

Inhibition of SARS-CoV-2 main protease: A repurposing study that targets the dimer interface of the protein

Hanife Pekel,^{†,‡,||} Metehan Ilter,^{¶,||} and Ozge Sensoy^{*,§,‡}

[†]*Department of Pharmacy Services, Vocational School of Health Services, Istanbul Medipol University, Istanbul, Turkey*

[‡]*Regenerative and Restorative Medicine Research Center (REMER), Research Institute for Health Sciences and Technologies (SABITA), Istanbul Medipol University, Istanbul, Turkey*

[¶]*Graduate School of Engineering and Natural Sciences, Istanbul Medipol University, Istanbul, Turkey*

[§]*Department of Computer Engineering, School of Engineering and Natural Sciences, Istanbul Medipol University, Istanbul, Turkey*

^{||}*These authors equally contributed.*

E-mail: osensoy@medipol.edu.tr

Abstract

Coronavirus disease-2019 (COVID-19) was firstly reported in Wuhan, China, towards the end of 2019, and, unfortunately, within a short period of time, emerged as a pandemic. The spread and lethality rates of the COVID-19 have ignited studies that focus on the development of therapeutics for either treatment or prophylaxis purposes. In parallel, drug repurposing studies have also come into prominence. In this study, we aimed at having a holistic understanding of conformational and dynamical changes

induced by an experimentally characterized inhibitor on main protease (M-pro) which would enable the discovery of novel inhibitors. To this end, we performed molecular dynamics simulations using crystal structures of *apo* and α -ketoamide-13b-bound M-pro homodimer. Analysis of trajectories pertaining to *apo* M-pro revealed a new target site, which is located at the homodimer interface, next to the catalytic dyad. Thereafter, we performed ensemble-based virtual screening by exploiting the ZINC and DrugBank databases and identified three candidate molecules, namely eluxadoline, diosmin, and ZINC02948810 that could invoke local and global conformational rearrangements which were also elicited by α -ketoamide-13b on the catalytic dyad of M-pro. Furthermore, ZINC23881687 was also discerned as a promising candidate due to its interaction with catalytically important residues Glu166 and Ser1. Last but not least, we could find another candidate, namely ZINC20425029, whose mode of action was different. It modulated the dynamical properties of catalytically important residue, Ala285 rather than the catalytic dyad. As such, this study presents valuable findings that might be used in the development of novel therapeutics against SARS-CoV-2 M-pro.

Introduction

People have been suffering from contagious coronavirus disease-2019 (COVID-19) since December 2019. It is caused by the acute respiratory syndrome coronavirus-2 (SARS-CoV-2) and spreads among people when an infected person makes close contact with uninfected people in the population.¹⁻⁴ Consequently, it has led to a severe pandemic since March 2020. Among the most frequently stated symptoms are dry cough, shortness of breath, nasal congestion, abnormal sputum production, and fever as well as debilitation of organs/systems, particularly occurring in the late-stage.⁵ Most of the deaths were mainly caused by pneumonia, acute respiratory distress syndrome, and multi-organ failure caused by an exorbitant increase in inflammatory cytokines.^{6,7}

The extensive genomic studies have shown that SARS-CoV-2 encompasses *ca.* 30,000

nucleotides and shares 82% sequence identity with the SARS-CoV, hence categorized as a member of Betacoronaviruses^{1-3,8,9}. Besides high sequence identity, the mechanism of pathogenesis is also similar, where spike protein binds to angiotensin-converting enzyme-2 of the host, which is expressed in the lung, nasopharynx, and intestinal epithelial tissues¹⁰⁻¹⁴, to enter the cell. After the cellular entry, viral RNA launches translation of large polyproteins, namely polyprotein-1a and -1ab in the ribosome^{2,3,15-17}, which are further processed by the M-pro and papain-like protease^{18,19}. Moreover, the M-pro of the SARS-CoV-2 shares 96.1% sequence identity with that of the SARS-CoV. It is also important to emphasize that known human proteases possess different cleavage specificity. Therefore, targeting the M-pro enzyme by therapeutics is less likely to pose toxic effects to humans. Consequently, the M-pro has come into prominence as a drug target for preventing viral replication.^{1,20-22} In parallel, substantial efforts have been made to unveil 3D structures of the M-pro enzyme^{23,24}. It has been revealed that M-pro is a homodimer and each subunit (A and B) is composed of three domains (Domain I, II, and III) along with an N-finger domain that encompasses the first nine residues of the protein. Domain I and II (residues 10 to 99 and 100-182, respectively) adopt a chymotrypsin fold and collectively make up the substrate-binding pocket (SBP). It is located at the interface of the two domains and its conformation is modulated by interactions formed between the N-finger and Glu166^{1,20}. The SBP is responsible for the recognition of peptide substrates and their cleavages as well and composed of four subsites, namely S1, S2, S3, and S4¹. Structural studies have shown that S1 and S2 subsites are well-conserved in Betacoronavirus family²⁵⁻²⁸. S1 subsite is built up of side-chain atoms of His163 and Phe140 and backbone atoms of Met165, Glu166, and His172^{21,29}. It is stabilized by hydrogen bonds formed between residues that are located on subunits A and B of M-pro such as Ser1A-Phe140B, Ser1A-Glu166B, Ser1B-Phe140A, and Ser1B-Glu166A, where A and B denote the name of the subunits^{17,30}. Moreover, SBPs of both subunits participate in the formation of the catalytic dyad, which is driven by interactions made between Cys145 and His41^{1,31}. Besides S1, S2 subsite is also crucial as it is responsible for the plasticity of the SBP, which

is required for recognition of the protein by its interacting partners^{1,32} .

Domain III (residues 198 to 303), which is made up of five α -helices, on the other hand, is mainly involved in dimer formation as similar to N-finger domain.^{1,33–35} In addition, it also plays a key role in the enzymatic activity of the protein as deletion of the domain caused loss of the enzymatic activity.^{1,36}

Elucidation of a cysteine residue (Cys145), whose side chain can be readily oxidized, in the catalytic dyad of the M-pro enzyme has prompted studies towards the development of therapeutics that can covalently bind to the residue to inhibit the activity of the enzyme^{37,38} . To investigate possible binding modes of inhibitors, SARS-CoV-2 M-pro was crystallized with α -ketoamide-11a, 11b, and 13b, as well as antineoplastic carmofur.^{1,17,30,39,40} The covalent Michael inhibitors bind to the M-pro by means of nucleophilic attacks made at Cys145. However, this might induce adverse effects, such as allergies, tissue damage, or carcinogenesis, thereby violating their drug safety profiles^{41,42} . Consequently, all these attempts have expedited studies that have focused on elaborating the mechanism of Cys145-driven catalytic activity. Kneller *et al.* has shown that when the distance between S γ atom of Cys145 and N ϵ 2 atom of His41 is 3.8 Å, a hydrogen bond between these atoms would not be formed in the ligand-free SARS-CoV-2 M-pro. In this state of the enzyme, the thiol group of the Cys145 is protonated, whereas the imidazole ring of His41 is neutral. This conceivably indicates that substrate binding or transition state of the enzyme would initiate proton transfer from Cys145 to His41, hence subsequent activation of the catalytic dyad³¹ .

It is also worth mentioning that the M-pro enzyme has been shown to display asymmetry in terms of activation states of the protomers¹ . That is to say, when both protomers of the M-pro are bound with α -ketoamide-13b, only Glu166B- not the one on the subunit A, was shown to be in an inactive state, which was characterized by loss of the H-bond formed between P1 moiety of the α -ketoamide-13b and Glu166B. Moreover, a similar finding was also obtained in a study where SARS-CoV M-pro was shown to display asymmetry upon activation^{43,44} .

The availability of plenty of structural data pertaining to the M-pro enzyme has triggered computer-aided drug design studies for either development of novel therapeutics or repurposing approved molecules. For the latter, virtual screening has been emerged as an effective computational technique for identifying compounds in libraries of small molecules that are most likely to bind to the target protein⁴⁵⁻⁵⁰.

In this study, we employed atomistic molecular dynamics (MD) simulations to investigate structural and dynamical properties of *apo* and α -ketoamide-13b-bound (*holo*), which is a well-characterized inhibitor, M-pro homodimer. Comparative analyses of the trajectories showed that α -ketoamide-13b increased mobility of the loops found on subunit A and B, on which Glu47 and Arg188 are located, but only disturbing the catalytic dyad of subunit A. On the other hand, it did not induce any change on subunit B. We also clustered the trajectories pertaining to *apo* M-pro to identify mostly populated conformational states therein. Consequently, we deciphered a possible binding site, which is located at the homodimer interface. The pharmacophore features of this region were modeled and possible candidates, which could bind to the region, were sought in the ZINC⁵¹ and DrugBank⁵²⁻⁵⁶ databases. Thereafter, candidate molecules were docked to the target site on the M-pro and those with favorable binding energies were further tested in molecular dynamics simulations (a total of 16.5 μ s) to evaluate their impact on the structure and dynamic of the M-pro. Afterward, structural and dynamical properties of ligand/M-pro complexes were compared to those of *holo* M-pro to identify compounds that could display similar impact as elicited by α -ketoamide-13b. Consequently, diosmin, eluxadoline, and ZINC02948810 were found to be successful candidates. Moreover, ZINC23881687 was considered as a successful candidate even though it did not invoke similar dynamic and structural properties to *holo* protein, since it formed hydrogen bonds with both Glu166A and Ser1B, which are involved in dimerization. Apart from the aforementioned molecules, we could also identify another candidate, ZINC20425029, which impacted the dynamics of Ala285, which is known to be responsible for the catalytic activity of the enzyme.^{1,20} Finally, in addition to local modulation of the

M-pro by the compounds, the global dynamic properties of the protein were also influenced.

Methods

Preparation of *apo* and *holo* M-pro homodimer systems

Crystal structures of *apo* (PDB ID: 6Y2E) and α -ketoamide-13b-bound M-pro monomer (*holo*) (PDB ID: 6Y2F) were obtained from the Protein Data Bank (PDB)^{1,23,24}. The crystal structure of the *apo* form of the protein contains all the residues. The homodimer was built up according to the BIOMT information provided in the PDB file using the Visual Molecular Dynamics (VMD) software⁵⁷. Missing residues, Glu47, Asp48, and Gln306, in the crystal structure of *holo* M-pro, were filled using the GalaxyFill⁵⁸. Afterwards, dimethyl sulfoxide was removed from the crystal, and subsequently, the system was dimerized according to the BIOMT data using the VMD⁵⁷ software. The ionization states of residues of *apo* and *holo* M-pro dimers were set at pH 7.0 using the ProteinPrep tool of Schrödinger Maestro software^{59,60}. Eventually, systems were solvated and neutralized using the Solution Builder of CHARMM-GUI^{61–63}. The dimensions of the water boxes were set as 104x104x104 Å and 105x105x105 Å for *apo* and *holo* systems, respectively. As a side note, the dimensions of the water boxes were determined according to information deposited in the COVID-19 archive of CHARMM-GUI⁶¹. The systems were neutralized with 0.15 M KCl based on the Monte-Carlo ion displacing method as it is more accurate than distance-based ion displacement⁶⁴. Proteins and water molecules were modeled using the CHARMM36m force-field⁶⁵ and TIP3P⁶⁶, respectively.

