JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

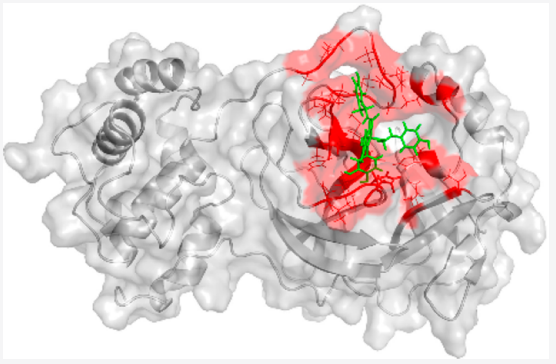
Computational Determination of Potential Inhibitors of SARS-CoV-2 Main Protease

Son Tung Ngo,* Ngoc Quynh Anh Pham, Ly Thi Le, Duc-Hung Pham,* and Van V. Vu*

Cite This: https://dx.doi.org/10.1021/acs.jcim.0c00491 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The novel coronavirus (SARS-CoV-2) has infected several million people and caused thousands of deaths worldwide since December 2019. As the disease is spreading rapidly all over the world, it is urgent to find effective drugs to treat the virus. The main protease (Mpro) of SARS-CoV-2 is one of the potential drug targets. Therefore, in this context, we used rigorous computational methods, including molecular docking, fast pulling of ligand (FPL), and free energy perturbation (FEP), to investigate potential inhibitors of SARS-CoV-2 Mpro. We first tested our approach with three reported inhibitors of SARS-CoV-2 Mpro, and our computational results are in good agreement with the respective experimental data. Subsequently, we applied our approach on a database of ∼4600 natural compounds, as well as 8 available HIV-1 protease (PR) inhibitors and an aza-peptide epoxide. Molecular docking resulted in a short list of 35 natural compounds, which was subsequently refined using the FPL scheme. FPL simulations resulted in five potential inhibitors, including three natural compounds and two available HIV-1 PR inhibitors. Finally, FEP, the most accurate and precise method, was used to determine the absolute binding free energy of these five compounds. FEP results indicate that two natural compounds, cannabisin A and isoacteoside, and an HIV-1 PR inhibitor, darunavir, exhibit a large binding free energy to SARS-CoV-2 Mpro, which is larger than that of **13b**, the most reliable SARS-CoV-2 Mpro inhibitor recently reported. The binding free energy largely arises from van der Waals interaction. We also found that Glu166 forms H-bonds to all of the inhibitors. Replacing Glu166 by an alanine residue leads to ∼2.0 kcal/mol decreases in the affinity of darunavir to SARS-CoV-2 Mpro. Our results could contribute to the development of potential drugs inhibiting SARS-CoV-2.



INTRODUCTION

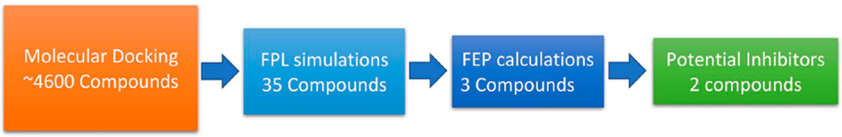
Members of the *Coronaviridae* virus family often cause mild respiratory syndrome in humans.[1] However, the severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV) are transfected from animals to humans and cause severe cases of respiratory syndromes and deaths.[2,3] In 2002, SARS-CoV was first recorded in Guandong, China, and linked to 8096 laboratory-confirmed cases of infection and 774 deaths.[3] The natural reservoir of SARS-CoV is Chinese horseshoe bats,[4] and intermediate hosts are civet cats and raccoon dogs.[5] This shows that Coronavirus can induce severe symptoms and potential pneumonia and death. In December 2019, a novel coronavirus (2019-nCoV or SARS-CoV-2) that has a similar sequence to SARS-CoV emerged in Wuhan, Hubei province, China.[6−8] The initial cluster of infection seemed to relate to Huanan seafood market. SARS-CoV-2 is thought to originate from bat. though the intermediate hosts are still unknown;[9] human-to-human transmission has been validated.[10] As of May 7, 2020, SARS-CoV-2 has infected more than 3,800,000 people and caused over 265,000 deaths worldwide.[11]

Coronaviruses have the largest genomes among all known RNA viruses, ranging from 26 to 32 kb in length, which encode structural and nonstructural proteins.[12,13] The SARS-CoV-2 genome encodes more than 20 proteins, which include the main protease (Mpro), a 3C-like protease (3CLP) that shares 96,1% similarity with 3CLP of SARS-CoV.[13,14] Mpro, a homodimeric cysteine protease, plays an important role in SARS virus replication and transcription. When the mRNA of the virus is translated polyproteins, Mpro is first autocleaved to become a mature enzyme, which in turn cleaves all of the 11 remaining downstream nonstructural proteins of the polyproteins to polypeptides, which are required for the replication process of the virus.[13] SARS-CoV-2 Mpro has thus been an attractive drug target.[14,15] Darunavir and ritonavir can potentially inhibit SARS-CoV-2 Mpro and have been put into clinical trials for COVID-19 treatment.[16,17]

Special Issue: COVID19 - Computational Chemists Meet the Moment

Received: May 7, 2020
Published: June 12, 2020

Scheme 1. Computational Strategy to Determine the Probable Natural Inhibitors of SARS-CoV-2 Mpro



Computer-aided drug design (CADD) is frequently used to estimate the probable inhibitors that could prevent the activity of an enzyme. This method significantly decreases the time and cost to develop a new drug.[18] Determination of the ligand-binding free energy is one of the most critical factors in CADD.[19] Many schemes were then developed to resolve this problem.[20] Typically, the ligand-binding affinity of several thousand ligands to a protein is frequently predicted via the molecular docking method;[21] a short list of these compounds would be then refined by using more computationally expensive binding free energy methods such as the molecular mechanism/Poisson−Boltzmann surface area (MM/PBSA),[22−24] linear interaction energy (LIE),[25,26] or fast pulling of ligand (FPL)[27] approaches. The top-lead potential inhibitors will be finally confirmed via an accurate binding free energy approach as well as the free energy perturbation (FEP),[28,29] thermodynamic integration (TI),[30,31] and nonequilibrium molecular dynamics simulations (NEMD).[32] Especially, in some cases, calculations can be carried out by using a combination of Hamiltonian/temperature replica exchange molecular dynamics (REMD) simulations and the perturbation method.[33−36] In this work, we carried out computational investigations according to Scheme 1 to evaluate the potential inhibitors for SARS-CoV-2 Mpro. The obtained results could help enhance the development of SARS-CoV-2 therapy.

MATERIALS AND METHODS

Structure of Complexes. The three-dimensional structure of SARS-CoV-2 Mpro was downloaded from https://innophore.com/.[37] This modeled structure well superimposes the recent experimental structure with a C_α RMSD smaller than 0.05 nm (Figure S1 of the Supporting Information).[14] The G166A mutant ofSARS-CoV-2 Mpro was generated using the PyMOL mutagensis tools.[38]

Molecular Docking Simulations. The molecular docking using the Autodock Vina package[39] was employed to rapidly determine the ligand-binding pose and affinity to SARS-CoV-2 Mpro with the exhaustiveness of 8 referring to the previous study.[40] The parameters of complexes were prepared using AutodockTools 1.5.6,[41] which were denoted in the PDBQT file. The atomic charges of both receptor and ligands were predicted via the Gasteiger−Marsili approach.[42,43] The receptor and ligands were represented via a united atom model with explicit polar hydrogens.[44] The best docking mode was selected as the lowest obtained binding energy result. The grid center was selected as the center of mass of aza-peptide epoxide, which bound to the active site of SARS-CoV Mpro.[45] The grid size was chosen as $26 \times 26 \times 26$ Å^3.

