

Computational Design of Potent Inhibitors for SARS-CoV-2's Main Protease

Abd Al-Aziz A. Abu-Saleh^{1*}, Ibrahim Awad¹, Arpita Yadav², Raymond A. Poirier^{1*}.

¹Chemistry Department, Memorial University, St. John's, NL A1B 3X7, Canada

²Department of Chemistry, University Institute of Engineering & Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur 208024, India

*Corresponding author

E-Mail: rpoirier@mun.ca; aabusaleh@mun.ca

ABSTRACT:

COVID-19 has caused lockdowns all over the world in early 2020, as a global pandemic. Both theoretical and experimental efforts are seeking to find an effective treatment to suppress the virus. *In silico* drug design can play a vital role in identifying promising drug candidates against COVID-19. Herein, we focused on the main protease of SARS-CoV-2 that plays crucial biological functions in the virus. We performed a ligand-based virtual screening followed by a docking screening for testing approved drugs and bioactive compounds listed in the DrugBank and ChEMBL databases. The top 8 docking results were advanced to all-atom MD simulations to study the relative stability of the protein-ligand interactions. MD simulations support that the catalytic residue, His41, has a neutral side chain with a protonated delta position. An absolute binding energy (ΔG) of -42 kJ mol^{-1} for the protein-ligand (M^{pro} -N3) complex has been calculated using the potential-of-mean-force (geometrical) approach. Furthermore, the relative binding energies were computed for the top docking results. Our results suggest several promising approved and bioactive inhibitors of SARS-CoV-2 M^{pro} as follows: a bioactive compound, ChEMBL275592, which has the best MM/GBSA binding energy; the second-best compound, montelukast, is an approved drug used in the treatment of asthma and allergic rhinitis; the third-best compound, ChEMBL288347, is a bioactive compound. Bromocriptine and saquinavir, are other approved drugs that also demonstrate stability in the active site of M^{pro} , albeit their relative binding energies are low compared to the N3 inhibitor. This study provides useful insights into de novo protein design and novel inhibitor development, which could reduce the cost and time required for the discovery of a potent drug to combat SARS-CoV-2.

1. INTRODUCTION:

Coronaviruses can provoke infectious diseases in humans and animals. Coronavirus disease 2019 (COVID-19) is caused by a novel severe acute respiratory syndrome, SARS-CoV-2, which has spread as a worldwide pandemic. Patients infected with this disease present major symptoms including, high fever, rhinorrhea, cough, sore throat, pneumonia and ultimately, death in severe cases.^{1,2} Globally, according to the world health organization (WHO), there have been 8,061,550 confirmed cases of COVID-19, including 440,290 deaths as of June 17, 2020. Most of these cases and deaths are from Europe and the Americas.

Coronaviruses were named due to their protein spikes that have a crown-like shape.³ The coronavirus genome encodes several proteins, such as the spike, M^{pro} (also known as 3-chymotrypsin-like cysteine protease, 3CL^{pro}), and RNA-dependent RNA polymerase (RdRp).⁴ Different approaches are being investigated for the development of useful drugs to fight against SARS-CoV-2. Developing inhibitors of viral M^{pro} is one of the promising approaches. M^{pro} is essential for viral replication by facilitating the cleavage of viral peptides into smaller functional units.⁵ Consequently, drugs that target the M^{pro} would cease the replication process and prevent viral infection. More recently, 15 drugs are being tested to cure COVID-19 including virus protease inhibitors.⁶ However, albeit the preliminary results are promising, there were issues in the design of the study. According to WHO, there is no specific medicine to treat COVID-19 yet.

In silico drug design has been playing a vital role in modern drug therapies against infectious diseases.⁷⁻⁹ Currently, enormous efforts and approaches are being pursued for the discovery of inhibitors against SARS-CoV-2 M^{pro} by employing high performance computational resources. Nutho et al studied two HIV-1 protease inhibitors specifically, lopinavir and ritonavir, by MD simulations and relative binding energy calculations.¹⁰ MD simulations, molecular docking, and structure-activity relationship were used for the discovery of new hydroxyethylamine analogs

against the M^{pro}.¹¹ Docking screening of approved drugs and drug candidates in clinical trials,¹² and medicinal plants¹³ were also conducted along with MD simulations. Molecular docking, fast pulling of ligand, and free energy perturbation calculations were also performed to investigate potential inhibitors of SARS-CoV-2 M^{pro}.¹⁴ Moreover, both MD simulations and molecular docking were employed to explore the M^{pro} inhibitory of 19 marketed drugs,¹⁵ and to repurpose protease inhibitor compounds.¹⁶

In this work, we will focus on the newly released protein crystal structure of the SARS-CoV-2 M^{pro}.¹⁷ This protein is a dimer that includes two protomers; each protomer has three domains. Domains I and II, residues 8-184, have an antiparallel β -barrel structure. Domain III, residues 201-303, has five α -helices. The extended loop, residues 185-200, connects Domains II and III. Furthermore, this protein has a catalytic dyad, His41 and Cys145, and the substrate binding site is in a cleft between Domains I and II. In this paper, we performed an integrated computational protocol including ligand-based and structure-based molecular screening, molecular dynamics simulations, and binding energy calculations to facilitate the identification of promising candidate drugs to treat COVID-19.

2. METHODS AND COMPUTATIONAL DETAILS

The initial coordinates of one protomer of the SARS-CoV-2 M^{pro} was retrieved from the RCSB [Protein Data Bank (PDB) entry 6LU7] determined at 2.16 Å resolution with no missing residues.¹⁷ The cocrystalized ligand (named N3) of SARS-CoV-2 M^{pro} was used to conduct the ligand-based virtual screening. The N3 ligand is a potent inhibitor and it was also found as a complex with earlier coronaviruses M^{pro} such as IBV,¹⁸ HCoV-HKU1,¹⁹ SARS-CoV,²⁰ HCoV-NL63,²¹ FIPV,²² and PEDV.²³

