

Binding of Inhibitors to the Monomeric and Dimeric SARS-CoV-2 Mpro

Nguyen Minh Tam,^{ab} Pham Cam Nam,^c Duong Tuan Quang,^d Nguyen Thanh Tung,^{ef} and Van V. Vu,^g and Son Tung Ngo^{bh*}

^aComputational Chemistry Research Group, Ton Duc Thang University, Ho Chi Minh City, Vietnam

^bFaculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam

^cDepartment of Chemistry, The University of Danang, University of Science and Technology, Danang, Vietnam

^dUniversity of Education, Hue University, Vietnam

^eInstitute of Materials Science, Vietnam Academy of Science and Technology, Hanoi, Vietnam

^fGraduate University of Science and Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam

^gNTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam

^hLaboratory of Theoretical and Computational Biophysics, Ton Duc Thang University, Ho Chi Minh City, Vietnam

ABSTRACT: SARS-CoV-2 rapidly infects millions of people worldwide since December 2019. There is still no effective treatment for the virus, resulting in the death of more than one million of patients. Inhibiting the activity of SARS-CoV-2 main protease (Mpro), 3C-like protease (3CLP), is able to block the viral replication and proliferation. Although the dimer was shown to be the biologically active form of the SARS-CoV-2 Mpro, in this context, our study has revealed that *in silico* screening for inhibitors of SARS-CoV-2 Mpro can be reliably done using the monomeric structure of the receptor. Docking and fast pulling of ligand (FPL) simulations for both monomeric and dimeric forms correlate well with the corresponding experimental binding affinity data of 30 compounds. In particular, the correlation coefficients between computational and experimental binding free energy of the monomeric SARS-CoV-2 Mpro are $R_{\text{Dock}}^{\text{Monomer}} = 0.59 \pm 0.11$ and $R_{\text{Work}}^{\text{Monomer}} = -0.66 \pm 0.08$. The metrics are approximately similar to the dimeric target with the coefficients of $R_{\text{Dock}}^{\text{Dimer}} = 0.52 \pm 0.10$ and $R_{\text{Work}}^{\text{Dimer}} = -0.70 \pm 0.09$. Moreover, the correlation coefficient between the rupture forces to binding free energy are roughly the same since $R_{\text{Force}}^{\text{Monomer}} = -0.64 \pm 0.08$ and $R_{\text{Force}}^{\text{Dimer}} = -0.63 \pm 0.10$. Furthermore, the correlation coefficient between calculated metrics of the monomeric and dimeric SARS-CoV-2 Mpro is $R_{\text{Monomer}}^{\text{Dimer}} = 0.74 \pm 0.09$. Our study results show that it is possible to speed up computer-aided drug design for SARS-CoV-2 Mpro by focusing on the monomeric form instead of the larger dimeric one.

virus initially appeared the first case since December 2019 in Wuhan, Hubei province, China.²⁻⁴ It shares more than 82% identical RNA genome to the SARS-CoV, SARS-CoV-2 severe cases of respiratory syndromes.⁵ Although the bat has been thought of as the original reservoir, the intermediate host is still unknown.⁶ Moreover, it is known that the SARS-CoV-2 can endure in aerosol for more than 3 hours,⁷ which may be a major factor behind the outbreak of COVID-19 pandemic, which has caused several hundred thousands of deaths worldwide.⁵ Therefore, the COVID-19 pandemic becomes an urgency for community health, which requires to develop an effective treatment or vaccine immediately.

Coronaviruses genomes occupy ca. 26-32 kb in length that is the largest sequence among RNA viruses.^{8, 9} The SARS-CoV-2 genome encodes more than 20 various structural and non-structural proteins. Particularly, the SARS-CoV-2 main protease (Mpro), 3C-like protease (3CLP), is one of the most important viral enzymes, having more than 96% similarity with SARS-CoV 3CLP.^{9, 10} SARS-CoV-2 Mpro cleaves nascent polyproteins, which are generated by the translation of the viral RNA. During this process, 11 non-structural polyproteins are auto-cleaved to become polypeptides, which are required for the viral replication and transcription.⁹ Therefore, SARS-CoV-2 Mpro turns out to be an attractive target for antiviral drug aiming since blocking viral protease can inhibit viral replication and proliferation.^{10, 11} Numerous investigations following this strategy have been carried out and shown some initial success.¹²⁻¹⁸ However, unfortunately, an effective drug for COVID-19 is still unavailable until the date.

In addition, currently, it should be noted that the time and cost to develop a drug has been significantly decreased by using the power of computational approaches.¹⁹⁻²² Normally, the binding free energy ΔG between a ligand and an enzyme

INTRODUCTION

The novel coronavirus (2019-nCoV or SARS-CoV-2), a member of the Coronaviridae virus family, has been reported to be able to transmit from human to human.¹ The

can be probed via computational approaches. The ΔG is associated with experimental inhibition constant k_i via formula $\Delta G_{\text{bind}} = RT \ln(k_i)$, where R is gas constant, T is absolute temperature, and k_i is a critical metric revealing the nature of binding between two biomolecules.¹⁹ Accurate determination of the ligand binding free energy is very important in computer-aided drug design (CADD) problem.²³

Moreover, it should be noted that the dimer was shown to be the biologically active form of the SARS-CoV-2 Mpro but the interface does not contain a ligand-binding pocket.¹² An important question which arises is that can we use monomeric form of SARS-CoV-2 Mpro as inhibitor-screening target instead of the dimeric one to reduce CPU time consumption? Therefore, in this context, the binding free energy of 30 available inhibitors¹²⁻¹⁸ to the monomeric and dimeric SARS-CoV-2 Mpro was examined via docking and FPL schemes. The similar of correlation coefficients between computational and experimental values of monomeric and dimeric systems suggests that we can use the monomeric form of SARS-CoV-2 Mpro as CADD target instead of the dimeric form. The obtained results can be beneficial to the COVID-19 therapy by speeding up CADD progression.