Molecular Dynamics Simulations. GROMACS version 5.1.5[46] was employed to simulate the structural change of the solvated complex SARS-CoV-2 Mpro + inhibitor. The parameters for MD simulations were referred to the previous works.[36] SARS-CoV-2 Mpro was parametrized using the Amber99SB-ILDN force field.[47] Water molecules were parameterized using the Tip3p water model. Ligand structures were downloaded from the PubChem database.[48] The ligands were parametrized with the general Amber force field (GAFF)[49] using the combination of AmberTools18[50] and ACPYPE[51] protocols. The atomic charges were assigned using the restrained electrostatic potential (RESP) method[52] computed with quantum chemical calculation at the B3LYP double-hybrid functional and the 6-31G(d,p) basis set. The details of complexed configuration upon applied free energy methods were reported in the next subsection. The time steps of MD simulations were set to 2 fs. The electrostatic interaction was mimicked via the fast smooth particle-mesh Ewald electrostatics method.[53] The cutoff of the van der Waals (vdW) interaction was set at 0.9 nm. The temperature and pressure couplings were calculated using the V-rescale and Parrinello−Rahman schemes, respectively. The solvated complex was minimized using the steepest descent method. The energy minimized system was relaxed over 100 ps of NVT and 2 ns of NPT ensembles at 310 K. During the NVT and NPT simulations, C_α atoms of the SARS-CoV-2 Mpro were softly retrained using a harmonic force. The coordinates of the solvated complexes were monitored over the atomistic simulations every 10 ps.

Free Energy Calculation. *Fast Pulling of Ligand (FPL) Approach.* The last snapshot of NPT simulations was used as the initial structure for SMD simulation.[27] Details of the computations were referred to the previous studies.[27] In particular, the (x, y, z) dimensions of the systems are (9.83, 5.92, 8.70) nm, as shown in Figure 1. The systems in FPL simulations consist of 1 SARS-CoV-2 Mpro, 1 ligand, 15,000 water molecules, and Na^+ ions for a total of ca. 50,000 atoms. The pulling speed (v) and spring constant cantilever (k) were set at 0.005 nm ps^{-1} and 600 kJ mol^{-1} nm^{-2}, respectively. During the simulations, the C_α atoms of Mpro were positionally restrained using a weak harmonic potential. A harmonic force was put on the center of mass of the inhibitor to disassociate it from the binding cavity of the SARS-CoV-2 Mpro (Figure 1). The pulling force value and displacement of the ligand along the unbinding direction were monitored every 0.1 ps. The FPL

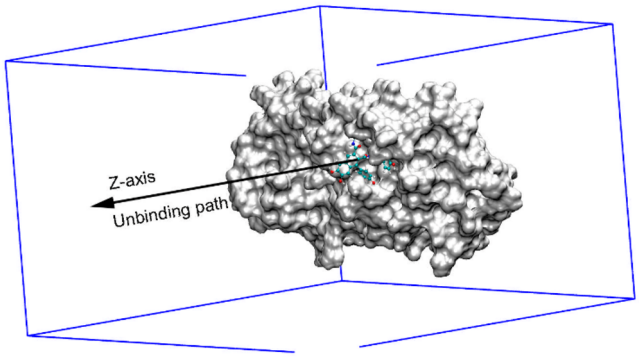


Figure 1. Computational model for FPL simulations of SARS-CoV-2 Mpro−ligand binding affinity.

https://dx.doi.org/10.1021/acs.jcim.0c00491
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

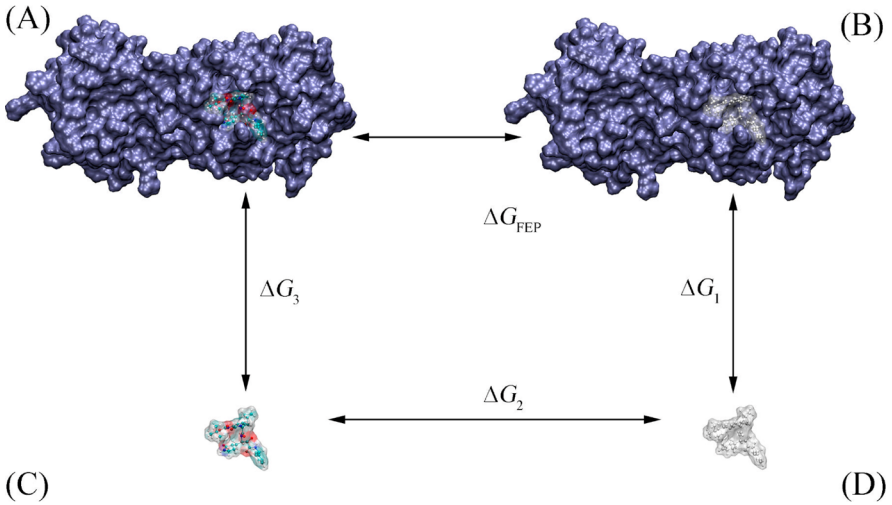


Figure 2. Thermodynamics diagram of determination of the absolute binding free energy between a ligand and SARS-CoV-2 Mpro. (A) The full-interaction state of an inhibitor with surrounding molecules, including the Mpro and solvent molecules. (B) A dummy inhibitor with the solvated protease. (C) The full-interaction state of an inhibitor with the solvent molecules. (D) A dummy inhibitor in solution. A dummy inhibitor is a molecule that has no nonbonded interaction with neighboring molecules. The solvent molecules are hidden for clarity.

simulations were repeated with eight independent trajectories to guarantee the sampling of simulations.

Free Energy Perturbation (FEP) Simulations.[54] The last snapshot of NPT simulations was used as the initial conformation for 20 ns long MD simulations. In particular, the SARS-CoV-2 Mpro + inhibitor complex was inserted into a dodecahedron periodic boundary condition (PBC) box with a volume of ca. 820 nm^3. The complexed system comprises 1 SARS-CoV-2 Mpro, 1 ligand, 25,280 water molecules, and 4 Na^+ ions for a total of ca. 80,600 atoms. Moreover, the isolated inhibitor was inserted into a dodecahedron PBC box with a volume of ca. 85 nm^3. The solvated ligand system consists of 1 ligand and ca. 2750 water molecules for a total of ca. 8300 atoms. The equilibrium conformation of MD simulations was then employed as the starting structure for FEP calculations according to the previous study.[36] During FEP simulations, the coupling parameter λ, which varies from 0 to 1, was employed to evaluate the free energy change ΔG of the system modification from the *full-interaction* state ($\lambda = 0$) to the *non-interaction* state ($\lambda = 1$) via the alteration of the systemic Hamiltonian between various circumstances. The change of a ligand from *full-interaction* to *non-interaction* states with surrounding molecules is called the ligand annihilation process (Figure 2). Eight values of λ_{cou}, including 0.00, 0.100, 0.20, 0.35, 0.50, 0.65, 0.80, and 1.00, were used to modify the Coulomb interactions. Nine values of λ_{vdW}, including 0.00, 0.10, 0.25, 0.35, 0.50, 0.65, 0.75, 0.90, and 1.00, were used to alter the vdW interactions. Sixteen alter-λ simulations were performed to demolish a ligand from a solvated system (Figure 2). The total energy change of the ligand annihilation process was then summed via the Bennet's acceptance ratio (BAR) method.[55] Finally, the absolute binding free energy between a ligand to SARS-CoV-2 Mpro was deduced as the different energy between two annihilation processes involving decoupling the ligand from the solvated ligand system and from the solvated protein−ligand system.