The simulation protocol

MD simulations of the *apo* and *holo* systems were carried out using the GRONINGEN MACHINE for Chemical Simulations (GROMACS) package⁶⁷. The systems were minimized by means of the steepest descent algorithm. Subsequently, the systems were equilibrated in the NVT

ensemble for 125 picoseconds (ps), where the Nose-Hoover thermostat was used to maintain the temperature at 310.15 K^{68,69}. Furthermore, production runs were done in the NPT ensemble using the Nose-Hoover thermostat and Parrinello-Rahman barostat⁷⁰ for a total of 4 μ s. In minimization, equilibration, and production steps, the LINCS algorithm was exploited to constraint H-bonds⁷¹. Two replicates, each of which started with a different velocity distribution, were used for the systems studied.

Clustering trajectories pertaining to *apo* M-pro homodimer and binding pocket identification

The distance between C α atoms of Glu47A and Arg188A was measured in *apo* M-pro trajectories to determine conformational states pertaining to the SBP. Thereafter, time-line data was converted into the probability plots by setting the frequency range to 1 Å. Consequently, two frames, which represented the most probable states of the SBP, were selected. Subsequently, the selected structures were prepared using the OPLS3e force-field available in the ProteinPrep tool of Schrödinger software^{59,60,72}. The possible binding pockets were sought by means of the SiteMap module of Schrödinger, where hydrophobicity, hydrophilicity, drug-gability, volume, exposure, and enclosure of each possible pocket were calculated and scored, which are collectively given as SiteScore^{60,73,74}. The binding sites having close proximity to the homodimer interface and relatively higher SiteScores were selected to be used in drug repurposing studies.

Determination of pharmacophore features and candidate molecule search in the ZINC and DrugBank databases

The residues, which are crucial for the function of the M-pro enzyme, were included in pharmacophore modeling. In addition, we also included hotspot residues that were determined using the Hotpoint server⁷⁵. To do so, we compared hotspot residues in both *apo* and

holo trajectories of M-pro homodimer and identified those that displayed lower frequency in *holo* than in the *apo* system. The pharmacophore features of each target structure were determined by considering geometrical and chemical properties of the residues Ala116A, Phe140A, Leu141A, Glu166A, His172A, Ser1B, Arg4B, and Gln299B via the Develop Pharmacophore Hypothesis module of Schrödinger^{76,77}. Thereafter, the pharmacophore models were assessed according to the relative orientations of the groups included in the model. The models which contained groups that were clustered far from each other were excluded to be searched in the ZINC⁵¹ and DrugBank⁵²⁻⁵⁶ databases. Moreover, molecules having less than 700 kDa molecular weight, and 15 rotatable bonds were retrieved from the ZINC and DrugBank databases. Finally, the candidates having a relatively higher match with the pharmacophore model were determined using the Hypothesis Validation tool of Phase^{60,78,79}.

Molecular docking

The grid files of each target structure were generated to compute interactions between the binding pocket and ligands. To this end, the dimensions of grid boxes were determined according to the centroid of the residues utilized in the pharmacophore model using the Receptor Grid Generation module of Glide^{60,80-82}. Before starting molecular docking, the ionization states of ligands were determined at pH 7.0 and stereochemical information was obtained from 3D geometry via the LigPrep module of Schrödinger⁶⁰. Eventually, the ligands were docked to the receptors by using the Glide standard-precision (SP) docking algorithm^{60,80-82}. The obtained ligand-protein complexes were assessed according to GScores and orientation of the ligands within the binding pocket. According to that, ligands, which had relatively lower GScores and interacted with at least one of Glu166A, Ser1B, and Arg4B residues, were considered as promising candidates.

Preparation of ligand-protein complexes and corresponding MD simulation protocols

The topology and parameter files of the ligands were generated using the Ligand Reader & Modeller of CHARMM-GUI⁸³. Subsequently, the obtained topology and parameter files were exploited to prepare the ligand-protein complexes for MD simulations utilizing the Solution Builder of CHARMM-GUI⁶¹⁻⁶³, where the systems were neutralized with 0.15 M KCl via the Monte Carlo ion displacement. The protein and water molecules were modeled using the CHARMM36m force-field⁶⁵ and TIP3P⁶⁶, respectively. The systems were simulated using three replicates per system, each of which started with a different velocity distribution for a total of 16.5 μ s following the same procedures explained above.

The root-mean-square fluctuation (RMSF) calculations

The RMSF of *apo* and *holo* systems were calculated using the 'gmx rmsf' module of GROMACS⁶⁷ and with following equation:

$$\text{RMSF} = \sqrt{(1/N) \sum_{n=1}^N (X_i(n) - \bar{X}_i)^2}$$

, where simulation time over which one wants to average, and coordinates of backbone atom X_i at time n were denoted by N , and $X_i(n)$, respectively. In this way, the fluctuation pattern of each amino acid was calculated and compared to point out regions where substantial differences were observed.

The probability distributions of residue-pair distances

The distance between certain residues was measured using the 'gmx distance' module of GROMACS⁶⁷. Then, corresponding time-line data was converted into probability plots. To this aim, the minimum and maximum distances were calculated for distance data, and in seriatim, the data was sampled by 2 Å from the minimum distance value to the maximum distance value. Lastly, the frequencies of each interval were calculated and divided by the

total frequency, and consequently, the probability of each sampled distance was computed. The range of sampling was set to 1 Å only for the selection of frames, which represent different states of the *apo* system, with higher sensitivity.

Calculation of percent of hydrogen bond occupation

In order to calculate the percentage of hydrogen bond occupancy of the ligands with Glu166A and Ser1B, the Hbond plugin of the VMD was utilized⁵⁷. For this, the cut-off distance and angle between acceptor and donor were set to 3.0 Å and 20 degrees, respectively.

Principal component analysis (PCA)

In order to explore impacts of candidate molecules on dynamical and structural properties of M-pro homodimer, principal components of the systems, which reflected the collective dynamics of trajectories, were calculated. To do so, the C α atom of each residue was picked up and the systems were aligned with respect to the C α atoms of the crystal structure. Subsequently, the eigenvalues and eigenvectors of the systems were calculated by diagonalizing the co-variance matrix of each system using the following equation.

$$C_{mn} = [M_{mn}\Delta r_m\Delta r_n]$$

, where covariance matrix is denoted by C_{mn} as well as $M_{mn}\Delta r_m\Delta r_n$ corresponds to a change in position from time-averaged structure for each coordinate of m and n atoms.

$$Cv = \delta^2 v$$

, where set of eigenvalues and eigenvectors of the diagonalized co-variance matrix are shown by δ^2 , and v, respectively.

Consequently, for calculating 2D PCAs, trajectories of each system were aligned with respect to the first two eigenvalues of *holo* in 2D space. To this aim, the 'gmxc covar' and 'gmxc anaeig' modules of the GROMACS were utilized to generate the diagonalized co-variance matrix pertaining to each system and calculate corresponding eigenvalues and eigenvectors⁶⁷. In addition, each system was projected in 2D space with respect to the first eigenvalue

to unveil the collective change in each C α atom throughout the trajectory. To do so, an open-source Python package, namely, ProDy, was used^{84,85}.

Results and discussion

In this study, we employed atomistic MD simulations using crystal structures of *apo* and α -ketoamide 13b-bound (*holo*) SARS-CoV-2 M-pro, where the inhibitor was bound to both subunits A and B of the protein. Trajectories of the systems were compared and conformational changes induced by α -ketoamide 13b were elaborated. Consequently, the most probable conformational states of *apo* M-pro enzyme were identified and ensemble-based virtual screening was performed on the binding pocket that was identified within the dimer interface. Finally, the capability of candidate molecules to induce conformational changes on the M-pro, which were similar to those induced by α -ketoamide 13b, was interrogated by means of MD simulations.

The root-mean-square fluctuation (RMSF) profiles of the systems revealed differences in the regions which are known to be responsible for function of SARS-CoV-2 M-pro

Comparison of RMSF profiles pertaining to the *apo* and *holo* systems unveiled differences in the regions that are responsible for the function of the protein. In the presence of α -ketoamide 13b, fluctuation of N-finger (residues 1 to 9) of subunit A decreased, whereas fluctuation of the short helix on Domain I (residues 54 to 60) and Domain II (residues 100 to 182) were slightly increased (See Figure 1). Herein, it can be thought that modulation of the fluctuation pattern of short helix might also affect the dynamics of the SBP due to their spatial arrangement on the protein. Consequently, this might lead to substantial distinctions in the dynamics of the catalytic dyad of the subunit. In correspondence with that, a recent study has pointed out that residues found in Domain I and II play a pivotal

role in modulating the conformation of the SBP as well as the activity of the enzyme¹.

As opposed to the above-mentioned dynamical changes in Domain II on subunit A, no change was observed in fluctuation patterns of the same domain of subunit B. On the other hand, the short helix of subunit B displayed increased fluctuation in the *holo* system. Moreover, the fluctuation pattern of the N-finger on subunit B was not also impacted by the inhibitor, except Ser1B. Since the interaction of Ser1B with Glu166A regulates the conformational state of the SBP^{1,20}, we also investigated fluctuation profiles of Glu166A and B, and showed that the former fluctuated more in the presence of the inhibitor while the difference was not pronounced for the latter.

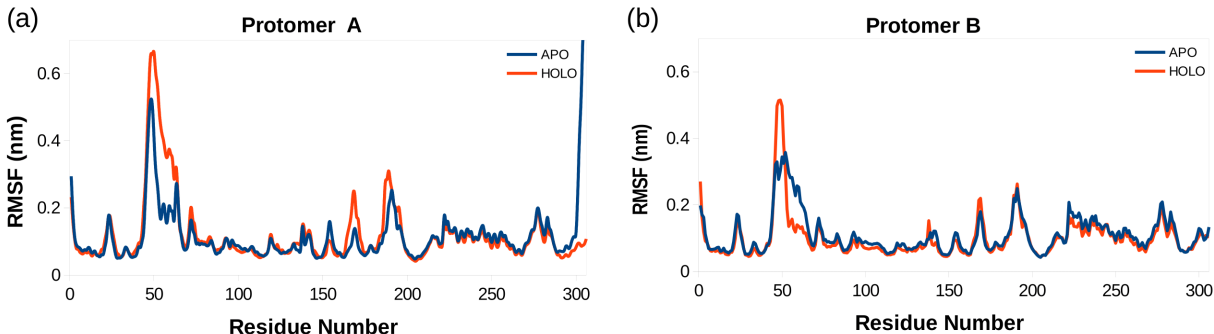


Figure 1: The backbone root-mean-square fluctuations of subunit (a) A and (b) B for *apo* (PDB ID: 6Y2E) (blue) and *holo* (PDB ID: 6Y2F) (red) forms were calculated from the obtained trajectories.