MATERIALS & METHODS

Structure of Ligands and SARS-CoV-2 Mpro

Three-dimensional structures of the monomeric and dimeric SARS-CoV-2 Mpro were copied from the Protein Data Bank with ID 6Y2F¹² and 6XBG,²⁴ respectively. Ligand structures were taken from PubChem database.²⁵ The protonation states of the ligands were predicted by using chemicalize webserver, an online tool of ChemAxon. The ligand structure was first optimized using quantum chemical calculation with the B3LYP functional at 6-31G(d) level of basis set.

Molecular Docking Simulations

The binding position and affinity of ligands to the monomeric and dimeric SARS-CoV-2 Mpro were predicted via the Autodock Vina package (cf. **Figure 1**).²⁶ The docking parameter was selected referring to the previous study,²⁷⁻²⁹ in which the exhaustiveness is of 8. The best docking result was chosen as the highest binding affinity conformations. The grid center was selected as the center of mass of α -ketoamide **13b** and UAW246 compounds, which correspond to the monomeric and dimeric Mpro, respectively.^{12, 24} The grid size was chosen as $24 \times 24 \times 24$ Å, which is large enough to cover the ligand-binding cleft of the Mpro.^{28, 29}

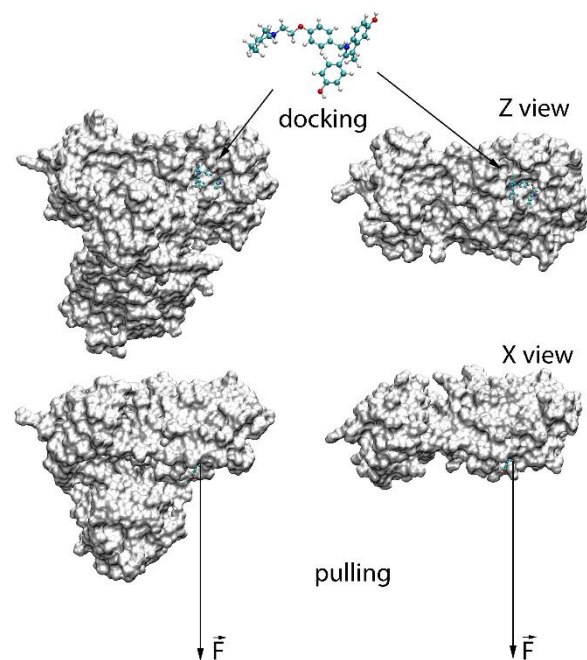


Figure 1. Computational scheme for evaluation of the ligand-binding affinity to the monomeric and dimeric SARS-CoV-2 Mpro.

Steered-Molecular Dynamics Simulations

GROMACS version 5.1.3³⁰ was employed to simulate the solvated complex involving the ligand and monomeric/dimeric SARS-CoV-2 Mpro. The Amber99SB-ILDN force field was utilized for the SARS-CoV-2 Mpro.³¹ The general Amber force field (GAFF) was employed to parameterize the ligand by using AmberTools18 and ACPYPE approaches.^{32, 33} In particular, atomic charges of the ligand were assigned via the Restrained Electrostatic Potential (RESP) method³⁴ through quantum chemical calculation using B3LYP functional at 6-31G(d,p) level of theory. The quantum chemical calculation was performed using implicit water model (with the dielectric constant of $\epsilon = 78.4$). The monomeric and dimeric SARS-CoV-2 Mpro + inhibitor were inserted into a rectangular periodic boundary (PBC) condition box with a size of (9.83, 5.92, 8.70) and (9.39, 8.96, 12.05) nm, respectively. The corresponding box volumes of the monomeric and dimeric systems are 506.28 and 1013.82 nm³, respectively. Therefore, the total atoms of these systems approximately are 50 000 and 100 000 atoms, respectively.

The MD simulations were carried out with the parameters denoted to the previous works.^{28, 29} Particularly, the MD time step is 2 fs. The nonbonded cutoff was set to 0.9 nm. The Coulomb interaction was computed using the fast Particle-Mesh Ewald electrostatics scheme.³⁵ The solvated system was then minimized and relaxed over EM, NVT, and NPT simulations. The NVT and NPT simulations were length of 0.1 and 2.0 ns, respectively. During NVT and NPT simulation, the SARS-CoV-2 Mpro C_{α} atoms were restrained via a small harmonic force with a value of 1000 kJ mol⁻¹ nm⁻² per dimensions. The relaxed conformation of the SARS-CoV-2 Mpro + inhibitor was then employed as initial structure of FPL simulation. During which, the inhibitor was pulled out of the binding cleft under effect of an externally harmonic force with parameters of $k = 0.005$ nm ps⁻¹ and $v = 600$ kJ mol⁻¹ nm⁻² for pulling speed and cantilever spring constant

(cf. **Figure 1**), respectively.^{29, 36} Totally, 8 independent trajectories were carried out to estimate the ligand-binding affinity.