2.1. Hierarchical Virtual Screening and Docking.

In the PDB crystal structure, the N3 ligand was covalently bonded to Cys145. We constructed the free ligand by breaking the covalent bond and making an α,β -unsaturated ketone using the molecular structure editor (maestro)²⁴. We minimized the N3 ligand using the MMFF94s force field.²⁵ We then screened the data sets from both the DrugBank²⁶ and ChEMBL²⁷ libraries based on the preprocessed N3 ligand. For ligand-based screening, a novel fully flexible high-throughput 3D molecular similarity approach (Screen3D algorithm)²⁸ was performed as implemented in the BRUSELAS server.²⁹

The top 200 compounds (the top 100 from each library) that are structurally similar to the N3 ligand were advanced to the next hierarchical filter, docking screening. OpenBabel software³⁰ was used to generate and minimize conformations from the top 200 screened compounds. These compounds were docked into a preprocessed protein pocket by utilizing the AutoDock Vina package.³¹ It has been shown that AutoDock Vina has an effective scoring function in terms of accuracy and performance.³² The grid cell of 18.0 Å, 21.3 Å, and 24.3 Å in the x, y, and z directions, respectively, was built for docking calculations. This grid cell is located around the active site of the M^{pro}, centroid to residues His41, Met49, Phe140, Leu141, Asn142, Gly143, Cys145, His164, Met165, Glu166, Leu167, Pro168, His172, Gln189, and Thr190.

2.2. Molecular Dynamics (MD) simulations.

The top 8 docking results were advanced to all-atom MD simulations to study the relative stability of the protein-ligand interactions, and to screen a set of compounds for further binding energy calculations. All the simulations were done using the NAMD 2.13 package^{33,34} and the CHARMM36 force field.³⁵ The parameters for the N3 ligand and the top 8 docking compounds were generated using the CHARMM general force field (CGenFF).³⁶ A protomer of the M^{pro} has a total of 306 residues.¹⁷ The protonation state of the titratable residues were assigned at pH 7.4

exploiting the H++ web server.³⁷ It should be pointed out that the catalytic residue His41 can adopt three different protonation states: neutral HSD (δ -nitrogen protonated), neutral HSE (ϵ -nitrogen protonated), and protonated HSP (both δ - and ϵ -nitrogens protonated). Therefore, three MD simulations were run to explore the effect of the protonation states of His41 on the stability of the active site. The TIP3P explicit solvation model was used, and the periodic boundary conditions were set with dimensions of 115.2 Å³. Afterward, the system was neutralized using four sodium (Na⁺) ions. The MD protocols involved minimization, annealing, equilibration, and production. The atoms of the protein backbone were restrained in the minimization and annealing simulations, while the C α atoms of the protein were restrained in the 1 ns equilibration simulation. However, no atoms were restrained in the 100 ns MD production simulation. The isothermal–isobaric (NPT) ensemble and a 2 fs time step of integration was chosen for all MD simulations. Through the 100 ns of MD production, the pressure was set at 1 atm using the Nosé-Hoover Langevin piston barostat^{38,39} with the Langevin piston decay of 0.2 ps and a period of 0.4 ps. The temperature was set at 298.15 K using the Langevin thermostat⁴⁰ with a damping frequency of 1 ps⁻¹. A distance cutoff of 10.0 Å was applied to short-range nonbonded interactions with a pair list distance of 12 Å, and Lennard Jones interactions were smoothly truncated at 8.0 Å. Long-range electrostatic interactions were treated using the particle-mesh Ewald (PME) method,^{41,42} where a grid spacing of 1.0 Å was used for all simulation cells. All covalent bonds involving hydrogen atoms were constrained using the SHAKE algorithm.⁴³ For consistency, we have applied the same protocol for all MD simulations.

2.3. Binding Energy Calculations.

Starting from the equilibrated protein–ligand (M^{pro}–N3) complex, we calculated the absolute binding energy of the M^{pro}–N3 complex using the geometrical energy approach. It has been shown

that the geometrical approach accurately predicts protein–ligand binding energies.^{44,45} We used the default set up of the binding free energy estimator as described by Chipot and coworkers.⁴⁶ More details on theoretical background of the geometrical (i.e., potential-of-mean-force) free energy calculations can be obtained elsewhere.⁴⁷ In addition, we collected a total of 1000 snapshots extracted consistently from the 100 ns of MD production to calculate the relative protein–ligand binding energy of the top docking results.

The one-average molecular mechanics generalized Born surface area (MM/GBSA) approach^{48,49} was used for the relative binding energy calculations, in which the ligand (L) binds to the protein receptor (R) to form the complex (RL),

$$\Delta G_{\text{bind}} = G_{RL} - G_R - G_L$$

which can be represented by contributions of different interactions,

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S = \Delta E_{MM} + \Delta G_{\text{sol}} - T\Delta S$$

where the changes in the gas phase molecular mechanics (ΔE_{MM}), solvation Gibbs energy (ΔG_{sol}), and conformational entropy ($-T\Delta S$) are determined as follows: ΔE_{MM} is the sum of the changes in the electrostatic energies ΔE_{ele} , the van der Waals energies ΔE_{vdW} , and the internal energies ΔE_{int} (bonded interactions); ΔG_{sol} is the total of both the polar solvation (calculated by generalized Born model) and the nonpolar solvation (calculated using the solvent-accessible surface area); $-T\Delta S$ is calculated by the normal mode analysis. The solvent dielectric constant of 78.5 and the surface tension constant of $0.021 \text{ kJ mol}^{-1} \text{ \AA}^{-2}$ were used for MM/GBSA calculations.

3. RESULTS AND DISCUSSION:

We performed a ligand-based virtual screening of approved drugs and bioactive compounds from both the DrugBank and ChEMBL databases. The top 100 screened compounds from each library

are advanced to flexible molecular docking using AutoDock Vina. Nine conformations from each compound were generated through a flexible docking. Docking score was chosen as the measure of the binding affinity to rank the poses of the 200 compounds. Our selection of the top 8 compounds were based on compounds that have at least five conformations with a docking score of $\leq -33 \text{ kJ mol}^{-1}$ (-8 kcal mol^{-1}). Docking results of the top 8 compounds for SARS-COV-2 M^{pro} are listed in Table 1. The structures of the top 8 compounds are shown in Figure 1.

