PAPER

Drug repurposing and computational modeling for discovery of inhibitors of the main protease (M^pro) of SARS-CoV-2†

José Rogério A. Silva, [iD] *[a] Hendrik G. Kruger [iD] [b] and Fábio A. Molfetta*[c]

The main protease (M^pro or 3CL^pro) is a conserved cysteine protease from the coronaviruses and started to be considered an important drug target for developing antivirals, as it produced a deadly outbreak of COVID-19. Herein, we used a combination of drug reposition and computational modeling approaches including molecular docking, molecular dynamics (MD) simulations, and the calculated binding free energy to evaluate a set of drugs in complex with the M^pro enzyme. Particularly, our results show that darunavir and triptorelin drugs have favorable binding free energy (-63.70 and -77.28 kcal mol^{-1}, respectively) in complex with the M^pro enzyme. Based on the results, the structural and energetic features that explain why some drugs can be repositioned to inhibit M^pro from SARS-CoV-2 were exposed. These features should be considered for the design of novel M^pro inhibitors.

Introduction

Coronaviruses (CoVs) are enveloped viruses with a diameter of 60 to 130 nm that contain a single-stranded positive ribonucleic acid (RNA) genome and belong to the Coronaviridae family, which is divided into four genera (α, β, γ e δ), in which SARS-CoV belongs to the genus Betacoronavirus.[1,2]

COVID-19 (COronaVIrus Disease 2019) is a disease caused by the SARS-CoV-2 coronavirus that has spread rapidly around the world since it was reported in late 2019 in Wuhan, China.[3] SARS-CoV-2 is the seventh coronavirus known to infect humans and that can cause serious diseases, such as SARS-CoV and the Middle East respiratory syndrome (MERS-CoV),[4] while HKU1,[5] NL63,[6] OC43 (ref. 7) and 229E[8] are associated with mild symptoms. SARS-CoV-2 is the one with the highest transmission rate, as it includes asymptomatic carriers, a long latency period, and a high infection rate.[9] COVID-19 causes fever, fatigue, dry cough, muscle pain, shortness of breath, and in some cases lead to pneumonia.[10] In addition, a small proportion of those infected (∼5%) develop diseases that can progress to severe acute respiratory syndrome (SARS), which can lead to death.[9] Since the initial SARS-CoV-2 outbreak, the World Health Organization (WHO) declared on March 11, 2020, COVID-19 as a pandemic disease due to the spread of the coronavirus.[11,12] Until April 26, 2021, the total number of confirmed cases of COVID-19 was 146 841 882 with over 3 104 743 deaths in at least 223 countries,[13] which indicates the epidemic is a serious public health problem.

The 3-chymotrypsin-like protease enzyme (3CL^pro), also known as the "main protease" (M^pro), is considered an attractive target in drug design for the treatment of coronavirus infection.[14,15] The role of this protease involves the proteolytic processing of viral polyproteins, being considered essential in the process of replication and maturation of the virus.[16] The M^pro cysteine protease contains three domains, domain I being formed by residues 1 to 101, and domain II consisting of antiparallel barrels that are formed by residues 102 to 184, and for enzyme activity domain III, composed mainly of alpha helices, formed by residues 201 to 301.[17,18] The substrate-binding region is located in the gap between domains I and II and consists of a catalytic dyad formed by the residues Hys41 and Cys145, in which cysteine acts as a nucleophile and histidine as a proton acceptor.[14,19] In addition, as M^pro does not have a homologous human enzyme, it presents itself as an ideal target for drug design.[15,19,20]

A traditional drug discovery process is expensive and time-consuming and can take decades to complete.[21] Currently, there are no drugs available for the treatment of COVID-19, and in the absence of effective treatment, the redirection of drugs becomes an attractive solution, due to the reduction in development time and cost, as the drugs have already been tested to what concerns safety in humans, which means they should quickly enter the most advanced clinical stages.[22,23]

Structure-Based Virtual Screening (SBVS) is a method used to identify and prioritize ligands for subsequent *in vitro* and *in vivo* profiling and, therefore, is an attractive way to identify new compounds for more effective treatment of COVID-19.[18,24]

*a*Laboratório de Planejamento e Desenvolvimento de Fármacos, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Belém, Pará 66075-110, Brazil. E-mail: rogerio@ufpa.br*

*b*Catalysis and Peptide Research Unit, University of KwaZulu-Natal, Durban 4000, South Africa*

*c*Laboratório de Modelagem Molecular, Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, CP 11101, 60075-110 Belém, PA, Brazil. E-mail: fabioam@ufpa.br*

Therefore, in this study, a virtual screen based on the structure of the receptor in a set of molecules of approved drugs was used to identify promising candidates that could be tested against the SARS-CoV-2 enzyme. Then, molecular dynamics and binding free energy calculations were carried out to evaluate the inhibition potential of the selected commercial drugs against the M^{pro} enzyme of SARS-CoV-2. Overall, this work shines a light on the M^{pro}–inhibitor complex interactions, which is paramount for the development of new drugs against COVID-19.

Materials and methods

Molecular docking, libraries description, and preparation

The crystallographic structure of the hydrolase enzyme was recovered from the Protein Data Bank (PDB) database under code 6LU7 with a resolution of 2.16 Å.[15] The enzyme has a chain with a total of 306 amino acid residues. Molecular docking can be used to explore possible conformations of the ligand inside the receptor's binding pocket as well as estimation of the strength of ligand–protein interaction.[25] The molecular docking simulations were performed using the GOLD 2020.1 program (Cambridge Crystallographic Data Centre – CCDC, Cambridge, UK).[26,27] GOLD program uses a genetic algorithm to generate and select conformations of flexible compounds that bind to a protein's receptor site.[28] Compounds were docked by applying the GoldScore scoring function with a search efficiency of 100%. The binding site was defined as a 10 Å sphere centered on the N3 crystallographic ligand. Before the preparation of the enzyme for docking, where the covalent bond between the Cys145 residue and the N3 crystallized ligand was removed. Finally, water molecules were removed from the structure of the enzyme, and hydrogen atoms were added by the GOLD program.