The inhibitor induced remarkable conformational rearrangements in the catalytic dyad and substrate-binding pocket of M-pro enzyme

To investigate conformational rearrangements induced by α -ketoamide 13b in the SBP and catalytic dyad, we compared distances between certain atoms: (i) C α atoms of Glu47 and Arg188 (See Figure 2.c), and (ii) N ϵ 2 atom of His41 and S γ atom of Cys145. The results showed that the inhibitor caused widening of the SBP on subunit A as evidenced by longer distances measured for Glu47-Arg188 pair (See Figure 2.a). In line with this, probability distribution pertaining to the distance measured between His41 and Cys145 showed that the

ligand caused a loss in the coordination of the catalytic dyad on subunit A, which might have stemmed from the opening of the loops that bear Glu47 and Arg188 on opposite sites of the SBP (See Figure 3.a). On the other hand, the inhibitor led to neither distortion of the SBP nor coordination loss in the catalytic dyad on subunit B (See Figure 3.b). Interestingly, Glu47 and Arg188 could come closer in the presence of the inhibitor on subunit B (See Figure 2.b).

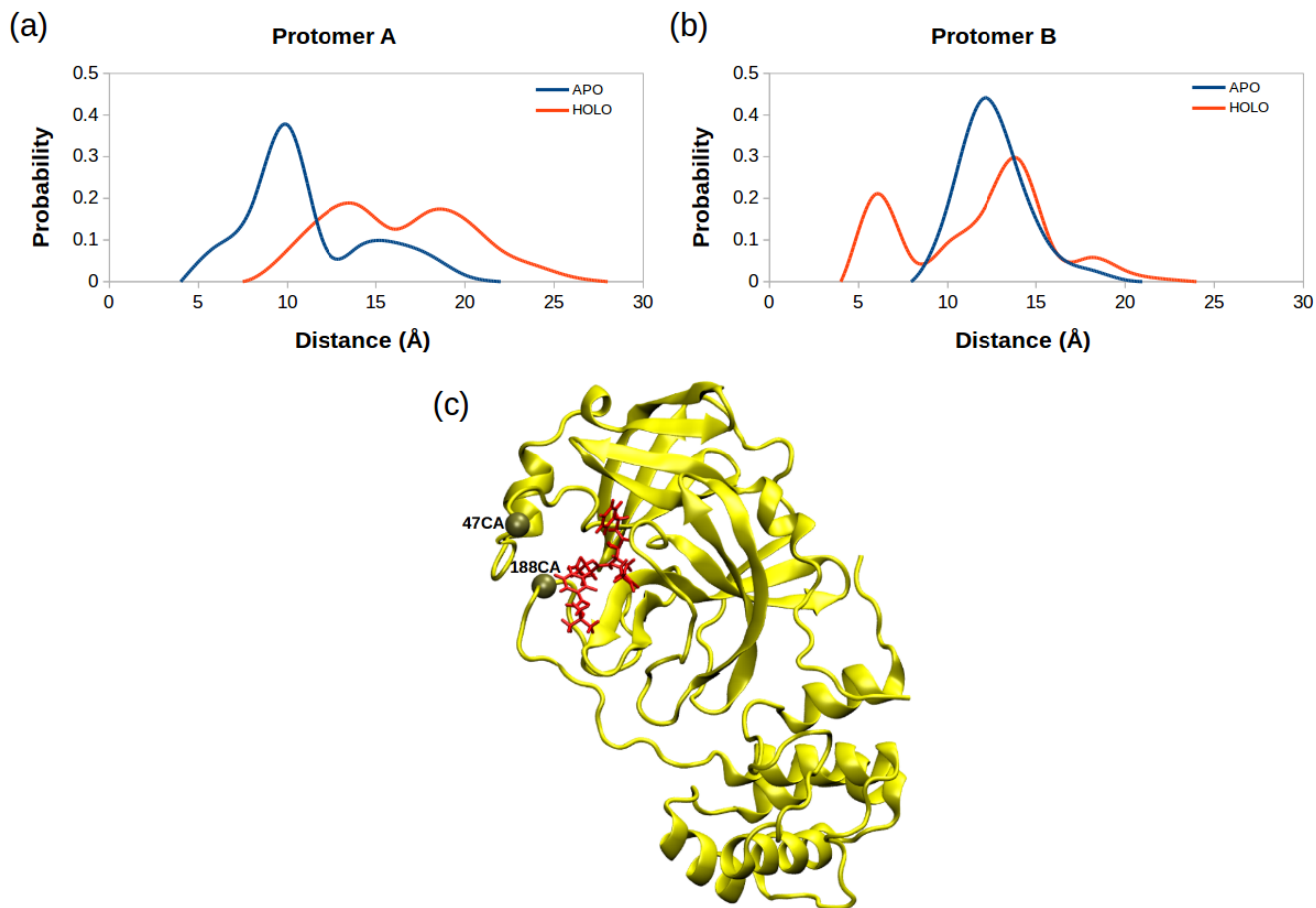


Figure 2: The probability distributions of the distance measured between C α atoms of Glu47 and Arg188 on subunit (a) A and (b) B. (c) Crystal structure of *holo* M-pro. α -ketoamide 13b is shown in licorice representation whereas the protein is shown in New Cartoon representation. The C α atoms of Glu47 and Arg188 are shown in van der Waals representation.

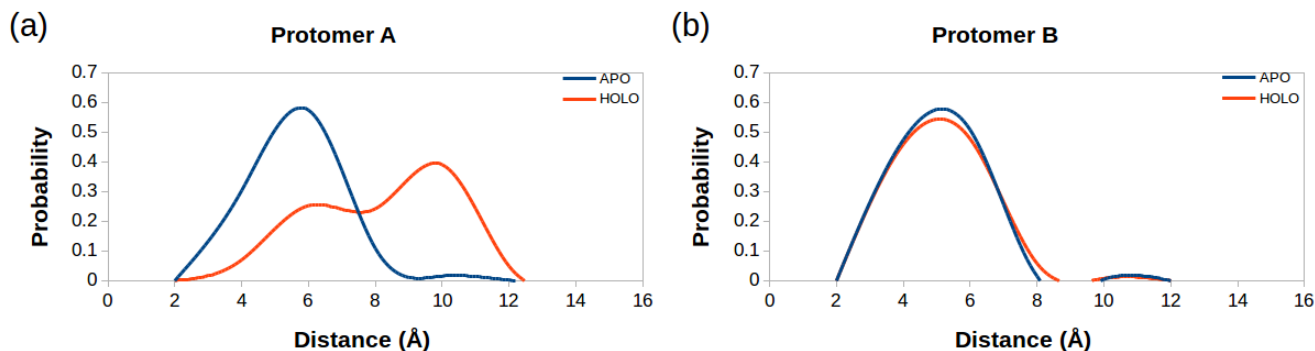


Figure 3: The probability distributions of the distance which was measured between N ϵ 2 atom of His41 and S γ atom of Cys145 for subunit (a) A and (b) B.

In addition to the above-mentioned residue-pairs, we also measured the distance between Glu166A and Ser1B since this interaction was shown to impact conformation of the SBP (See Figure 4)¹. The results showed that the ligand could slightly weaken the interaction between these two atoms as evidenced by longer distances sampled in the presence of the inhibitor (See Figure 4(Compare blue and red lines)).

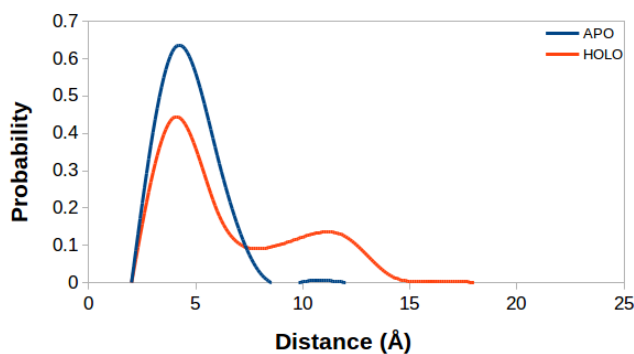


Figure 4: The probability distribution of the distance which was measured between OE2 atom of Glu166 on subunit A and N atom of Ser1 on subunit B.

Clustering of trajectories pertaining to *apo* M-pro revealed a possible binding site which is located at the dimer interface

The trajectories pertaining to *apo* M-pro were clustered depending on the conformational state of the SBP on subunit A, which was described by the distance between Glu47-Arg188.

The results showed that the SBP could be found in three conformational states, namely State 1, 2, and 3 (See Figure 5). State 1 and 2 were used in the subsequent steps of the study.

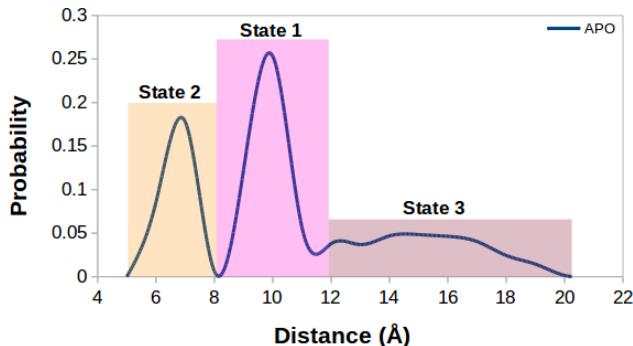


Figure 5: The probability distribution of the distance which was measured between C α atom of Glu47 and Arg188 on subunit A.

First, conformations pertaining to State 1 and 2 were assessed according to their SiteMap scores, and those with close proximity to the dimer interface and higher score were identified to have a possible binding site on the M-pro (See Figure S1). For pharmacophore modeling, we included residues not only those that were in close proximity to the identified pocket and had a central role in the catalytic activity of the enzyme and dimerization such as Glu166A, and Ser1B^{1,15,86,87}, but also the ones that showed lower frequency as a hotspot in the presence of the inhibitor, namely, Ala116A, Phe140A, Leu141A, Arg4B, and Gln299B. Moreover, we also included His172 in the pharmacophore model as it forms a hydrogen bond with Glu166, hence, presumably indirectly modulates the conformation of the SBP^{1,43}. Therefore, Ala116A, Phe140A, Leu141A, Glu166A, His172A, Ser1B, Arg4B, and Gln299B were included in the pharmacophore models (See Figure S2). Subsequently, pharmacophore features of the binding pockets were evaluated according to the organization of the groups included in the models. If the groups in the models were oriented far from each other it was not considered in the subsequent database search step. Furthermore, the ZINC⁵¹ and DrugBank^{52–56} databases were sought to find molecules that could match the properties of the pharmacophore model created. Consequently, a total of 4,174 molecules was retrieved

from the databases and docked to the identified pocket on *apo* M-pro using the Glide SP Docking scheme (See Figure 6.a and 7.a)^{60,80-82} . Candidate molecules that could interact with at least one of Glu166A, Ser1B, and Arg4B residues, were considered as promising ones due to substantial roles of these residues in dimerization and enzymatic activity^{1,15,86-89} . In this regard, six molecules were selected, whose GScores varied between -4.540 to -7.615 (See Table 1 and Figure 6 and 7).