Insert Table 1

Insert Figure 1

Next, the top 8 docking results were advanced for MD simulations. Many studies have validated the role of MD simulations for the improvement of docking results.^{50,51} Before launching MD simulations on the top docking results, three MD simulations of SARS-COV-2 M^{pro} complexed with the N3 ligand were carried out to explore the plausible protonation states of the catalytic His41, as illustrated earlier in the methods section. His41 is a well-preserved residue amongst various viruses including hepatitis C virus (HCV), MERS-CoV, SARS-CoV, and SARS-CoV-2.¹² The three MD simulations were labeled as follows: neutral HSD (Model 1), neutral HSE (Model 2), and protonated HSP (Model 3). Control of the structure stability of the three models was achieved by the root-mean-square deviation (RMSD) over protein backbone atoms, and by the root-mean-square fluctuation (RMSF). In addition, the M^{pro}–N3 hydrogen bond interactions were also analyzed during the 100 ns NPT ensemble. Figure 2 shows the RMSD, RMSF, and hydrogen bond analyses of the protein–ligand complex for the three models.

Insert Figure 2

Unambiguously, amongst the three models, Model 1 shows a modest structural stability in terms of the RMSD of both the M^{pro} and N3 ligand, and the RMSF analysis. Moreover, the intermolecular interaction of the M^{pro}–N3 complex has an average of five hydrogen bonds (see the top histogram in Figure 1). The stability of the M^{pro}–N3 complex is mainly due to the hydrogen bonds between the N3 ligand and residues His41, Gly143, Glu166, and Gln189 of the M^{pro}, besides the hydrophobic interactions. These results are in concert with a recent experimental study.¹⁷ Figure 3 represents the pose of the N3 ligand inside the pocket of the M^{pro}.

Insert Figure 3

The RMSD and the RMSF of the protein backbone of Model 2 is comparable with Model 1, however, the N3 ligand in Model 2 was less strongly bound inside the M^{pro} pocket than Model 1 (see Figure 2). Model 2 has an average of three hydrogen bonds between the N3 ligand and the M^{pro}. In Model 3, the N3 ligand has left the binding pocket of the M^{pro} during the MD simulations (see the green line of the RMSD analysis at the bottom of Figure 1). Thus, Model 1 is considered as a starting point for MD simulations for the top docking results. In other words, His41 is set as a neutral form with a protonated delta position (HSD type of CHARMM format), and this model agrees with the common reaction mechanism of cysteine protease.⁵² The optimal binding pose of the M^{pro}–N3 complex of Model 1 was further assessed by absolute binding energy calculations using the geometrical approach. It should be mentioned that the N3 is a Michael acceptor inhibitor, which means that the inhibitor forms a reversible complex (M^{pro}–N3) under the equilibrium binding constant K_{eq} . It then undergoes a nucleophilic attack by Cys145 of the M^{pro} active site, forming a stable covalent bond. The later step is controlled by the inactivation rate constant, which is beyond the scope of this paper. To calculate K_{eq} , the different contributions arising from the geometrical restraints were computed, either in the bound state, or in the unbound state.^{46,47} The

final absolute binding energy, $\Delta G = -1/\beta \ln K_{eq}$, has a value of -42 kJ mol^{-1} . Yet, no experimental binding energy was reported for the SARS-CoV-2 M^{pro}. Nevertheless, our results are comparable to experimental binding energies of similar protease–inhibitor systems. Based on the experimental inhibitory constant (K_i), HCoV-NL63²¹, SARS-CoV⁵³, and HCoV-229E⁵³ have the M^{pro}–N3 binding energies of -28 , -29 , and -33 kJ mol^{-1} , respectively. Therefore, our predicted results conclude that the SARS-CoV-2 M^{pro} has relatively a higher binding affinity to the N3 inhibitor than the other Coronaviruses.

Using the same protocol as Model 1, MD simulations were carried out for the top 8 docking compounds to investigate their stability inside the M^{pro} pocket. The RMSD analyses of the top 8 docking compounds are provided in the Supporting Information (Figure S1). The RMSD values of ChEMBL275592, montelukast, ChEMBL288347, bromocriptine, and saquinavir systems increased at the beginning of the simulation then remained stable until the end of the simulation. These compounds have reliable MMGBSA binding energies relative to the known N3 inhibitor that has a value of -150 kJ mol^{-1} calculated at MMGBSA method. The rest of the docking compounds, zafirlukast, bosentan, and doismin, lack the stability inside the M^{pro} pocket, consequently, these compounds have poor binding energies relative to the known N3 inhibitor. Table 2 lists the MD results and the MMGBSA relative binding energies for the top docking compounds.

Insert Table 2

The averaged contact area analysis was calculated based on the surface area of a ligand that is exposed to residues of a protein.⁵⁴ Table 3 summarizes the contact area analysis for compounds that have good stability inside the binding site of the M^{pro} through the 100 ns MD simulations. Snapshots of the last frame of simulated systems are depicted in Figure 4. Analyses of the time-

evolution of the contact area are also provided in the Supporting Information (Figure S2). We found that residues Thr25, His41, Ser46, Met49, Asn142, Cys145, Met165, Glu166, Pro168, and Gln189 have a good contact area with the top five compounds. Wang studied the common significant hot spot residues of the M^{pro}; these residues are His41, Met49, Asn142, His164, Met165, Glu166, and Gln189, which are in agreement with our results.¹²

Insert Figure 4

Insert Table 3

CONCLUSION:

In this study, the ligand-based screening, structure-based docking screening, MD simulations, and binding energy calculations were conducted based on the M^{pro} as a drug target. The M^{pro}–N3 complex has an absolute binding energy of – 42 kJ mol⁻¹. The predicted results suggest that montelukast, ChEMBL275592, and ChEMBL288347 (top 3 compounds) show good inhibitory efficiency on the focused M^{pro} target. Montelukast is an approved drug that is used to control and prevent breathing problems. ChEMBL275592 is a bioactive compound that shows the inhibitory activity of 2.8 nM against HIV-1 protease. ChEMBL288347 is also a bioactive compound that shows the inhibitory activity of 1.4 nM through in vitro inhibition of purified human renal renin. In addition, bromocriptine and saquinavir may also be candidates for M^{pro} inhibition. Our findings pave the way for further optimizations and designs of M^{pro}'s inhibitors.