Before starting the virtual screen with all compounds chosen from the database, a re-docking of the N3 crystallographic ligand was carried out. Thus, the best conformation obtained by the program is compared with that of the ligand co-crystallized in the enzyme. In this way, an analysis is made by comparing the values of the root-mean-square deviation (RMSD). RMSD values below 2 Å show that the re-docking was successful according to data in the literature.[29] The PoseView online server[30] was used to analyze the interactions of hydrogens as well as hydrophobic interactions between the selected ligands and the amino acids of the enzyme, derived from molecular docking.

The two-dimensional structures of the 2998 compounds that were retrieved from the Repurposing hub (http://www.broadinstitute.org/repurposing) and built-in the MarvinSketch program,[31] as well as all structures, were optimized in the OpenBabel program.[32] For the selection of compounds, the total drug-score parameter calculated by the OSIRIS Property Explorer program was also evaluated.[33] This measurement is a score generated by the combination of the parameters of drug-similarity, hydrophobicity ($c \log P$), aqueous solubility ($\log S$), molar mass, and toxicity risk data.[34]

System setup for MD simulations

The molecular docking structures were submitted to QM optimization at the HF/6-31G* level by using the Gaussian09

program.[35] Then, at the same QM level, partial atomic charges for all docked compounds were obtained by applying the RESP method[36] as implemented in the antechamber module of the AmberTools19 package.[37] The general AMBER force field (GAFF)[38] and AMBER ff14SB[39] were used in describing the MM parameters for the ligands and enzyme, respectively. The PROPKA3.0 (ref. 40) was used to compute the protonation states of the amino acid residues at neutral pH. Particularly, the catalytic dyad (Cys145 and His41) was considered in neutral form, as reported for the SARS-CoV-2 M^{pro}.[41] Then, the absent protons were added by using *tleap* module of the AmberTools18 package.[37] Each system was solvated in a truncated octahedron box of TIP3P[42] water molecules with a buffer region of at least 10 Å from the solute atoms. Then, Na+ ions were added to neutralize the overall charge of the complexes. The SHAKE algorithm[43] was applied to constraint the covalent bonds involving hydrogen atoms. The particle mesh Ewald approach[44] was adopted to compute the long-range electrostatic interactions, a cut-off distance of 10 Å was used.

Each resulting system was minimized using 10 000 steps of the steepest descent method followed by 5000 steps of the conjugate-gradient method until the root-mean-square of the gradient was below 10^{-4} kcal (mol^{-1}·Å^{-1}). After, each minimized system was heated up from 0 to 300 K with solute restrained using a harmonic potential with a force constant of 100 kcal (mol^{-1}·Å^{-2}) during 500 ps. Along the equilibration stage, the positional restraint force was reduced from 50 to 2 kcal (mol^{-1}·Å^{-2}) during 1.2 ns of MD simulations. Then, 200 ps of density equilibration without any restraint under the NPT was performed. Finally, 100 ns of NPT simulation at 300 K was performed with a 2 fs time step. The GPU version of the *pmemd*[45] module of the Amber18 package was used to run all MM MD simulations. The structural analysis of each system was performed by computing the root-mean-square deviations (RMSD) of the main chain atoms concerning average coordinates structures. Similarly, the root-mean-square fluctuations (RMSF) of Cα atoms of each amino acid residue were computed. Besides, MD runs for M^{pro}-triptorelin system were also performed with a different starting structure (random seed). In addition to that, additional replicas starting from different sets of atomic coordinates and velocities were also chosen to generate completely independent trajectories. Details are presented in the ESI† files. The enabled us to calculate the error for these free energy results.

Binding free energy and residual decomposition analysis

For the binding free energy calculations, an average ensemble of structures from MD simulations, as described above, was extracted by using the CPPTRAJ module[46] of the AmberTools18 package. A total of 1000 snapshots were extracted from the last 10 ns trajectory of the production MD stage (at 10 ps intervals). Next, MM/GBSA[47,48] approach was employed to determine the binding free energy) M^{pro}–ligand complexes using the MMPBSA.py,[49] according to the eqn (1):

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{SOLV} - T\Delta S \qquad (1)$$

where, ΔE_{MM}, gas-phase MM energy is computed by eqn (2):

$$\Delta E_{\text{MM}} = \Delta E_{\text{int}} + \Delta E_{\text{ele}} + \Delta E_{\text{vdW}} \qquad (2)$$

where, ΔE_{int}, ΔE_{ele} and ΔE_{vdW} are the changes in the internal (bond, angles, and dihedral energies), electrostatic, and van der Waals energies, respectively.

The ΔG_{SOLV} is the sum of the electrostatic solvation energy (ΔG_{GB}) (polar contribution) and the nonpolar contribution (ΔG_{SA}), according to eqn (3):

$$\Delta G_{\text{SOLV}} = \Delta G_{\text{GB}} + \Delta G_{\text{SA}} \qquad (3)$$

Particularly, ΔG_{SA} is estimated by using the solvent-accessible surface area (SASA) method.[50] As the entropic term ($-T\Delta S$) has a heavy computational cost, it has been neglected for binding free energy calculations.[51]

Finally, since the MM/PBSA[48,50] approach is computationally time-consuming,[52] MM/GBSA has been used to compute the contributions from individual residues by free energy decomposition analysis.[51] This method has been frequently applied in drug design studies.[53–58]

Results and discussion

Virtual screening and molecular docking calculation

It should be highlighted that SARS-CoV-2 M$^{\text{pro}}$ is a crucial target for potential repurposing of known clinical drugs[41,59–64] as well as for designing specific protease inhibitors.[15,65–67] Although clinical drugs are not available for use against SARS-CoV-2 M$^{\text{pro}}$, others

protease inhibitors have been designed to inhibit the very closely related SARS-CoV M$^{\text{pro}}$.[1,68–70] Therefore, in this study, a computational protocol was applied to provide new insights for structure-assisted and computational drug design against SARS-CoV-2 M$^{\text{pro}}$.