Structural Analysis. A hydrogen bond (HB) is determined when an acceptor (A)−hydrogen (H)−donor (D) angle is larger than 135° and the A−D distance is less than 0.35 nm. A side-chain (SC) contact between inhibitors and SARS-CoV-2 Mpro is counted when the distance between non-hydrogen atoms of two molecules is less than 0.45 nm. The two-dimensional interaction diagram between a protein and a ligand was generated using the LigPlot++ program.[56] Moreover, the pharmacokinetics of the top-lead compounds were predicted using the PreADME server.[57]

RESULTS AND DISCUSSION

Potential Inhibitor Screening Using Molecular Docking. Autodock Vina[39] is one of the most popular free packages to roughly and rapidly estimate the binding affinity and binding pose of a trial inhibitor to an enzyme or a protein. The successful-docking rate of the package was up to 81% according to our previous benchmark study on over 800 protein−ligand complexes.[40] We used Autodock Vina to dock three previously reported ligands[14] to SARS-CoV-2 Mpro and obtained binding energies reasonably consistent with experimentally determined values (Table 1). Therefore, in this project, Autodock Vina[39] was employed to rapidly evaluate the binding affinity of ca. 4600 natural compounds from the Vietherb database.[58] Because some current HIV-1 PR inhibitors, such as darunavir[16] or ritonavir,[17] have been tested for SARS-CoV-2 inhibition, eight drugs inhibiting HIV-1 PR, including amprenavir, atazanavir, darunavir, indinavir, lopinavir, nelfinavir, ritonavir, and saquinavir, were also investigated. Moreover, the binding of aza-peptide epoxide was also redocked to SARS-CoV-2 Mpro in order to compare with other ligands. The binding affinity of top-lead compounds to SARS-CoV-2 Mpro is provided in Table S1 of the Supporting Information, respectively. The obtained

Table 1. Recently Reported Inhibitors of SARS-CoV-2 Mpro

N^0	compound name	ΔG_{Dock}[a]	ΔG_{EXP}[b]
1	**11r**	−7.1	−9.23
2	**13a**	−6.7	−7.70
3	**13b**	−6.9	−8.45

[a]Docking binding free energy obtained by Autodock Vina.[39] [b]Experimental binding free energy ΔG_{EXP} roughly estimated based on the IC50 value reported recently,[14] assuming that the inhibition constant (k_i) is equal to the IC50 value. The unit of ΔG is kcal/mol.

https://dx.doi.org/10.1021/acs.jcim.0c00491
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

docking energies fall in the range from −1.2 to −9.8 kcal/mol with a median of −6.22 ± 0.02 kcal/mol (the computed error is the standard error of the mean) (Figure 3).

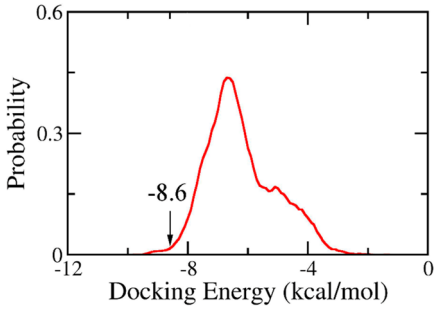


Figure 3. Distribution of docking energy between 4663 natural compounds and SARS-CoV-2 Mpro.

There are 35 natural compounds from the Vietherb database exhibiting a large ligand affinity to SARS-CoV-2 Mpro. The affinity of these compounds ranges from −8.6 to −9.8 kcal/mol (Table S1 in the Supporting Information), which is significantly larger than that found in the range of −6.4 to −7.6 kcal/mol for the eight HIV-1 PR inhibitors and aza-peptide epoxide. The 35 natural compounds form more HBs to SARS-CoV-2 Mpro than these nine compounds (Figure 4 and Table S3 of the Supporting Information).

Refining Docking Results Using FPL Simulations. The obtained docking results were refined using the FPL method.[27] The FPL scheme is a very efficient technique to rapidly explore the binding affinity of a ligand to a protein, when the protein binding cavity is accessible to the exogenous ligand without sizable conformational change during the binding/unbinding process. The FPL approach requires a small amount of computing resource, but it could provide results with high accuracy and precision.[27] The maximum pulling force (F_{max}), called the rupture force, and the recorded pulling work (W) were used as a criterion to rank the ligand affinity.[27,59] However, as mentioned in the previous work,[27] the pulling work is more appropriate than the rupture force, as it directly associates with the ligand-binding free energy via isobaric−isothermal Jarzynski equality.[60,61]

In this work, we carried out FPL simulations to rank the affinity to SARS-CoV-2 Mpro of 44 compounds screened with docking studies. The FPL calculations for **11r**, **13a**, and **13b**[14] were also carried out for comparison. The equilibrated snapshot obtained from 2 ns of NPT simulations was used as an initial structure for the FPL simulations. The maximum pulling force, called the rupture force, and the pulling work were obtained from eight independent trajectories. The obtained results are provided in Table 2. The mean of recorded rupture forces F_{max} ranges from 416.9 ± 35.4 to 901.0 ± 59.2 pN. The time-dependent pulling forces of these 47 systems are provided in Figure S3 of the Supporting Information. The form of pulling force curves are in good agreement with the previous studies,[27] in which the pulling forces continuously increase to maximum values before rapidly dropping to zero after the nonbonded contacts between the ligand to the protein were terminated. Here, the pulling work was selected as a criterion to rank the ligand affinity (Figure 5). The average pulling work W ranges from 36.1 ± 4.5 to 104.0 ± 5.6 kcal/mol (Table 2). The FPL-derived pulling work for **11r**, **13a**, and **13b** is 43.3 ± 3.9, 94.6 ± 5.0, and 91.9 ± 3.6 kcal/mol, respectively, which is consistent with respecttive experiments.[14] This result supports our approach in using FPL to refine the docking results.

A short list of potential inhibitors of SARS-CoV-2 Mpro was obtained and shown in Table 1. The pulling work W for darunavir and ritonavir is 83.9 and 85.9 kcal/mol, respectively, which is >11% larger than that of the other HIV-1 PR inhibitors (Table 2). Previous computational investigations suggested that lopinavir was able to prevent SARS-CoV-2 Mpro.[62] However, FPL results show otherwise, which is consistent with the recent clinical research.[63] Two natural compounds, cannabisin A and isoacteoside, have larger W values than that of ritonavir. Cannabisin A, Pubchem ID of 15086398, adopts the largest values of both W and F_{max}, which are 104.0 kcal/mol and 901.0 pN, respectively. Isoacteoside, Pubchem ID of 6476333, has a pulling work W of 92.2 ± 4.4 kcal/mol. Besides these compounds, quercetin 7-*O*-rutinoside was also included in the list of potential inhibitors for the SARS-CoV-2 Mpro, because it adopts a W value of 79.4 ± 6.4 kcal/mol, which is only 5% smaller than that of darunavir. These five compounds adopt an appropriate pulling work W in comparison with that obtained for **13b** (91.9 ± 3.6 kcal/mol), the most reliable SARS-CoV-2 Mpro inhibitor recently reported.[14]

Validation of FPL Results Using FEP Calculations. Accurate and precise determination of the ligand-binding free energy probably reduces drug discovery cost.[64] Therefore, in order to validate the FPL results, the absolute binding free energy between five ligands was computed using the FEP