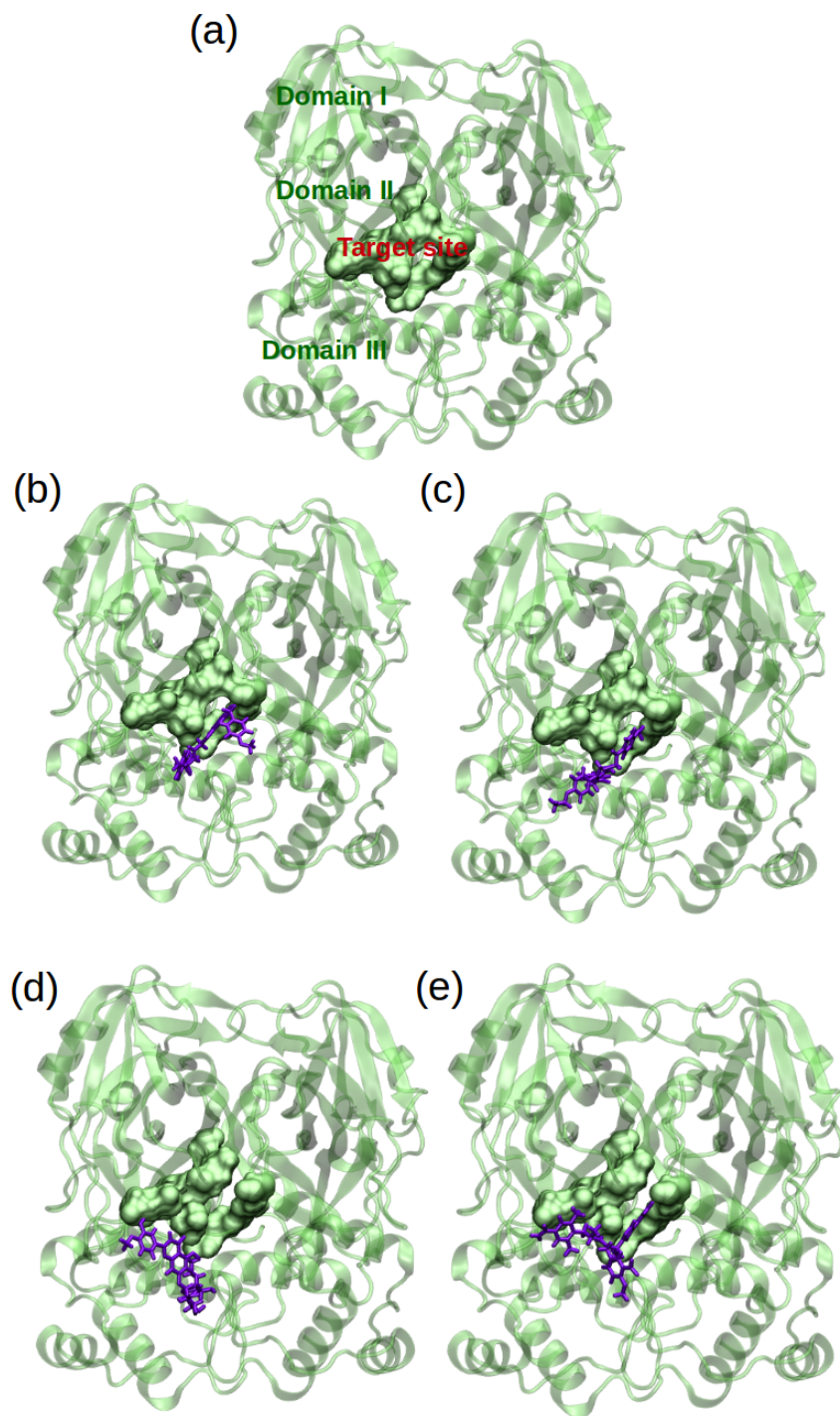


Figure 6: (a) The identified binding pocket was shown on 3D structure of M-pro in surface representation, which adopted State 1. The protein was shown in New Cartoon representation. The binding poses of (b) ZINC02948810, (c) ZINC39362669, (d) diosmin, and (e) eluxadoline within the pocket were also shown.⁹⁰

Table 1: The selected candidate molecules and their GScores

Molecule	GScore
ZINC02948810	-7.615
ZINC39362669	-7.550
Diosmin	-7.283
Eluxadoline	-7.145
ZINC20425029	-6.301
ZINC23881687	-4.540

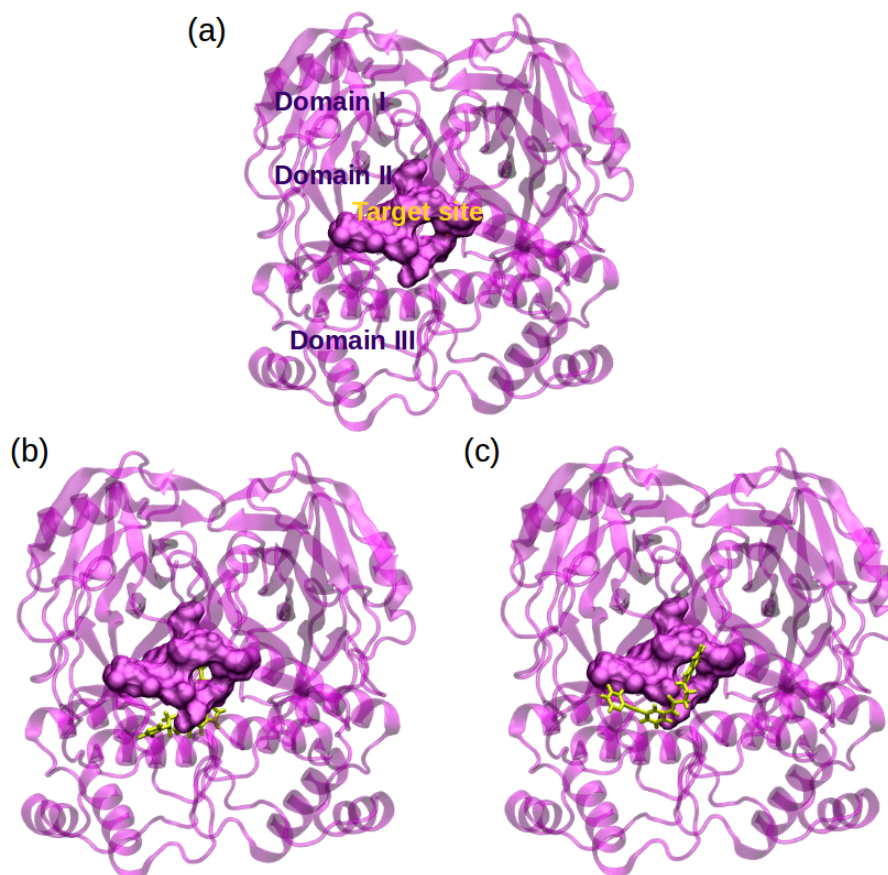


Figure 7: (a) The identified binding cavity was shown on M-pro adopting State 2. The pocket was depicted in surface representation while the protein was shown in New Cartoon representation. The binding poses of (b) ZINC20425029, and (c) ZINC23881687 were depicted.⁹⁰

Comparison of trajectories pertaining to ligand/M-pro and α -ketoamide-13b/M-pro complexes reveals similar conformational rearrangements on the substrate-binding pocket and catalytic dyad

We sought for molecules that could induce similar dynamical and structural rearrangements which were elicited by α -ketoamide-13b by means of atomistic MD simulations. To this end, we calculated and compared probability distribution plots pertaining to distances, which were measured between (i) C α atoms of Glu47 and Arg188, and (ii) N ϵ 2 atom of His41 and S γ atom of Cys145, and could identify molecules with similar propensities for mediating conformational changes that were induced by α -ketoamide-13b. Here, it is important to emphasize that the successful molecules bound to the novel target site on M-pro without need of any covalent bond formation and influenced the dynamics of the catalytic dyad as opposed to those which were covalently bound to the catalytic dyad^{17,30,91}.

The analysis of the trajectories showed that eluxadoline, which has been recently approved to be used in the treatment of Irritable Bowel Syndrome with diarrhea⁹², could induce similar conformational rearrangements induced by α -ketoamide-13b. Importantly, it was also proposed as a promising candidate for inhibition of SARS-CoV-2 M-pro in a recent molecular docking study⁹³. Specifically, eluxadoline could enable sampling of longer distances by the Glu47-Arg188 pair located on subunit A (See Figure 8.a), thereby facilitating distortion of the catalytic dyad therein, which was evidenced by relatively longer distances sampled by His41-Cys145 pair (See Figure 8.c). Interestingly; however, although it did not impact dynamics of the Glu47-Arg188 pair on subunit B (See Figure 8.b), it could slightly distort the catalytic dyad which was evidenced by right shift in the distance probability distribution (See Figure 8.d).