Structure-based virtual screening was carried out in the gold program 2020.1, which uses a genetic algorithm to generate and select conformers of the database of compounds. From the results of re-docking, it was found that re-docking of the crystallographic ligand with the gold program gives an RMSD value of 2 Å. Thus, it is observed that the GOLD program was able to reproduce the conformation of the crystallographic ligand and the main interactions obtained at the M$^{\text{pro}}$ enzyme site.[29] Besides, molecular docking calculations have been successfully applied in recent SARS-CoV-2 M$^{\text{pro}}$ studies.[71–75]

Initially, 75 of the nearly 3000 drug molecules that had the best scores in the GOLD program were selected. It is noteworthy that the selection criterion was based on the score of remdesivir, as this has a potential effect in the treatment of COVID-19.[9,29,76] After, a visual inspection of the docking results was carried out on the Poseview online server,[30] taking into account mainly the interactions with the catalytic dyad residues of the M$^{\text{pro}}$ enzyme, in addition to the drug-score values to qualify the best drug candidates against COVID-19.[77] Thus, from the 75 structures, ten hit compounds were selected for further analysis (Table 1), considering the highest gold score values with M$^{\text{pro}}$ enzyme. Besides, they were selected using as criteria hydrogen interactions, complementarity through visual analysis at the active site of the M$^{\text{pro}}$ enzyme and the drug-score values.

Table 1 The selected compounds, GOLD score values, hydrogen, and hydrophobic interactions, drug score values, and pharmacological activity

Compound	Gold score	Hydrogen interactions	Hydrophobic interactions	Drug score	Pharmacological activity
Afamelanotide	111.02	Asn119, 2 × (Asn142), His164, Gln189	**His41**, Leu141, Asn142, Gly143, Met165, Gln189	0.31	Erythropoietic protoporphyria
Triptorelin	110.38	2 × (Asn142), Glu166, Pro168, Asp187, Gln189	Thr25, Ser46, Met49, Asn142, **Cys145**, His163, Met165, Glu166, Gln189	0.33	Antineoplastic, synthetic analog of gonadotropin-releasing hormone palliative treatment of advanced prostate cancer. Stimulates the release of luteinizing hormone
Pentagastrin	98.06	Gly143, Ser144, His163, Glu166, Gln189	Asn142, Met165, Pro168, Gln189, Thr190, Ala191	0.18	Evaluation of gastric acid secretory function
Terlipressin	97.51	Thr24, Ser139, Phe140, 2 × (Asn142), Gly143, Met165, 2 × (Glu166), Gly170	Thr25, Ser46, Met49, His172	0.46	Vasoactive drug in the management of hypotension
Adaptavir	92.78	Thr24, Thr26, Ser46, Thr190, Gly143, 2 × (**Cys145**), Thr190, Gln192	Met165	0.50	Chemokine receptor antagonist
Pepstatin	90.27	Thr24, 2 × (Thr26), **His41**, Cys145, Gln189	**Cys145**	0.44	Inhibitor of aspartic proteinases
Octreotide	88	Ser46, Met49, **Cys145**, Glu166	**His41**, Ser46, Met49, Glu166, Pro168, Gln189, Thr190,	0.21	Potent inhibitor of growth hormone, glucagon, and insulin
Darunavir	85.46	**Cys145**, 2 × (Ser144), 2 × (Glu166), Gln189	**His41**, Met49, Leu141, Asn142, Met165, Gln189	0.29	Protease inhibitor used as a treatment of human immunodeficiency virus (HIV)
Afatinib	84.43	His41, 2 × (Glu166), Gln189	**His41**, **Cys145**, Met165, Asp187	0.24	Therapy of selected forms of metastatic non-small cell lung cancer
Foretinib	80.31	Thr26, **Cys145**	Thr25, Thr26, **His41**, Met49, Leu50, Gly143, Met165, Gln189	0.25	Treatment of breast cancer

In addition to the dyad, the enzyme M^{pro} has a catalytic region formed by the residues Phe140, Asn142, Gly143, Ser144, Cys145, Met165, Glu166, Gln189 and Thr190.[78] As a protease, the catalytic site contains different subsites where interactions occur with the amino acids of the substrate whose peptide bond will be hydrolyzed. The S1 subsite is formed by the residues Phe140, Leu141, Asn142, His163 and Glu166, the S2 subsite is formed by the residues Met49, Tyr54, His164, Asp187 and Arg188, the S3 subsite is formed by the residues Met165, Leu167, Gln189, Thr190 and Gln192, in addition to the S1′, formed by the residues His41, Gly143, Ser144 and Cys145.[75]

Analyzing Table 1, it was possible to observe that the selected ligands interact with the enzyme through hydrogen bonds with the Thr24, Thr26, Hys41, Ser46, Met49, Asn119, Ser139, Phe140, Asn142, Gly143, Ser144, Cys145, His163, His164, Met165, Glu166, Pro168, Gly170, Asp187, Gln189, Thr190 and Gln192. In addition, among the ten main compounds selected, the interactions of residues His41, Asn142, Cys145, Glu166 and Gln189 can be highlighted, in which they make two, three, five, six and six hydrogen bonds, respectively. As mentioned earlier, His41 and Cys145 are part of the catalytic dyad, and a computational study of Wang's drug repurposing showed that His41 and Gln189 residues were considered important.[79] These residues are mainly part of the S1 subsite (Asn142, Cys145 e Glu166) and S2 (His41 e Gln189), in addition, studies by Koulgi and collaborators suggest that the inhibitory effect is observed when interactions with these key residues occur at the active site of the enzyme.[80]

Apart from remdesivir, application of darunavir against Covid-19 should be highlighted, especially since it has been studied as a possible treatment for SARS-CoV-2, and due to *in vitro* evidence that supports its ability to fight the disease.[81]

Stability of MD simulations

Initially, it should be noted that experimental evidence reveals that in the biological environment, M^{pro} enzyme acts in the dimer form instead of the monomeric one.[82] However, for computational proposals, it is possible to speed up computer-aided drug design by using only the monomeric form of SARS-CoV-2 M^{pro}.[73] Therefore, a total of 100 ns production simulations were performed for each repurposed drug on the M^{pro} enzyme from molecular docking procedures. Initially, we compare the representative snapshots chosen from the last 10 ns MD trajectory of each system. As illustrated in Fig. 1, the binding modes of all repurposed drugs in the active sites of the M^{pro} from MD simulated results are nearly the same place.