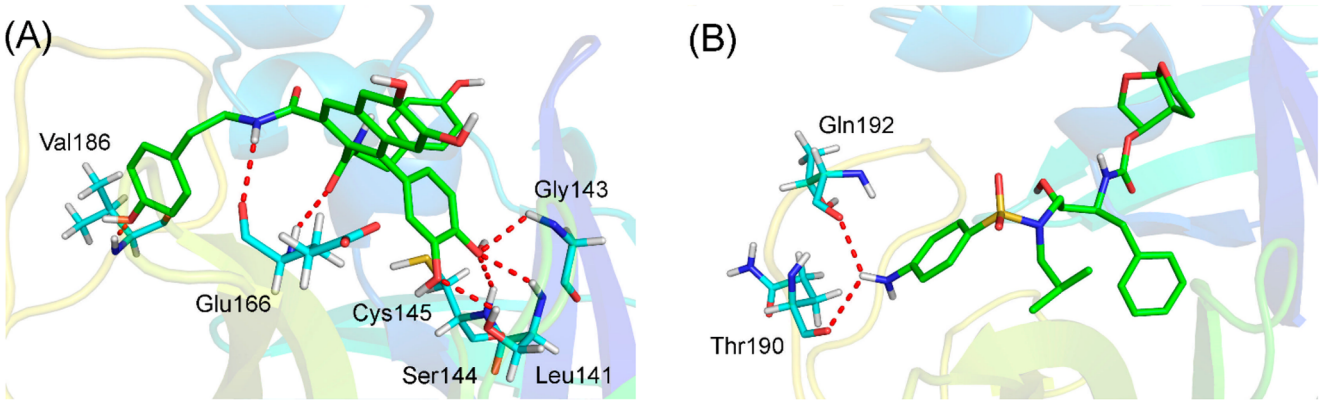


Figure 4. Docked conformations of SARS-CoV-2 Mpro + cannabisin A (A) and SARS-CoV-2 Mpro + darunavir (B) complexes.

https://dx.doi.org/10.1021/acs.jcim.0c00491
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 2. FPL Results of Top-Lead Compounds Screened with Molecular Docking

N^0	Pubchem	compound name	ΔF_{Max} [a]	W [b]	ΔG_{EXP} [c]
1		**11r**	857.5 ± 38.7	94.6 ± 5.0	−9.23
2		**13a**	496.0 ± 32.5	43.3 ± 3.9	−7.70
3		**13b**	884.3 ± 36.5	91.9 ± 3.6	−8.45
4	10621	hesperidin	575.6 ± 46.2	62.7 ± 4.6	
5	73330	strictinin	633.2 ± 27	67.9 ± 3.8	
6	83489	eriocitrin	588.7 ± 26.8	71.0 ± 4.8	
7	114777	CHEMBL346119	721.2 ± 38.1	72.6 ± 4.5	
8	122738	procyanidin B2	668.8 ± 20	77.8 ± 3.7	
9	124356	physalin F	614.2 ± 23.5	52.6 ± 1.8	
10	156766	kihadanin B	500.3 ± 30.2	45.0 ± 3.0	
11	179651	limonin	516.3 ± 31.9	45.1 ± 1.8	
12	183905	6,8-di-*C*-beta-D-arabinopyranosyl apigenin	672.6 ± 38.9	67.4 ± 4.8	
13	190799	stephasubine	807.4 ± 54.4	78.4 ± 7.3	
14	196583	mulberrofuran G	674.2 ± 52.4	71.7 ± 4.9	
15	442431	narirutin	535.8 ± 45.3	54.4 ± 5.1	
16	480819	albanol B	546.6 ± 27.7	49.7 ± 3.8	
17	5281600	amentoflavone	710.6 ± 50.9	74.3 ± 6.3	
18	5281613	diosmin	714.0 ± 47.5	77.5 ± 5.6	
19	5281627	hinokiflavone	645.1 ± 51.0	67.6 ± 4.2	
20	5317025	linarin	548.3 ± 21.9	58.5 ± 3.0	
21	5319276	marchantin K	567.6 ± 13.3	50.0 ± 1.7	
22	5319278	marchantin L	616.1 ± 34.0	53.9 ± 3.2	
23	5319933	mulberrofuran Q	539.4 ± 16.4	54.2 ± 2.7	
24	5458744	physalin B 5,6-epoxide	476.2 ± 32.9	38.4 ± 3.3	
25	6476333	isoacteoside	730.7 ± 40.1	92.2 ± 4.4	
26	6711179	hypopistephanine	707.0 ± 34.5	65.1 ± 3.8	
27	9851181	isorhoifolin	567.7 ± 36.7	57.9 ± 4.8	
28	10456516	cinchonain-Ib	547.3 ± 33.7	53.5 ± 3.1	
29	10461109	luteolin-7-*O*-beta-rutinoside	608.1 ± 63.0	65.0 ± 6.0	
30	11827970	diosgenin glucoside	514.8 ± 41.2	49.1 ± 5.6	
31	15086398	cannabisin A	901.0 ± 59.3	104 ± 5.6	
32	16760075	didymin	574.5 ± 50.9	63.7 ± 6.0	
33	21123844	gamma-chaconine	416.9 ± 35.4	36.1 ± 4.5	
34	44558930	anabsinthin	589.3 ± 57.8	56.4 ± 5.4	
35	71437113	2,3-dihydrohinokiflavone	546.5 ± 36.1	61.5 ± 3.3	
36	71448965	cannabisin D	733.1 ± 32.9	70.4 ± 4.1	
37	90473381	N/A	564.9 ± 53.0	50.0 ± 7.1	
38	101764560	quercetin 7-*O*-rutinoside	737.9 ± 47.8	79.4 ± 6.4	
39	65016	amprenavir	607.6 ± 29.9	55.4 ± 3.7	
40	148192	atazanavir	647.7 ± 37.9	74.1 ± 3.3	
41		aza-peptide epoxide	586.4 ± 48.2	61.5 ± 6.4	
42	213039	darunavir	817.8 ± 32.0	83.9 ± 4.3	
43	5362440	indinavir	456.3 ± 33.0	48.5 ± 1.7	
44	92727	lopinavir	684.8 ± 44.5	71.2 ± 3.9	
45	64143	nelfinavir	607.9 ± 31.5	58.1 ± 3.0	
46	392622	ritonavir	764.8 ± 54.0	85.9 ± 7.8	
47	441243	saquinavir	601.3 ± 41.6	66.4 ± 4.4	

[a]Mean rupture force ΔF_{Max}. [b]Mean pulling work W obtained from eight independent trajectories of SMD simulations. [c]Experimental binding free energy ΔG_{EXP} roughly estimated based on the IC50 value reported recently,[14] assuming that the inhibition constant (k_i) is equal to the IC50 value. The error is the standard error of the mean. The unit of energy and work is kcal/mol; the unit of force is pN.

method (Table 3), one of the most accurate and precise methods known to date.[20,65] FEP is often used in CADD, as it often provides results consistent with experiments.[66−68] The binding free energy of three recently reported inhibitors of SARS-CoV-2 Mpro, including **11r**, **13a**, and **13b**, was also calculated. The good agreement between computational and experimental values[14] indicates that the FEP method is reliable in calculating the binding free energy of ligands to SARS-CoV-2 Mpro.

The equilibrium snapshots of the SARS-CoV-2 Mpro + inhibitor systems generated in NPT simulations were treated as the initial conformations for MD simulations. These MD simulations were set to run for 20 ns, in which all-atom RMSD of the complex was recorded every 10 ps (Figure S4 of the Supporting Information). During the MD simulations, the binding pose between the SARS-CoV-2 Mpro and the inhibitor was refined under the effects of the interaction among them. The number of HBs between protein−ligand fluctuates from the

https://dx.doi.org/10.1021/acs.jcim.0c00491
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

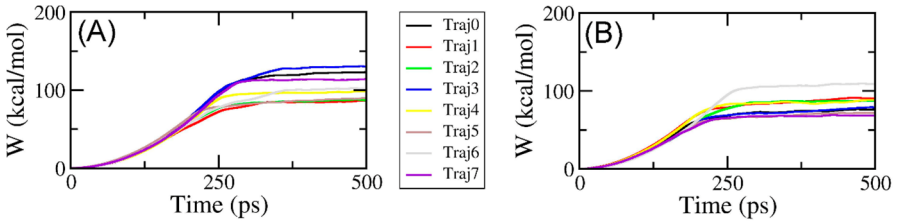


Figure 5. Recorded pulling works during FPL simulations of SARS-CoV-2 Mpro + cannabisin A (A) and SARS-CoV-2 Mpro + darunavir (B) complexes.