ASSOCIATED CONTENT

Supporting Information Available:

RMSD and the contact surface area analyses of the top docking compounds complexed individually with the M^{pro} are available free of charge via the Internet at <http://pubs.acs.org>.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge Compute Canada for computer time. We are also immensely grateful to Shahidul M. Islam for his comments on an earlier version of the manuscript. R.A.P. is grateful to the Natural Sciences and Engineering Council of Canada (NSERC) for financial support. A.A.A thanks Chen Graduate Scholarship and the School of Graduate Studies of the Memorial University of Newfoundland for funding. A.Y. thanks Science and Engineering Research Board, New Delhi for research grant number EMR/2016/000769. A.Y is also grateful to Shastri Indo Canadian Institute, New Delhi for Shastri Scholar Travel Subsidy Grant in the year 2016.

References

- (1) Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; Qiu, Y.; Wang, J.; Liu, Y.; Wei, Y.; et al. Epidemiological and Clinical Characteristics of 99 Cases of 2019 Novel Coronavirus Pneumonia in Wuhan, China: A Descriptive Study. *Lancet* **2020**, *395*, 507–513.
- (2) Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497–506.
- (3) Stadler, K.; Maignani, V.; Eickmann, M.; Becker, S.; Abrignani, S.; Klenk, H.-D.; Rappuoli, R. SARS — Beginning to Understand a New Virus. *Nat. Rev. Microbiol.* **2003**, *1*, 209–218.
- (4) Liu, C.; Zhou, Q.; Li, Y.; Garner, L. V.; Watkins, S. P.; Carter, L. J.; Smoot, J.; Gregg, A. C.; Daniels, A. D.; Jervy, S.; et al. Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases. *ACS Cent. Sci.* **2020**, *6*, 315–331.
- (5) Goetz, D. H.; Choe, Y.; Hansell, E.; Chen, Y. T.; McDowell, M.; Jonsson, C. B.; Roush, W. R.; McKerrow, J.; Craik, C. S. Substrate Specificity Profiling and Identification of a New Class of Inhibitor for the Major Protease of the SARS Coronavirus † , ‡. *Biochemistry* **2007**, *46*, 8744–8752.
- (6) Shaffer, L. 15 Drugs Being Tested to Treat COVID-19 and How They Would Work. *Nat. Med.* **2020**.

- (7) Macalino, S. J. Y.; Gosu, V.; Hong, S.; Choi, S. Role of Computer-Aided Drug Design in Modern Drug Discovery. *Arch. Pharm. Res.* **2015**, *38*, 1686–1701.
- (8) Njogu, P. M.; Guantai, E. M.; Pavadai, E.; Chibale, K. Computer-Aided Drug Discovery Approaches against the Tropical Infectious Diseases Malaria, Tuberculosis, Trypanosomiasis, and Leishmaniasis. *ACS Infect. Dis.* **2016**, *2*, 8–31.
- (9) Aminpour, M.; Montemagno, C.; Tuszynski, J. A. An Overview of Molecular Modeling for Drug Discovery with Specific Illustrative Examples of Applications. *Molecules* **2019**, *24*, 1693.
- (10) Nutho, B.; Mahalapbutr, P.; Hengphasatporn, K.; Pattarangoon, N. C.; Simanon, N.; Shigeta, Y.; Hannongbua, S.; Rungrotmongkol, T. Why Are Lopinavir and Ritonavir Effective against the Newly Emerged Coronavirus 2019? Atomistic Insights into the Inhibitory Mechanisms. *Biochemistry* **2020**, *59*, 1769–1779.
- (11) Kumar, S.; Sharma, P. P.; Shankar, U.; Kumar, D.; Joshi, S. K.; Pena, L.; Durvasula, R.; Kumar, A.; Kempaiah, P.; Poonam, *; et al. Discovery of New Hydroxyethylamine Analogs Against 3CLpro Protein Target of SARS-CoV-2: Molecular Docking, Molecular Dynamics Simulation and Structure-Activity Relationship Studies. *J. Chem. Inf. Model.* **2020**.
- (12) Wang, J. Fast Identification of Possible Drug Treatment of Coronavirus Disease-19 (COVID-19) through Computational Drug Repurposing Study. *J. Chem. Inf. Model.* **2020**.
- (13) Tahir ul Qamar, M.; Alqahtani, S. M.; Alamri, M. A.; Chen, L.-L. Structural Basis of SARS-CoV-2 3CLpro and Anti-COVID-19 Drug Discovery from Medicinal Plants. *J. Pharm. Anal.* **2020**.
- (14) 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. *J. Chem. Inf. Model.* **2020**, acs.jcim.0c00491.
- (15) Huynh, T.; Wang, H.; Luan, B. In Silico Exploration of the Molecular Mechanism of Clinically Oriented Drugs for Possibly Inhibiting SARS-CoV-2's Main Protease. *J. Phys. Chem. Lett.* **2020**, *11*, 4413–4420.
- (16) Havranek, B.; Islam, S. M. An in Silico Approach for Identification of Novel Inhibitors as Potential Therapeutics Targeting COVID-19 Main Protease. *J. Biomol. Struct. Dyn.* **2020**, 1–12.
- (17) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; et al. Structure of Mpro from SARS-CoV-2 and Discovery of Its Inhibitors. *Nature* **2020**.
- (18) Xue, X.; Yu, H.; Yang, H.; Xue, F.; Wu, Z.; Shen, W.; Li, J.; Zhou, Z.; Ding, Y.; Zhao, Q.; et al. Structures of Two Coronavirus Main Proteases: Implications for Substrate Binding and Antiviral Drug Design. *J. Virol.* **2008**, *82*, 2515–2527.
- (19) Zhao, Q.; Li, S.; Xue, F.; Zou, Y.; Chen, C.; Bartlam, M.; Rao, Z. Structure of the Main Protease from a Global Infectious Human Coronavirus, HCoV-HKU1. *J. Virol.* **2008**, *82*, 8647–8655.