The root-mean-square deviations (RMSDs) of each system concerning the average structure were plotted in Fig. 2. As the plots show, the RMSDs of each system tend to converge after 20 ns simulation time, indicating the system is stable and equilibrated. Besides, all systems show very small RMSDs, which change from 0.97 ± 0.23 Å (pentagastrin) to 1.40 ± 0.20 Å (octeotride), which suggests that they are in a similar conformation. Additionally, for the M^{pro}-triptorelin, which shows the best binding free energy (as shown below), two more 100 ns of MD runs were performed using random seed with different
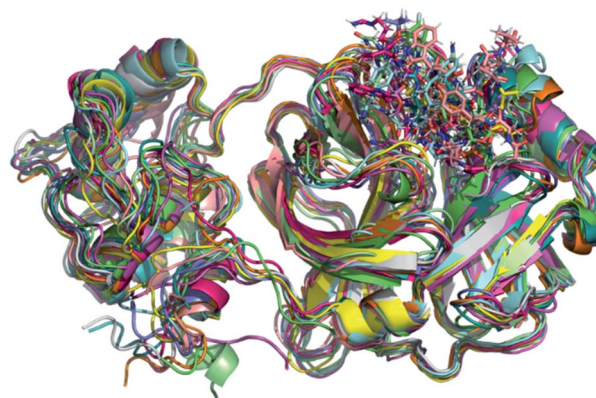

Fig. 1 Overlay of all M^{pro} systems after 100 ns of MD simulations. M^{pro} is shown in the cartoon model while repurposed drugs are in the stick model. All 3D structures of representative snapshots of M^{pro} systems are available as ESI† (PDB format).

atomic velocities. These results suggest that their average structures have very high 3D conformations and similar RMSD plots (Fig. S1 in the ESI† file).

To explore the fluctuations of the residue of system-wise, the root-mean-square fluctuation (RMSF) analysis was carried out. The RMSF of the Cα atoms of the protein for each residue during whole MD simulations for the complexes is shown in Fig. 3. The highest RMSF values are attributed to the N- and C-terminals, and loop regions. As shown in Fig. 3, the region closer to the active site formed by the Cys44–Asn53 residues (represented in red), that presented greater fluctuation to the terlipressin compound. Based on this observation, Bzówka *et al.* suggested that the presence of an inhibitor in this region might stabilize the loops surrounding the active site.[72] Another highlight is that residues are formed by the Asp153–Cys156 (represented in orange). In addition, a study by Shekh and colleagues describes the importance of Cys156 in the accessibility of external agents in the M^{pro} enzyme.[83] Another region that
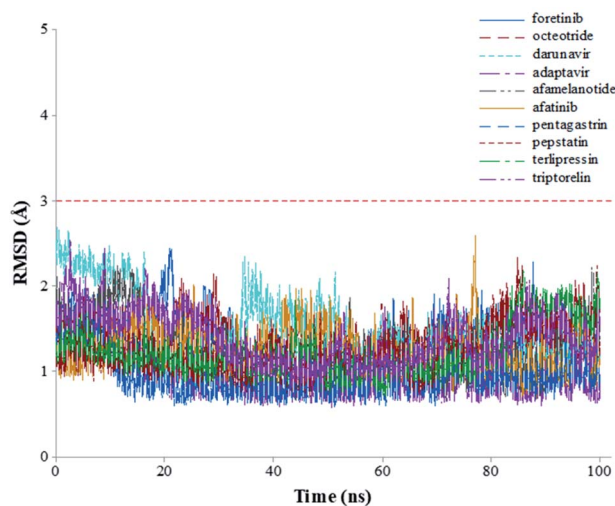

Fig. 2 Graphical representation of the RMSD values (Å) during 100 ns of MD simulation time for the selected repurposed drug complexes.
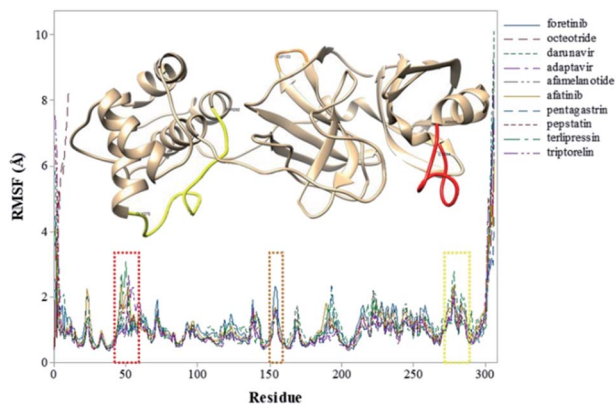
Fig. 3 RMSF values (in Å) of M^{pro}–inhibitor complexes.

presented high values of RMSF (Fig. 3), for all compounds, was the region formed by the Gly275–Thr292 residues (represented in yellow).

Binding free energy analysis

The binding free energies were computed using MM/GBSA method[48] as implemented in MMPBSA.py[49] are presented in Table 2. As a single MD trajectory of the bound complexes was used to compute ΔG_{bind}, the ΔE_{int} term can be canceled once the energy difference between complex systems and their components are computed using the same configurations.[51]

The energy components of the binding free energies were also listed in Table 2. As can be seen from Table 2, the ΔE_{ele}, ΔE_{vdW} and ΔG_{SA} terms are favorable contributors to most of the repurposed drugs, except for the pentagastrin and pepstatin drugs. Interestingly, in these systems, the term ΔG_{GB} is favorable, which does not occur on other systems. Then, in most of the systems, the favorable electrostatic interactions are equilibrated by the unfavorable polar solvation upon binding. As a result, the total electrostatic interaction contributions have detrimental effects on the binding of repurposed drugs. Except for the pentagastrin and pepstatin, the positive solvation contribution indicates that the repurposed drugs were exposed to water interaction in their respective complexes. If we consider the contributions of different binding free energy components of M^{pro}–ligand

binding, the most important determinant of difference in the binding affinity are ΔE_{vdW} and ΔG_{GB} contributions.