Table 3. Computationally Determined Potential Inhibitors for Wild Type (WT) and E166A Mutants of SARS-CoV-2 Mpro

N^0	Pubchem ID	complex	herb name	ΔG_{cou}	ΔG_{vdW}	ΔG_{FEP} [a]	ΔG_{EXP} [b]
1		WT + **11r**		−6.02	−7.30	−13.31 ± 2.58	−9.23
2		WT + **13a**		−0.59	−7.59	−8.18 ± 2.20	−7.70
3		WT + **13b**		−1.97	−7.22	−9.18 ± 2.48	−8.45
4	101764560	WT + quercetin 7-*O*-rutinoside	*Platycodon grandiflorum*	−3.82	−9.33	−5.52 ± 1.18	
5	15086398	WT + cannabisin A	*Cannabis sativa*	−2.57	−10.20	−12.76 ± 1.37	
6	6476333	WT + isoacteoside	*Fernandoa adenophylla*	−2.06	−7.34	−9.40 ± 2.64	
7	213039	WT + darunavir		−3.44	−8.52	−11.96 ± 1.99	
8	392622	WT + ritonavir		2.10	−9.83	−7.73 ± 1.77	
9	213039	E166A + darunavir		−1.58	−8.32	−9.90 ± 2.48	

[a]Absolute binding free energy ΔG_{FEP} obtained using three independent FEP calculations. [b]Experimental binding free energy ΔG_{EXP} roughly estimated based on the IC50 value reported recently,[14] assuming that the inhibition constant (k_i) is equal to the IC50 value. The error is the standard error of the mean values. The unit of energy and work is kcal/mol.

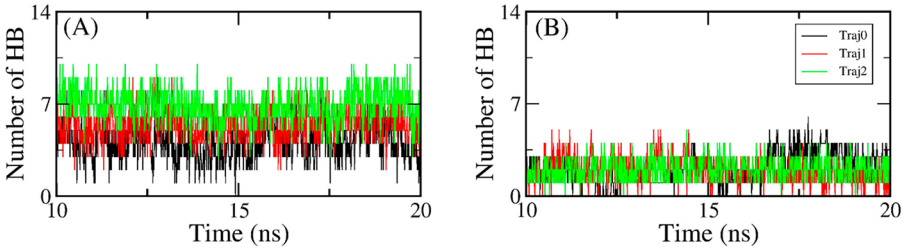


Figure 6. Number of HBs between SARS-CoV-2 Mpro and cannabisin A (A) and darunavir (B) over the equilibrium region of MD simulations.

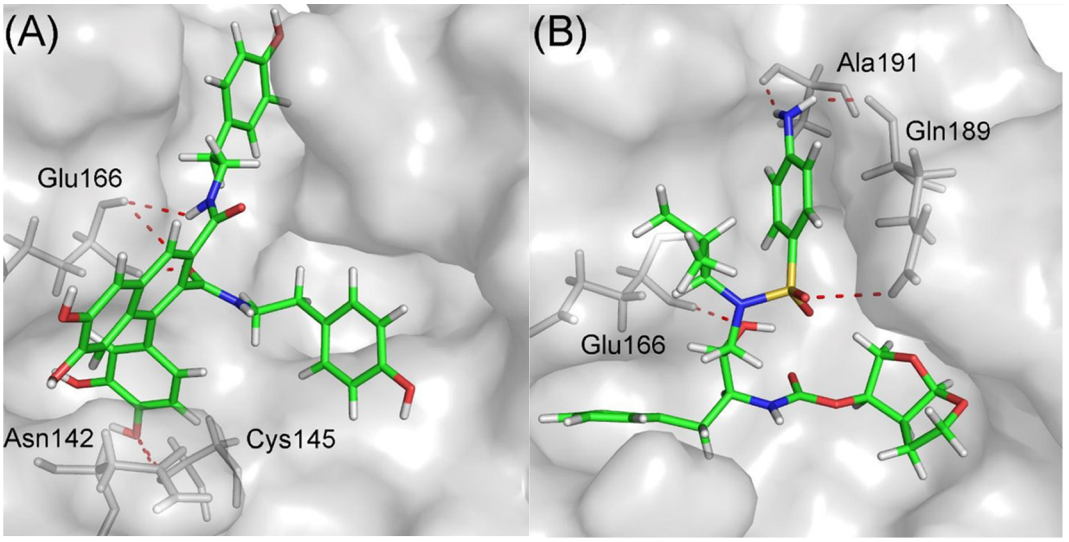


Figure 7. Binding pose of the SARS-CoV-2 Mpro + cannabisin A (A) and SARS-CoV-2 Mpro + darunavir (B) systems, obtained by all-atom clustering with a cutoff of 0.3 nm using 3000 equilibrium snapshots.

beginning of MD simulations and becomes stable after 10 ns of MD simulations (Figure 6). It is consistent with all-atom RMSD of the complex over the MD simulations (Figure S4 of the Supporting Information). Figure 7 shows the dominant structures and binding poses of cannabisin A and darunavir with SARS-CoV-2 Mpro, respectively. The poses for SARS-

https://dx.doi.org/10.1021/acs.jcim.0c00491
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

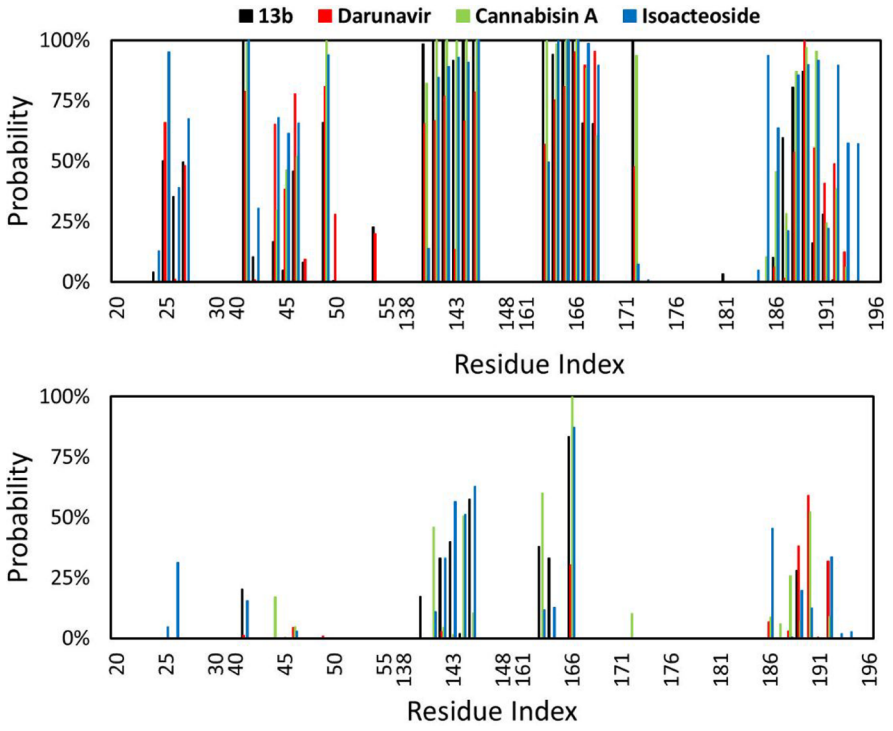


Figure 8. Probabilities of SC (top) and HB (bottom) contacts between high-affinity inhibitors and active site residues of SARS-CoV-2 Mpro over 3000 MD equilibrium snapshots.

CoV-2 Mpro + isoacteoside, quercetin 7-*O*-rutinoside, and ritonavir complexes are provided in Figures S5−S7 of the Supporting Information. Interestingly, Glu166 appears to be an important residue involved in the binding of the inhibitors to SARS-CoV-2 Mpro, as it forms HBs with all of these inhibitors. The mutation of Glu166 could possibly alter the affinity of inhibitors to SARS-CoV-2 Mpro.