- (20) Xue, X.; Yang, H.; Shen, W.; Zhao, Q.; Li, J.; Yang, K.; Chen, C.; Jin, Y.; Bartlam, M.; Rao, Z. Production of Authentic SARS-CoV Mpro with Enhanced Activity: Application as a Novel Tag-Cleavage Endopeptidase for Protein Overproduction. *J. Mol. Biol.* **2007**, *366*, 965–975.
- (21) Wang, F.; Chen, C.; Tan, W.; Yang, K.; Yang, H. Structure of Main Protease from Human Coronavirus NL63: Insights for Wide Spectrum Anti-Coronavirus Drug Design. *Sci. Rep.* **2016**, *6*, 22677.
- (22) Wang, F.; Chen, C.; Liu, X.; Yang, K.; Xu, X.; Yang, H. Crystal Structure of Feline Infectious Peritonitis Virus Main Protease in Complex with Synergetic Dual Inhibitors. *J. Virol.* **2016**, *90*, 1910–1917.
- (23) Wang, F.; Chen, C.; Yang, K.; Xu, Y.; Liu, X.; Gao, F.; Liu, H.; Chen, X.; Zhao, Q.; Liu, X.; et al. Michael Acceptor-Based Peptidomimetic Inhibitor of Main Protease from Porcine Epidemic Diarrhea Virus. *J. Med. Chem.* **2017**, *60*, 3212–3216.
- (24) Maestro, Schrödinger Release 11.8.0; Schrödinger, LLC: New York, NY, 2019.
- (25) Halgren, T. A. MMFF VI. MMFF94s Option for Energy Minimization Studies. *J. Comput. Chem.* **1999**, *20*, 720–729.
- (26) 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 Res.* **2014**, *42*, D1091–D1097.
- (27) Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrián-Uhalte, E.; et al. The ChEMBL Database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945–D954.
- (28) Kalászi, A.; Szisz, D.; Imre, G.; Polgár, T. Screen3D: A Novel Fully Flexible High-Throughput Shape-Similarity Search Method. *J. Chem. Inf. Model.* **2014**, *54*, 1036–1049.
- (29) Banegas-Luna, A. J.; Cerón-Carrasco, J. P.; Puertas-Martín, S.; Pérez-Sánchez, H. BRUSELAS: HPC Generic and Customizable Software Architecture for 3D Ligand-Based Virtual Screening of Large Molecular Databases. *J. Chem. Inf. Model.* **2019**, *59*, 2805–2817.
- (30) O’Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, *3*, 33.
- (31) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461.
- (32) Wang, Z.; Sun, H.; Yao, X.; Li, D.; Xu, L.; Li, Y.; Tian, S.; Hou, T. Comprehensive Evaluation of Ten Docking Programs on a Diverse Set of Protein–Ligand Complexes: The Prediction Accuracy of Sampling Power and Scoring Power. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12964–12975.
- (33) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J.*

Comput. Chem. **2005**, *26*, 1781–1802.

- (34) Ribeiro, J. V.; Bernardi, R. C.; Rudack, T.; Stone, J. E.; Phillips, J. C.; Freddolino, P. L.; Schulten, K. QwikMD — Integrative Molecular Dynamics Toolkit for Novices and Experts. *Sci. Rep.* **2016**, *6*, 26536.
- (35) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain X1 and X2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257–3273.
- (36) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell, A. D. Extension of the CHARMM General Force Field to Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33*, 2451–2468.
- (37) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating PK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res.* **2012**, *40*, 537–541.
- (38) Nosé, S.; Klein, M. L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50*, 1055–1076.
- (39) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **2002**, *100*, 191–198.
- (40) Grest, G. S.; Kremer, K. Molecular Dynamics Simulation for Polymers in the Presence of a Heat Bath. *Phys. Rev. A* **1986**, *33*, 3628–3631.
- (41) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An $N \cdot \log(N)$ Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (42) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (43) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. . Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (44) Gumbart, J. C.; Roux, B.; Chipot, C. Standard Binding Free Energies from Computer Simulations: What Is the Best Strategy? *J. Chem. Theory Comput.* **2013**, *9*, 794–802.
- (45) Kötter, A.; Mootz, H. D.; Heuer, A. Standard Binding Free Energy of a SIM–SUMO Complex. *J. Chem. Theory Comput.* **2019**, *15*, 6403–6410.
- (46) Fu, H.; Gumbart, J. C.; Chen, H.; Shao, X.; Cai, W.; Chipot, C. BFEE: A User-Friendly Graphical Interface Facilitating Absolute Binding Free-Energy Calculations. *J. Chem. Inf. Model.* **2018**, *58*, 556–560.
- (47) Fu, H.; Cai, W.; Hénin, J.; Roux, B.; Chipot, C. New Coarse Variables for the Accurate Determination of Standard Binding Free Energies. *J. Chem. Theory Comput.* **2017**, *13*, 5173–5178.
- (48) Genheden, S.; Ryde, U. Comparison of End-Point Continuum-Solvation Methods for the

Calculation of Protein-Ligand Binding Free Energies. *Proteins Struct. Funct. Bioinforma.* **2012**, *80*, 1326–1342.

- (49) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478–9508.
- (50) Salmaso, V.; Moro, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Front. Pharmacol.* **2018**, *9*.
- (51) Guterres, H.; Im, W. Improving Protein-Ligand Docking Results with High-Throughput Molecular Dynamics Simulations. *J. Chem. Inf. Model.* **2020**, *60*, 2189–2198.
- (52) Otto, H.-H.; Schirmeister, T. Cysteine Proteases and Their Inhibitors. *Chem. Rev.* **1997**, *97*, 133–172.
- (53) 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.
- (54) Scheurer, M.; Rodenkirch, P.; Siggel, M.; Bernardi, R. C.; Schulten, K.; Tajkhorshid, E.; Rudack, T. PyContact: Rapid, Customizable, and Visual Analysis of Noncovalent Interactions in MD Simulations. *Biophys. J.* **2018**, *114*, 577–583.

Table 1. Top 8 Docking Results for SARS-COV-2 M^{pro}.

Compound ID	Generic Name	Number of Conformations that Possess Binding Affinity of ≤ -33 kJ mol ⁻¹
ChEMBL275592	NA	9
DB01232	saquinavir	9
DB00471	montelukast	7
DB00549	zafirlukast	7
ChEMBL288347	NA	6
DB00559	bosentan	6
DB01200	bromocriptine	5
DB08995	doismin	5

Table 2. MD Results and Relative Binding Energies of the Top 8 Docking Compounds.