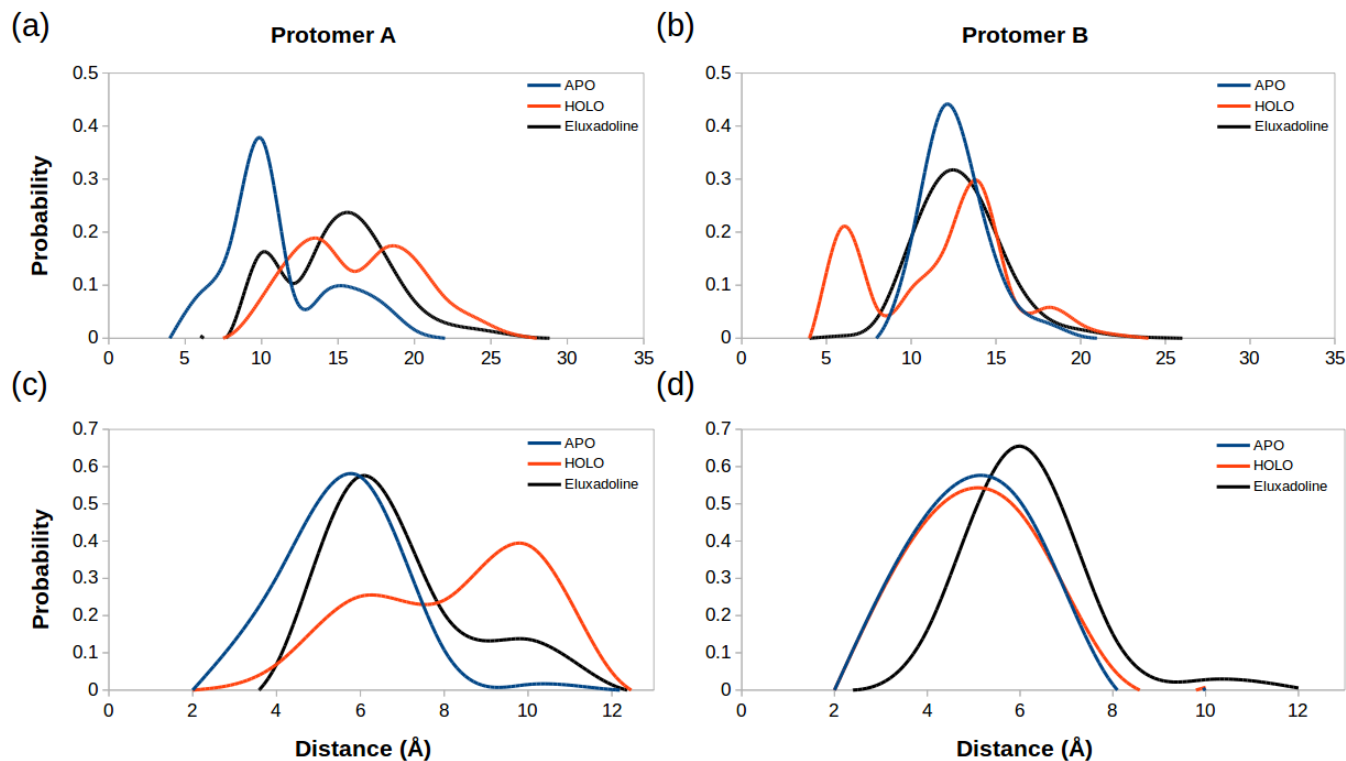


Figure 8: The probability distributions pertaining to distance which was measured between C α atoms of Glu47 and Arg188 on subunit (a) A and (b) B. The probability distributions pertaining to distance which was measured between N ϵ 2 atom of His41 and S γ atom of Cys145 on subunit (c) A and (d) B.

Another compound identified in this study was diosmin, and like eluxadoline, it has been proposed as a candidate molecule in the treatment of SARS-CoV-2 in a couple of recent *in silico* studies^{8,93,94}. It is also important to emphasize that diosmin has been currently tested to be used in a combinatorial therapy with hesperidin for prophylaxis and treatment of COVID-19⁹⁵. Analysis of the trajectories showed that diosmin caused widening of the loops on which Glu47 and Arg188 are located on subunit A. On the other hand, this rearrangement did not cause a significant change on the structure of the catalytic dyad as seen for α -ketoamide 13b on subunit A (See Figure 9.a and c). The expansion of the loops on subunit B caused a slight distortion in the catalytic dyad as opposed to α -ketoamide 13b (See Figure 9.b and d).

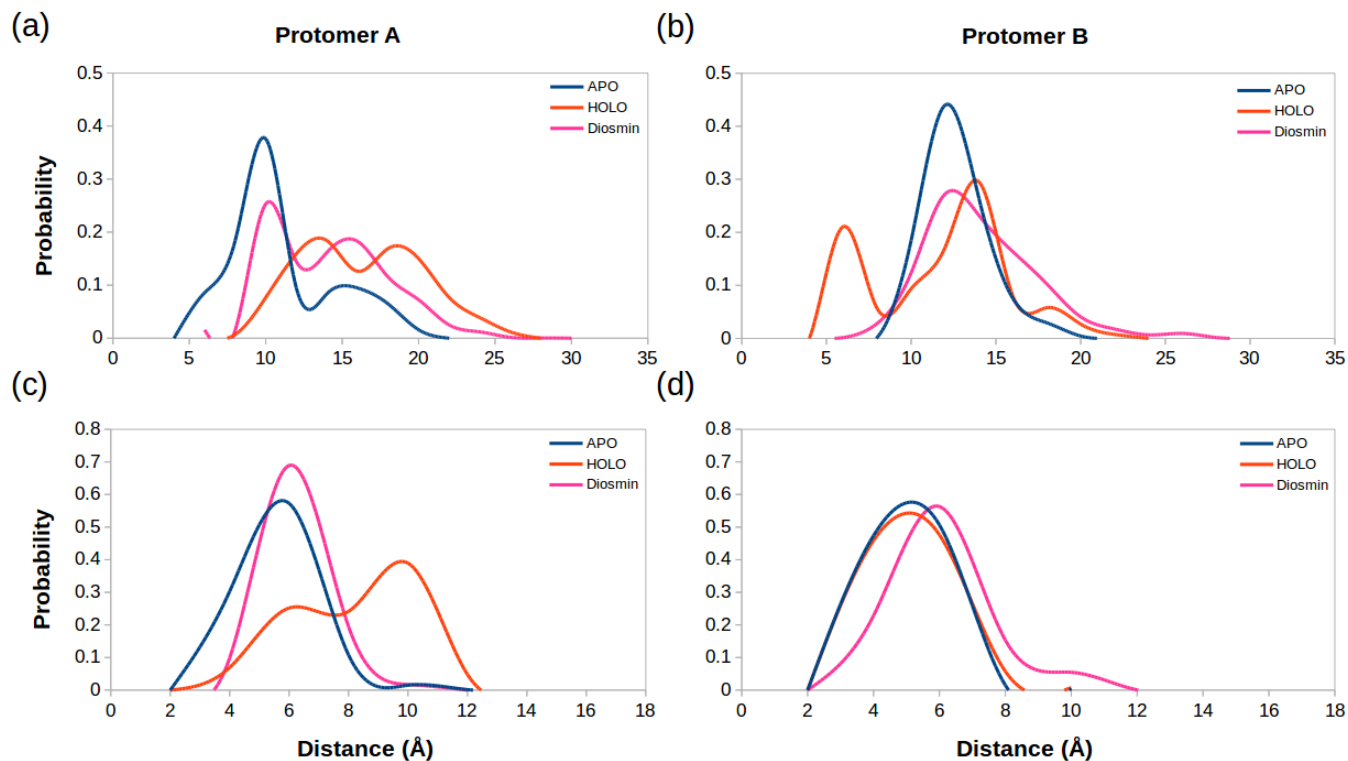


Figure 9: The probability distributions pertaining to distance which was measured between $C\alpha$ atoms of Glu47 and Arg188 on subunit (a) A and (b) B. The probability distribution pertaining to distance which was measured between $N\epsilon 2$ atom of His41 and $S\gamma$ atom of Cys145 on subunit (c) A and (d) B.

Moreover, ZINC029498810 also led to an increase in the distances measured between Glu47 and Arg188 on both subunits (See Figure 10.a and b); however, this conformational rearrangement did not distort the catalytic dyad on subunit A (See Figure 10.c), whereas it remarkably distorted the catalytic dyad on subunit B (See Figure 10.d).

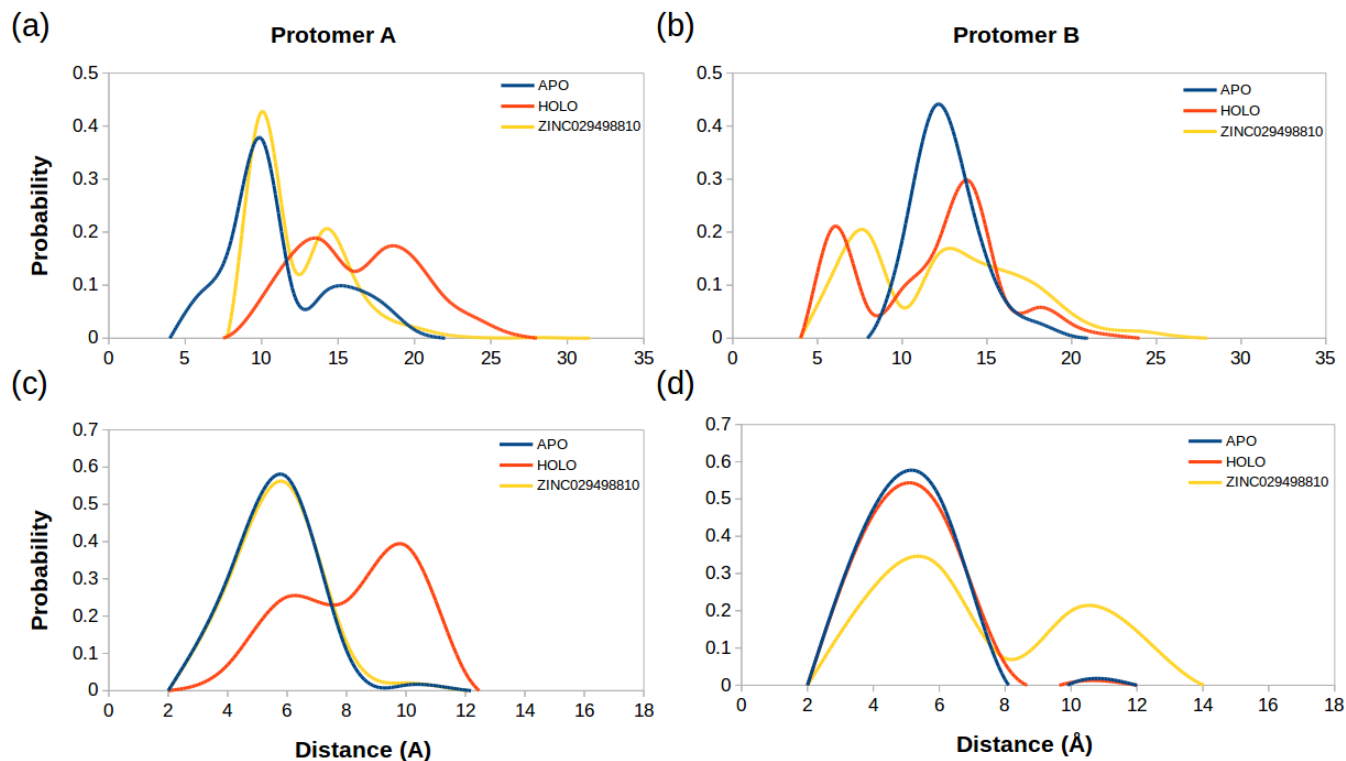


Figure 10: The probability distribution pertaining to distance which was measured between $C\alpha$ atoms of Glu47 and Arg188 on subunit (a) A and (b) B. The probability distribution pertaining to distance which was measured between $N\epsilon 2$ atom of His41 and $S\gamma$ atom of Cys145 on subunit (c) A and (d) B.