The work values of the decoupling ligand from the solvated system are used to determine the free energy change over the annihilation ligand process (Figure 2). Sixteen λ-alteration MD simulations with a length of 2 ns were employed to calculate the work values. The soft-core potentials were used to characterize the altered Hamiltonian system. However, the soft-core potentials cost much more CPU time than the traditional one. Therefore, the performance of MD simulation is significantly reduced during λ-alteration simulations. In total, MD simulations were carried out in 219 ns over three trajectories. The free energy terms were then computed using the BAR method[55] over the interval 1−2 ns of the λ-alteration simulations with a period of 100 ps. Overall, the absolute binding free energy between five potential inhibitors to SARS-CoV-2 Mpro was then obtained (Table 3).

According to the results shown in Table 3, the vdW free energy interaction energy dominates over the electrostatic free interaction energy during a ligand binding to SARS-CoV-2 Mpro. Moreover, darunavir adopts a stronger binding free energy (−11.96 ± 1.99 kcal/mol) to SARS-CoV-2 Mpro than ritonavir (−7.73 ± 1.77 kcal/mol). This result is consistent with recent clinical research[63] that ritonavir only has a weak inhibitory effect on SARS-CoV-2.

Quercetin 7-*O*-rutinoside, a compound from *Platycodon grandiflorum*,[58] exhibits poor binding affinity to SARS-CoV-2 Mpro (−5.52 ± 1.18 kcal/mol). Isoacteoside, a compound from *Fernandoa adenophylla*,[58] has a binding affinity of −9.40 ± 2.64 kcal/mol, which falls between that of ritonavir and darunavir. Cannabisin A, a compound from *Cannabis sativa*,[69] adopts the strongest binding affinity to SARS-CoV-2 Mpro (−12.76 ± 1.37 kcal/mol).

The FEP result suggests that canabisin A, isoacetoside, and darunavir are the potential inhibitors for SARS-CoV-2 Mpro, since their binding free energies are larger than that of compound **13b**, which has a computational binding free energy of −9.18 ± 2.48 kcal/mol (Table 3). In addition, the cell membrane crossing ability was then predicted for these compounds using the preADMET server. The logP value predicted for canabisin A is 5.18, which is similar to that of ritonavir (5.59) and higher than that of darunavir (2.22). Having a logP value similar to an approved drug further supports canabisin A as a potential drug for SARS-CoV-2. In addition, the predicted logP value of isoacetoside is relatively small but is still on the positive side and should still be included in future study.

Potential Key Residue in SARS-CoV-2 Mpro−Ligand Binding. The potential important residues in SARS-CoV-2 Mpro−ligand binding could be probed via estimating the probabilities of SC and HB contacts between high-affinity inhibitors and individual residues of SARS-CoV-2 Mpro (Figure 8). Particularly, the residues His41, Met49, Leu141, Asn142, Ser144, Cys145, His164, Met165, Glu166, Leu167, and Gln189 formed SC contacts to the inhibitors in more than 80% of MD equilibrium snapshots. Interestingly, the inhibitors only adopt HB contacts to Glu166 over 81% of MD equilibrium snapshots. The mutation of this residue could possibly alter the binding affinity of a ligand to SARS-CoV-2 Mpro. To test this hypothesis, we have replaced Glu166 with an alanine residue and carried FEP calculation for the E166A SARS-CoV-2 Mpro + darunavir complex (Table 1). Upon E166A mutation, the calculated binding affinity of darunavir to SARS-CoV-2 Mpro changes from −11.96 ± 1.99 to −9.90 ± 2.48 kcal/mol. This ∼2

https://dx.doi.org/10.1021/acs.jcim.0c00491
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

kcal/mol decrease in binding affinity mostly arises from the weakening in coulomb interaction (ΔG_{cou}) due to the loss of HBs formed by the residue 166. This result suggests that G166 is potentially a key residue in ligand binding to SARS-CoV-2, and its mutation to a hydrophobic residue could lead to decreases in the inhibitory effect of ligands.

CONCLUSIONS

In this work, we utilized a rigorous computational approach to determine potential inhibitors of SARS-CoV-2 Mpro. First, we tested our approach on three recently reported inhibitors of SARS-CoV-2 Mpro and obtained computational results consistent with the experimental data. Subsequently, we investigated a database of 4600 natural compounds found in Vietnamese plants (Vietherbs). Eight HIV-1 PR inhibitors and an aza-peptide epoxide were added to the database. The database was first short-listed to 44 by molecular docking, which was further refined using FPL simulations to 5 compounds. The refined compounds were validated with the FEP method, the most accurate binding free energy estimation method. We found that two two natural compounds, cannabisin A and isoacteoside, in the Vietherbs database and an HIV-1 PR inhibitor, darunavir, can potentially inhibit SARS-CoV-2 Mpro, since their affinities are significantly larger than that of compound **13b**, the most reliable SARS-CoV-2 inhibitor from the recent work.[14] Moreover, the other HIV-1 PR inhibitors are unlikely to prevent SARS-CoV-2 Mpro, consistent with a recent clinical study.[63] Furthermore, we also found that the residue Glu166 possibly plays an important role in the binding of a ligand to SARS-CoV-2 Mpro, which could be a target for future study.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.0c00491.

List of top-lead inhibitors for SARS-CoV-2 Mpro obtained via molecular docking; two-dimensional protein−ligand interaction diagram of top-lead compounds; superposition of modeled and experimental SARS-CoV-2 Mpro; superposition of experimental and docked conformations of compound **13b**; FPL results; all-atom RMSD of SARS-CoV-2 Mpro + inhibitor systems; binding pose between SARS-CoV-2 Mpro + isoacteoside/quercetin 7-*O*-rutinoside/ritonavir obtained via MD simulations; and a docking list of all compounds to SARS-CoV-2 Mpro (PDF)

Videos S1 and S2 describing the dynamics of SARS-CoV-2 Mpro + **13b** or cannabisin A during MD simulations (MPG)

AUTHOR INFORMATION

Corresponding Authors

Son Tung Ngo − *Laboratory of Theoretical and Computational Biophysics and Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City 700000, Vietnam;* orcid.org/0000-0003-1034-1768; Email: ngosontung@tdtu.edu.vn

Duc-Hung Pham − *Division of Immunobiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio 45229, United States*; Email: duchung.pham@cchmc.org

Van V. Vu − *NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City 700000, Vietnam;* orcid.org/0000-0003-0009-6703; Email: vanvu@ntt.edu.vn

Authors

Ngoc Quynh Anh Pham − *Faculty of Chemical Engineering, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City 700000, Vietnam*

Ly Thi Le − *School of Biotechnology, International University, Ho Chi Minh Ciy 700000, Vietnam*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.0c00491

Author Contributions

S.T.N., D.-H.P., and V.V.V. designed the studies and wrote the manuscript. S.T.N. and N.Q.A.P. performed and analyzed data. L.T.L. collected the structures of natural compounds.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by Vietnam National Foundation for Science & Technology Development (NAFOSTED), grant no. 104.99-2019.57.

REFERENCES

(1) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271−280.e8.

(2) Fehr, A. R.; Channappanavar, R.; Perlman, S. Middle East Respiratory Syndrome: Emergence of a Pathogenic Human Coronavirus. *Annu. Rev. Med.* **2017**, *68*, 387−399.

(3) de Wit, E.; van Doremalen, N.; Falzarano, D.; Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **2016**, *14*, 523−534.

(4) Lau, S. K. P.; Woo, P. C. Y.; Li, K. S. M.; Huang, Y.; Tsoi, H.-W.; Wong, B. H. L.; Wong, S. S. Y.; Leung, S.-Y.; Chan, K.-H.; Yuen, K.-Y. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 14040−14045.

