** 75 μM, condition H6: 0.1M DL-Glutamic acid monohydrate, 0.1M DL-Alanine 0.1M Glycine 0.1M DL-Lysine monohydrochloride 0.1M DL-Serine, 0.1M Hepes/MOPS pH 7.5, 20% v/v Ethylene glycol 10 % w/v PEG 8000.
- Mpro:**X77** 150 μM, condition D6: 0.12M 1,6-Hexanediol 0.12M 1-Butanol 0.12M 1,2-Propanediol 0.12M 2-Propanol 0.12M 1,4-Butanediol 0.12M 1,3-Propanediol, 0.1M Hepes/MOPS pH 7.5, 20% v/v Ethylene glycol 10 % w/v PEG 8000.
- Mpro:**MG-132** 75 μM, condition E2: 0.12 M Diethylene glycol 0.12M Triethylene glycol 0.12M Tetraethylene glycol 0.12M Penta-ethylene glycol, 0.1M imidazole/MES pH 6.5, 20% v/v Ethylene glycol 10 % w/v PEG 8000.
- Mpro:**MG-132** 150 μM, condition D1: 0.12M 1,6-Hexanediol 0.12M 1-Butanol 0.12M 1,2-Propanediol 0.12M 2-Propanol 0.12M 1,4-Butanediol 0.12M 1,3-Propanediol, 0.1 M imidazole/MES pH 6.5, 20% v/v PEG 500 MME 10 % w/v PEG 20000.
- For Mpro:**X77** 500 μM enantiomer 1/R, condition D10: 0.12M 1,6-Hexanediol 0.12M 1-Butanol 0.12M 1,2-Propanediol 0.12M 2-Propanol 0.12M 1,4-Butanediol 0.12M 1,3-Propanediol, 0.1 M Tris/Bicine pH 8.5, 20% v/v Ethylene glycol 10% w/v PEG 8000.
- For Mpro:**X77** 500 μM enantiomer 2/S and Mpro:**X77** 5 mM enantiomer 2/S condition G4: 0.1M Sodium formate 0.1M Ammonium acetate 0.1M Sodium citrate

tribasic dihydrate 0.1M Potassium sodium tartrate tetrahydrate 0.1M Sodium oxamate, 0.1M imidazole/MES pH 6.5, 12.5% v/v MPD 12.5% PEG 1000 12.5% w/v PEG 3350.

- For Mpro:**X77** 5 mM enantiomer 1/R, condition E10: 0.12 M Diethylene glycol 0.12M Triethylene glycol 0.12M Tetraethylene glycol 0.12M Penta-ethylene glycol, 0.1 M Tris/bicine pH 8.5, 20% v/v Ethylene glycol, 10% w/v PEG 8000.
- For Mpro:**MG-132** 5 mM "2-months old" crystal condition A2: 0.06 M Magnesium chloride hexahydrate, 0.06 M Calcium chloride dihydrate, 0.1 M Hepes/MOPS pH 7.5, 20% v/v PEG 500 MME, 10% w/v PEG 20000
- For M-pro:**X77** 5 mM "2-months old" crystal condition G6: 0.1 M Sodium formate, 0.1M Ammonium acetate, 0.1M Sodium citrate tribasic dihydrate, 0.1M Potassium sodium tartrate tetrahydrate, 0.1M Sodium oxamate, 0.1 M Hepes/MOPS pH 7.5, 20% v/v Ethylene glycol, 10 % w/v PEG 8000

Data Collection, Data Reduction, Structure Determination, Refinement and Final Model Analysis. Xray diffraction measurements were performed at 100K at the XRD2 beamline of the Elettra synchrotron (Trieste, Italy) using a 1.000 Å wavelength. Crystals were flash frozen in the original crystallization solution, with no further addition of cryoprotectants. The collected datasets were processed using XDS [43] and Aimless[44] from the CCP4 suite [45].

Structures were solved with Phaser[46] by molecular replacement with 7BB2 (PDBid) as a search model. Refinement was carried out by alternating cycles of manual model building in COOT [47,48] and automatic refinement using Phenix [49] (version 1.19.2_4158). Data collection and refinement statistics are reported in Table S1. Figures were prepared using Pymol [50].

Data availability. Coordinates and structure factors were deposited to the Protein Data Bank with accession numbers 7PHZ (Mpro:**X77** in space group P212121), 8P57 (Mpro:**X77** at 75 μM), 8P56 (Mpro:**X77** at 150 μM), 8P55(Mpro:**MG-132** at 75 μM), 8P54 (Mpro:MG132 at 150 μM), 8P58 (Mpro: R-**X77** at 500 μM), 8P5A (Mpro: R-**X77** at 5 mM), 8P5B (Mpro: S-**X77** at 500 μM), 8P5C (Mpro: S-**X77** at 5 mM), 8P86 (Mpro: **MG-132** at 5 mM, "2-months old" crystal), 8P87 (Mpro: **X77** at 5 mM, "2-months old" crystal).