Compound	Structure Stability in the M ^{pro} Pocket	Averaged Contact Area (Å ²) ^a	MMGBSA Relative Binding Energies (kJ mol ⁻¹)
ChEMBL275592	Stable	395	– 158
montelukast	Stable	433	– 154
ChEMBL288347	Stable	411	– 144
bromocriptine	Stable	354	– 121
saquinavir	Stable	362	– 117
zafirlukast	Unstable		– 73
bosentan	Unstable		– 59

^aAveraged contact area between compounds and the M^{pro} during the 100 ns NPT ensemble.

Table 3. Averaged Contact Area (Å²) between Compounds and Active Site Residues of the M^{pro} during the 100 ns NPT Ensemble.

Residue	Contact Area (Å ²)					
	N3	ChEMBL275592	montelukast	ChEMBL288347	bromocriptine	saquinavir
Thr25	14	19	16	14	12	12
His41	24	15	17	19	6	11
Ser46	2	20	8	32	31	29
Met49	5	27	7	30	24	17
Asn142	44	3	30	16	69	32
Gly143	25	31	19	6	28	6
Cys145	13	14	23	5	18	4
Met165	19	24	55	52	39	14
Glu166	50	49	43	30	20	26
Pro168	53	18	30	40	9	33
Gln189	56	40	43	70	40	83
Ala191	23	16	19	12	0	4

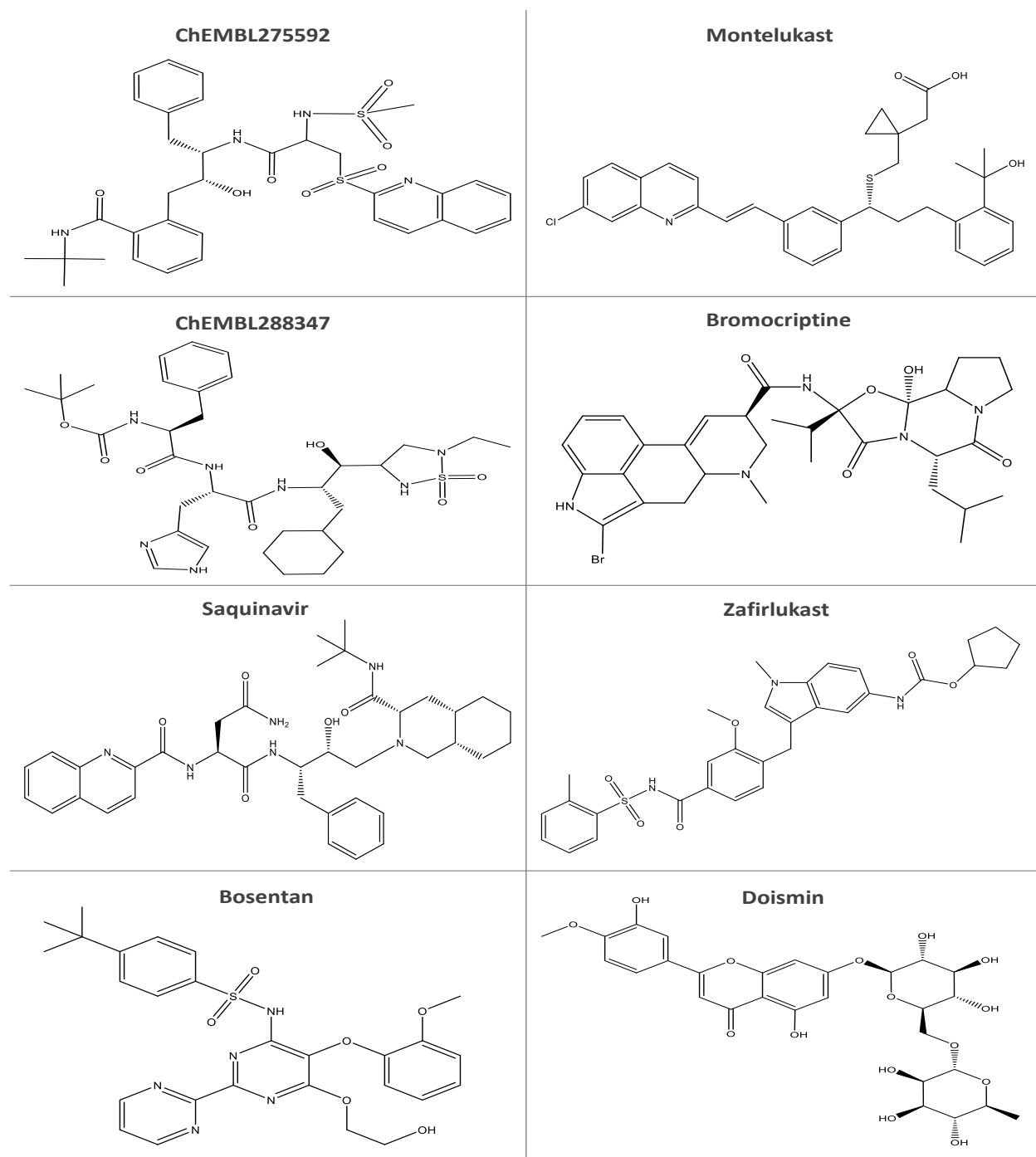


Figure 1. Two-dimensional structures of the selected top 8 compounds.