(5) Guan, Y.; Zheng, B. J.; He, Y. Q.; Liu, X. L.; Zhuang, Z. X.; Cheung, C. L.; Luo, S. W.; Li, P. H.; Zhang, L. J.; Guan, Y. J.; Butt, K. M.; Wong, K. L.; Chan, K. W.; Lim, W.; Shortridge, K. F.; Yuen, K. Y.; Peiris, J. S. M.; Poon, L. L. M. Isolation and Characterization of Viruses Related to the SARS Coronavirus from Animals in Southern China. *Science* **2003**, *302*, 276.

(6) Huang, C. L.; Wang, Y. M.; Li, X. W.; Ren, L. L.; Zhao, J. P.; Hu, Y.; Zhang, L.; Fan, G. H.; Xu, J. Y.; Gu, X. Y.; Cheng, Z. S.; Yu, T.; Xia, J. A.; Wei, Y.; Wu, W. J.; Xie, X. L.; Yin, W.; Li, H.; Liu, M.; Xiao, Y.; Gao, H.; Guo, L.; Xie, J. G.; Wang, G. F.; Jiang, R. M.; Gao, Z. C.; Jin, Q.; Wang, J. W.; Cao, B. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497−506.

(7) Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, *395*, 470−473.

(8) Yu Wai, C.; Chin-Pang, Y.; Kwok-Yin, W. Prediction of the SARS-CoV-2 (2019-nCoV) 3C-like Protease (3CLpro) Structure: Virtual Screening Reveals Velpatasvir, Ledipasvir, and Other Drug Repurposing Candidates. *F1000Research* **2020**, *9*, 129.

(9) Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; Chen, H.-D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R.-D.; Liu, M.-Q.; Chen, Y.; Shen, X.-R.; Wang, X.; Zheng, X.-S.; Zhao, K.; Chen, Q.-J.; Deng, F.; Liu, L.-L.; Yan, B.; Zhan, F.-X.; Wang, Y.-Y.; Xiao, G.-F.; Shi, Z.-L. A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin. *Nature* **2020**, *579*, 270−273.

(10) Chan, J. F. W.; Yuan, S. F.; Kok, K. H.; To, K. K. W.; Chu, H.; Yang, J.; Xing, F. F.; Liu, J. L.; Yip, C. C. Y.; Poon, R. W. S.; Tsoi, H. W.; Lo, S. K. F.; Chan, K. H.; Poon, V. K. M.; Chan, W. M.; Ip, J. D.; Cai, J.

P.; Cheng, V. C. C.; Chen, H. L.; Hui, C. K. M.; Yuen, K. Y. A Familial Cluster of Pneumonia Associated with the 2019 Novel Coronavirus Indicating Person-to-Person Transmission: a Study of a Family Cluster. *Lancet* **2020**, *395*, 514−523.
(11) WHO Coronavirus disease 2019 (COVID-19) Situation Report - 52.
(12) Schoeman, D.; Fielding, B. C. Coronavirus envelope protein: current knowledge. *Virol. J.* **2019**, *16*, 69.
(13) Fauquet, C. M.; Fargette, D. International Committee on Taxonomy of Viruses and the 3,142 unassigned species. *Virol. J.* **2005**, *2*, 64.
(14) 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.
(15) Nukoolkarn, V.; Lee, V. S.; Malaisree, M.; Aruksakulwong, O.; Hannongbua, S. Molecular Dynamic Simulations Analysis of Ritronavir and Lopinavir as SARS-CoV 3CLpro Inhibitors. *J. Theor. Biol.* **2008**, *254*, 861−867.
(16) Efficacy and Safety of Darunavir and Cobicistat for Treatment of Pneumonia Caused by 2019-nCoV (DACO-nCoV). https://clinicaltrials.gov/ct2/show/NCT04252274 (accessed Mar 8, 2020).
(17) A Randomized,Open,Controlled Clinical Study to Evaluate the Efficacy of ASC09F and Ritonavir for 2019-nCoV Pneumonia. https://clinicaltrials.gov/ct2/show/NCT04261270 (accessed Mar 8, 2020).
(18) Marshall, G. R. Computer-Aided Drug Design. *Annu. Rev. Pharmacol. Toxicol.* **1987**, *27*, 193−213.
(19) Yu, W.; MacKerell, A. D. Computer-Aided Drug Design Methods. In *Antibiotics: Methods and Protocols*; Sass, P., Ed.; Springer: New York, 2017; pp 85−106.
(20) Ryde, U.; Soderhjelm, P. Ligand-Binding Affinity Estimates Supported by Quantum-Mechanical Methods. *Chem. Rev.* **2016**, *116*, 5520−5566.
(21) Gehlhaar, D. K.; Verkhivker, G.; Rejto, P. A.; Fogel, D. B.; Fogel, L. J.; Freer, S. T. Docking Conformationally Flexible Small Molecules into a Protein Binding Site through Evolutionary Programming. In Proceedings of the Fourth International Conference on Evolutionary Programming: 1−3 March 1995; San Diego; McDonnell, J. R., Reynolds, R. G., Fogel, D. B., Eds.; MIT Press: 1995.
(22) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(23) Kuhn, B.; Kollman, P. A. Binding of a siverse set of ligands to avidin and streptavidin: an accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. *J. Med. Chem.* **2000**, *43*, 3786−3791.
(24) Wang, W.; Kollman, P. A. Computational study of protein specificity: the molecular basis of HIV-1 protease drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 14937−14942.
(25) Aqvist, J.; Medina, C.; Samuelsson, J.-E. A New Method for Predicting Binding Affinity in Computer-Aidded Drug Design. *Protein Eng., Des. Sel.* **1994**, *7*, 385−391.
(26) Jones-Hertzog, D. K.; Jorgensen, W. L. Binding Affinities for Sulfonamide Inhibitors with Human Thrombin Using Monte Carlo Simulations with a Linear Response Method. *J. Med. Chem.* **1997**, *40*, 1539−1549.
(27) Ngo, S. T.; Hung, H. M.; Nguyen, M. T. Fast and Accurate Determination of the Relative Binding Affinities of Small Compounds to HIV-1 Protease using Non-Equilibrium Work. *J. Comput. Chem.* **2016**, *37*, 2734−2742.
(28) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(29) Beveridge, D. L.; DiCapua, F. M. Free energy via molecular simulation: applications to chemical and biomolecular systems. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431−492.
(30) Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3*, 300−313.
(31) Kollman, P. Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395−2417.
(32) Jarzynski, C. Equilibrium free-energy differences from non-equilibrium measurements: A master-equation approach. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1997**, *56*, 5018−5035.
(33) Jiang, W.; Roux, B. Free Energy Perturbation Hamiltonian Replica-Exchange Molecular Dynamics (FEP/H-REMD) for Absolute Ligand Binding Free Energy Calculations. *J. Chem. Theory Comput.* **2010**, *6*, 2559−2565.
(34) Meng, Y.; Sabri Dashti, D.; Roitberg, A. E. Computing Alchemical Free Energy Differences with Hamiltonian Replica Exchange Molecular Dynamics (H-REMD) Simulations. *J. Chem. Theory Comput.* **2011**, *7*, 2721−2727.
(35) Jiang, W.; Thirman, J.; Jo, S.; Roux, B. Reduced Free Energy Perturbation/Hamiltonian Replica Exchange Molecular Dynamics Method with Unbiased Alchemical Thermodynamic Axis. *J. Phys. Chem. B* **2018**, *122*, 9435−9442.
(36) Ngo, S. T.; Nguyen, T. H.; Tung, N. T.; Nam, P. C.; Vu, K. B.; Vu, V. V. Oversampling Free Energy Perturbation Simulation in Determination of the Ligand-Binding Free Energy. *J. Comput. Chem.* **2020**, *41*, 611−618.
(37) Christian C, G.; Georg, S. Wuhan coronavirus 2019-nCoV − what we can find out on a structural bioinformatics level 2020. https://figshare.com/articles/innophore-com-2019-ncov_2020-01-29_pdf/11752749.
(38) *The PyMOL Molecular Graphics System*, version 1.3r1; Schrödinger, LLC.
(39) Trott, O.; Olson, A. J. Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455−461.
(40) Nguyen, N. T.; Nguyen, T. H.; Pham, T. N. H.; Huy, N. T.; Bay, M. V.; Pham, M. Q.; Nam, P. C.; Vu, V. V.; Ngo, S. T. Autodock Vina Adopts More Accurate Binding Poses but Autodock4 Forms Better Binding Affinity. *J. Chem. Inf. Model.* **2020**, *60*, 204−211.
(41) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785−2791.
(42) Gasteiger, J.; Marsili, M. New Model for Calculating Atomic Charges in Molecules. *Tetrahedron Lett.* **1978**, *19*, 3181.
(43) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity—A Rapid Cccess to Atomic Charges. *Tetrahedron* **1980**, *36*, 3219−3228.
(44) Forli, S.; Huey, R.; Pique, M. E.; Sanner, M. F.; Goodsell, D. S.; Olson, A. J. Computational protein−ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **2016**, *11*, 905.
(45) Lee, T.-W.; Cherney, M. M.; Huitema, C.; Liu, J.; James, K. E.; Powers, J. C.; Eltis, L. D.; James, M. N. G. Crystal Structures of the Main Peptidase from the SARS Coronavirus Inhibited by a Substrate-like Aza-peptide Epoxide. *J. Mol. Biol.* **2005**, *353*, 1137−1151.
(46) 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**, *1−2*, 19−25.
(47) Aliev, A. E.; Kulke, M.; Khaneja, H. S.; Chudasama, V.; Sheppard, T. D.; Lanigan, R. M. Motional Timescale Predictions by Molecular Dynamics Simulations: Case Study using Proline and Hydroxyproline Sidechain Dynamics. *Proteins: Struct., Funct., Genet.* **2014**, *82*, 195−215.
(48) Wang, Y.; Xiao, J.; Suzek, T. O.; Zhang, J.; Wang, J.; Zhou, Z.; Han, L.; Karapetyan, K.; Dracheva, S.; Shoemaker, B. A.; Bolton, E.; Gindulyte, A.; Bryant, S. H. PubChem's BioAssay Database. *Nucleic Acids Res.* **2012**, *40*, D400−D412.
(49) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.
(50) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson,