Analysis of previously-deposited structures. A tabular report and corresponding structures were downloaded for 696 SARS-CoV-2 Mpro entries, deposited in the PDB database between 5th February 2020 and 26th April 2023. Among these, 497 structures were found to contain non-solvent ligands with molecular weight >= 100 Da. An in-house python script was used to check for covalent bonds between protein and ligands, which were present in 333 structures out of 497. Percentages reported in text are derived from the results summarized in Table 1. The classification of structures in this table has been manually curated. For instance, a structure in which Mpro active site interacts with another protein classified as "apo" required a manual correction. Additionally, some of the submitted structures might not contain all domains of Mpro, i.e. six asymmetric structures with a ligand do not contain the whole dimer within the unit cell. Among the 10

apo structures with non-cyclic symmetry, the distribution of space groups is the following: P1: 3,$P2_1$ 2_1 2_1: 2, $P12_11$: 2, $P2_12_12$: 2, $P4_32_12$: 1.

Table 1. Ligand binding and symmetry in previously-deposited Mpro structures

Ligands	Covalent	Symmetry	Number	Percentage
No	No	Cyclic	186	26.7%
No	No	Non-Cyclic	10	1.4%
Yes	No	Cyclic	163	23.4%
Yes	No	Asymmetric	4	0.6%
Yes	Yes	Cyclic	328	47.1%
Yes	Yes	Asymmetric	5	0.7%

4. Simulations

*Molecular dynamics***.** The systems 6W63, 7PHZ, 8P57 were studied in 500-ns unbiased MD simulations, using GROMACS 2019.2 [51] and the Amber14SB force field [52]. The TIP3P model was used for the water molecules, while the ligand was parameterized using the General AMBER Force Field (GAFF) with AM1-BCC charges [53]. The protein was pre-processed using Schrodinger's Protein Preparation Wizard [54] and the protonation state of residues in the active site was compared and confirmed with the output of the VirginiaTech H++ Web Server [55]. N-terminal acetyl and C-terminal amide capping groups were added to the 7PHZ and 8P57 structures. The protein and the ligand were then placed at the center of a 16 x 16 x 16 cubic nanometers box, and solvated with water and 0.15M NaCl. The systems were minimized with 50,000 steps of steepest descent, 50,000 steps of conjugate gradient, and then heated from 5K to 310K over the course of 5 ns, followed by a 1-ns equilibration stage in an NPT ensemble. During the annealing and NPT equilibration, 1000 kJ/mol restraints were applied on the C alpha atoms and on the ligand, along all three coordinates. The restraints were then released for the 500-ns unbiased simulation conducted with a timestep of 2fs, Parrinello-Rahman barostat, Velocity Rescale thermostat, and LINCS constraints on all bonds. Long-range electrostatics interactions were handled with Particle Mesh Ewald (PME) using 1.6 Å of grid spacing. The cutoff radius of van der Waals interaction and short-range electrostatics was set to 1.2 Å.

*Water analysis***.** The analysis was conducted using an in-house python (v 3.10.6) [56] script with the packages MDtraj (v 1.9.7) [57] and alphashape (v 1.3.1) [58] . We investigated the change of number of water molecules within the region around the binding pocket S1 during MD simulations. This region was defined by a convex hull bordered by the alpha carbon atoms (CA) of residues VAL114, ALA116, GLY138, PHE140, ASN142, GLY146, HIS164, HIS172, GLY174, the carbonyl carbon atom of residue Thr135, and the carbonyl

oxygen of residue CYS117. The analysis was performed on 5000 frames of 500-ns MD trajectory for each system.

Principal Component Analysis. The analysis was performed on the last 400 ns of simulation time, with a sampling timestep of 0.1 ns. The two subunits in each simulation were analyzed separately and only the alpha carbon atoms were considered. The standard GROMACS tools *gmx covar* and *gmx anaeig* were used for the analysis and for the generation of the protein structures deformed along the first eigenvector.