RESULTS AND DISCUSSION

The binding pose and affinity of the trial inhibitors to the monomeric and dimeric SARS-CoV-2 Mpro were initially estimated by a molecular docking method. Autodock Vina,²⁶ a very efficient molecular docking approach with a successful-docking rate up to 81 %, ²⁷ would be able to complete this task. We have thus docked 30 available inhibitors to the monomeric and dimeric SARS-CoV-2 Mpro using Autodock Vina referring to the previous study.^{28, 29} By using exhaustiveness 8 as suggested in the previous work,²⁷ the results were rapidly obtained in few hours (**Table 1** and **Table S1** of the Supplementary – SI file). Interestingly, the correlation coefficient between docking and experimental affinities of the monomeric target, $R_{\text{Dock}}^{\text{Monomer}} = 0.59 \pm 0.11$, is slightly larger than that of the dimeric target, $R_{\text{Dock}}^{\text{Dimer}} = 0.52 \pm 0.10$ (cf. **Figure 2**). The root mean square error (RMSE) between calculated and experimental values also indicates good consistency between monomer and dimer docking results. In particular, the monomeric system gives a value of $RMSE_{\text{Dock}}^{\text{Monomer}} = 1.17 \pm 0.15$ kcal mol⁻¹ and the dimeric system adopts a metric of $RMSE_{\text{Dock}}^{\text{Dimer}} = 1.35 \pm 0.18$ kcal mol⁻¹. It should be noted that the computed error bars was obtained via 1000 rounds of the bootstrapping method.³⁷

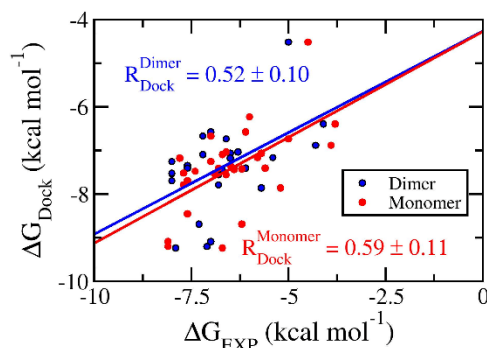


Figure 2. Correlation between docking and experimental binding free energy. Computational results were obtained using Autodock Vina. The experimental binding free energies were estimated using IC₅₀ value¹²⁻¹⁸ as an approximation for the inhibition constant k_i . The computed error was attained via 1000 rounds of the bootstrapping method.³⁷

Table 1. Computed values of docking energy in comparison with experiments.

Nº	Name	$\Delta G_{\text{Dock}}^{\text{Monomer}}$			$\Delta G_{\text{Dock}}^{\text{Dimer}}$			ΔG_{EXP}^a
		short	medium	long	short	medium	long	
1	11r	-6.7	-6.4	-6.3	-7.9	-8.1	-8.3	-9.23
2	13a	-7.6	-7.6	-7.6	-8.0	-7.8	-7.8	-7.70
3	13b	-7.6	-7.8	-7.8	-7.6	-7.1	-7.8	-8.45
4	Bazedoxifene	-7.4	-7.5	-7.4	-7.4	-7.4	-7.5	-7.48
5	Calpain inhibitor XII	-6.2	-6.3	-6.3	-7.3	-7.3	-7.2	-8.69
6	Carmofur	-5.2	-5.5	-5.6	-5.7	-5.8	-6.1	-7.86
7	Chloroquine	-5.0	-5.3	-5.1	-6.6	-6.6	-6.6	-6.74
8	Cyclosporine	-5.8	-5.7	-5.7	-5.4	-5.4	-5.4	-7.17
9	Digitoxin	-8.1	-8.1	-8.2	-7.0	-7.0	-7.2	-9.09
10	Digoxin	-8.1	-8.1	-8.1	-7.1	-7.2	-7.2	-9.20
11	Dihydrogambogic Acid	-7.0	-7.0	-7.0	-7.2	-7.2	-7.2	-6.67
12	Disulfiram	-3.9	-3.8	-3.9	-4.3	-4.1	-4.1	-6.89
13	Ebastine	-5.7	-6.5	-6.1	-6.5	-6.3	-6.4	-7.06
14	Favipiravir	-4.5	-4.8	-4.8	-5.0	-5.0	-5.0	-4.52
15	Fluspirilene	-6.9	-7.2	-7.3	-8.0	-7.7	-7.6	-7.53
16	Isoosajin	-7.7	-7.7	-7.7	-8.0	-8.0	-8.0	-7.52
17	Ivacaftor	-6.7	-6.7	-6.7	-7.2	-7.6	-7.5	-7.10
18	Lusutrombopag	-6.2	-6.1	-6.8	-6.4	-6.5	-6.3	-7.42
19	Mefloquine	-6.5	-6.5	-6.5	-7.6	-7.7	-7.6	-7.34
20	Mequitazine	-6.6	-6.6	-6.6	-6.3	-6.3	-6.3	-7.03
21	MG-132	-5.6	-6.2	-6.2	-6.1	-5.8	-6.2	-7.41
22	Narlaprevir	-7.8	-7.5	-7.4	-6.5	-6.9	-6.8	-7.18
23	Osajin	-6.8	-6.9	-6.8	-7.6	-8.0	-8.0	-7.41
24	Oxyclozanide	-6.4	-6.4	-6.4	-6.7	-6.7	-6.7	-7.44
25	Penfluridol	-7.0	-6.9	-6.9	-8.0	-8.2	-8.2	-7.26
26	Phenazopyridine	-6.0	-6.0	-6.0	-6.0	-6.0	-6.0	-6.23
27	Proscillaridin	-7.7	-7.7	-7.7	-6.8	-7.3	-7.3	-7.79
28	PX-12	-3.8	-3.8	-3.8	-4.1	-4.2	-4.5	-6.39
29	Shikonin	-6.1	-6.1	-6.1	-7.0	-6.9	-6.9	-6.58
30	Tetrandrine	-6.6	-6.6	-6.6	-6.8	-6.8	-6.8	-7.56