As opposed to the above-mentioned ligands, analysis of the trajectories pertaining to ZINC20425029-bound M-pro showed that the ligand did not have a significant impact on the distances measured between Glu47 and Arg188 as well as His41 and Cys145 (See Figure S3). Herein, we thought to further consider this ligand in the candidate short-list as it increased the distance between Ala285 residues (See Figure 11) located on both subunits as a result of occupation of a hydrophobic cavity, which was formed by Met6A, Ala7A, Val125A, Tyr126A, Met6B, Ala7B, Val125B, and Tyr126B, by para-fluorophenylmethyl moiety of the ligand. The importance of the Ala285 pair is due to its role in the regulation of catalytic activity of the dimer. Moreover, Ala285A and Ala285B were oriented closer to each other in SARS-CoV-2 than in SARS-CoV due to Thr to Ala mutation^{1,20}. Therefore, modulation of the Ala285 pair might prevent formation of a catalytically competent dimer.

Recalling the GScore values provided in Table 1, that of ZINC23881687 turned out to be smaller than those of other candidates. On the other hand, the ligand was included in the subsequent molecular dynamics studies as its docking pose was found to make contacts with crucial residues such as Glu166, Ser1 and Arg4. Detailed analysis of the trajectories showed that the ligand did not induce significant changes in the distances measured between Glu47 and Arg188 as well as His41 and Cys145 (See Figure S4) similar to ZINC20425029. On the other hand, it could make H-bonds with main and side chains of the Ser1B (See Figure 12). Considering the fact that Ser1B is involved in the dimerization of M-pro and modulation of the SBP the perturbation of the residue might help interfere with the function of the M-pro enzyme.

Finally, it is also crucial to stress that although ZINC39362669 and ZINC02948810 had close GScore values, the former did not have a significant impact on the structural properties discussed for other candidates (See Figure S5).

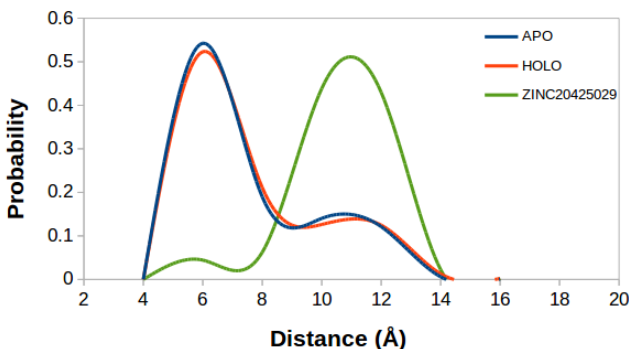


Figure 11: The probability distributions of distances measured between C α atoms of Ala285 on both subunits.

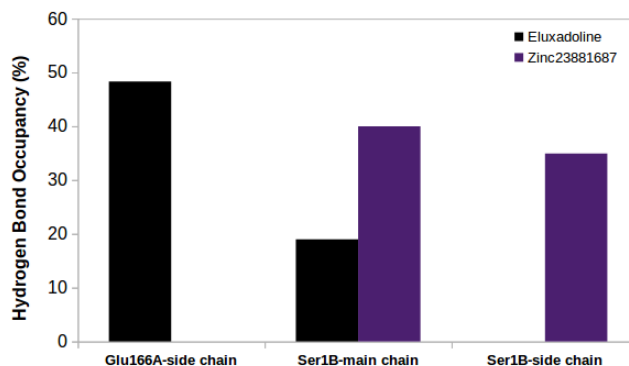


Figure 12: Hydrogen bond occupancies of the ligands with Glu166A and Ser1B.

The global dynamics of M-pro is differently modulated by the ligands and α -ketoamide 13b

In addition to investigation of local structural/dynamical features of the systems, we also carried out an essential dynamics analysis to explore dominant collective motions of the systems studied. Firstly, we presented 2D projections of the trajectories with respect to the first two eigenvectors of the *holo* system, which cumulatively accounted for 36% and 33% of the overall motions of subunit A and B, respectively. The results showed that all the ligands, which were found to be successful in terms of local modulation of the M-pro, caused confinement of the conformational space available to subunit A (See Figure 13 and S7). For subunit B, it was shown that all the candidates could confine the spanned space except diosmin and ZINC02948810 (See Figure 13 and S7).

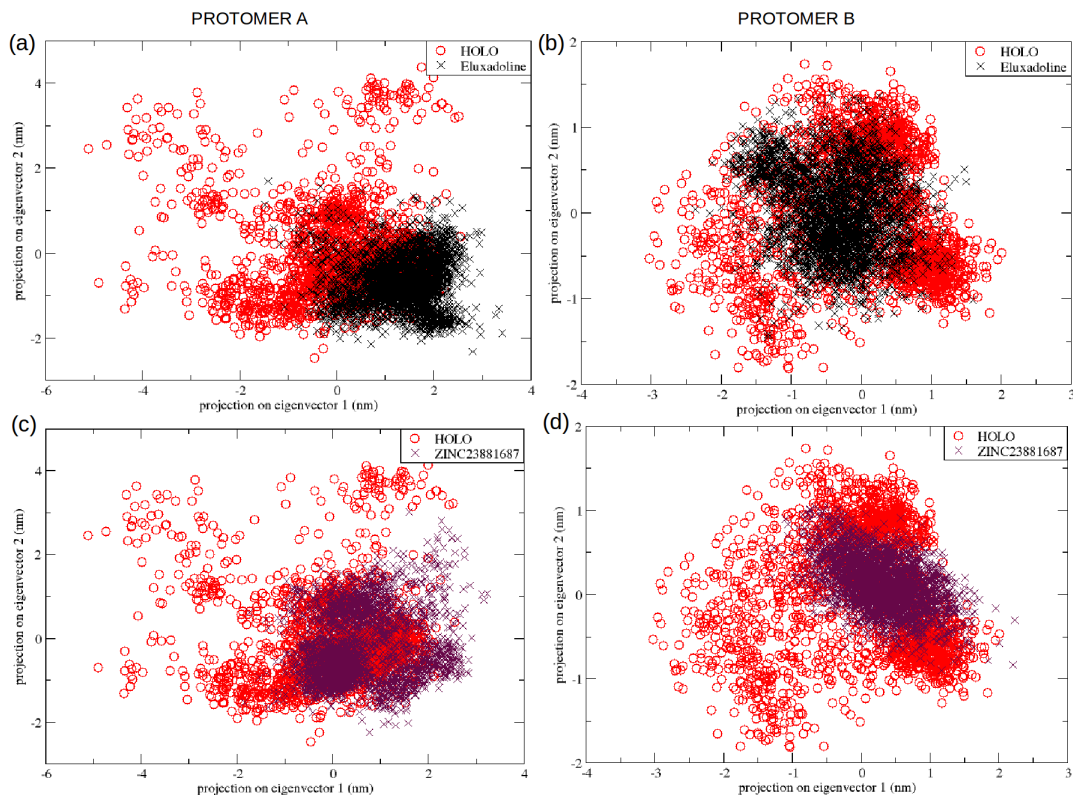


Figure 13: 2D projections of subunit A of (a) eluxadoline- and (c) ZINC23881687-bound systems as well as subunit B of the (b) eluxadoline- and (d) ZINC23881687-bound systems with respect to the first two eigenvectors of *holo* M-pro. The first and second principal components of subunit A count for 20.368% and 16.045%, whereas those of subunit B account for 20.715% and 12.158%, respectively.

Moreover, we also projected trajectories of ligand/M-pro systems on 3D structure of M-pro along with their first eigenvectors and compared them with that of *holo* (See Figure 14.a) to shed some light on regions that were influenced by the ligand. We observed that α -ketoamide-13b mainly affected dynamics of the loops on which Glu47 and Arg188 were located on both subunits of the protein (See Figure 14.a). In line with this, eluxadoline impacted the same loops (See Figure 14.b) on both subunits. Similarly, ZINC23881687 also affected the same loops on both protomers in addition to Domain III of subunit B, which is extensively involved in dimerization to form a functional M-pro homodimer (See Figure 14.c)^{1,35,86}. Diosmin substantially altered the collective dynamics of the above-mentioned loops located on subunit B as well as Domain III of subunit A (See Figure S8.a).

As it is consistent with the mentioned local alterations on the catalytic dyads, the first principal component pertaining to ZINC02948810 showed that the ligand caused changes on the loops, on which Glu47 and Arg188 were located, on subunit B, while it did not impact the corresponding loop on subunit A (See Figure S8.b).

Lastly, ZINC20425029 displayed its effect on the loops found on both protomers as well as Domain IIIs, thus thought of as a promising candidate like other mentioned molecules (See Figure S8.c).

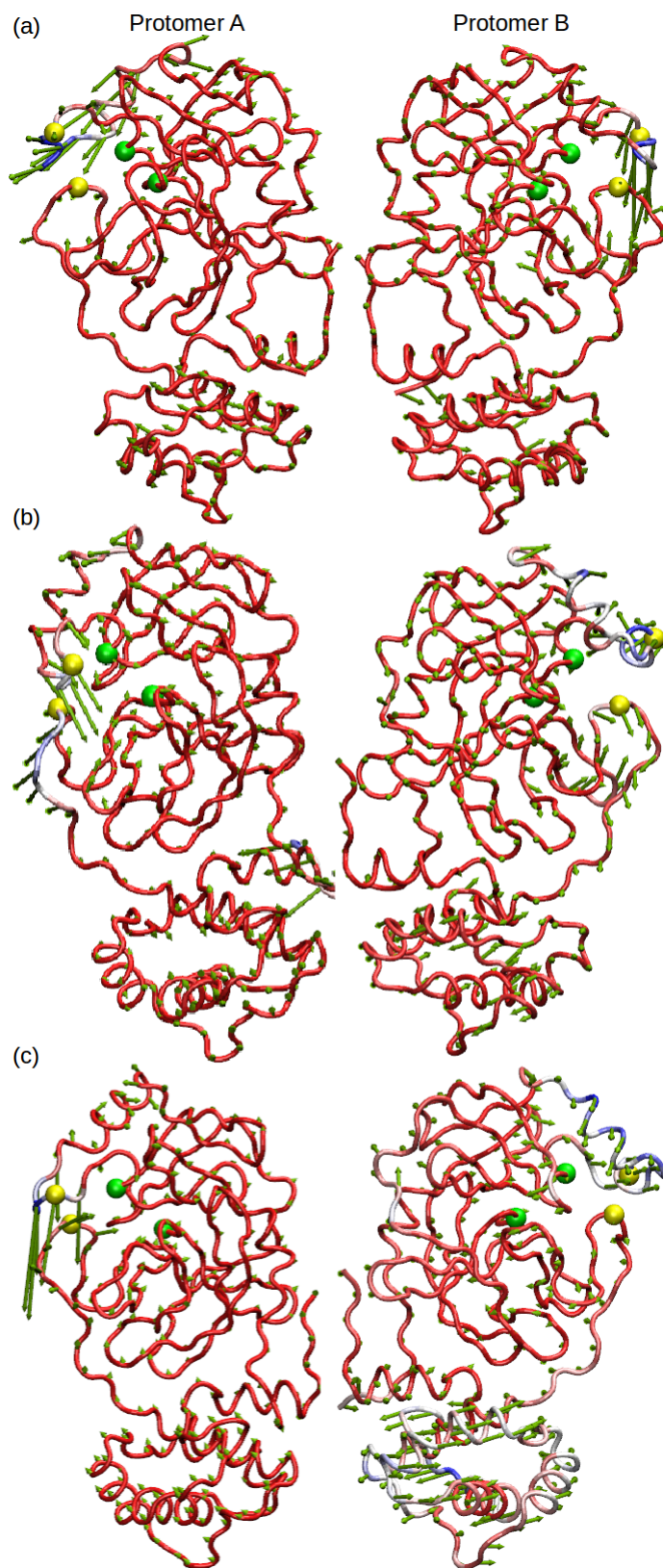


Figure 14: The projections of (a) α -ketoamide-13b-, (b) eluxadoline-, and (c) ZINC23881687-bound M-pro complexes with respect to the first eigenvectors of the systems