Interestingly, according to molecular docking results, the triptorelin was pointed with the second-highest value of gold score (Table 1) and after 100 ns of MD simulations has become the most promising M^{pro} inhibitor among the selected compounds. On the other hand, terlipressin, which has the fourth-highest value of gold score (Table 1), has been identified as the weakest M^{pro} inhibitor. This evidence can be observed by comparing the binding free energy values of the selected repurposed drugs (Table 2) it is possible to observe two compounds with distinct energy values. The calculated values of the terlipressin and triptorelin compounds have ΔG_{bind} (−25.39 ± 0.21 and −77.28 ± 0.20 kcal mol^{-1}, respectively) and ΔE_{vdW} contribution (−54.24 ± 0.30 and −92.42 ± 0.19 kcal mol^{-1}, respectively). These results suggest the importance of applied computational strategies as highlighted in previous studies.[25]

To better understand the individual contribution of each residue for the binding free energy for the lowest ΔG_{bind}, a residual decomposition analysis of ΔG_{bind} was performed using MM/GBSA approach.[51] Here, any residue that contributes to the binding free energy values below −1.50 kcal mol^{-1} was considered an important residue to the binding process. The energy contribution of each residue for the complexes formed by terlipressin (weakest binding free energy) and triptorelin (strongest binding free energy) obtained by the MM/GBSA method are shown in Fig. 4, once they have shown the different binding free energy values. The residue contributions to binding energies using the plugin CHEWD[84] with Chimera[85] software are shown on the right side of Fig. 4. From this, the favourable interactions are at the blue of the colour scale, that is, those that contribute to the stabilization of ligand into complex, while the red colour represents interaction by residue with positive values and corresponds to unfavourable interaction values.

For the M^{pro}–terlipressin complex (Fig. 4A) the residues that most significantly contribute to the total energy and, therefore, to the stabilization of the complex are His41, Cys44, Met49, Met165, Pro168, Pro188, Gln189, Thr190 and Ala191, that displayed energy values of −2.46, −1.79, −1.65, −1.53, −1.77, −1.86, −1.75, −3.25 and −1.82 kcal mol^{-1}, respectively. The binding free energy decomposition of M^{pro}–triptorelin complex (Fig. 4B) revealed that the residues Thr25, His41, Ser46, Asn142, Cys145, Met165, Glu166, Pro168, Asp187, Gln189 and Ala191

Table 2 MM/GBSA binding free energies (ΔG_{bind}) and its components for the repurposed drugs in the M^{pro} enzyme. The errors were computed by using single-trajectory protocol.[48] All values are reported in kcal mol^{-1}

| Inhibitor	ΔE_{ele}	ΔE_{vdW}	ΔG_{GB}	ΔG_{SA}	ΔG_{bind}
Triptorelin	−121.9 ± 0.3	−92.4 ± 0.2	147.9 ± 0.3	−10.8 ± 0.01	−77.3 ± 0.2
Darunavir	−207.7 ± 0.7	−67.7 ± 0.2	221.9 ± 0.6	−10.2 ± 0.01	−63.7 ± 0.3
Foretinib	−104.1 ± 0.4	−63.6 ± 0.2	122.2 ± 0.4	−7.0 ± 0.01	−52.5 ± 0.2
Pentagastrin	24.8 ± 0.5	−57.9 ± 0.2	−7.9 ± 0.4	−7.3 ± 0.02	−48.2 ± 0.2
Adaptavir	−34.5 ± 0.3	−53.3 ± 0.2	51.6 ± 0.3	−6.1 ± 0.02	−42.2 ± 0.2
Afamelanotide	−185.0 ± 1.4	−57.8 ± 0.2	211.2 ± 1.3	−7.7 ± 0.03	−39.4 ± 0.3
Afatinib	−115.7 ± 0.5	−51.3 ± 0.1	134.6 ± 0.5	−6.6 ± 0.01	−38.9 ± 0.1
Octeotride	−243.1 ± 1.0	−43.1 ± 0.2	261.7 ± 0.9	−6.1 ± 0.02	−30.5 ± 0.2
Pepstatin	42.9 ± 0.4	−49.8 ± 0.2	−15.9 ± 0.3	−6.4 ± 0.02	−29.2 ± 0.2
Terlipressin	−220.2 ± 1.4	−54.2 ± 0.3	256.5 ± 1.6	−7.4 ± 0.04	−25.4 ± 0.2
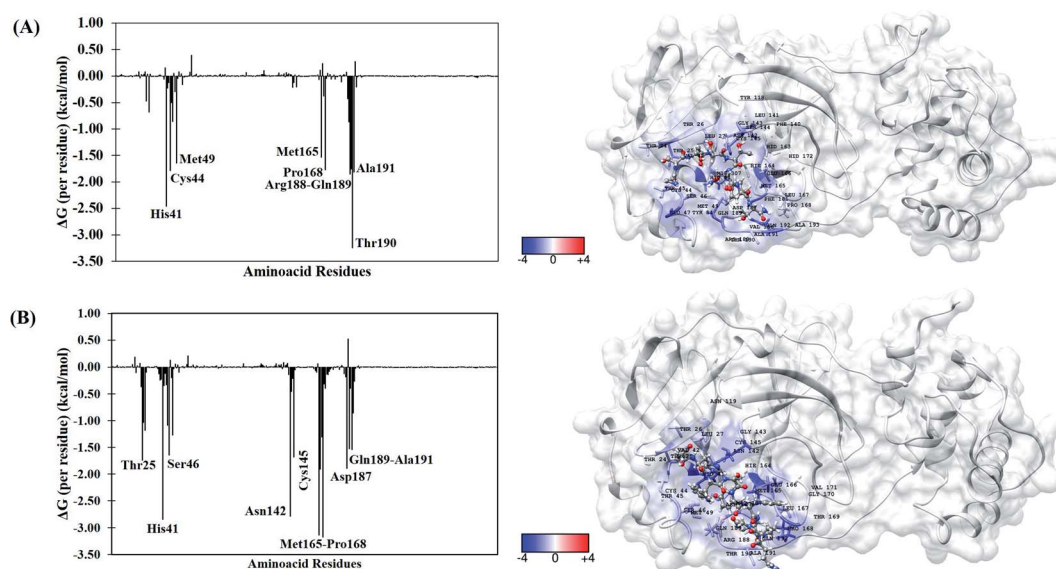
Fig. 4 Per-residue binding free energy decomposition (in kcal mol^{-1}) (left) and representative snapshot (right) of M^{pro} systems after 100 ns of MD simulations. (A) Terlipressin and (B) triptorelin. The carbon atoms of each inhibitor are shown in gray color. The per-residue results for other systems are available in the ESI file (Fig. S2†).

contribute greatly to the complex stability, exhibiting the energy values −1.74, −2.85, −1.65, −2.78, −1.69, −3.14, −1.91, −3.18, −1.90, −1.53 and −1.54 kcal mol^{-1}, respectively. From these results, we can observe that the enzyme–triptorelin complex present more favorable contributions to the binding free energy when compared to those formed by the terlipressin complex. Moreover, we should pay more attention to those residues with a relatively large difference in the contribution to the total binding free energies. For the terlipressin system, we can highlight the residues His41 and Thr190, while for the triptorelin system, the most relevant residues are His41, Asn142, Met165, and Pro168.