^aThe experimental binding free energies were gained based on IC₅₀ value,¹²⁻¹⁸ approximating that the one equals to the inhibition constant k_i . The unit is of kcal mol⁻¹

The molecular docking with larger exhaustiveness, which selected as 56 and 400 according to the previous study,²⁷ were also performed in order to validate the convergence of the docking scheme. In total we used three different values of exhaustiveness including 400, 56, and 8 which are denoted as *long*, *medium*, and *short* options, respectively. The accuracies of the docking simulations for monomer and dimer with respect to experiment are shown in Figure 3 **Figure 3**. Interestingly, changing the docking exhaustiveness parameter from *short* to *medium* and/or *long* does not have a significant impact on the correlation coefficient and RMSE, which is consistent with the prior benchmark²⁷. In particular, the correlation coefficients slightly change to $R_{\text{Dock}}^{\text{Monomer}} = 0.57 \pm 0.11$ and $R_{\text{Dock}}^{\text{Dimer}} = 0.50 \pm 0.11$ matching with the *medium* option Figure 3A(cf. **Figure 3A**). The metrics are of $R_{\text{Dock}}^{\text{Monomer}} = 0.58 \pm 0.12$ and $R_{\text{Dock}}^{\text{Dimer}} = 0.55 \pm 0.10$ resembling the *long* option (**Figure 3A**). Moreover, the calculated accuracy is also associated with the RMSE value. Absolutely, within computed error, the RMSE was unchanged over the docking options *short*, *medium*, and *long* with amounts of $RMSE_{\text{Dock}}^{\text{Dimer}} = 1.17 \pm 0.15$, $RMSE_{\text{Dock}}^{\text{Dimer}} = 1.18 \pm 0.15$, and $RMSE_{\text{Dock}}^{\text{Dimer}} = 1.10 \pm 0.15$ kcal mol⁻¹ for dimeric systems and $RMSE_{\text{Dock}}^{\text{Monomer}} = 1.35 \pm 0.18$, $RMSE_{\text{Dock}}^{\text{Monomer}} = 1.29 \pm 0.19$, and $RMSE_{\text{Dock}}^{\text{Monomer}} = 1.28 \pm 0.19$ kcal mol⁻¹ for monomeric systems, respectively (cf. **Figure 3B**). Overall, the docking simulations provide slightly accurate results for monomeric systems than for the dimeric systems.

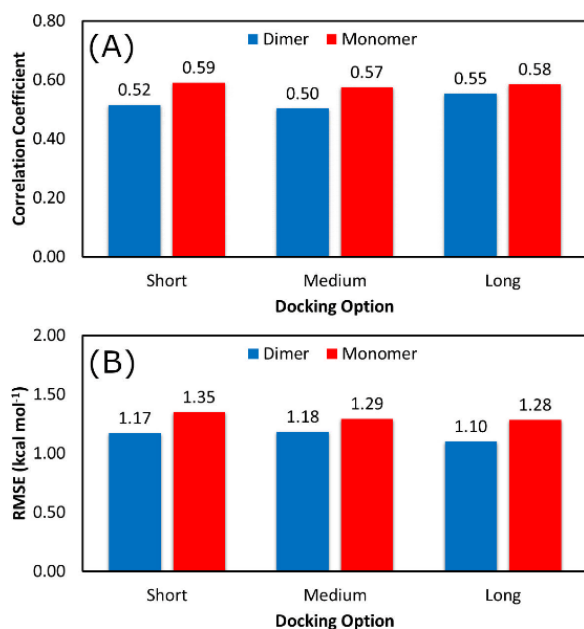


Figure 3. Correlation and RMSE values between calculated and experimental binding affinity.

As mentioned above, the binding affinity of 30 available inhibitors¹²⁻¹⁸ to the monomeric and dimeric SARS-CoV-2 Mpro was appropriately probed using molecular docking calculations. However, it should be noted that the dynamics of receptors were not considered in docking simulations, and the number of trial docking poses was small. To overcome this limitation we have performed MD

simulations which serve as a validation for the docking results³⁸⁻⁴⁰. Moreover, FPL is an efficient computational approach to assess ligand-binding affinity with a suitable time-consuming calculation.^{41, 42} Furthermore, the scheme was successfully applied to the monomeric SARS-CoV-2 Mpro system recently.^{28, 29} The FPL approach is thus used to probe the binding affinity of 30 available inhibitors to the monomeric and dimeric SARS-CoV-2 Mpro. In the simulations, the ligand binding pose was optimized over short canonical and isothermal-isobaric simulations. The equilibrated ligand was then pulled to translocate from *bound* to *unbound* states. The maximum of pulling force, called rupture force, and pulling work are typically assumed to correlate with ligand-binding affinity. It should be noted that the rupture force corresponds to the point that the non-covalent bond between a ligand and a receptor was terminated.

