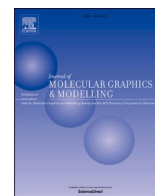




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Computational investigation of potent inhibitors against SARS-CoV-2 2'-O-methyltransferase (nsp16): Structure-based pharmacophore modeling, molecular docking, molecular dynamics simulations and binding free energy calculations

Liying Shi^a, Zeyu Wen^a, Yu Song^a, Jian Wang^{b,**}, Dayong Yu^{a,*}

^a The School of Life Science and Biotechnology, Dalian University, Dalian, 116622, PR China

^b Key Laboratory of Structure-Based Drug Design & Discovery of Ministry of Education, School of Pharmaceutical Engineering, Shenyang Pharmaceutical University, 103 Wenhua Road, Shenhe District, Shenyang, 110016, PR China

ARTICLE INFO

Keywords:

SARS-CoV-2
SARS-CoV-2 2'-O-methyltransferase(nsp16)
Pharmacophore
Molecular dynamics
MM/GB (PB) SA

ABSTRACT

The Coronavirus Disease 2019 (COVID-19), caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, has created unprecedented public health and economic crises around the world. SARS-CoV-2 2'-O-methyltransferase (nsp16) adds a “cap” to viral RNA to maintain the stability of viral RNA, and inhibition of nsp16 activity may reduce viral proliferation, making this protein an attractive drug target. Here, we report the identification of several small molecule inhibitors of nsp16 by virtual screening. First, the nsp16-sinefungin complex (PDB ID: 6WKQ) was selected from the protein data bank. Asp6912, Cys6913, Asp6897 and Asp6928 were determined to be the key amino acids for sinefungin binding in the crystal structure of nsp16-sinefungin complex by molecular dynamics simulation. The complex structures in the stable binding trajectory of nsp16-sinefungin were then clustered through molecular dynamics RMSD analysis. Six clusters were generated, and six representative structures were selected to construct the pharmacophore based on the structure. These six pharmacophores were superimposed on the binding pocket to simplify and pick the common characteristics. The compounds obtained by the pharmacophore screening from Bionet and Chemdiv databases were docked into the nsp16 active pocket. The candidate compounds were selected according to the molecular docking score and then screened by MM/GBSA. Finally, four candidate compounds were obtained. Four sets of 150ns molecular dynamics simulations were performed to determine whether candidate compounds could maintain stable interactions with key amino acids. The results of MD and MM/PBSA energy decomposition indicated that C1 and C2 could form a stable complex system with nsp16, and could form strong hydrogen bonds and salt bridges with the key amino acid Asp6897 and Asp6928. This study thus identifies and attempts to validate for the first time the potential inhibitory activities of C1 and C2 against nsp16, allowing the development of potent anti-COVID-19 drugs and unique treatment strategies.

1. Introduction

Since December 2019, disease caused by severe acute respiratory syndrome (SARS-CoV-2) virus has been declared as an accessible disease by the World Health Organization (WHO) [1,2]. As of April 26, 2021, the virus has had an extremely dangerous impact on more than 200 countries, with more than 140 million confirmed cases and more than 3 million deaths [3]. However, few drugs are available for the treatment of

COVID-19 infection, and the maximum efficacy of these drugs may only reduce the symptoms of infection [4]. Therefore, there is an urgent need to develop effective antiviral therapies to combat the pandemic.

The 5' end of the coronavirus RNA is modified by a cap structure. Modification of the 5' end of RNA virus genome is very important for the replication of the virus genome [5], the translation and synthesis of virus protein, the assembly of virus particles and the elusion of the host innate immune system from the recognition of virus pathogen molecular

* Corresponding author.

** Corresponding author.

E-mail addresses: jianwang@syphu.edu.cn (J. Wang), yudayong@dlu.edu.cn (D. Yu).