It is worth highlighting that His41, Met165, Pro168, Gln189 and Ala191 residues were important for both ligands. Interestingly, that triptorelin compound has interactions with catalytic dyad formed by His41 and Cys145 residues.[15] Besides, the interactions with residuals Thr25, Cys145, Met165, Asp187, Gln189 and Gln192 had already been observed in the molecular docking results (Table 1).

Particularly, we can highlight the results for the darunavir, which shows a moderate binding free energy value (−63.70 ± 0.25 kcal mol^{-1}). This drug has shown evidence *in vitro* studies involving a possible treatment for SARS-CoV-2.[86] Therefore, our results suggest that M^{pro} would be a potential target for the darunavir drug.

Conclusion

Drug repurposing is the well-established safety strategy of the development of novel therapeutic compounds, where the principal advantage is reducing the time and cost of new candidates in a future clinical trial. The COVID-19 caused by the SARS-CoV-2 has become a pandemic health crisis and the discovery of effective treatment against the virus remains an outstanding challenge. The results of this study showed that the use of

molecular modelling approaches, such as molecular docking and molecular dynamics simulations, can be used to design new drug leads that specifically target the M^{pro} enzyme from SARS-CoV-2. The molecular docking results allowed the identification of key interactions in the M^{pro} active site that make hydrogens bonds involving His41, Asn142, Cys145, Glu166 and Gln189 residues. The MD simulations of the best top 10 compounds indicated that systems were stable and equilibrated along the trajectory in the active site of the M^{pro} enzyme, as well as a favourable value for the prediction of the binding energy with the MM-GBSA method. It is worth mentioning that the darunavir and triptorelin drugs exhibited the strongest binding free energies in complex with M^{pro} enzyme. Thus, darunavir and triptorelin are potential inhibitors of SARS-CoV-2. In summary, all of the repurposed compounds reported here may provide insights for structure-assisted optimization to for potential application against the COVID-19 pandemic.

Author contributions

The manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 D. Hamre and J. J. Procknow, *Exp. Biol. Med.*, 1966, **121**, 190–193.

2 D. Schoeman and B. C. Fielding, *Virol. J.*, 2019, **16**, 69.

3 N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao and W. Tan, *N. Engl. J. Med.*, 2020, **382**, 727–733.

4 C. Drosten, S. Günther, W. Preiser, S. van der Werf, H.-R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. M. Fouchier, A. Berger, A.-M. Burguière, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J.-C. Manuguerra, S. Müller, V. Rickerts, M. Stürmer, S. Vieth, H.-D. Klenk, A. D. M. E. Osterhaus, H. Schmitz and H. W. Doerr, *N. Engl. J. Med.*, 2003, **348**, 1967–1976.

5 P. C. Y. Woo, S. K. P. Lau, C. Chu, K. Chan, H. Tsoi, Y. Huang, B. H. L. Wong, R. W. S. Poon, J. J. Cai, W. Luk, L. L. M. Poon, S. S. Y. Wong, Y. Guan, J. S. M. Peiris and K. Yuen, *J. Virol.*, 2005, **79**, 884–895.

6 L. van der Hoek, K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. M. Berkhout, K. C. Wolthers, P. M. E. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren and B. Berkhout, *Nat. Med.*, 2004, **10**, 368–373.

7 D. Hamre and J. J. Procknow, *Exp. Biol. Med.*, 1966, **121**, 190–193.

8 K. McIntosh, J. H. Dees, W. B. Becker, A. Z. Kapikian and R. M. Chanock, *Proc. Natl. Acad. Sci. U. S. A.*, 1967, **57**, 933–940.

9 R. K. Guy, R. S. DiPaola, F. Romanelli and R. E. Dutch, *Science*, 2020, **368**, 829–830.

10 W. Guan, Z. Ni, Y. Hu, W. Liang, C. Ou, J. He, L. Liu, H. Shan, C. Lei, D. S. C. Hui, B. Du, L. Li, G. Zeng, K.-Y. Yuen, R. Chen, C. Tang, T. Wang, P. Chen, J. Xiang, S. Li, J. Wang, Z. Liang, Y. Peng, L. Wei, Y. Liu, Y. Hu, P. Peng, J. Wang, J. Liu, Z. Chen, G. Li, Z. Zheng, S. Qiu, J. Luo, C. Ye, S. Zhu and N. Zhong, *N. Engl. J. Med.*, 2020, **382**, 1708–1720.

11 S. Arshad Ali, M. Baloch, N. Ahmed, A. Arshad Ali and A. Iqbal, *Journal of Infection and Public Health*, 2020, **13**, 644–646.

12 I. Chakraborty and P. Maity, *Sci. Total Environ.*, 2020, **728**, 138882.

13 WHO, *Coronavirus Disease (COVID-19) Outbreak Situation*, https://covid19.who.int/.

14 H. Yang, M. Yang, Y. Ding, Y. Liu, Z. Lou, Z. Zhou, L. Sun, L. Mo, S. Ye, H. Pang, G. F. Gao, K. Anand, M. Bartlam, R. Hilgenfeld and Z. Rao, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 13190–13195.

15 Z. Jin, X. Du, Y. Xu, Y. Deng, M. Liu, Y. Zhao, B. Zhang, X. Li, L. Zhang, C. Peng, Y. Duan, J. Yu, L. Wang, K. Yang, F. Liu, R. Jiang, X. Yang, T. You, X. Liu, X. Yang, F. Bai, H. Liu, X. Liu, L. W. Guddat, W. Xu, G. Xiao, C. Qin, Z. Shi, H. Jiang, Z. Rao and H. Yang, *Nature*, 2020, **582**, 289–293.

16 B. Nutho, P. Mahalapbutr, K. Hengphasatporn, N. C. Pattaranggoon, N. Simanon, Y. Shigeta, S. Hannongbua and T. Rungrotmongkol, *Biochemistry*, 2020, **59**, 1769–1779.