The computed values of the rupture force and pulling work were shown in **Table 2**. The denoted pulling force and work profiles were described in **Tables S2** and **S3** of the SI file. The shape of both pulling force and work appear reliable when compared to the previous exertion.^{41, 42} In particular, starting at zero, the pulling force quickly increases to the maximum value, then suddenly drops to zero due to the loss of the non-covalent bond to the receptor. During this process, recorded-pulling work speedily rises from zero value to a stable value, corresponding to the distance at which the contact between protein and inhibitor is vanished. Moreover, the rupture force $F_{\text{Max}}^{\text{Monomer}}$ of monomeric Mpros diffuses in the range from 295.0 to 977.6 pN corresponding with the spreading of pulling work W^{Monomer} from 13.7 to 106.1 kcal mol⁻¹. Besides that, the matching metrics of dimeric Mpros forms in the range from 336.1 to 769.6 pN and 20.5 to 84.7 kcal mol⁻¹, correspondingly. It should be noted that the computed works are significantly larger than the magnitude of experimental binding affinity, which diffuses in the range from 4.52 to 9.23 kcal mol⁻¹, since applied large cantilever and high pulling velocity.⁴¹ Although the discrepancy can be reduced to zero by using a small cantilever and an extremely low pulling velocity, it is not appropriate since it requires to perform several trajectories with hundred nanoseconds each.⁴³ Furthermore, previous investigations revealed that although reducing the magnitude of cantilever spring constant and pulling velocity was able to enlarge the accuracy of the estimations, the observed results are approximately the equivalent as those at high pulling velocity.^{36, 41}

Table 2. Computed values of rupture force and pulling work in comparison with experiments.

Nº	Name	$F_{\text{Max}}^{\text{Monomer}}$	W^{Monomer}	$F_{\text{Max}}^{\text{Dimer}}$	W^{Dimer}	ΔG_{EXP}^a
1	11r	724.8 ± 57.7	77.6 ± 7.1	636.6 ± 28.2	71.5 ± 2.9	-9.23
2	13a	526.9 ± 56.4	54.4 ± 7.3	769.6 ± 16.3	84.7 ± 3.2	-7.70
3	13b	977.6 ± 33.7	106.1 ± 4.6	739.1 ± 28.4	81.6 ± 3.0	-8.45
4	Bazedoxifene	460.3 ± 26	41.2 ± 3.1	471.1 ± 20.0	47.5 ± 3.6	-7.48
5	Calpain inhibitor XII	491.6 ± 20.5	46 ± 2.3	693.6 ± 50.7	63.5 ± 4.8	-8.69
6	Carmofur	485.5 ± 34.2	36.2 ± 2.7	436.9 ± 16.3	33.6 ± 1.8	-7.86
7	Chloroquine	363.4 ± 32.1	28.5 ± 2.8	410.9 ± 12.5	36.0 ± 1.6	-6.74
8	Cyclosporine	638.8 ± 33.4	67.7 ± 5.4	426.5 ± 41.6	44.1 ± 4.7	-7.17
9	Digitoxin	667.4 ± 17.7	70.9 ± 2.1	502.6 ± 65	55.3 ± 8.3	-9.09
10	Digoxin	637.0 ± 30.3	75.0 ± 2.5	573.1 ± 42.3	59.4 ± 4.9	-9.20
11	Dihydrogambogic Acid	542.8 ± 37.7	59.6 ± 3.2	487.5 ± 29.9	44.0 ± 3.3	-6.67
12	Disulfiram	364.7 ± 24.7	22.7 ± 1.9	526.2 ± 30.3	40.1 ± 1.9	-6.89
13	Ebastine	447.5 ± 40.1	40.2 ± 3.5	389.8 ± 25.0	32.8 ± 2.8	-7.06
14	Favipiravir	364.9 ± 26.2	21.3 ± 2.9	336.1 ± 19.1	20.5 ± 2.5	-4.52
15	Fluspirilene	490.1 ± 23.6	43.8 ± 2.0	544.6 ± 36.3	58.0 ± 3.2	-7.53
16	Isoosajin	393.1 ± 32.8	28.9 ± 3.2	454.4 ± 19.7	40.4 ± 2.5	-7.52
17	Ivacaftor	347.9 ± 34.8	22.3 ± 4.4	477.5 ± 22.1	41.0 ± 2.1	-7.10
18	Lusutrombopag	540.6 ± 37.5	59.1 ± 3.7	396.8 ± 24.3	41.8 ± 2.2	-7.42
19	Mefloquine	523.7 ± 23.5	41.5 ± 2.3	509.6 ± 43.3	46.3 ± 3.3	-7.34
20	Mequitazine	392.5 ± 51.3	29.5 ± 4.0	384.9 ± 24.4	29.0 ± 2.2	-7.03
21	MG-132	543.2 ± 22.2	49.8 ± 2.1	505.7 ± 41.1	47.5 ± 6.0	-7.41
22	Narlaprevir	601.8 ± 31.9	64.8 ± 2.8	522.0 ± 38.3	54.7 ± 4.3	-7.18
23	Osajin	367.9 ± 20.4	30.8 ± 2.9	471.4 ± 23.9	39.8 ± 1.8	-7.41
24	Oxyclozanide	463.7 ± 31.7	33.6 ± 3.2	468.1 ± 13.3	39.2 ± 3.5	-7.44
25	Penfluridol	542.3 ± 33.1	53.3 ± 2.7	444.5 ± 25.0	48.0 ± 3.9	-7.26
26	Phenazopyridine	391.7 ± 36.2	25.6 ± 2.8	384.8 ± 22.7	32.4 ± 1.4	-6.23
27	Proscillaridin	485.6 ± 37.2	45.8 ± 3.3	512.8 ± 18.9	58.0 ± 1.6	-7.79
28	PX-12	295.0 ± 17.4	13.7 ± 1.2	382.0 ± 25.5	27.2 ± 2.0	-6.39
29	Shikonin	321.8 ± 29.7	19.7 ± 3.0	504.5 ± 22.8	39.1 ± 1.2	-6.58
30	Tetrandrine	485.6 ± 37.2	45.8 ± 3.3	401.5 ± 18.5	31.6 ± 1.8	-7.56

^aThe experimental binding free energies were gained based on IC₅₀ value,¹²⁻¹⁸ approximating that the one equals to the inhibition constant k_i . The unit of force and energy/work are in pN and kcal mol⁻¹, respectively.