In conclusion, analysis of trajectories pertaining to *apo* M-pro enzyme revealed a novel binding pocket at the dimer interface, which can be targeted by small therapeutic molecules. Indeed, we showed that occupation of the pocket by eluxadoline, diosmin, and ZINC02948810 could alter conformational properties of both SBP and the catalytic dyad similarly as elicited by α -ketoamide-13b. On the other hand, although ZINC20425029 and ZINC23881687 did not impact the SBP and the catalytic dyad, they induced perturbations on regions that were important for the functional activity and dimerization of the enzyme. Specifically, eluxadoline and ZINC23881687 could make stable H-bonds with Glu166A and Ser1B, respectively two of which are involved in the dimerization. Hence, perturbation of inter-molecular hydrogen bonding network of these two residues might help prevent formation of functional M-pro dimers. In parallel to this, it was revealed in global dynamics of ZINC23881687-bound M-pro as higher distortion observed in Domain III of the protein. Interestingly, ZINC20425029 impacted interaction between Ala285 residues located both subunits. Since the Ala285 pair has been shown to be important for dimer formation, perturbation of the interaction in the Ala285 pair might prevent formation of the catalytically competent dimers. Moreover, despite having similar GScore values with ZINC02948810, ZINC39362669 did not impact properties of either the SBP or the catalytic dyad emphasizing the fact that binding score values, which were obtained from molecular docking studies, might not be enough for precise assessment of successful candidates. As opposed to similar local conformational changes induced by ligands and α -ketoamide-13b on the M-pro enzyme, we observed differences in the global dynamics of the protein. Notably, the ligands confined conformational space available to subunit A and B, while α -ketoamide-13b allowed the protein to sample a wider conformational space compared to that of *apo* M-pro. This might be due to the fact M-pro might be stuck in one of the minima on the energy surface by means of ligands that help maintain M-pro in the catalytically incompetent state.

Acknowledgement

HP, MI, and OS would like to thank Istanbul Medipol University for providing access to the High-Performance Computing System so as to run some of the MD simulations. The rest of the numerical calculations reported in this paper were performed at TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources).

Supporting Information Available

The following materials are available free of charge.

Figures S1-S8

References

- (1) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α -ketoamide inhibitors. *Science* **2020**, *368*, 409–412.
- (2) Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L., et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *nature* **2020**, *579*, 270–273.
- (3) Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y., et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265–269.
- (4) Sohrabi, C.; Alsafi, Z.; O'Neill, N.; Khan, M.; Kerwan, A.; Al-Jabir, A.; Iosifidis, C.; Agha, R. World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International Journal of Surgery* **2020**,

- (5) Lai, C.-C.; Ko, W.-C.; Lee, P.-I.; Jean, S.-S.; Hsueh, P.-R. Extra-respiratory manifestations of COVID-19. *International Journal of Antimicrobial Agents* **2020**, 106024.
- (6) Wu, C.; Chen, X.; Cai, Y.; Zhou, X.; Xu, S.; Huang, H.; Zhang, L.; Zhou, X.; Du, C.; Zhang, Y., et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA internal medicine* **2020**,
- (7) Gao, Y.-M.; Xu, G.; Wang, B.; Liu, B.-C. Cytokine storm syndrome in coronavirus disease 2019: A narrative review. *Journal of internal medicine* **2020**,
- (8) Ngo, S. T.; Quynh Anh Pham, N.; Thi Le, L.; Pham, D.-H.; Vu, V. V. Computational determination of potential inhibitors of SARS-CoV-2 main protease. *Journal of chemical information and modeling* **2020**,
- (9) Schoeman, D.; Fielding, B. C. Coronavirus envelope protein: current knowledge. *Virology journal* **2019**, *16*, 1–22.
- (10) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A., et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **2020**,
- (11) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veasler, D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **2020**,
- (12) Uhal, B. D.; Li, X.; Xue, A.; Gao, X.; Abdul-Hafez, A. Regulation of alveolar epithelial cell survival by the ACE-2/angiotensin 1–7/Mas axis. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **2011**, *301*, L269–L274.
- (13) Mossel, E. C.; Wang, J.; Jeffers, S.; Edeen, K. E.; Wang, S.; Cosgrove, G. P.; Funk, C. J.;

- Manzer, R.; Miura, T. A.; Pearson, L. D., et al. SARS-CoV replicates in primary human alveolar type II cell cultures but not in type I-like cells. *Virology* **2008**, *372*, 127–135.
- (14) Xu, H.; Zhong, L.; Deng, J.; Peng, J.; Dan, H.; Zeng, X.; Li, T.; Chen, Q. High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral mucosa. *International journal of oral science* **2020**, *12*, 1–5.
- (15) Hilgenfeld, R. From SARS to MERS: crystallographic studies on coronaviral proteases enable antiviral drug design. *The FEBS journal* **2014**, *281*, 4085–4096.
- (16) Liu, W.; Morse, J. S.; Lalonde, T.; Xu, S. Learning from the past: possible urgent prevention and treatment options for severe acute respiratory infections caused by 2019-nCoV. *Chembiochem* **2020**,
- (17) Jin, Z.; Zhao, Y.; Sun, Y.; Zhang, B.; Wang, H.; Wu, Y.; Zhu, Y.; Zhu, C.; Hu, T.; Du, X., et al. Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur. *Nature structural & molecular biology* **2020**, *27*, 529–532.
- (18) Hegyi, A.; Ziebuhr, J. Conservation of substrate specificities among coronavirus main proteases. *Journal of general virology* **2002**, *83*, 595–599.
- (19) Liang, J.; Pitsillou, E.; Karagiannis, C.; Darmawan, K. K.; Ng, K.; Hung, A.; Karagianis, T. C. Interaction of the prototypical α -ketoamide inhibitor with the SARS-CoV-2 main protease active site in silico: Molecular dynamic simulations highlight the stability of the ligand-protein complex. *Computational Biology and Chemistry* **2020**, 107292.
- (20) Anand, K.; Palm, G. J.; Mesters, J. R.; Siddell, S. G.; Ziebuhr, J.; Hilgenfeld, R. Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra α -helical domain. *The EMBO journal* **2002**, *21*, 3213–3224.

- (21) Yang, H.; Yang, M.; Ding, Y.; Liu, Y.; Lou, Z.; Zhou, Z.; Sun, L.; Mo, L.; Ye, S.; Pang, H., et al. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proceedings of the National Academy of Sciences* **2003**, *100*, 13190–13195.
- (22) Pillaiyar, T.; Manickam, M.; Namasivayam, V.; Hayashi, Y.; Jung, S.-H. An overview of severe acute respiratory syndrome–coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy. *Journal of medicinal chemistry* **2016**, *59*, 6595–6628.
- (23) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic acids research* **2000**, *28*, 235–242.
- (24) Burley, S. K.; Berman, H. M.; Bhikadiya, C.; Bi, C.; Chen, L.; Di Costanzo, L.; Christie, C.; Dalenberg, K.; Duarte, J. M.; Dutta, S., et al. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic acids research* **2019**, *47*, D464–D474.
- (25) Su, H.-x.; Yao, S.; Zhao, W.-f.; Li, M.-j.; Liu, J.; Shang, W.-j.; Xie, H.; Ke, C.-q.; Hu, H.-c.; Gao, M.-n., et al. Anti-SARS-CoV-2 activities in vitro of Shuanghuanglian preparations and bioactive ingredients. *Acta Pharmacologica Sinica* **2020**, *41*, 1167–1177.
- (26) Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J., et al. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol* **2005**, *3*, e324.
- (27) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. Coronavirus main

- proteinase (3CLpro) structure: basis for design of anti-SARS drugs. *Science* **2003**, *300*, 1763–1767.
- (28) Muramatsu, T.; Takemoto, C.; Kim, Y.-T.; Wang, H.; Nishii, W.; Terada, T.; Shirouzu, M.; Yokoyama, S. SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. *Proceedings of the National Academy of Sciences* **2016**, *113*, 12997–13002.
- (29) Li, Q.; Kang, C. Progress in developing inhibitors of SARS-CoV-2 3C-like protease. *Microorganisms* **2020**, *8*, 1250.
- (30) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C., et al. Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, 1–5.
- (31) Kneller, D. W.; Phillips, G.; O'Neill, H. M.; Jedrzejczak, R.; Stols, L.; Langan, P.; Joachimiak, A.; Coates, L.; Kovalevsky, A. Structural Plasticity of the SARS-CoV-2 3CL Mpro Active Site Cavity Revealed by Room Temperature X-ray Crystallography. **2020**,
- (32) Zhang, L.; Lin, D.; Kusov, Y.; Nian, Y.; Ma, Q.; Wang, J.; Von Brunn, A.; Leyssen, P.; Lanko, K.; Neyts, J., et al. α -Ketoamides as broad-spectrum inhibitors of coronavirus and enterovirus replication: Structure-based design, synthesis, and activity assessment. *Journal of medicinal chemistry* **2020**,
- (33) Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L. 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry* **2004**, *43*, 4568–4574.
- (34) He, J.; Hu, L.; Huang, X.; Wang, C.; Zhang, Z.; Wang, Y.; Zhang, D.; Ye, W. Potential of coronavirus 3C-like protease inhibitors for the development of new anti-SARS-CoV-

- 2 drugs: Insights from structures of protease and inhibitors. *International Journal of Antimicrobial Agents* **2020**, 106055.
- (35) Shi, J.; Song, J. The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain. *The FEBS journal* **2006**, *273*, 1035–1045.
- (36) Shi, J.; Wei, Z.; Song, J. Dissection Study on the Severe Acute Respiratory Syndrome 3C-like Protease reveals the critical role of the extra domain in dimerization of the enzyme defining the extra domain as a new target for design of highly specific protease inhibitors. *Journal of Biological Chemistry* **2004**, *279*, 24765–24773.
- (37) Paasche, A.; Zipper, A.; Schäfer, S.; Ziebuhr, J.; Schirmeister, T. Evidence for substrate binding-induced zwitterion formation in the catalytic Cys-His dyad of the SARS-CoV main protease. *Biochemistry* **2014**, *53*, 5930–5946.
- (38) Hosseini-Zare, M. S.; Thilagavathi, R.; Selvam, C. Targeting severe acute respiratory syndrome-coronavirus (SARS-CoV-1) with structurally diverse inhibitors: a comprehensive review. *RSC Advances* **2020**, *10*, 28287–28299.
- (39) Dai, W.; Zhang, B.; Jiang, X.-M.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Peng, J.; Liu, F., et al. Structure-based design of antiviral drug candidates targeting the SARS-CoV-2 main protease. *Science* **2020**, *368*, 1331–1335.
- (40) Ullrich, S.; Nitsche, C. The SARS-CoV-2 main protease as drug target. *Bioorganic & Medicinal Chemistry Letters* **2020**, 127377.
- (41) Frecer, V.; Miertus, S. Antiviral agents against COVID-19: structure-based design of specific peptidomimetic inhibitors of SARS-CoV-2 main protease. *RSC Advances* **2020**, *10*, 40244–40263.
- (42) Bauer, R. A. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. *Drug discovery today* **2015**, *20*, 1061–1073.