17 D. Gentile, V. Patamia, A. Scala, M. T. Sciortino, A. Piperno and A. Rescifina, *Mar. Drugs*, 2020, **18**, 225.

18 L. Mittal, A. Kumari, M. Srivastava, M. Singh and S. Asthana, *J. Biomol. Struct. Dyn.*, 2020, 1–19.

19 W. Dai, B. Zhang, X.-M. Jiang, H. Su, J. Li, Y. Zhao, X. Xie, Z. Jin, J. Peng, F. Liu, C. Li, Y. Li, F. Bai, H. Wang, X. Cheng, X. Cen, S. Hu, X. Yang, J. Wang, X. Liu, G. Xiao, H. Jiang, Z. Rao, L.-K. Zhang, Y. Xu, H. Yang and H. Liu, *Science*, 2020, **368**, 1331–1335.

20 M. Bzówka, K. Mitusińska, A. Raczyńska, A. Samol, J. A. Tuszyński and A. Góra, *Int. J. Mol. Sci.*, 2020, **21**, 3099.

21 H. P. Rang and R. G. Hill, in *Drug Discovery and Development*, ed. R. G. Hill and H. P. Rang, Churchill Livingstone, 2nd edn, 2013, pp. 203–209.

22 Z. Liu, H. Fang, K. Reagan, X. Xu, D. L. Mendrick, W. Slikker and W. Tong, *Drug Discovery Today*, 2013, **18**, 110–115.

23 C. H. Schein, *Med. Res. Rev.*, 2020, **40**, 586–605.

24 S. C. Araujo, V. G. Maltarollo, M. O. Almeida, L. L. G. Ferreira, A. D. Andricopulo and K. M. Honorio, *Molecules*, 2020, **25**, 264.

25 B. Honarparvar, T. Govender, G. E. M. Maguire, M. E. S. Soliman and H. G. Kruger, *Chem. Rev.*, 2014, **114**, 493–537.

26 G. Jones, P. Willett, R. C. Glen, A. R. Leach and R. Taylor, *J. Mol. Biol.*, 1997, **267**, 727–748.

27 M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray and R. D. Taylor, *Proteins: Struct., Funct., Bioinf.*, 2003, **52**, 609–623.

28 G. Jones, P. Willett and R. C. Glen, *J. Mol. Biol.*, 1995, **245**, 43–53.

29 K. E. Hevener, W. Zhao, D. M. Ball, K. Babaoglu, J. Qi, S. W. White and R. E. Lee, *J. Chem. Inf. Model.*, 2009, **49**, 444–460.

30 K. Stierand and M. Rarey, *J. Cheminf.*, 2010, **2**, P50.

31 *ChemAxon*, 2021.

32 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminf.*, 2011, **3**, 33.

33 T. Sander, J. Freyss, M. von Korff, J. R. Reich and C. Rufener, *J. Chem. Inf. Model.*, 2009, **49**, 232–246.

34 F. Azam, A. Amer, A. Abulifa and M. Elzwawi, *Drug Des., Dev. Ther.*, 2014, **8**, 2045–2059.

35 G. M. J. Frisch, W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, G. Cheeseman, J. R. Scalmani, V. Barone, B. Mennucci, H. Petersson, G. A. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, K. Hada, M. Ehara, M. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven,

J. E. Montgomery, J. A. Peralta Jr, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, R. Staroverov, V. N. Kobayashi, K. Normand, J. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, S. Dannenberg, J. J. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision E. 01*, Gaussian, 2009.

36 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.

37 D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York and P. A. Kollman, *Amber 2018*, University of California, San Francisco, 2018.

38 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

39 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.

40 M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525–537.

41 D. W. Kneller, G. Phillips, K. L. Weiss, Q. Zhang, L. Coates and A. Kovalevsky, *J. Med. Chem.*, 2021, **64**, 4991–5000.

42 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

43 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.

44 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

45 R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878–3888.

46 D. R. Roe and T. E. Cheatham, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.

47 J. Srinivasan, T. E. Cheatham, P. Cieplak, P. A. Kollman and D. A. Case, *J. Am. Chem. Soc.*, 1998, **120**, 9401–9409.

48 E. Wang, H. Sun, J. Wang, Z. Wang, H. Liu, J. Z. H. Zhang and T. Hou, *Chem. Rev.*, 2019, **119**, 9478–9508.

49 B. R. Miller, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, *J. Chem. Theory Comput.*, 2012, **8**, 3314–3321.

50 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889–897.

51 S. Genheden and U. Ryde, *Expert Opin. Drug Discovery*, 2015, **10**, 449–461.

52 S. Genheden and U. Ryde, *J. Chem. Theory Comput.*, 2011, **7**, 3768–3778.

53 S. M. Ahmed, G. E. M. Maguire, H. G. Kruger and T. Govender, *Chem. Biol. Drug Des.*, 2014, **83**, 472–481.

54 T. Ntombela, Z. Fakhar, C. U. Ibeji, T. Govender, G. E. M. Maguire, G. Lamichhane, H. G. Kruger and B. Honarparvar, *J. Comput.-Aided Mol. Des.*, 2018, **32**, 687–701.

55 Z. K. Sanusi, M. M. Lawal, T. Govender, G. E. M. Maguire, B. Honarparvar and H. G. Kruger, *J. Phys. Chem. B*, 2019, **123**, 6389–6400.

56 A. Souza, F. Cardoso, L. Martins, C. Alves, J. Silva and F. Molfetta, *J. Braz. Chem. Soc.*, 2021, **32**, 83–97.

57 P. R. M. Pereira, J. d. O. Araújo, J. R. A. Silva, C. N. Alves, J. Lameira and A. H. Lima, *J. Chem. Inf. Model.*, 2020, **60**, 738–746.

58 L. S. Martins, J. Lameira, H. G. Kruger, C. N. Alves and J. R. A. Silva, *Int. J. Mol. Sci.*, 2020, **21**, 4783.

59 W. R. Ferraz, R. A. Gomes, A. L. S Novaes and G. H. Goulart Trossini, *Future Med. Chem.*, 2020, **12**, 1815–1828.

60 C. N. Cavasotto and J. I. Di Filippo, *Mol. Inf.*, 2021, **40**, 2000115.

61 F. Touret, M. Gilles, K. Barral, A. Nougairède, J. van Helden, E. Decroly, X. de Lamballerie and B. Coutard, *Sci. Rep.*, 2020, **10**, 13093.