ACKNOWLEDGMENT

Cornelia Vermeeren (RWTH Aachen University) is kindly acknowledged for carrying out the preparative CSP-HPLC separation. The support of Susanne Pohlmann (RWTH Aachen University) in compound purification is highly appreciated. TTN is supported by the Vietnam National University - Hanoi, grant number QG.20.82. GR, PC and MP acknowledge the Helmholtz European Partnering fundings for the project 'Innovative high-performance computing approaches for molecular neuromedicine'. GR and PC acknowledge the Joint Lab 'Supercomputing and Modeling for the Human Brain' of the Helmholtz Association.

References

1. Levy, E. D., Pereira-Leal, J. B., Chothia, C. & Teichmann, S. A. 3D Complex: A Structural Classification of Protein Complexes. *PLOS Comput. Biol.* **2**, e155 (2006).
2. Shima, S. *et al.* Lyotropic-salt-induced changes in monomer/dimer/tetramer association equilibrium of formyltransferase from the hyperthermophilic Methanopyrus kandleri in relation to the activity and thermostability of the enzyme. *Eur. J. Biochem.* **258**, 85–92 (1998).
3. Boggetto, N. & Reboud-Ravaux, M. Dimerization Inhibitors of HIV-1 Protease. **383**, 1321–1324 (2002).
4. Abdalla, A.-M., Bruns, C. M., Tainer, J. A., Mannervik, B. & Stenberg, G. Design of a monomeric human glutathione transferase GSTP1, a structurally stable but catalytically inactive protein. *Protein Eng. Des. Sel.* **15**, 827–834 (2002).
5. Shimba, N., Nomura, A. M., Marnett, A. B. & Craik, C. S. Herpesvirus Protease Inhibition by Dimer Disruption. *J. Virol.* **78**, 6657–6665 (2004).
6. Lookene, A., Zhang, L., Hultin, M. & Olivecrona, G. Rapid Subunit Exchange in Dimeric Lipoprotein Lipase and Properties of the Inactive Monomer *. *J. Biol. Chem.* **279**, 49964–49972 (2004).
7. Lee, K. N. *et al.* Antiplasmin-cleaving enzyme is a soluble form of fibroblast activation protein. *Blood* **107**, 1397–1404 (2006).
8. Nussinov, R. Introduction to Protein Ensembles and Allostery. *Chem. Rev.* **116**, 6263–6266 (2016).
9. Koshland, D. E. & Hamadani, K. Proteomics and Models for Enzyme Cooperativity *. *J.*

Biol. Chem. **277**, 46841–46844 (2002).
10. Bush, E. C. *et al.* Modeling the Role of Negative Cooperativity in Metabolic Regulation and Homeostasis. *PLOS ONE* **7**, e48920 (2012).
11. Yin, H. & Hamilton, A. D. Strategies for Targeting Protein–Protein Interactions With Synthetic Agents. *Angew. Chem. Int. Ed.* **44**, 4130–4163 (2005).
12. Thiel, P., Kaiser, M. & Ottmann, C. Small-Molecule Stabilization of Protein–Protein Interactions: An Underestimated Concept in Drug Discovery? *Angew. Chem. Int. Ed.* **51**, 2012–2018 (2012).
13. Bruzzese, F. J. & Connelly, P. R. Allosteric Properties of Inosine Monophosphate Dehydrogenase Revealed through the Thermodynamics of Binding of Inosine 5‘-Monophosphate and Mycophenolic Acid. Temperature Dependent Heat Capacity of Binding as a Signature of Ligand-Coupled Conformational Equilibria. *Biochemistry* **36**, 10428–10438 (1997).
14. Bronowska, A. K. Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design. in *Thermodynamics - Interaction Studies - Solids, Liquids and Gases* (IntechOpen, 2011). doi:10.5772/19447.
15. Ha, S. H. & Ferrell, J. E. Thresholds and ultrasensitivity from negative cooperativity. *Science* **352**, 990–993 (2016).
16. Sevlever, F., Bella, J. P. D. & Ventura, A. C. Discriminating between negative cooperativity and ligand binding to independent sites using pre-equilibrium properties of binding curves. *PLOS Comput. Biol.* **16**, e1007929 (2020).
17. Altszyler, E., Ventura, A. C., Colman-Lerner, A. & Chernomoretz, A. Ultrasensitivity in signaling cascades revisited: Linking local and global ultrasensitivity estimations. *PLOS ONE* **12**, e0180083 (2017).
18. Weiss, J. N. The Hill equation revisited: uses and misuses. *FASEB J.* **11**, 835–841 (1997).
19. Monod, J., Wyman, J. & Changeux, J.-P. On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **12**, 88–118 (1965).
20. Goldbeter, A. & Koshland, D. E. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci.* **78**, 6840–6844 (1981).
21. Seydoux, F., Malhotra, O. P., Bernhard, S. A. & Stark, G. Half-Site Reactivit. *CRC Crit. Rev. Biochem.* **2**, 227–257 (1974).
22. Swapna, L. S., Srikeerthana, K. & Srinivasan, N. Extent of Structural Asymmetry in Homodimeric Proteins: Prevalence and Relevance. *PLOS ONE* **7**, e36688 (2012).
23. Brown, J. H. Breaking symmetry in protein dimers: Designs and functions. *Protein Sci.* **15**, 1–13 (2006).
24. Koshland, D. E. Jr., Némethy, G. & Filmer, D. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits*. *Biochemistry* **5**, 365–385 (1966).
25. Silvestrini, L. *et al.* The dimer-monomer equilibrium of SARS-CoV-2 main protease is affected by small molecule inhibitors. *Sci. Rep.* **11**, 9283 (2021).
26. Zhang, L. *et al.* Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **368**, 409–412 (2020).
27. Adachi, S., Koma, T., Doi, N., Nomaguchi, M. & Adachi, A. Commentary: Origin and