- (43) Chen, H.; Wei, P.; Huang, C.; Tan, L.; Liu, Y.; Lai, L. Only one protomer is active in the dimer of SARS 3C-like proteinase. *Journal of Biological Chemistry* **2006**, *281*, 13894–13898.
- (44) Hill, T. L.; Levitzki, A. Subunit neighbor interactions in enzyme kinetics: half-of-the-sites reactivity in a dimer. *Proceedings of the National Academy of Sciences* **1980**, *77*, 5741–5745.
- (45) Onawole, A. T.; Sulaiman, K. O.; Kolapo, T. U.; Akinde, F. O.; Adegoke, R. O. COVID-19: CADD to the rescue. *Virus Research* **2020**, 198022.
- (46) Kapetanovic, I. Computer-aided drug discovery and development (CADD): in silico-chemico-biological approach. *Chemico-biological interactions* **2008**, *171*, 165–176.
- (47) Leelananda, S. P.; Lindert, S. Computational methods in drug discovery. *Beilstein journal of organic chemistry* **2016**, *12*, 2694–2718.
- (48) Macalino, S. J. Y.; Gosu, V.; Hong, S.; Choi, S. Role of computer-aided drug design in modern drug discovery. *Archives of pharmacal research* **2015**, *38*, 1686–1701.
- (49) Manas, E. S.; Green, D. V. CADD medicine: design is the potion that can cure my disease. *Journal of computer-aided molecular design* **2017**, *31*, 249–253.
- (50) Melo-Filho, C. C.; Braga, R. C.; Muratov, E. N.; Franco, C. H.; Moraes, C. B.; Freitas-Junior, L. H.; Andrade, C. H. Discovery of new potent hits against intracellular *Trypanosoma cruzi* by QSAR-based virtual screening. *European Journal of Medicinal Chemistry* **2019**, *163*, 649–659.
- (51) Koes, D. R.; Camacho, C. J. ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic acids research* **2012**, *40*, W409–W414.
- (52) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.;

- Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z., et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic acids research* **2018**, *46*, D1074–D1082.
- (53) Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A. C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V., et al. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic acids research* **2014**, *42*, D1091–D1097.
- (54) Knox, C.; Law, V.; Jewison, T.; Liu, P.; Ly, S.; Frolkis, A.; Pon, A.; Banco, K.; Mak, C.; Neveu, V., et al. DrugBank 3.0: a comprehensive resource for ‘omics’ research on drugs. *Nucleic acids research* **2010**, *39*, D1035–D1041.
- (55) Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic acids research* **2008**, *36*, D901–D906.
- (56) Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research* **2006**, *34*, D668–D672.
- (57) Humphrey, W.; Dalke, A.; Schulten, K., et al. VMD: visual molecular dynamics. *Journal of molecular graphics* **1996**, *14*, 33–38.
- (58) Coutsiias, E. A.; Seok, C.; Jacobson, M. P.; Dill, K. A. A kinematic view of loop closure. *Journal of computational chemistry* **2004**, *25*, 510–528.
- (59) Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of computer-aided molecular design* **2013**, *27*, 221–234.
- (60) Release, S. 4: Glide. *Schrödinger, LLC, New York, NY* **2018**,
- (61) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry* **2008**, *29*, 1859–1865.

- (62) Brooks, B. R.; Brooks III, C. L.; Mackerell Jr, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S., et al. CHARMM: the biomolecular simulation program. *Journal of computational chemistry* **2009**, *30*, 1545–1614.
- (63) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y., et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Journal of chemical theory and computation* **2016**, *12*, 405–413.
- (64) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M., et al. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *Journal of computational chemistry* **2014**, *35*, 1997–2004.
- (65) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nature methods* **2017**, *14*, 71–73.
- (66) Mark, P.; Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *The Journal of Physical Chemistry A* **2001**, *105*, 9954–9960.
- (67) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*, 19–25.
- (68) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of chemical physics* **1984**, *81*, 511–519.
- (69) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A* **1985**, *31*, 1695.

- (70) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics* **1981**, *52*, 7182–7190.
- (71) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry* **1997**, *18*, 1463–1472.
- (72) Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L., et al. OPLS3e: Extending force field coverage for drug-like small molecules. *Journal of chemical theory and computation* **2019**, *15*, 1863–1874.
- (73) Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chemical biology & drug design* **2007**, *69*, 146–148.
- (74) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *Journal of chemical information and modeling* **2009**, *49*, 377–389.
- (75) Tuncbag, N.; Keskin, O.; Gursoy, A. HotPoint: hot spot prediction server for protein interfaces. *Nucleic acids research* **2010**, *38*, W402–W406.
- (76) Salam, N. K.; Nuti, R.; Sherman, W. Novel method for generating structure-based pharmacophores using energetic analysis. *Journal of chemical information and modeling* **2009**, *49*, 2356–2368.
- (77) Loving, K.; Salam, N. K.; Sherman, W. Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation. *Journal of computer-aided molecular design* **2009**, *23*, 541–554.
- (78) Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. A. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *Journal of computer-aided molecular design* **2006**, *20*, 647–671.

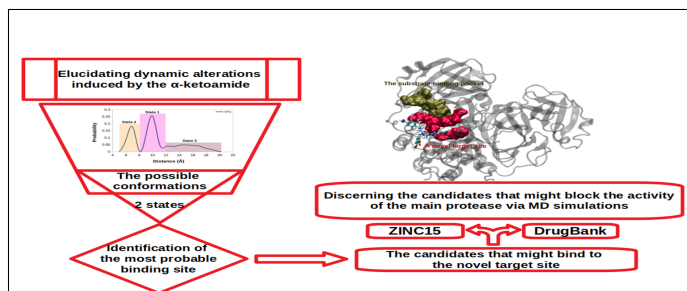
- (79) Dixon, S. L.; Smondryev, A. M.; Rao, S. N. PHASE: A Novel Approach to Pharmacophore Modeling and 3D Database Searching. *Chemical Biology & Drug Design* **2006**, *67*, 370–372.
- (80) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. *Journal of medicinal chemistry* **2006**, *49*, 6177–6196.
- (81) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of medicinal chemistry* **2004**, *47*, 1750–1759.
- (82) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K., et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* **2004**, *47*, 1739–1749.
- (83) Kim, S.; Lee, J.; Jo, S.; Brooks III, C. L.; Lee, H. S.; Im, W. CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *Journal of Computational Chemistry* **2017**, *38*, 1879–1886.
- (84) Bakan, A.; Meireles, L. M.; Bahar, I. ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* **2011**, *27*, 1575–1577.
- (85) Bakan, A.; Dutta, A.; Mao, W.; Liu, Y.; Chennubhotla, C.; Lezon, T. R.; Bahar, I. Evol and ProDy for bridging protein sequence evolution and structural dynamics. *Bioinformatics* **2014**, *30*, 2681–2683.
- (86) Douangamath, A.; Fearon, D.; Gehrtz, P.; Krojer, T.; Lukacik, P.; Owen, C. D.; Resnick, E.; Strain-Damerell, C.; Ábrányi-Balogh, P.; Brandaõ-Neto, J., et al. Crys-

- tallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. *bioRxiv* **2020**,
- (87) Cheng, S.-C.; Chang, G.-G.; Chou, C.-Y. Mutation of Glu-166 blocks the substrate-induced dimerization of SARS coronavirus main protease. *Biophysical journal* **2010**, *98*, 1327–1336.
 - (88) Chou, C.-Y.; Chang, H.-C.; Hsu, W.-C.; Lin, T.-Z.; Lin, C.-H.; Chang, G.-G. Quaternary structure of the severe acute respiratory syndrome (SARS) coronavirus main protease. *Biochemistry* **2004**, *43*, 14958–14970.
 - (89) Goyal, B.; Goyal, D. Targeting the dimerization of main protease of coronaviruses: A potential broad-spectrum therapeutic strategy. *ACS Combinatorial Science* **2020**,
 - (90) Stone, J. *An Efficient Library for Parallel Ray Tracing and Animation*. M.Sc. thesis, Computer Science Department, University of Missouri-Rolla, 1998.
 - (91) Vuong, W.; Khan, M. B.; Fischer, C.; Arutyunova, E.; Lamer, T.; Shields, J.; Safra, H. A.; McKay, R. T.; van Belkum, M. J.; Joyce, M., et al. Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication. *bioRxiv* **2020**,
 - (92) Simrén, M.; Tack, J. New treatments and therapeutic targets for IBS and other functional bowel disorders. *Nature Reviews Gastroenterology & Hepatology* **2018**, *15*, 589–605.
 - (93) Chen, Y. W.; Yiu, C.-P. B.; Wong, K.-Y. Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like protease (3CL pro) structure: virtual screening reveals velpatasvir, ledipasvir, and other drug repurposing candidates. *F1000Research* **2020**, *9*.
 - (94) Arun, K.; Sharanya, C.; Abhithaj, J.; Francis, D.; Sadasivan, C. Drug repurposing against SARS-CoV-2 using E-pharmacophore based virtual screening, molecular dock-

ing and molecular dynamics with main protease as the target. *Journal of Biomolecular Structure and Dynamics* **2020**, 1–12.

- (95) Haggag, Y. A.; El-Ashmawy, N. E.; Okasha, K. M. Is hesperidin essential for prophylaxis and treatment of COVID-19 Infection? *Medical Hypotheses* **2020**, 109957.

For Table of Contents Use Only



Inhibition of SARS-CoV-2 main protease: A repurposing study that target dimer interface of the protein

Hanife Pekel, Metehan Ilter, and Ozge Sensoy*