62 F.-J. Meyer-Almes, *Comput. Biol. Chem.*, 2020, **88**, 107351.

63 L. Riva, S. Yuan, X. Yin, L. Martin-Sancho, N. Matsunaga, L. Pache, S. Burgstaller-Muehlbacher, P. D. De Jesus, P. Teriete, M. V Hull, M. W. Chang, J. F.-W. Chan, J. Cao, V. K.-M. Poon, K. M. Herbert, K. Cheng, T.-T. H. Nguyen, A. Rubanov, Y. Pu, C. Nguyen, A. Choi, R. Rathnasinghe, M. Schotsaert, L. Miorin, M. Dejosez, T. P. Zwaka, K.-Y. Sit, L. Martinez-Sobrido, W.-C. Liu, K. M. White, M. E. Chapman, E. K. Lendy, R. J. Glynne, R. Albrecht, E. Ruppin, A. D. Mesecar, J. R. Johnson, C. Benner, R. Sun, P. G. Schultz, A. I. Su, A. García-Sastre, A. K. Chatterjee, K.-Y. Yuen and S. K. Chanda, *Nature*, 2020, **586**, 113–119.

64 C. Ma, M. D. Sacco, B. Hurst, J. A. Townsend, Y. Hu, T. Szeto, X. Zhang, B. Tarbet, M. T. Marty, Y. Chen and J. Wang, *Cell Stress*, 2020, **30**, 678–692.

65 A. Douangamath, D. Fearon, P. Gehrtz, T. Krojer, P. Lukacik, C. D. Owen, E. Resnick, C. Strain-Damerell, A. Aimon, P. Ábrányi-Balogh, J. Brandão-Neto, A. Carbery, G. Davison, A. Dias, T. D. Downes, L. Dunnett, M. Fairhead, J. D. Firth, S. P. Jones, A. Keeley, G. M. Keserü, H. F. Klein, M. P. Martin, M. E. M. Noble, P. O'Brien, A. Powell, R. N. Reddi, R. Skyner, M. Snee, M. J. Waring, C. Wild, N. London, F. von Delft and M. A. Walsh, *Nat. Commun.*, 2020, **11**, 5047.

66 L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, S. Becker, K. Rox and R. Hilgenfeld, *Science*, 2020, **368**, 409–412.

67 Z. Jin, Y. Zhao, Y. Sun, B. Zhang, H. Wang, Y. Wu, Y. Zhu, C. Zhu, T. Hu, X. Du, Y. Duan, J. Yu, X. Yang, X. Yang, K. Yang, X. Liu, L. W. Guddat, G. Xiao, L. Zhang, H. Yang and Z. Rao, *Nat. Struct. Mol. Biol.*, 2020, **27**, 529–532.

68 S. Konno, P. Thanigaimalai, T. Yamamoto, K. Nakada, R. Kakiuchi, K. Takayama, Y. Yamazaki, F. Yakushiji, K. Akaji, Y. Kiso, Y. Kawasaki, S.-E. Chen, E. Freire and Y. Hayashi, *Bioorg. Med. Chem.*, 2013, **21**, 412–424.

69 J. Jacobs, V. Grum-Tokars, Y. Zhou, M. Turlington, S. A. Saldanha, P. Chase, A. Eggler, E. S. Dawson, Y. M. Baez-Santos, S. Tomar, A. M. Mielech, S. C. Baker, C. W. Lindsley, P. Hodder, A. Mesecar and S. R. Stauffer, *J. Med. Chem.*, 2013, **56**, 534–546.

70 Y. Liu, C. Liang, L. Xin, X. Ren, L. Tian, X. Ju, H. Li, Y. Wang, Q. Zhao, H. Liu, W. Cao, X. Xie, D. Zhang, Y. Wang and Y. Jian, *Eur. J. Med. Chem.*, 2020, **206**, 112711.

71 B. Shah, P. Modi and S. R. Sagar, *Life Sci.*, 2020, **252**, 117652.

72 M. Bzówka, K. Mitusińska, A. Raczyńska, A. Samol, J. A. Tuszyński and A. Góra, *Int. J. Mol. Sci.*, 2020, **21**, 3099.

73 N. M. Tam, P. C. Nam, D. T. Quang, N. T. Tung, V. V Vu and S. T. Ngo, *RSC Adv.*, 2021, **11**, 2926–2934.

74 N. M. Tam, M. Q. Pham, N. X. Ha, P. C. Nam and H. T. T. Phung, *RSC Adv.*, 2021, **11**, 17478–17486.

75 A. Gimeno, J. Mestres-Truyol, M. J. Ojeda-Montes, G. Macip, B. Saldivar-Espinoza, A. Cereto-Massagué, G. Pujadas and S. Garcia-Vallvé, *Int. J. Mol. Sci.*, 2020, **21**, 3793.

76 G. Li and E. De Clercq, *Nat. Rev. Drug Discovery*, 2020, **19**, 149–150.

77 J. Gao, L. Liang, Y. Zhu, S. Qiu, T. Wang and L. Zhang, *Int. J. Mol. Sci.*, 2016, **17**, 1141–1150.

78 S. A. Khan, K. Zia, S. Ashraf, R. Uddin and Z. Ul-Haq, *J. Biomol. Struct. Dyn.*, 2021, **39**, 2607–2616.

79 J. Wang, *J. Chem. Inf. Model.*, 2020, **60**, 3277–3286.

80 S. Koulgi, V. Jani, M. Uppuladinne, U. Sonavane, A. K. Nath, H. Darbari and R. Joshi, *J. Biomol. Struct. Dyn.*, 2020, 1–21.

81 V. Mody, J. Ho, S. Wills, A. Mawri, L. Lawson, M. C. C. J. C. Ebert, G. M. Fortin, S. Rayalam and S. Taval, *Commun. Biol.*, 2021, **4**, 93.

82 H. Chen, P. Wei, C. Huang, L. Tan, Y. Liu and L. Lai, *J. Biol. Chem.*, 2006, **281**, 13894–13898.

83 S. Shekh, K. K. A. Reddy and K. H. Gowd, *J. Sulfur Chem.*, 2021, **42**, 109–120.

84 S. Raza, K. E. Ranaghan, M. W. van der Kamp, C. J. Woods, A. J. Mulholland and S. S. Azam, *J. Comput.-Aided Mol. Des.*, 2019, **33**, 461–475.

85 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.

86 C. Harrison, *Nat. Biotechnol.*, 2020, **38**, 379–381.