evolution of pathogenic coronaviruses. *Front. Immunol.* **11**, (2020).
28. Nashed, N. T., Aniana, A., Ghirlando, R., Chiliveri, S. C. & Louis, J. M. Modulation of the monomer-dimer equilibrium and catalytic activity of SARS-CoV-2 main protease by a transition-state analog inhibitor. *Commun. Biol.* **5**, 1–9 (2022).
29. Lee, J. *et al.* Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. *Nat. Commun.* **11**, 5877 (2020).
30. Vuong, W. *et al.* Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *Nat. Commun.* **11**, 4282 (2020).
31. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
32. Ansari, N., Rizzi, V., Carloni, P. & Parrinello, M. Water-Triggered, Irreversible Conformational Change of SARS-CoV-2 Main Protease on Passing from the Solid State to Aqueous Solution. *J. Am. Chem. Soc.* **143**, 12930–12934 (2021).
33. Gossen, J. *et al.* A Blueprint for High Affinity SARS-CoV-2 Mpro Inhibitors from Activity-Based Compound Library Screening Guided by Analysis of Protein Dynamics. *ACS Pharmacol. Transl. Sci.* **4**, 1079–1095 (2021).
34. Chen, H. *et al.* Only one protomer is active in the dimer of SARS 3C-like proteinase. *J. Biol. Chem.* **281**, 13894–13898 (2006).
35. Parmar, M. *et al.* Structural differences in 3C-like protease (Mpro) from SARS-CoV and SARS-CoV-2: molecular insights revealed by Molecular Dynamics Simulations. *Struct. Chem.* (2022) doi:10.1007/s11224-022-02089-6.
36. Costanzi, E. *et al.* Structural and Biochemical Analysis of the Dual Inhibition of MG-132 against SARS-CoV-2 Main Protease (Mpro/3CLpro) and Human Cathepsin-L. *Int. J. Mol. Sci.* **22**, 11779 (2021).
37. Kuzikov, M. *et al.* Identification of Inhibitors of SARS-CoV-2 3CL-Pro Enzymatic Activity Using a Small Molecule in Vitro Repurposing Screen. *ACS Pharmacol. Transl. Sci.* **4**, 1096–1110 (2021).
38. Arya, R. *et al.* Structural insights into SARS-CoV-2 proteins. *J. Mol. Biol.* **433**, 166725 (2021).
39. Arutyunova, E. *et al.* N-Terminal Finger Stabilizes the S1 Pocket for the Reversible Feline Drug GC376 in the SARS-CoV-2 Mpro Dimer. *J. Mol. Biol.* **433**, 167003 (2021).
40. Kneller, D. W. *et al.* Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat. Commun.* **11**, 3202 (2020).
41. Ferrell, J. E. Q&A: Cooperativity. *J. Biol.* **8**, 53 (2009).
42. Goyal, B. & Goyal, D. Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy. *ACS Comb. Sci.* **22**, 297–305 (2020).
43. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
44. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
45. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
46. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

47. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
48. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
49. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. Struct. Biol.* **75**, 861–877 (2019).
50. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
51. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
52. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
53. Sousa da Silva, A. W. & Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes* **5**, 367 (2012).
54. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234 (2013).
55. Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40**, W537–W541 (2012).
56. Python Release Python 3.10.6. *Python.org* https://www.python.org/downloads/release/python-3106/.
57. McGibbon, R. T. *et al.* MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **109**, 1528–1532 (2015).
58. Bellock, K., Godber, N. & Kahn, P. bellockk/alphashape: v1.3.1 Release. (2021) doi:10.5281/zenodo.4697576.