In practice, the rupture force has been used as a predictor of ligand-binding affinity based on the assumption that a ligand binds with a higher affinity requires a stronger pulling force to dissociate it from binding cleft.⁴⁴ Using the rupture force as a proxy to ligand-binding affinity, numerous investigations were successful in predicting the ligand-binding affinity to various targets.^{44, 45} Here, the average of rupture forces were estimated over 8 independent FPL trajectories (cf. **Table 2**). The correlation coefficient, obtained results of monomeric systems, is $R_{\text{Force}}^{\text{Monomer}} = -0.64 \pm 0.08$; while the analogous value of dimeric forms is $R_{\text{Force}}^{\text{Dimer}} = -0.63 \pm 0.10$ as sketched in **Figure 4**. Clearly, the accuracy of the FPL technique is significantly larger than that of molecular docking calculation. Moreover, because the correlation coefficients appear to be the same within the

error range, we may conclude that there is no difference when using monomer or dimer as a CADD target.

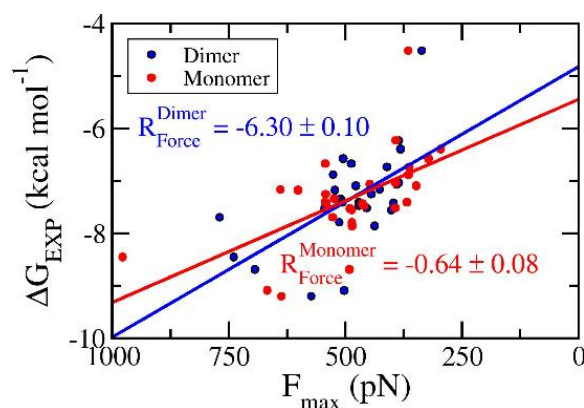


Figure 4. Relationship between rupture force and experimental binding free energy. Rupture forces were obtained via FPL calculations. The binding free energies were gained based on IC₅₀ value,¹²⁻¹⁸ approximating that the one equals to the inhibition constant k_i . The computed error was attained via 1000 rounds of the bootstrapping method.³⁷

The work of pulling force was assessed via formula $W = v \int_0^t F(t) dt$, where v is pulling velocity and $F(t)$ is pulling force. In isothermal-isobaric simulations, W is related to the experimental binding affinity via Jarzynski equality.⁴⁶ Therefore, utilizing W to estimate the ligand-binding affinity commonly acquires a better accurate result in comparison to rupture force.^{38, 41, 45} The obtained results reaffirmed this statement. The correlation coefficients of the monomeric and dimeric SARS-CoV-2 Mpro are $R_{\text{Work}}^{\text{Monomer}} = -0.66 \pm 0.09$ and $R_{\text{Work}}^{\text{Dimer}} = -0.70 \pm 0.09$ as shown in **Figure 5**, respectively. Although, the computational accuracy targeting the SARS-CoV-2 Mpro dimer is slightly larger than that of the monomeric system, the difference in correlation coefficients is small implying that the monomeric form of SARS-CoV-2 Mpro can be used as CADD target instead of the dimeric one.

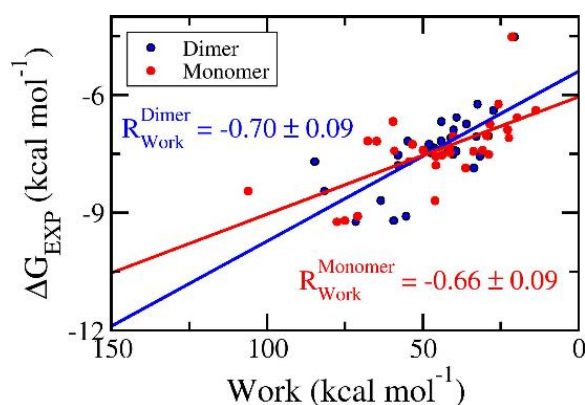


Figure 5. Association between pulling work and experimental binding free energy. Pulling works were obtained via FPL calculations. The binding free energies were gained based on IC₅₀ value,^{12–18} approximating that the one equals to the inhibition constant k_i . The computed error was attained via 1000 rounds of the bootstrapping method.³⁷

In addition, the association of computed pulling works of the monomeric and dimeric SARS-CoV-2 Mpro was probed and shown in **Figure 6**. Over the bootstrapping examination, the correlation coefficient is $R_{\text{Monomer}}^{\text{Dimer}} = 0.74 \pm 0.09$ confirming the observation above. We can manipulate the inhibitor screening for SARS-CoV-2 Mpro

with smaller computing resources since targeting the monomeric form.

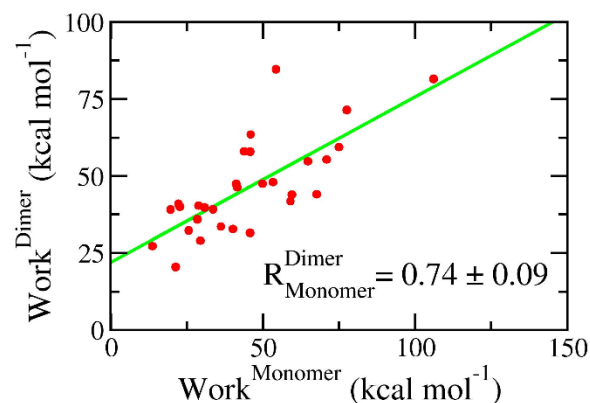


Figure 6. Association between calculated pulling work of the monomeric and dimeric SARS-CoV-2 Mpro. The computed error was attained via 1000 rounds of the bootstrapping method.³⁷