<https://doi.org/10.1016/j.jmglm.2022.108306>

Received 24 September 2021; Received in revised form 6 August 2022; Accepted 8 August 2022

Available online 18 August 2022

1093-3263/© 2022 Elsevier Inc. All rights reserved.

pattern [6,7]. Therefore, 5' terminal modification of the RNA virus genome serves as a potential target for antiviral drug design. Because the replication life cycles of coronavirus takes place in the cytoplasm, the capping enzyme of the host cannot complete the capping of coronavirus itself [8,9], so coronavirus completes the capping of viral RNA cap1 through its own protein [10–12]. Additionally ribose 2'-O-methylation plays an important role in the evasion of innate immune recognition. In SARS-CoV-2 virus, 2'-O-methyltransferase (2'-O-Mtase) consists of two subunits: catalytic subunit nsp16 and stimulating subunit nsp16 [13, 14]. SARS CoV-2 2'-O-methyltransferase (nsp16) is an important enzyme for virus survival [15–17]. The role of nsp16 is to protect viral RNA from the cellular innate immunity by participating in the formation of a special arrangement at the 5' end of RNA molecule. Therefore, nsp16 inhibitor can effectively inhibit the infection of COVID-19 [18, 19].

The use of computer-aided drug design (CADD) experiments can accelerate the process of drug development. Currently, a multitude of simulation experiments have been conducted to develop inhibitors for nsp16. Mahmoud et al. [20] screened the zinc database by constructing the nsp16-sinefungin complex pharmacophore model, and evaluated the candidate compounds by molecular dynamics and binding free energy in order to identify potential nsp16 inhibitors. Aldahham et al. [21] screened naphthyridine and quinoline derivatives as potential nsp16-nsp10 complex inhibitors by similar methods. In addition to these, other screening efforts were directed towards natural product development studies. Anuradha et al. [22] screened 10 natural products for nsp16 inhibitors by molecular docking and concluded that theaflavins and catechins have the potential to inhibit the interaction of viral molecules. In another study, Fatoki et al. [23] used gene network analysis, molecular docking, and sequence and structural dynamics simulations to determine potential medicinal phytochemicals to modulate gene expression networks associated with the pathology of SARS-CoV-2 in the human host. Malik et al. [24] identified the plant compound Withanolide and the anti-HIV drug Dolutegravir as Nsp16 inhibitors after analyzing 128 plant compounds and 10 FDA-approved anti-HIV drugs through AutoDock Vina docking and Gromacs dynamics simulations analysis.

In this study, based on the reported nsp16 and sinefungin crystal complex structure (PDB ID: 6WKQ) [25], we applied a virtual screening workflow based on a combination of structure-based pharmacophore modeling with molecular dynamics simulation techniques in order to discover potent SARS-CoV-2 inhibitors.

2. Materials and methods

2.1. Protein preparation

The crystal structure of the nsp16 in complex with the pan-methyltransferase inhibitor sinefungin was retrieved from the protein data bank (PDB ID: 6WKQ, resolution: 1.98 Å). The crystal structure of nsp 16 in complex with sinefungin was prepared by Schrödinger 2018. First, water molecules and other cofactors were removed. Hydrogen atoms and amino acid residues of the crystal structure were subsequently added via “Protein Preparation Wizard” in the Schrödinger [26]. At pH 7.4, the protonated and tautomeric of amino acids were adjusted to the corresponding state. Finally, through the atomic force field OPLS 2005, the hydrogen atom energy of the crystal structure was minimized and the heavy atom converges to an RMSD of 0.3 Å [27].