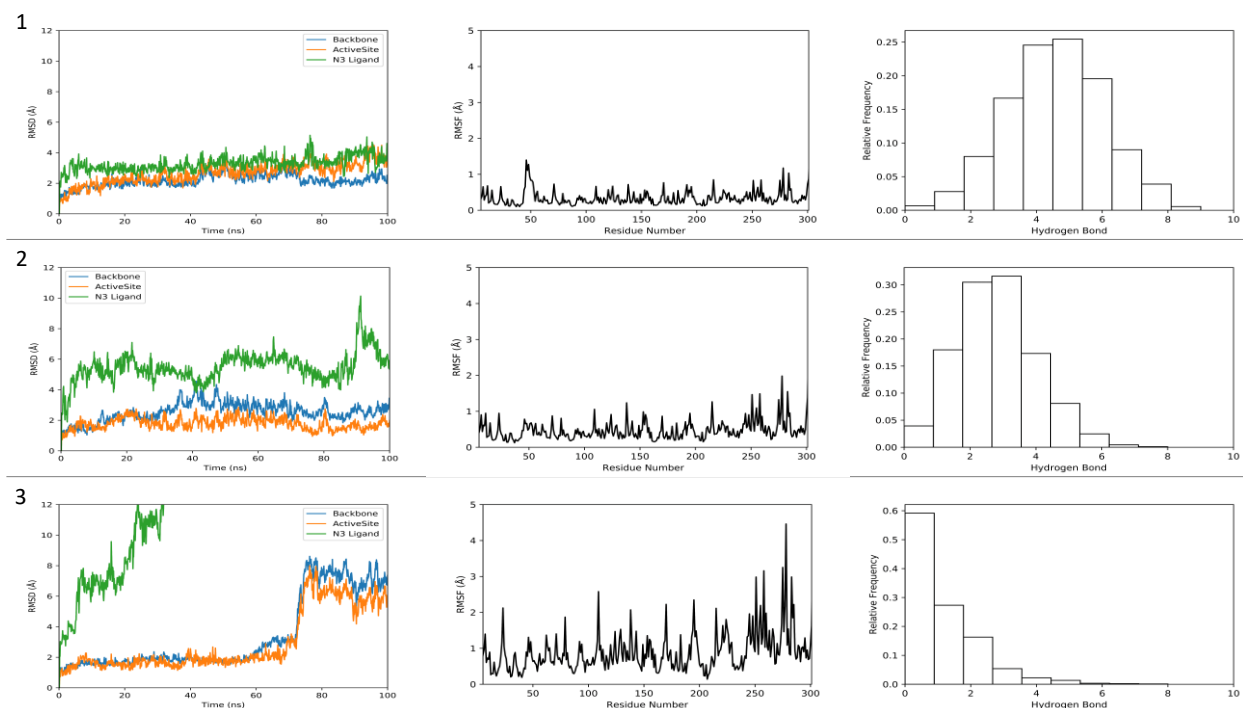


Figure 2. From left to right: RMSD, RMSF, and hydrogen bond interactions of the M^{pro} -N3 complex through 100 ns MD simulations. Model 1 (top), Model 2 (middle), and Model 3 (bottom).

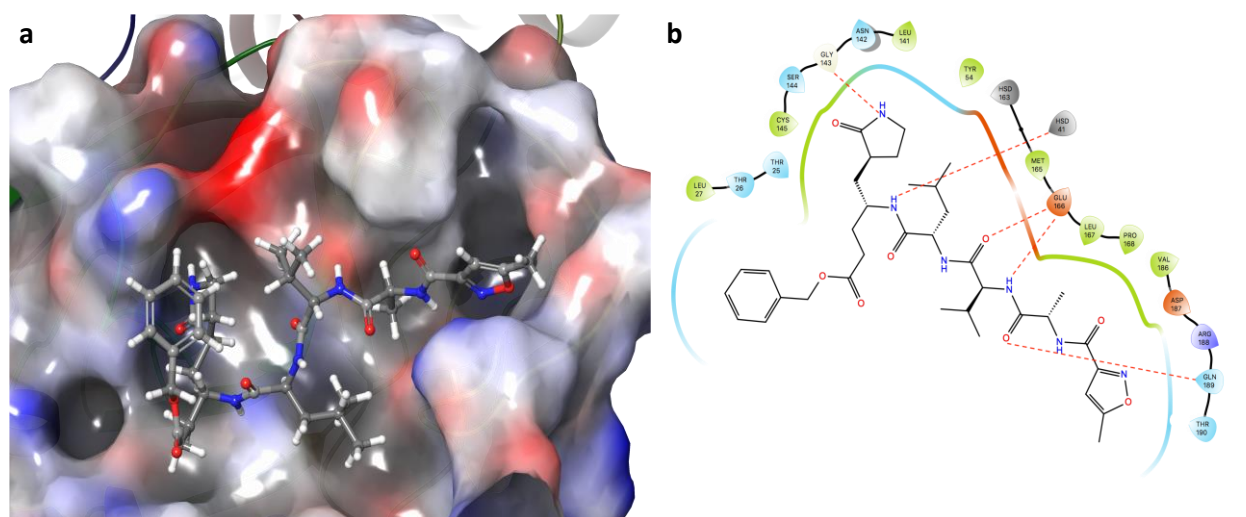


Figure 3. (a) Representative pose of the N3 ligand in the M^{pro} pocket. The protein surface is colored based on the electrostatic potential. (b) Hydrogen bond interactions (red dashed line) of the N3 ligand in complex with the M^{pro} .