CONCLUSIONS

Both of Autodock Vina and FPL simulations were confirmed to be able to appropriately estimate the ligand-binding affinity of the SARS-CoV-2 Mpro in both monomeric and dimeric forms. The assessed results suggested that the monomeric form of SARS-CoV-2 Mpro can be used as a CADD target instead of the dimeric form. Moreover, in good agreement with the previous observation,²⁷ the molecular docking by Vina package rapidly converged since the correlation coefficient between computed and experimental values did not change when the docking option was altered. The RMSE of docking results also unchanged upon these alterations. Furthermore, it may be concluded that for SARS-CoV-2 Mpro system the pulling work is better than rupture force in predicting the ligand-binding affinity. It is well compatible with earlier probe various protein-ligand complexes.^{38, 41, 45}

ASSOCIATED CONTENT

Supporting Information

Supporting Information Available: The pulling forces and works over 8 independent trajectories of SMD simulations. The material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Email: ngosontung@tdtu.edu.vn

Author Contributions

All authors designed the studies, collected and analyzed the data, and wrote the manuscript

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

This work was supported by Vietnam National Foundation for Science & Technology Development (NAFOSTED) grant #104.99-2019.57. We thank Dr. Trung Hai Nguyen for helpful discussions and comments on the manuscript.

REFERENCES

- 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 (10223), 514-523.
- 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 (10223), 497-506.
- Wang, C.; Horby, P. W.; Hayden, F. G.; Gao, G. F., A novel coronavirus outbreak of global health concern. *Lancet* **2020**, 395 (10223), 470-473.
- 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. *F1000Res* **2020**, 9, 129.
- WHO Coronavirus disease 2019 (COVID-19) Situation Report - 52.
- 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 (7798), 270-273.
- van Doremalen, N.; Bushmaker, T.; Morris, D. H.; Holbrook, M. G.; Gamble, A.; Williamson, B. N.; Tamin, A.; Harcourt, J. L.; Thornburg, N. J.; Gerber, S. I.; Lloyd-Smith, J. O.; de Wit, E.; Munster, V. J., Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1. *N Engl J Med* **2020**, 382, 1564-1567.
- Schoeman, D.; Fielding, B. C., Coronavirus envelope protein: current knowledge. *Virology* **2019**, 16 (1), 69.
- Fauquet, C. M.; Fargette, D., International Committee on Taxonomy of Viruses and the 3,142 unassigned species. *Virology* **2005**, 2 (1), 64.
- Alex, Z.; Vladimir, A.; Alexander, Z.; Bogdan, Z.; Victor, T.; Dmitry S., B.; Daniil, P.; Rim, S.; Andrey, F.; Philipp, O.; Yilin, Y.; Olga, P.; Quentin, V.; Alex, A.; Yan, I., *Potential COVID-2019 3C-like Protease Inhibitors Designed Using Generative Deep Learning Approaches*. 2020.
- 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 (4), 861-867.
- 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.
- Jin, Z.; Zhao, Y.; Sun, Y.; Zhang, B.; Wang, H.; Wu, Y.; Zhu, Y.; Zhu, C.; Hu, T.; Du, X.; Duan, Y.; Yu, J.; Yang, X.; Yang, X.; Yang, K.; Liu, X.; Guddat, L. W.; Xiao, G.; Zhang, L.; Yang, H.; Rao, Z., Structural Basis for the Inhibition of SARS-CoV-2 Main protease by Antineoplastic Drug Carmofur. *Nat Struct Mol Biol* **2020**, 27, 529-532.
- Dai, W.; Zhang, B.; Su, H.; Li, J.; Zhao, Y.; Xie, X.; Jin, Z.; Liu, F.; Li, C.; Li, Y.; Bai, F.; Wang, H.; Cheng, X.; Cen, X.; Hu, S.; Yang, X.; Wang, J.; Liu, X.; Xiao, G.; Jiang, H.; Rao, Z.; Zhang, L.-K.; Xu, Y.; Yang, H.; Liu, H., Structure-based Design of Antiviral Drug Candidates Targeting the SARS-CoV-2 Main Protease. *Science* **2020**, 368, 1331-1335.
- Gao, K.; Nguyen, D. D.; Chen, J.; Wang, R.; Wei, G.-W., Repositioning of 8565 Existing Drugs for COVID-19. *J. Phys. Chem. Lett* **2020**, 11 (13), 5373-5382.
- Ma, C.; Sacco, M. D.; Hurst, B.; Townsend, J. A.; Hu, Y.; Szeto, T.; Zhang, X.; Tarbet, B.; Marty, M. T.; Chen, Y.; Wang, J., Boceprevir, GC-376, and calpain inhibitors II, XII inhibit SARS-CoV-2 viral replication by targeting the viral main protease. *Cell Res* **2020**, 30 (8), 678-692.
- Weston, S.; Coleman, C. M.; Haupt, R.; Logue, J.; Matthews, K.; Li, Y.; Reyes, H. M.; Weiss, S. R.; Frieman, M. B., Broad anti-coronaviral activity of FDA approved drugs against SARS-CoV-2 *in vitro* and SARS-CoV *in vivo*. *J. Virol.* **2020**, JVI.01218-20.
- Jeon, S.; Ko, M.; Lee, J.; Choi, I.; Byun, S. Y.; Park, S.; Shum, D.; Kim, S., Identification of Antiviral Drug Candidates against SARS-CoV-2 from FDA-Approved Drugs. *Antimicrob. Agents Chemother.* **2020**, 64 (7), e00819-20.
- Marshall, G. R., Computer-Aided Drug Design. *Ann. Rev. Pharmacol. Toxicol.* **1987**, 27, 193-213.
- 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 (8), 3331-3344.
- Decherchi, S.; Cavalli, A., Thermodynamics and Kinetics of Drug-Target Binding by Molecular Simulation. *Chem. Rev.* **2020**.
- 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 (7), 611-618.
- Yu, W.; MacKerell, A. D., Computer-Aided Drug Design Methods. In *Antibiotics: Methods and Protocols*, Sass, P., Ed. Springer New York: New York, NY, 2017; pp 85-106.
- Ma, C.; Sacco, M.; Chen, Y.; Wang, J., Crystal structure of the SARS-CoV-2 (COVID-19) main protease in complex with inhibitor UAW246. **2020**.
- Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H., PubChem Substance and Compound databases. *Nucleic Acids Res.* **2016**, 44 (D1), D1202-D1213.
- 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.
- 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 (1), 204-211.
- 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**.
- Pham, M. Q.; Vu, K. B.; Han Pham, T. N.; Thuy Huong, L. T.; Tran, L. H.; Tung, N. T.; Vu, V. V.; Nguyen, T. H.; Ngo, S. T., Rapid prediction of possible inhibitors for SARS-CoV-2 main protease using docking and FPL simulations. *RSC Adv* **2020**, 10 (53), 31991-31996.
- 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.

31. 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., Bioinf.* **2014**, 82 (2), 195–215.

32. Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E. C., III, V.W.D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, 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.; SalomonFerrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; D.M., Y.; P.A., a. K., AMBER 18. *University of California, San Francisco* **2018**.

33. Sousa da Silva, A. W.; Vranken, W. F., ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes* **2012**, 5 (1), 1–8.

34. 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 (9), 1157–1174.

35. 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 (12), 10089–10092.

36. Truong, D. T.; Nguyen, M. T.; Vu, V. V.; Ngo, S. T., Fast Pulling of Ligand Approach for the Design of β -Secretase 1 Inhibitors. *Chem. Phys. Lett.* **2017**, 671, 142–146.

37. Efron, B., Bootstrap Methods: Another Kook at the Jackknife. *Ann. Stat.* **1979**, 7, 1–26.

38. Mai, N. T.; Lan, N. T.; Vu, T. Y.; Mai Duong, P. T.; Tung, N. T.; Thu Phung, H. T., Estimation of the Ligand-Binding Free Energy of Checkpoint kinase 1 via Non-

equilibrium MD Simulations. *J. Mol. Graph. Modell.* **2020**, 100, 107648.

39. Mansour, M. A.; AboulMagd, A. M.; Abdel-Rahman, H. M., Quinazoline-Schiff base conjugates: in silico study and ADMET predictions as multi-target inhibitors of coronavirus (SARS-CoV-2) proteins. *RSC Adv* **2020**, 10 (56), 34033–34045.

40. Dan, N. T.; Quang, H. D.; Van Truong, V.; Huu Nghi, D.; Cuong, N. M.; Cuong, T. D.; Toan, T. Q.; Bach, L. G.; Anh, N. H. T.; Mai, N. T.; Lan, N. T.; Van Chinh, L.; Quan, P. M., Design, Synthesis, Structure, *in vitro* Cytotoxic Activity Evaluation and Docking Studies on Target Enzyme GSK-3 β of New Indirubin-3'-Oxime Derivatives. *Sci Rep* **2020**, 10 (1), 11429.

41. 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 (31), 2734–2742.

42. Ngo, S. T.; Nguyen, M. T.; Nguyen, M. T., Determination of the absolute binding free energies of HIV-1 protease inhibitors using non-equilibrium molecular dynamics simulations. *Chem. Phys. Lett.* **2017**, 676, 12–17.

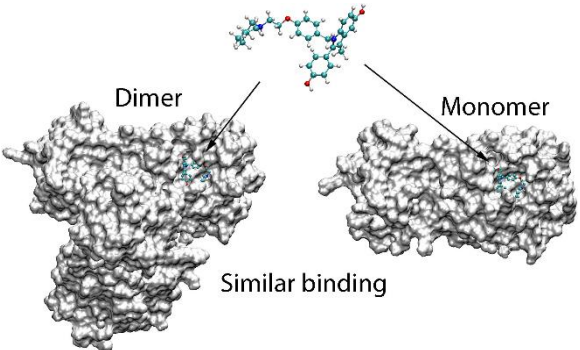
43. Li, D.-C.; Ji, B.-H., Free Energy Calculation of Single Molecular Interaction using Jarzynski's Identity Method: the Case of HIV-1 Protease Inhibitor System. *Acta Mech. Sin.* **2012**, 28 (3), 891–903.

44. 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 (12), 2236–2247.

45. Tam, N. M.; Vu, K. B.; Vu, V. V.; Ngo, S. T., Influence of various force fields in estimating the binding affinity of acetylcholinesterase inhibitors using fast pulling of ligand scheme. *Chem. Phys. Lett.* **2018**, 701, 65–71.

46. Park, S.; Schulten, K., Calculating Potentials of Mean Force from Steered Molecular Dynamics Simulations. *J. Chem. Phys.* **2004**, 120 (13), 5946–5961.

Table of Contents Only

<p>Nguyen Minh Tam, Pham Cam Nam, Duong Tuan Quang, Nguyen Thanh Tung, and Van V. Vu, and Son Tung Ngo</p>		<p>Binding of Inhibitors to the Monomeric and Dimeric SARS- CoV-2 Mpro</p>
---	--	---