2.2. Molecular dynamics simulations and clustering

Using the “System Builder” module in the Schrödinger, the solvent system was constructed for the crystal structure of nsp 16 in complex with sinefungin. The system was solvated in an orthorhombic box, with SPC as the solvent system, and then two Cl⁻ were added for neutralization. The 0.15 M NaCl acts to reproduce physiological conditions. At

the same time, the system was given OPLS_2005 force field. At 310K and normal pressure (NPT) for 50ns simulation, the prepared system was simulated by “Desmond 3.7” [28]. The energy and trajectory atomic coordinate data were recorded at 1.2ps and 100ps, respectively, and the data were used for statistical analysis. The most stable path in simulation was clustered. According to RMSD, “Desmond trajectory clustering” was used to cluster frames that were collected, and the frequency value was set to 1, resulting in a maximum of 6 clusters. The MD simulation for the complexes generated by docking studies were executed employing the aforementioned protocol.

2.3. Pharmacophore construction and virtual screening

The “Phase” in the Schrödinger was used for pharmacophore construction and virtual screening [29–31]. The representative conformations from 6 clusters were selected to construct the pharmacophore model, and the exclusion volume was adjusted according to the conformation of the binding pocket. Combined with molecular dynamics analysis, several representative pharmacophore models were merged and simplified to obtain the screening pharmacophore model.

The molecules in Bionet [35] and ChEMBL (<https://www.ebi.ac.uk/chembl/>) databases were matched with the screening pharmacophores. In the screening process, up to 50 conformations were generated for molecules and the energy minimization was utilized when the small molecules were exported. The compound matched at least three of the characteristics in the pharmacophore. The screening results were ranked by phase screen score. The pharmacophore-based virtual screening was performed by “Phase Ligand Screening” in the Schrödinger.

2.4. Docking-based virtual screening

2.4.1. Ligand Preparation

In the “LigPrep”, the molecules matched with the screening pharmacophore were optimized using OPLS3 force field. All molecules retain specified chirality, generate tautomers, and result in the corresponding ionized state at pH 7.0 ± 2.0 by using Epik (Empirical pKa Prediction) [32]. The co-crystalline compound performs the same treatment as described above.

2.4.2. Molecular docking

The docking simulation was performed in Glide. Re-docking was implemented to verify the accuracy of Glide docking protocol. The two docking precisions “standard precision (SP)” and “extra precision (XP)” were used in the docking simulation [33]. Therefore, both docking methods were used in the re-docking of the exfoliated co-crystalline compound. The root mean square deviation (RMSD) between crystal and re-dock pose was calculated. The RMSD value less than 2 Å can confirm that the conformation produced by the docking can reproduce the X-ray results accurately [34]. The receptor grid file was defined according to the sinefungin position, and the parameter coordinates are shown in Table 1. Docking score and glide score were used to evaluate the screening results.

2.5. MM/GBSA calculation

The stability of the ligand-protein complex system generated by the docking simulation was verified by the calculation of the free binding energy in Prime MM/GBSA. The free binding energy calculation abide by the following formulas:

$$\Delta G_{bind} = \Delta G_{complex} - (\Delta G_{protein} + \Delta G_{ligand}) \quad (2.1)$$

$$\Delta G_{bind} = \Delta H - (\Delta G_{solvation} + T\Delta S) \quad (2.2)$$

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{GB} + \Delta G_{SA} - T\Delta S \quad (2.3)$$

ΔG_{bind} is the binding energy; $\Delta G_{complex}$ is the free energy of the

Table 1
Parameters of Glide docking pocket definition.

Receptor	Native Ligand	Coordinates of Grid Center	Dimensions of the Outer Box (Å)	Dimensions of the Inner Box (Å)
SARS COV-2 2'-O-methyltransferase (nsp16)	Sinefungin	X = 2.73 Y = -9.70 Z = -5.90	X = 20 Y = 20 Z = 20	X = 10 Y = 10 Z = 10

complex system; $\Delta G_{\text{protein}}$ and ΔG_{ligand} are the free energy of protein and ligand in the complex system; ΔG_{GB} and ΔG_{SA} are the contribution of polarity and nonpolarity to solvent free energy in solvent environment. ΔS represents the entropy change of ligand structure during sampling. ΔE_{MM} is the Gas phase free energy. The parameters in MM/GBSA are the default settings: solvation model is VSGB, and sampling method is Minimize. Check “use constraints on flexible residues” to set the residues around the receptor pocket as flexible conformation during calculation.

2.6. MM/PBSA and energy decomposition calculation

Protein-ligand complexes were then subjected to a molecular dynamics study and MM-PBSA binding free energy calculations by Gromacs2021. The RESP charges were calculated for ligands by Gaussian16, and the GAFF force field was subsequently assigned by the Bio2Byte online server (<https://www.bio2byte.be/>). The AMBER99SB force field was assigned to the protein in Gromacs. The position of the ligand in the complex system was restricted. The complex system was then placed in a dodecahedral volume (the edge of the volume box was 10 Å from the molecular edge of the complex system), and the TIP3P solvent model was added to the volume box. The “genion” module in Gromacs adds counter ions (Na + or Cl-) for the system to maintain electrical neutrality. Before the molecular dynamics simulation, the energy of the system was minimized, 100ps equilibration was then carried out under the NVT and NPT systems respectively. The equilibrated system was finally subjected to 100ns dynamics simulation, and the trajectories were sampled every 100ps. The “trjconv” module in Gromacs removes the coordinates for the trajectory and corrects the periodicity. The binding free energy of the ligand-protein complex system and the energy contribution of the amino acids in the binding pocket were analyzed from the four terms $\Delta G_{\text{van der Waals}}$, $\Delta G_{\text{Electrostatic}}$, $\Delta G_{\text{Polar Solvation}}$ and $\Delta G_{\text{Non-Pol Solvation}}$ by the MMPBSA.py script in the Amber Tool 21 software.

3. Results and discussion

3.1. Molecular dynamics simulation

We analyzed the interaction between sinefungin and nsp16. Crystal structure 6WKQ interactions between sinefungin and nsp16 in PDB database revealed that sinefungin forms multiple hydrogen bonds with amino acid. As shown in Fig. 1a, the carboxylic acid on sinefungin forms hydrogen bonds with Asn6841 and Gly6879, Asp6897 as the hydrogen acceptor forms hydrogen bonds with two hydroxyl groups. The amino group and the nitrogen on the purine act as the donor and acceptor to form hydrogen bonds with Asp6928, Gly6869, Leu6898, Asp6912 and Cys6913. In addition, Asp6928 forms a salt bridge with amino cations.

In order to verify the stability of the interaction between sinefungin and nsp16, a 50ns molecular dynamics simulation was performed for the crystal structure of the nsp16-sinefungin complex, and the stability of the ligand in the binding pocket and the stability of the protein were evaluated by RMSD during the simulation (Fig. S1 in Supplementary Materials). The results showed that sinefungin maintained a relatively stable state in the binding pocket after 30 ns. In order to explore the main interaction between the ligand and receptor in the simulation process, the 50% of ligand-amino acid interaction was analyzed. As shown in Fig. 1b, the main interaction is dominated by hydrogen bonds. The proportions of hydrogen bonds between sinefungin and Asp6912, Asp6897, Cys6913 and Asp6928 are 98%, 97%, 86%, and 78%, respectively.

Based on molecular dynamics simulation analysis, sinefungin can stably exist in the binding pocket. The analysis of the interaction of the whole trajectory revealed that the hydrogen bonds formed by Asp6912, Cys6913, Asp6897 and Asp6928 with sinefungin were quite conservative (Fig. 1b; Fig. S2 in Supplementary Materials), and these amino acids were critical for the stable binding of sinefungin as an inhibitor.