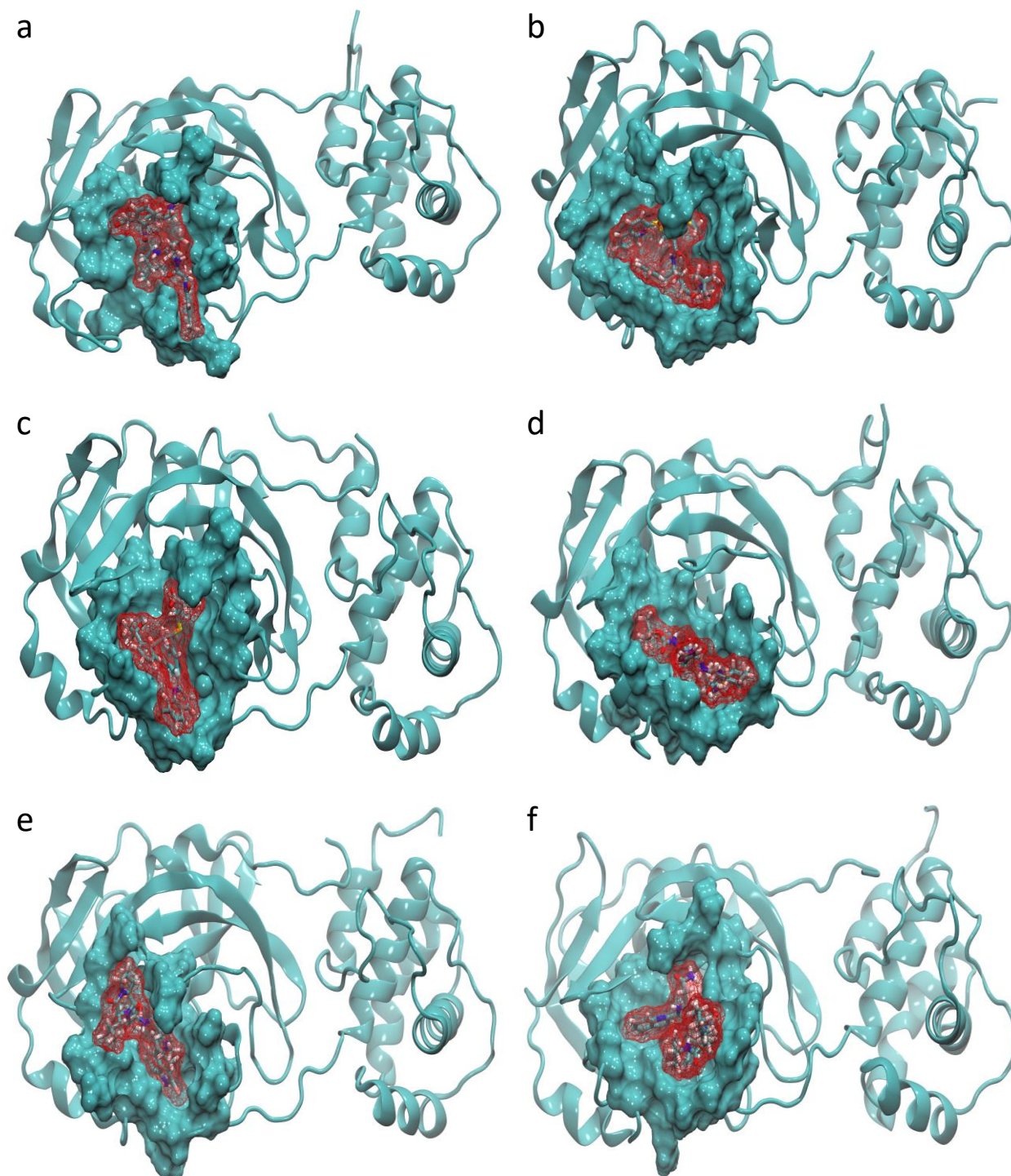
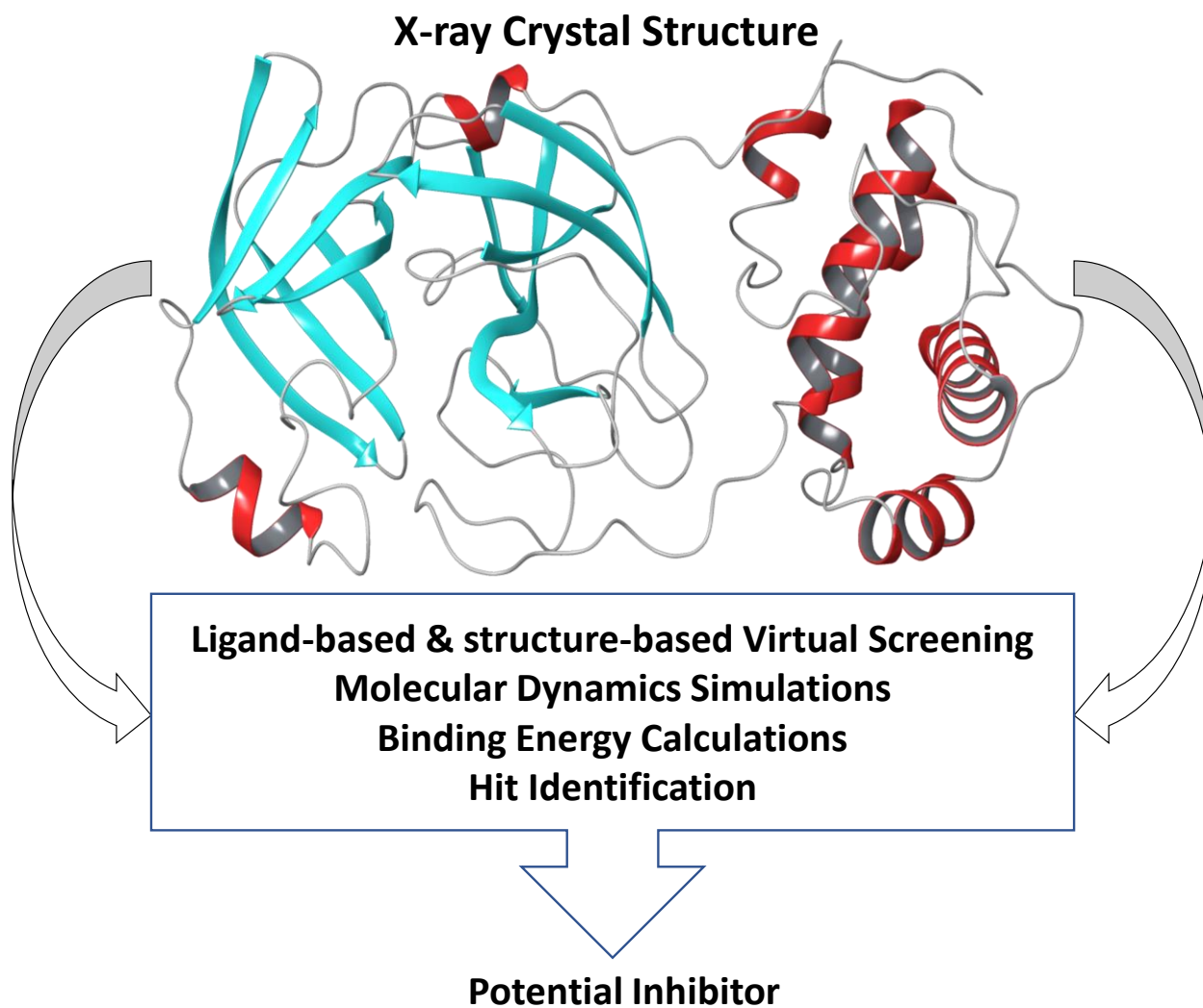


Figure 4. Last frame of 100 ns MD simulations of ligands in the M^{pro} pocket. (a) N3 inhibitor, (b) ChEMBL275592, (c) montelukast, (d) ChEMBL288347, (e) bromocriptine, and (f) saquinavir.



TOC Figure. Computational investigation of novel inhibitors for SARS-CoV-2 M^{pro}.