https://dx.doi.org/10.1021/acs.jcim.0c00491
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER 18*; University of California: San Francisco, CA, 2018.

(51) Sousa da Silva, A. W.; Vranken, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes* **2012**, *5*, 367.

(52) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

(53) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(54) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.

(55) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **1976**, *22*, 245−268.

(56) Laskowski, R. A.; Swindells, M. B. LigPlot+: Multiple Ligand−Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model.* **2011**, *51*, 2778−2786.

(57) Lee, S. K.; Lee, I. H.; Kim, H. J.; Chang, G. S.; Chung, J. E.; No, K. T. "The PreADME approach: Web-based program for rapid prediction of physico-chemical, drug absorption and drug-like properties", EuroQSAR 2002 Designing Drugs and Crop Protectants: Processes, Problems and Solutions; Blackwell Publishing: Maldenh, MA, 2003; pp 418−420. DOI: 10.1002/cphc.201701384.

(58) Nguyen-Vo, T.-H.; Le, T.; Pham, D.; Nguyen, T.; Le, P.; Nguyen, A.; Nguyen, T.; Nguyen, T.-N.; Nguyen, V.; Do, H.; Trinh, K.; Duong, H. T.; Le, L. VIETHERB: A Database for Vietnamese Herbal Species. *J. Chem. Inf. Model.* **2019**, *59*, 1−9.

(59) Mai, B. K.; Viet, M. H.; Li, M. S. Top-Leads for Swine Influenza A/H1N1 Virus Revealed by Steered Molecular Dynamics Approach. *J. Chem. Inf. Model.* **2010**, *50*, 2236−2247.

(60) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.

(61) Park, S.; Schulten, K. Calculating Potentials of Mean Force from Steered Molecular Dynamics Simulations. *J. Chem. Phys.* **2004**, *120*, 5946−5961.

(62) Zhavoronkov, A.; Aladinskiy, V.; Zhebrak, A.; Zagribelnyy, B.; Terentiev, V.; Bezrukov, D. S.; Polykovskiy, D.; Shayakhmetov, R.; Filimonov, A.; Orekhov, P.; Yan, Y.; Popova, O.; Vanhaelen, Q.; Aliper, A.; Ivanenkov, Y. Potential COVID-2019 3C-like Protease Inhibitors Designed Using Generative Deep Learning Approaches. 2020. ChemRxiv. https://doi.org/10.26434/chemrxiv.11829102.v2.

(63) Cao, B.; Wang, Y.; Wen, D.; Liu, W.; Wang, J.; Fan, G.; Ruan, L.; Song, B.; Cai, Y.; Wei, M.; Li, X.; Xia, J.; Chen, N.; Xiang, J.; Yu, T.; Bai, T.; Xie, X.; Zhang, L.; Li, C.; Yuan, Y.; Chen, H.; Li, H.; Huang, H.; Tu, S.; Gong, F.; Liu, Y.; Wei, Y.; Dong, C.; Zhou, F.; Gu, X.; Xu, J.; Liu, Z.; Zhang, Y.; Li, H.; Shang, L.; Wang, K.; Li, K.; Zhou, X.; Dong, X.; Qu, Z.; Lu, S.; Hu, X.; Ruan, S.; Luo, S.; Wu, J.; Peng, L.; Cheng, F.; Pan, L.; Zou, J.; Jia, C.; Wang, J.; Liu, X.; Wang, S.; Wu, X.; Ge, Q.; He, J.; Zhan, H.; Qiu, F.; Guo, L.; Huang, C.; Jaki, T.; Hayden, F. G.; Horby, P. W.; Zhang, D.; Wang, C. A Trial of Lopinavir−Ritonavir in Adults Hospitalized with Severe Covid-19. *N. Engl. J. Med.* **2020**, *382*, 1787.

(64) Homeyer, N.; Stoll, F.; Hillisch, A.; Gohlke, H. Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. *J. Chem. Theory Comput.* **2014**, *10*, 3331−3344.

(65) Nascimento, É. C. M.; Oliva, M.; Świderek, K.; Martins, J. B. L.; Andrés, J. Binding Analysis of Some Classical Acetylcholinesterase Inhibitors: Insights for a Rational Design Using Free Energy Perturbation Method Calculations with QM/MM MD Simulations. *J. Chem. Inf. Model.* **2017**, *57*, 958−976.

(66) Ngo, S. T.; Fang, S.-T.; Huang, S.-H.; Chou, C.-L.; Huy, P. D. Q.; Li, M. S.; Chen, Y.-C. Anti-Arrhythmic Medication Propafenone a Potential Drug for Alzheimer's Disease Inhibiting Aggregation of Aβ: *In Silico* and *In Vitro* Studies. *J. Chem. Inf. Model.* **2016**, *56*, 1344−1356.

(67) Ciordia, M.; Pérez-Benito, L.; Delgado, F.; Trabanco, A. A.; Tresadern, G. Application of Free Energy Perturbation for the Design of BACE1 Inhibitors. *J. Chem. Inf. Model.* **2016**, *56*, 1856−1871.

(68) Tran, P.-T.; Hoang, V.-H.; Lee, J.; Hien, T. T. T.; Tung, N. T.; Ngo, S. T. In vitro and in silico determination of glutaminyl cyclase inhibitors. *RSC Adv.* **2019**, *9*, 29619−29627.

(69) Sakakibara, I.; Katsuhara, T.; Ikeya, Y.; Hayashi, K.; Mitsuhashi, H. Cannabisin A, an arylnaphthalene lignanamide from fruits of Cannabis sativa. *Phytochemistry* **1991**, *30*, 3013−3016.

https://dx.doi.org/10.1021/acs.jcim.0c00491
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX