Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Molecular Graphics and Modelling 124 (2023) 108535

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/jmgm

Upgrading nirmatrelvir to inhibit SARS-CoV-2 Mpro via DeepFrag and free energy calculations

Nguyen Minh Tam [a], Trung Hai Nguyen [b,c], Minh Quan Pham [d,e], Nam Dao Hong [f], Nguyen Thanh Tung [e,g,**], Van V. Vu [h], Duong Tuan Quang [i,***], Son Tung Ngo [b,c,*]

[a] Faculty of Basic Sciences, University of Phan Thiet, Phan Thiet City, Binh Thuan, Viet Nam
[b] Laboratory of Biophysics, Institute for Advanced Study in Technology, Ton Duc Thang University, Ho Chi Minh City, Viet Nam
[c] Faculty of Pharmacy, Ton Duc Thang University, Ho Chi Minh City, Viet Nam
[d] Institute of Natural Products Chemistry, Vietnam Academy of Science and Technology, Hanoi, Viet Nam
[e] Graduate University of Science and Technology, Vietnam Academy of Science and Technology, Hanoi, Viet Nam
[f] University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Viet Nam
[g] Institute of Materials Science, Vietnam Academy of Science and Technology, Hanoi, Viet Nam
[h] NTT Hi-Tech Institute, Nguyen Tat Thanh University, Ho Chi Minh City, Viet Nam
[i] Department of Chemistry, Hue University, Thua Thien Hue Province, Hue City, Viet Nam

A B S T R A C T

The first oral drug for the treatment of COVID-19, Paxlovid, has been authorized; however, nirmatrelvir, a major component of the drug, is reported to be associated with some side effects. Moreover, the appearance of many novel variants raises concerns about drug resistance, and designing new potent inhibitors to prevent viral replication is thus urgent. In this context, using a hybrid approach combining machine learning (ML) and free energy simulations, 6 compounds obtained by modifying nirmatrelvir were proposed to bind strongly to SARS-CoV-2 Mpro. The structural modification of nirmatrelvir significantly enhances the electrostatic interaction free energy between the protein and ligand and slightly decreases the vdW term. However, the vdW term is the most important factor in controlling the ligand-binding affinity. In addition, the modified nirmatrelvir might be less toxic to the human body than the original inhibitor.

1. Introduction

The outbreak of SARS-CoV-2 caused the COVID-19 pandemic, which has created a huge issue for community health and the global economy since December 2019 [1,2]. The virus rapidly transmits among humans through aerosols [3,4]. Some vaccines were authorized for preventing infection [5]; however, the vaccine effectiveness was significantly reduced with the appearance of SARS-CoV-2 novel variants [6–11]. Discovering novel compounds to treat SARS-CoV-2 is thus of great interest. SARS-CoV-2 main protease (Mpro or 3CL pro) is a high-profile drug target for preventing viral replication and proliferation [12–20]. Nirmatrelvir, which was approved for use as an oral drug for COVID-19 treatment by the U.S. Food and Drug Administration (FDA) [21], is a SARS-CoV-2 Mpro inhibitor. Although the drug is highly effective, there are concerns about its effectiveness against novel SARS-CoV-2 variants.

Currently, searching for potential inhibitors for an enzymatic target can be rapidly and accurately completed via computational work [22–28]. The binding free energy between protein–ligand, ΔG, is an important quantity to obtain because this metric relates to the experimental inhibition constant k_i via formula $\Delta G = RTln(k_i)$, where T is the absolute temperature and R is the gas constant. Several approaches have been designed to calculate the ΔG value. Among them, the double-annihilation binding free method [29] is known as one of the most accurate methods. The approach is based on free energy perturbation (FEP) calculations [30] in which the ligand is annihilated twice in two systems, including the solvated complex and ligand in solution. However, FEP calculation is very expensive in terms of computing resources. In addition, the fast pulling of ligand (FPL) approach [31], an affordable method, is employed to rapidly rank the binding affinity of SARS-CoV-2 Mpro inhibitors with a high correlation coefficient [32,33]. A combination of FPL and FEP can rapidly and accurately evaluate a potential inhibitor for SARS-CoV-2 Mpro. Moreover, traditionally, modifying the chemical structure of a current inhibitor can lead to a stronger binding compound. In computational work, the chemical

* Corresponding author. Laboratory of Biophysics, Institute for Advanced Study in Technology, Ton Duc Thang University, Ho Chi Minh City, Viet Nam.
** Corresponding author. Graduate University of Science and Technology, Vietnam Academy of Science and Technology, Hanoi, Viet Nam.
*** Corresponding author. Department of Chemistry, Hue University, Thua Thien Hue Province, Hue City, Viet Nam.
E-mail addresses: tungnt@ims.vast.ac.vn (N.T. Tung), dtquang@hueuni.edu.vn (D.T. Quang), ngosontung@tdtu.edu.vn (S.T. Ngo).

https://doi.org/10.1016/j.jmgm.2023.108535
Received 29 March 2023; Received in revised form 23 May 2023; Accepted 31 May 2023
Available online 1 June 2023

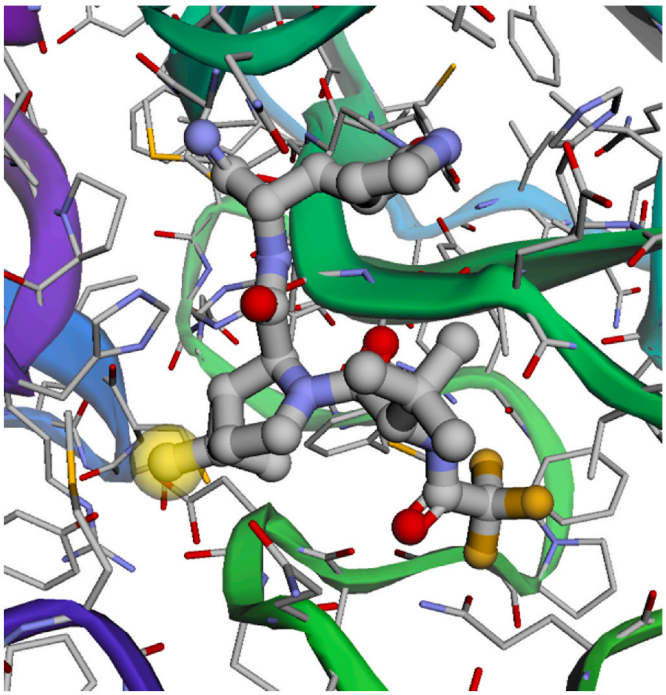

Fig. 1. Binding pose of nirmatrelvir to SARS-CoV-2 Mpro (PDB ID: 7VH8) shown by the Deepfrag web application. The yellow ball indicates the selected carbon that will be estimated for the potential of chemical replacement. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

structure alteration of an inhibitor in binding mode with a protein target can be predicted via a deep learning model [34] with the expectation that the ligand-binding affinity will be improved.

It should be noted that DeepFrag, FPL, and FEP are often used independently to construct and evaluate compounds, but combining them to design and evaluate inhibitors for the SARS-CoV-2 Mpro was not previously done. Therefore, in this context, we proposed to use a combination of a deep learning model and atomistic simulations to modify the chemical structure of nirmatrelvir in the binding mode with SARS-CoV-2 Mpro with the hope of finding a modified compound with a stronger ligand-binding affinity. DeepFrag [34,35], a deep learning model based on the convolutional neural network, was thus employed to modify nirmatrelvir. The ligand-binding affinity of the altered compounds was then refined via the FPL approach [31]. Ligands with a stronger ligand-binding affinity than nirmatrelvir were continuously modified via DeepFrag. The ligand-binding affinity to SARS-CoV-2 Mpro was refined via the FPL approach [31]. The process was repeated until no improved ligand was found. The stronger ligand-binding affinity was finally confirmed by using FEP calculations [30,36]. Overall, 6/83 modified compounds were found to be able to bind to SARS-CoV-2 Mpro with a stronger binding affinity than nirmatrelvir.

2. Materials and methods

2.1. Modifying the ligand structure

DeepFrag [34,35], a deep learning model based on a convolutional neural network, was utilized to change the ligand structure such that the ligand-binding affinity was improved. In particular, the complex structure of SARS-CoV-2 Mpro with nirmatrelvir was obtained from the Protein Data Bank (PDB) with ID 7VH8 [37] and MD simulations, respectively. The complexed structure was employed as an initial conformation of the DeepFrag model. The PDB files of SARS-CoV-2 Mpro and nirmatrelvir/modified nirmatrelvir were uploaded to the DeepFrag webserver (https://durrantlab.pitt.edu/deepfrag/), as shown in Fig. 1. Among these, the inhibitor atoms were designated to estimate that the ligand-binding affinity can be enhanced if they are replaced by another chemical group. A probable modification was kept if the score of the DeepFrag model was larger than 0.90 [38].

2.2. Molecular dynamics simulation

The conventional MD simulations were performed using GROMACS version 2019.6 with NVIDIA GPUs supported via CUDA [39]. The Amber99SB-iLDN force field [40], TIP3P [41], and GAFF [42] were employed to parameterize for protein + ion, water molecule, and ligand, respectively. The ligand parameter was obtained using AmberTools18 [43] and ACPYPE [43] approaches. In particular, ligand information was provided by quantum chemical calculations using the hybrid functional B3LYP with 6–31G (d,p) at the level of theory. The quantum calculation was performed using the implicit solvent, $\varepsilon = 78.4$. The atomic charges of the ligand were obtained from the outcome of

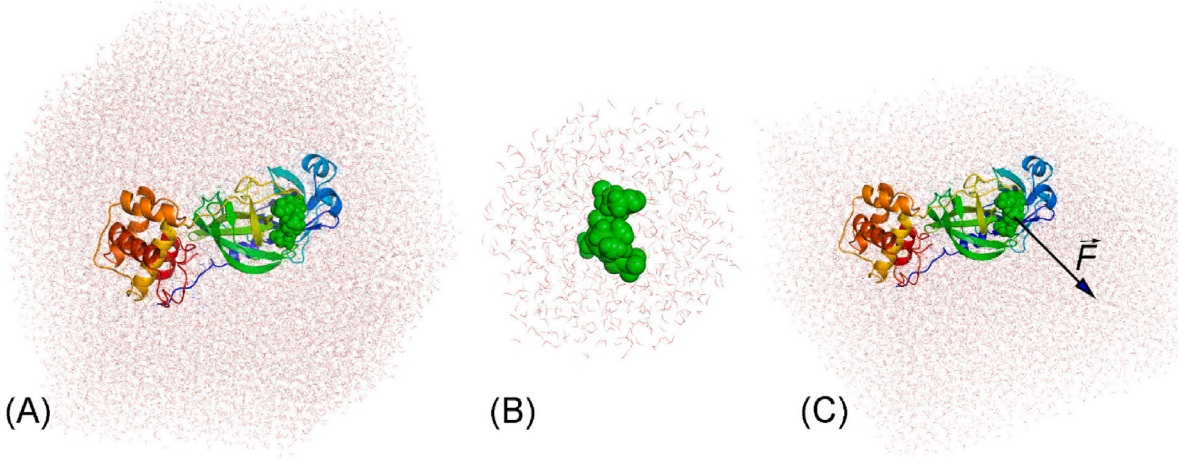

Fig. 2. Initial conformation of simulated systems. (A) The SARS-CoV-2 Mpro + inhibitor complex was inserted into the dPBC box. (B) The inhibitor was inserted into the dPBC box. The equilibrium structures of the SARS-CoV-2 Mpro + inhibitor and individual inhibitor in dPBC box systems were then used as the initial conformation of the FEP calculations. (C) The SARS-CoV-2 Mpro + inhibitor complex was inserted into the rPBC box. The equilibrium shapes of the SARS-CoV-2 Mpro + inhibitor in the rPBC box were then used as the initial conformation of the FPL calculations. In particular, the ligand was then forced to mobilize out of the binding cavity via external force $\vec{F}$ during SMD simulations.

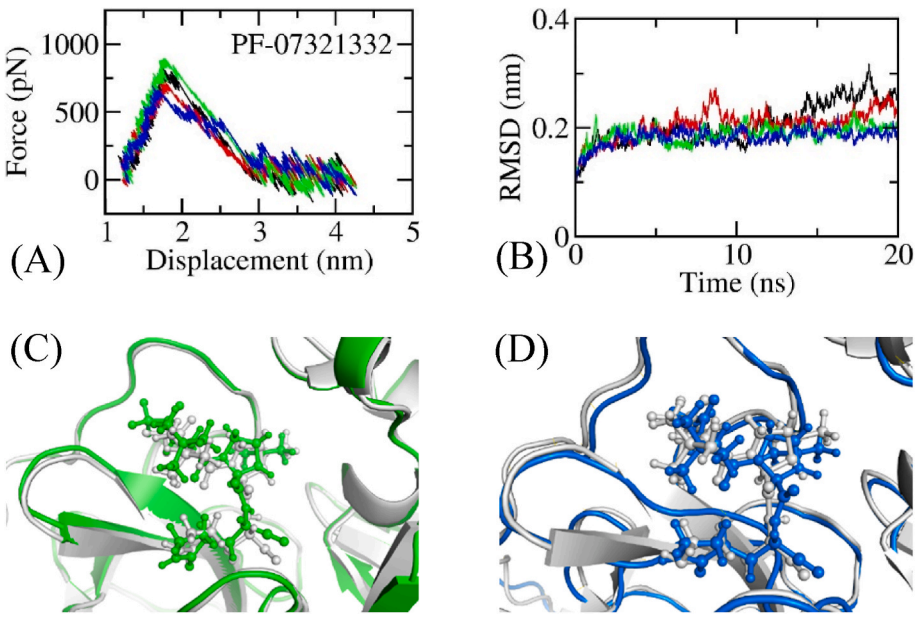

Fig. 3. The equilibrium structure of the SARS-CoV-2 Mpro + inhibitor in the rPBC box was used as the initial conformation of the FPL calculation. In particular, pulling force versus displacement of nirmatrelvir over unbinding pathways (A), in which four colors imply the various trajectories; The SARS-CoV-2 Mpro + inhibitor in dPBC box was relaxed over MD simulations, in which all-atom RMSD of the complex SARS-CoV-2 Mpro + nirmatrelvir over 4 MD trajectories (B); (C) Superposition between experimental (gray) and SMD-initial (green) structures of SARS-CoV-2 Mpro + nirmatrelvir; (D) Superposition between experimental (gray) and MD-refined (blue) structures of SARS-CoV-2 Mpro + nirmatrelvir. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

quantum calculation via the restrained electrostatic potential method [42].

The SARS-CoV-2 Mpro + inhibitor complex was inserted into a dodecahedron and rectangular periodic boundary condition (d/rPBC) box with volumes of ca. 673 and 503 nm3, respectively. The corresponding systems consist of ca. 66 700 and 50 200 atoms (cf. Fig. 2A and C), respectively. Moreover, the inhibitor was also inserted into a dPBC box with a volume of ca. 25 nm3 (cf. Fig. 2B). The solvated ligand system comprises ca. 2400 atoms. The MD simulations were carried out with the parameters referred to in previous work [32].

In the first step, the solvated system was minimized via the steepest descent method. In the second step, the minimized systems in the d/rPBC box were then relaxed via 100 ps of NVT and 100 ps/20 ns of NPT simulations. The C_α atoms were positionally restrained during relaxation simulations. Finally, MD simulations with lengths of 20 ns and 5 ns were then performed to obtain equilibrium snapshots of the SARS-CoV-2 Mpro + inhibitor in the dPBC box and isolated inhibitor, respectively. The simulations were repeated 4 times to guarantee sufficient sampling.

2.3. Fast pulling of ligand (FPL) scheme

The last snapshot of the SARS-CoV-2 Mpro + inhibitor in the rPBC box was utilized as the initial shape of the steered-MD simulations. An inhibitor was dissociated from the binding site of SARS-CoV-2 Mpro via an external harmonic force. Among these, the force cantilever spring constant v and pulling velocity k were set to 600 kJ mol-1 nm-2 and 0.005 nm ps-1, respectively, according to previous work [33]. The pulling work W was calculated for predicting the ligand-binding affinity because it is related to the binding free energy ΔG through the Jarzynski equality [44]. W is calculated as follows:

$$W = v\int_0^t F(t)dt \tag{1}$$

The predicted binding free energy can be estimated via the following formula:

$$\Delta G_{\mathrm{FPL}} = -0.056 \times W - 5.512 \tag{2}$$

where the slope and the intersection were obtained using a linear regression for 11 complexes in the previous work [33].

2.4. FEP calculation

The last snapshot of the solvated complex in the dPBC box and ligand in solution systems were used as initial conformations for FEP calculations according to previous work [32]. In particular, the ligand was removed from two systems including the solvated complex and ligand in solution via λ-alteration simulations [29,30]. The coupling parameter λ changes from 0, corresponding to the bound state, to 1, corresponding to the unbound state. In particular, 8 values of λ_{cou}, including 0.00, 0.10, 0.20, 0.35, 0.50, 0.65, 0.80, and 1.00, were used to modify the electrostatic interactions. Nine values of λ_{vdW}, involving 0.00, 0.10, 0.25, 0.35, 0.50, 0.65, 0.75, 0.90, and 1.00, were used to alter the van der Waals (vdW) interactions. The free energy alteration, $\Delta G_{\lambda=0\to1} = -k_{\mathrm{B}}T\ln\langle e^{-\frac{\Delta H}{k_{\mathrm{B}}T}}\rangle_{\lambda=0}$, matches the work of the ligand-demolishing process. The value can be calculated using the Bennett acceptance ratio (BAR) method [45]. The ΔG_{FEP} between SARS-CoV-2 Mpro and the inhibitor is thus obtained as follows:

$$\Delta G_{\mathrm{FEP}} = \Delta G^{\mathrm{com}}_{\lambda=0\to1} - \Delta G^{\mathrm{lig}}_{\lambda=0\to1} \tag{3}$$

2.5. Analysis tools

The ChemAxon webserver, www.chemicalize.com, was used to predict the ligand protonation states [46]. RMSD value and clustering calculation were computed through the GROMACS tool "gmx rms" and "gmx cluster", respectively [39]. The 2D ligand interaction diagram was generated via the free version of Maestro [47]. The toxicity of ligands was predicted via the PreADMET webserver [48].

3. Results and discussion

As mentioned above, altering the chemical structure of the current inhibitor can lead to a stronger binding compound. We thus used computational approaches, including physics- and knowledge-based methods, to change the chemical structure of nirmatrelvir with the expectation that the ligand-binding affinity would be enhanced. Atomistic simulations were employed to study the binding process of nirmatrelvir to SARS-CoV-2 Mpro (Fig. 3), and its calculation results were used as a positive control. Initially, nirmatrelvir was dissociated from the binding cavity of the protease via the FPL scheme to assess the ligand-binding affinity (Fig. 3A). The FEP calculations were then performed

Table 1
Top-lead compounds showed the largest binding affinity to SARS-CoV-2 Mpro by FPL, FEP, and toxicity calculations.a

N0	Name	F_{Max}	W	ΔG_{FPL} [b]	ΔG_{cou}	ΔG_{vdW}	ΔG_{FEP}	Toxicity
1	PF_3_4_62	1032.2 ± 33.5	146.5 ± 5.2	−13.71	−2.33	−11.65	−13.98 ± 1.07	Low
2	PF_3_4_32	978.9 ± 57.6	144.5 ± 4.1	−13.60	−5.48	−10.60	−16.08 ± 1.33	Low
3	PF_3_4_33_32	927.2 ± 39.3	142.2 ± 6.2	−13.47	−4.44	−10.46	−14.91 ± 1.08	Low
4	PF_3_4_33_31[a]	908.0 ± 63.5	134.7 ± 8.1	−13.05	−2.73	−9.78	−12.51 ± 0.34	Low
5	PF_12_9_31	940.2 ± 15.4	133.5 ± 2.2	−12.99	−1.53	−10.34	−11.86 ± 0.85	Low
6	PF_3_4_7	957.6 ± 44.0	131.2 ± 3.0	−12.86	−4.30	−8.87	−13.17 ± 0.38	Low
7	PF_9a_33	882.6 ± 56.9	130.1 ± 7.3	−12.80	−4.15	−10.67	−14.82 ± 1.13	Low
8	PF_3_4_33	922.9 ± 28.1	128.5 ± 3.4	−12.70	−2.65	−10.39	−13.04 ± 1.44	Low
9	PF_12_9_31_32	836.6 ± 28.1	126.8 ± 3.9	−12.61	−3.80	−10.52	−14.32 ± 0.95	Low
10	PF_3_4_32_32	911.3 ± 24.8	126.3 ± 3.8	−12.58	−4.34	−11.07	−15.41 ± 0.21	Low
11	PF_12_9_3	906.7 ± 34.0	125.8 ± 4.6	−12.55	−5.39	−12.69	−18.08 ± 1.30	Low
12	PF_12_33_3	882.7 ± 12.2	125.4 ± 5.5	−12.53	−7.07	−10.19	−17.26 ± 0.64	Low
13	PF_3_4_31	908.5 ± 30.7	124.5 ± 4.6	−12.48	−2.04	−9.70	−11.74 ± 1.21	Low
14	PF_3_4_31_32	863.1 ± 45.1	123.3 ± 6.1	−12.42	0.78	−9.96	−9.18 ± 1.55	Low
15	PF-07321332	788.3 ± 38.1	103.7 ± 3.6	−11.32	−1.41	−11.21	−12.62 ± 1.96	Medium

[a] The units of force and energy are pN and kcal mol-1, respectively.
[b] Predicted binding free energy can be calculated via Eq. (2).

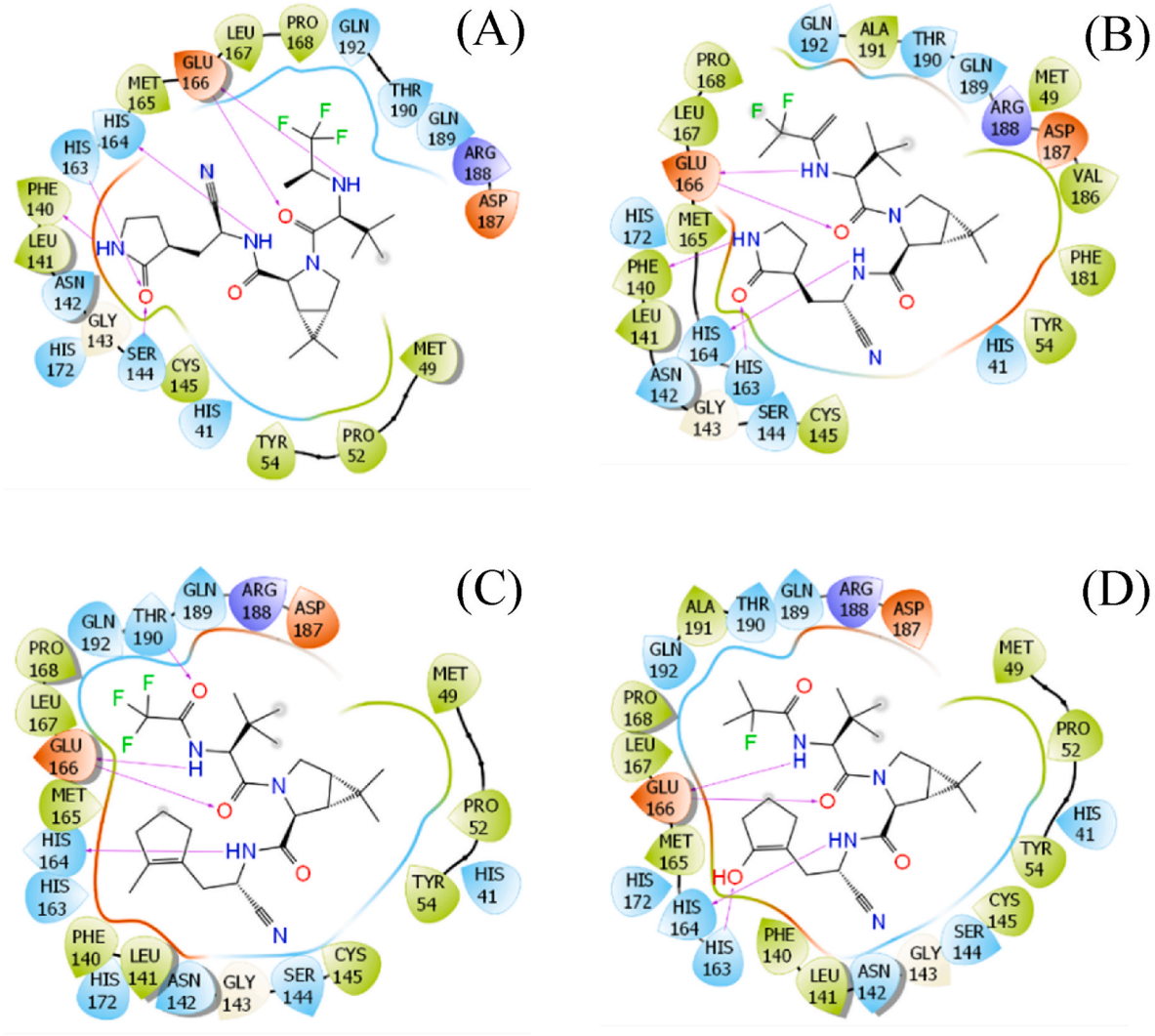

Fig. 4. Ligand interaction diagrams of MD-refined structures of SARS-CoV-2 Mpro + PF_9b (A), PF_9a_33 (B), PF_3_4_62 (C), and PF_3_4_33_32 (D). The diagram was prepared via Maestro free version.

to calculate the absolute binding free energy, in which the MD simulations were carried out to reach the equilibrium states of the solvated complex (Fig. 3B). The complex structure changed only slightly during atomistic simulations (Fig. 3C + D), implying the stability of the simulation results. The calculated ligand-binding affinities were reported in Table 1. In particular, the binding free energy of nirmatrelvir to the protease predicted by FPL simulations was −11.32 kcal mol-1 via Eq. (2) [33], whereas the rupture force and pulling work over 4

independent trajectories were 788.3 ± 38.1 pN and 103.7 ± 3.6 kcal mol-1, respectively. Moreover, the absolute binding free energy ΔG_{FEP} was also calculated as −12.62 ± 1.96 kcal mol-1 via FEP calculation. The obtained results are in good agreement with the experimental binding free energy of nirmatrelvir to SARS-CoV-2 Mpro, which is −13.53 kcal mol-1 ($k_i = 0.271$ nM) [49]. Furthermore, ΔG_{FPL} is underestimated because the linear fit of Eq. (2) was estimated by using an approximation that IC_{50} is equal to k_i.

DeepFrag calculations were designed to predict the structural change of the inhibitor that may enhance the ligand-binding affinity [34,35]. In particular, the ML model was successfully applied to the SARS-CoV-2 Mpro + inhibitor [38]. The ligand-binding affinity was then confirmed via FPL simulations [38], which revealed a correlation coefficient with the respective experiment of $R = -0.74 \pm 0.11$ [32]. The hybrid approaches were thus utilized to modify the structure of nirmatrelvir, which may lead to a novel compound forming stronger binding affinity to the protease. In the first round, the experimental structure of SARS-CoV-2 Mpro + nirmatrelvir was used as an initial shape for DeepFrag calculation. Nine chemical groups were suggested to be modified (Tables S1 and S7 of the Supporting Information - SI file). The ligand-binding affinity of nine compounds was then computed over 4 independent FPL trajectories. The obtained results are reported in Table S1 of the SI file, in which the recorded pulling force over the ligand displacement is described in Table S6 of the SI file. Five compounds were found to be able to form a larger binding affinity, which ranges from −11.52 to −12.35 kcal mol-1, compared to the original inhibitor. Moreover, the clustering method was employed to characterize the representative conformation of the complex over 4 final snapshots of NPT simulations with a cutoff of 0.2 nm. The 2D interaction diagram between the protease and ligand was then prepared and is shown in Fig. 4 and Table S7 of the SI file. The diagrams reveal that the vdW interaction may dominate over electrostatic interactions in the ligand binding process because a larger number of residues formed vdW interactions than hydrogen bonds (HBs).

Based on five modified compounds, DeepFrag/FPL calculations were continuously performed, and 18 modified compounds were thus generated (cf. Tables S2 and S6). The interaction diagram of these ligands with SARS-CoV-2 Mpro was also analyzed and is described in Table S7. The interaction picture is consistent with the 9 compounds above. Interestingly, 5/18 compounds showed a strong binding affinity, ranging from −12.03 to −12.80 kcal mol-1 (Table S2). These ligands were selected as templates for the next round of DeepFrag/FPL calculations. Therefore, in the next step, 23 altered ligands were estimated to be able to adopt a strong binding affinity to SARS-CoV-2 Mpro. The results of FPL simulations were obtained and mentioned in Tables S3 and S6 of the SI file. Among these, 8/23 compounds adopted strong binding affinity with ΔG_{FPL}, which ranges from −12.48 to −13.71 kcal mol-1. Furthermore, the representative conformation of the complex was also characterized, and then the ligand-interaction diagram was produced (cf. Table S7) to clarify the nature of binding. The interaction picture was not changed in comparison with the ligands considered above. The ligand-interaction maps of the representative ligands are shown in Fig. 4, in which the electrostatic interaction is only a weak driving force in the ligand-binding process.

In the next step, 29 modified compounds (Tables S4 and S6) were generated based on 8 compounds obtained in the previous step (Table S3). The ligand-binding affinities of these ligands were also calculated via FPL simulations. Stable conformations of the complexes during simulations were characterized via the clustering method. The 2D ligand-interaction diagrams of these compounds were then produced and are described in Table S7. There are only 2/29 modified compounds that can form a strong binding affinity to SARS-CoV-2 Mpro compared to nirmatrelvir (PF-07321332) (Table S4). Four altered compounds were suggested by the DeepFrag model in the next steps; however, FPL simulations imply that these compounds adopted a weaker binding affinity (Tables S5 and S6 of the SI file). The binding pose of 33 compounds to

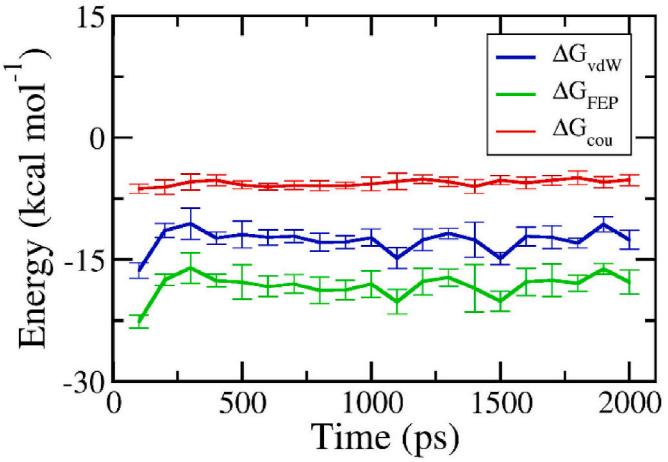

Fig. 5. The interaction free energy between PF_12_9_3 and SARS-CoV-2 Mpro via FEP calculations. Each circle corresponds to the average over 4 independent trajectories of the free energy computed via the BAR method over 100 ps of λ-alteration simulations. The absolute binding free energy ΔG_{FEP} is the sum of ΔG_{cou} and ΔG_{vdW}. The error is the standard error of the average.

SARS-CoV-2 Mpro was also analyzed and is shown in Table S7. Therefore, the improvement process was stopped. In total, 14 out of 83 modified compounds were generated, which can inhibit SARS-CoV-2 Mpro with a value of ΔG_{FPL} ranging from −12.42 to −13.71 kcal mol-1.

Although the FPL simulations formed an appropriate outcome compared with the respective experiment since $R = -0.74 \pm 0.11$ [32], perturbation simulations were also performed to confirm the obtained outcome. It should be noted that the binding free energy obtained by FEP calculation correlates strongly with the respective experiment with a value of $R = 0.85 \pm 0.06$ [32]. As mentioned above, the solvated complex and ligand in solution systems were mimicked over 20 and 5 ns of MD simulations. Both systems reach equilibrium after half of an MD trajectory, and the final snapshots of the simulations were thus used as the starting shape for the λ-alteration simulations. The ligand was changed from the bound state, fully interacting with surrounding molecules, to the unbound state, having no interaction with neighboring atoms. The free energies of two annihilation processes were thus obtained and described in Fig. 5. Because the system changed from the full-interaction state to the modified-interaction state, the annihilation free energy fluctuates greatly over the beginning of the λ-alteration simulations. However, the free energy reaches a stable point within half of the MD simulations. The binding free energy ΔG_{FEP} is thus calculated over the equilibrium interval of the λ-alteration simulations.

The binding free energy ΔG_{FEP} of the top-lead compounds ranges from −9.18 to −18.08 kcal mol-1 (cf. Table 1). In particular, the average electrostatic and vdW interaction free energies are −3.39 and −10.54 kcal mol-1, respectively. The obtained results are consistent with the data of nirmatrelvir, with corresponding values of −1.41 and −11.21 kcal mol-1. Moreover, the absolute binding free energy of nirmatrelvir to the protease is −12.62 ± 1.96 kcal mol-1, which is in good agreement with previous work ($\Delta G_{FEP} = -14.35 \pm 0.04$ kcal mol-1) [49]. Furthermore, among these, 6 compounds adopted a stronger binding affinity to SARS-CoV-2 Mpro than the original nirmatrelvir. Over these compounds, the average electrostatic and vdW interaction energies are −5.15 and −10.95 kcal mol-1, respectively. The structural change in the proposed ligands significantly increases the electrostatic interaction between the protein and ligand. The vdW interaction among molecules was slightly decreased with an amount of 0.26 kcal mol-1. Although the vdW interaction still dominates over coulomb interactions, the structural change of ligands reduces the energy gap. Furthermore, the increase in electrostatic interaction energy can be explained by the increase in the number of HBs. HB is normally adopted between a

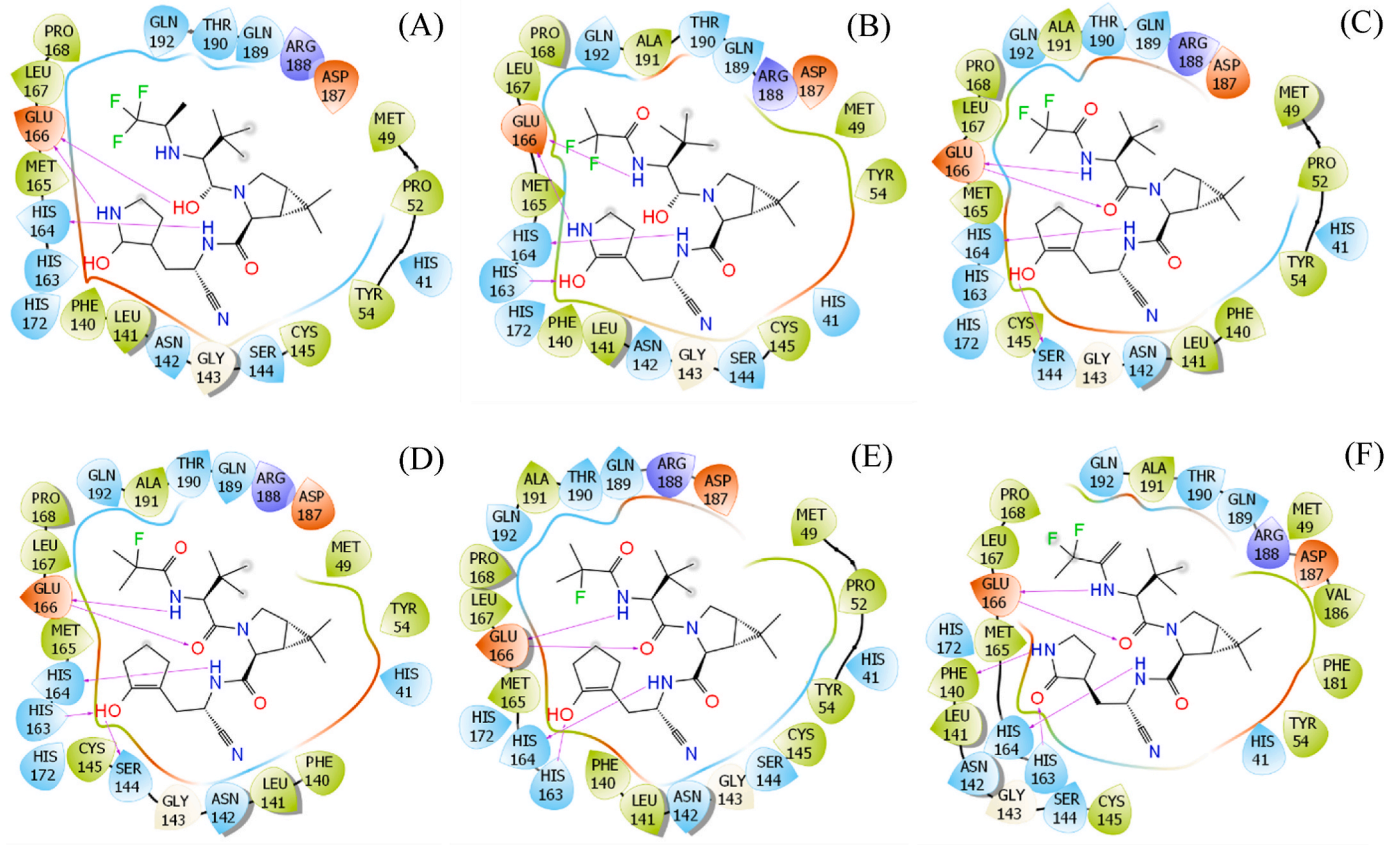

Fig. 6. The ligand-interaction diagram between PF_12_9_3 (A), PF_12_33_3 (B), PF_3_4_32 (C), PF_3_4_32_32 (D), PF_3_4_33_32 (E), and PF_9a_33 (F). The diagram was prepared via Maestro free version.

partially negatively charged atom and a partially positively charged atom. Therefore, an increase in the number of HBs implies that the electrostatic interaction becomes stronger [50]. Indeed, the 6 modified compounds form a 4.17 ± 0.28 HBs to the protease, which is significantly larger than that of the original nirmatrelvir. This observation is in good agreement with the ligand-interaction diagram (cf. Fig. 6).

Although nirmatrelvir is very effective in decreasing death or hospitalization compared to placebo, it can cause severe side effects. Consistently, nirmatrelvir was shown to be medium in terms of predicted toxicity (Table 1). Fortunately, all of the top-lead compounds might be less toxic to the human body than the original one (Table 1). More details on the toxicity prediction of the modified and original nirmatrelvir can be accessed in Table S8 of the SI file.

4. Conclusions

In this work, a combined DeepFrag/atomistic simulation approach was proposed to predict the structural change of the compound PF-07321332 to increase the ligand-binding affinity to SARS-CoV-2 Mpro. In particular, the DeepFrag model was used to predict the structural change with the expectation that the ligand-binding affinity would be enhanced. The FPL simulations were then applied to refine the ML predictions. The FEP calculations were finally utilized to confirm the ML/FPL results.

Our ML/FPL calculations suggested that 14/83 modified compounds could inhibit SARS-CoV-2 Mpro with a stronger ligand-binding affinity than nirmatrelvir. The obtained results for ΔG_{FPL} were the range of −12.42 to −13.71 kcal mol-1. FEP calculations then confirmed that 6 modified compounds exhibit a smaller binding free energy, including PF_12_9_3, PF_12_33_3, PF_3_4_32, PF_3_4_32_32, PF_3_4_33_32, and PF_9a_33. The ΔG_{FEP} ranges from −14.82 to −18.08 kcal mol-1, with an average value of −16.09 kcal mol-1.

The structural modification of nirmatrelvir significantly enhances the electrostatic interaction free energy between the protein and ligand from −1.41 to −5.15 kcal mol-1 and slightly decreases the vdW term by 0.26 kcal mol-1. However, the vdW term is the most important factor in controlling the ligand-binding affinity.

In addition, the proposed compounds might be less toxic to the human body than the original one, according to PreADMET prediction.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This research was funded by Vietnam National Foundation for Science & Technology Development (NAFOSTED) grant #02/2020/ĐX.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmgm.2023.108535.

References

[1] WHO, Coronavirus disease (COVID-19) outbreak, in: Nat Med, World Health Organization, 2021, 2021.
[2] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, Lancet 395 (2020) 470–473.
[3] N. van Doremalen, T. Bushmaker, D.H. Morris, M.G. Holbrook, A. Gamble, B. N. Williamson, et al., Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1, N. Engl. J. Med. 382 (2020) 1564–1567.
[4] J.F.W. Chan, S.F. Yuan, K.H. Kok, K.K.W. To, H. Chu, J. Yang, et al., A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster, Lancet 395 (2020) 514–523.
[5] COVID-19 Vaccines. FDA, FDA, 2021, Vol. 2021.
[6] P. Wang, M.S. Nair, L. Liu, S. Iketani, Y. Luo, Y. Guo, et al., Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7, Nature 593 (2021) 130–135.
[7] M. Hoffmann, H. Hofmann-Winkler, N. Krüger, A. Kempf, I. Nehlmeier, L. Graichen, et al., SARS-CoV-2 variant B.1.617 is resistant to Bamlanivimab and evades antibodies induced by infection and vaccination, bioRxiv (2021) 2021, 05.04.442663.
[8] S. Cele, I. Gazy, L. Jackson, S.-H. Hwa, H. Tegally, G. Lustig, et al., Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma, Nature (2021).
[9] R.E. Chen, X. Zhang, J.B. Case, E.S. Winkler, Y. Liu, L.A. VanBlargan, et al., Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies, Nat. Med. 27 (2021) 717–726.
[10] C.K. Wibmer, F. Ayres, T. Hermanus, M. Madzivhandila, P. Kgagudi, B. Oosthuysen, et al., SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma, Nat. Med. 27 (2021) 622–625.
[11] S.M.-C. Gobeil, K. Janowska, S. McDowell, K. Mansouri, R. Parks, V. Stalls, et al., Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity, Science (2021) eabi6226.
[12] Z. Jin, Y. Zhao, Y. Sun, B. Zhang, H. Wang, Y. Wu, et al., Structural basis for the inhibition of SARS-CoV-2 main protease by antineoplastic drug carmofur, Nat. Struct. Mol. Biol. 27 (2020) 529–532.
[13] L. Zhang, D. Lin, X. Sun, U. Curth, C. Drosten, L. Sauerhering, et al., Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors, Science 368 (2020) 409–412.
[14] H.T.H. Chan, M.A. Moesser, R.K. Walters, T.R. Malla, R.M. Twidale, T. John, et al., Discovery of SARS-CoV-2 Mpro peptide inhibitors from modelling substrate and ligand binding, Chem. Sci. 12 (2021) 13686–13703.
[15] A.D. Rathnayake, J. Zheng, Y. Kim, K.D. Perera, S. Mackin, D.K. Meyerholz, et al., 3C-like protease inhibitors block coronavirus replication in vitro and improve survival in MERS-CoV–infected mice, Sci. Transl. Med. 12 (2020) eabc5332.
[16] W. Vuong, M.B. Khan, C. Fischer, E. Arutyunova, T. Lamer, J. Shields, et al., Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus peplication, Nat. Commun. 11 (2020) 4282.
[17] A. Hegyi, J. Ziebuhr, Conservation of substrate specificities among coronavirus main proteases, J. Gen. Virol. 83 (2002) 595–599.
[18] K. Anand, J. Ziebuhr, P. Wadhwani, R. Mesters Jeroen, R. Hilgenfeld, Coronavirus main proteinase (3CLpro) structure: basis for design of anti-SARS drugs, Science 300 (2003) 1763–1767.
[19] R. Cannalire, C. Cerchia, A.R. Beccari, F.S. Di Leva, V. Summa, Targeting SARS-CoV-2 proteases and polymerase for COVID-19 treatment: state of the art and future opportunities, J. Med. Chem. 65 (2022) 2716–2746.
[20] M. Chauhan, V.K. Bhardwaj, A. Kumar, V. Kumar, P. Kumar, M.G. Enayathullah, et al., Theaflavin 3-gallate inhibits the main protease (Mpro) of SARS-CoV-2 and reduces its count in vitro, Sci. Rep. 12 (2022), 13146.
[21] Coronavirus (COVID-19) Update, FDA Authorizes First Oral Antiviral for Treatment of COVID-19, U.S. Food and Drug Administration, 2021.
[22] Z.T. Muhseen, A.R. Hameed, H.M.H. Al-Hasani, S. Ahmad, G. Li, Computational determination of potential multiprotein targeting natural compounds for rational drug design against SARS-COV-2, Molecules 26 (2021) 674.
[23] C.-H. Zhang, E.A. Stone, M. Deshmukh, J.A. Ippolito, M.M. Ghahremanpour, J. Tirado-Rives, et al., Potent noncovalent inhibitors of the main protease of SARS-CoV-2 from molecular sculpting of the drug perampanel guided by free energy perturbation calculations, ACS Cent. Sci. 7 (2021) 467–475.
[24] C.A. Ramos-Guzmán, J.J. Ruiz-Pernía, I. Tuñón, Computational simulations on the binding and reactivity of a nitrile inhibitor of the SARS-CoV-2 main protease, Chem. Commun. 57 (2021) 9096–9099.
[25] Z. Li, X. Li, Y.-Y. Huang, Y. Wu, R. Liu, L. Zhou, et al., Identify potent SARS-CoV-2 main protease inhibitors via accelerated free energy perturbation-based virtual screening of existing drugs, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 27381–27387.
[26] T.H. Nguyen, H.-X. Zhou, D.D.L. Minh, Using the fast fourier transform in binding free energy calculations, J. Comput. Chem. 39 (2018) 621–636.
[27] N.M. Tam, M.Q. Pham, N.X. Ha, P.C. Nam, H.T.T. Phung, Computational estimation of potential inhibitors from known drugs against the main protease of SARS-CoV-2, RSC Adv. 11 (2021) 17478–17486.
[28] S.T. Ngo, T.H. Nguyen, N.T. Tung, V.V. Vu, M.Q. Pham, B.K. Mai, Characterizing the ligand-binding affinity toward SARS-CoV-2 Mpro via physics- and knowledge-based approaches, Phys. Chem. Chem. Phys. 24 (2022) 29266–29278.
[29] H. Fujitani, Y. Tanida, M. Ito, G. Jayachandran, C.D. Snow, Direct calculation of the binding free energies of FKBP ligands, J. Chem. Phys. 123 (2005), 084108.
[30] R.W. Zwanzig, High-temperature equation of state by a perturbation method. I. Nonpolar gases, J. Chem. Phys. 22 (1954) 1420–1426.
[31] S.T. Ngo, H.M. Hung, M.T. Nguyen, Fast and accurate determination of the relative binding affinities of small compounds to HIV-1 protease using non-equilibrium work, J. Comput. Chem. 37 (2016) 2734–2742.
[32] S.T. Ngo, N.M. Tam, M.Q. Pham, T.H. Nguyen, Benchmark of popular free energy approaches revealing the inhibitors binding to SARS-CoV-2 Mpro, J. Chem. Inf. Model. 61 (2021) 2302–2312.
[33] M.Q. Pham, K.B. Vu, T.N. Han Pham, L.T. Thuy Huong, L.H. Tran, N.T. Tung, et al., Rapid prediction of possible inhibitors for SARS-CoV-2 main protease using docking and FPL simulations, RSC Adv. 10 (2020) 31991–31996.
[34] H. Green, D.R. Koes, J.D. Durrant, DeepFrag: a deep convolutional neural network for fragment-based lead optimization, Chem. Sci. 12 (2021) 8036–8047.
[35] H. Green, J.D. Durrant, DeepFrag: an open-source browser app for deep-learning lead optimization, J. Chem. Inf. Model. 61 (2021) 2523–2529.
[36] S.T. Ngo, T.H. Nguyen, N.T. Tung, P.C. Nam, K.B. Vu, V.V. Vu, Oversampling free energy perturbation simulation in determination of the ligand-binding free energy, J. Comput. Chem. 41 (2020) 611–618.
[37] Y. Zhao, C. Fang, Q. Zhang, R. Zhang, X. Zhao, Y. Duan, et al., Crystal Structure of SARS-CoV-2 Main Protease in Complex with Protease Inhibitor PF-07321332, Protein Cell, 2021.
[38] N.M. Tam, D.-H. Pham, D.M. Hiep, P.-T. Tran, D.T. Quang, S.T. Ngo, Searching and designing potential inhibitors for SARS-CoV-2 Mpro from natural sources using atomistic and deep-learning calculations, RSC Adv. 11 (2021) 38495–38504.
[39] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, et al., GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25.
[40] A.E. Aliev, M. Kulke, H.S. Khaneja, V. Chudasama, T.D. Sheppard, R.M. Lanigan, Motional timescale predictions by molecular dynamics simulations: case study using proline and hydroxyproline sidechain dynamics, Proteins: Struct., Funct., Bioinf. 82 (2014) 195–215.
[41] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[42] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (2004) 1157–1174.
[43] D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E.C. Cheatham III, V.W. D. Darden III, et al., AMBER 18, University of California, San Francisco, 2018.
[44] S. Park, K. Schulten, Calculating potentials of mean force from steered molecular dynamics simulations, J. Chem. Phys. 120 (2004) 5946–5961.
[45] C.H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, J. Comput. Phys. 22 (1976) 245–268.
[46] Chemicalize Was Used for Prediction of Chemical Properties.
[47] P. Schrödinger Llc, Schrödinger Release 2020-4, Maestro, 2020.
[48] S.K. Lee, I.H. Lee, H.J. Kim, G.S. Chang, J.E. Chung, K.T. No, 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.
[49] R. Owen Dafydd, M.N. Allerton Charlotte, S. Anderson Annaliesa, L. Aschenbrenner, M. Avery, S. Berritt, et al., An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19, Science 374 (2021) 1586–1593.
[50] S.C.C. van der Lubbe, C. Fonseca Guerra, The nature of hydrogen bonds: a delineation of the role of different energy components on hydrogen bond strengths and lengths, Chem. Asian J. 14 (2019) 2760–2769.