3.2. Construction of pharmacophore

The key feature of sinefungin stable in the binding pocket can be characterized by pharmacophores based on structure, which can be used in screening simulation to greatly improve the probability of screening

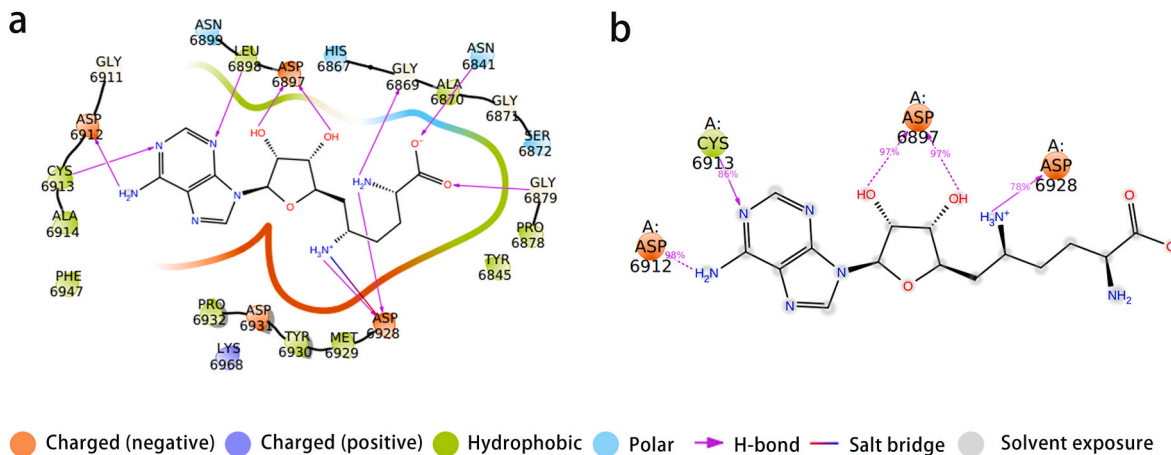


Fig. 1. (a) Interactions between sinefungin and SARS-COV-2 2'-O-methyltransferase (nsp16) from the crystal structure (PDB code 6WKQ). (b) Interactions between sinefungin and nsp16 during molecular dynamics simulation.

compound with inhibitory activity. In the 30–45ns interval of molecular dynamics simulation, sinefungin maintains a relatively stable state at the binding pocket (Fig. S1 in Supplementary Materials). The 300 frame structure conformations generated by the selected trajectory was used to cluster according to RMSD value, which is helpful for us to investigate the representative pose that maintains the ligand stability in the pocket and construct the screening pharmacophore with key characteristics. A total of 6 clusters were obtained by cluster analysis. The representative conformations of these 6 clusters were selected for pharmacophore construction. The pharmacophore characteristics included hydrogen bond receptor (A), hydrogen bond donor (D), hydrophobic (H), aromatic ring (R), positive charge (P) and negative charge (N). Based on the structural characteristics of receptor-ligand complexes, 6 pharmacophore models as shown in Fig. 2 were obtained.

The characteristic number of 6 pharmacophores varied from 4 to 7. In order to screen common characteristics of six pharmacophores, 6 pharmacophore models were superimposed and placed in binding pocket. As shown in Fig. 3a, the common features are D10, A2, D8, P14 and N13. Among them, the superposition outcomes of D10, A2, D8 and P14 shows that the deviation of the four features in both displacement and direction is very small. These results indicate that the interaction corresponding to the four features (hydrogen bonds with Asp6912, Cys6913, Asp6897 and Asp6928) is highly conservative in the 6 clusters which was consistent with the results of molecular dynamics simulation analysis. N13 has an obvious position shift, N13 is mapped on the carboxylic group at the end of sinefungin structure. The last frame of the dynamics simulation was superimposed with the initial frame. The deflection amplitude and direction during the dynamics simulation are shown in Fig. 3b. Compared with the initial conformation, the overall deviation of the stabilized sinefungin is smaller, but the terminal carboxylic acid part has a significant rotational displacement change. In order to systematically analyze the existing angle deviation, the change of the distance between the hydroxyl group on carboxylic acid and Asn6841 in the molecular dynamics simulation trajectory was monitored. The data (Fig. S3 in Supplementary Materials) showed that the distance between the monitored objects changed memorably ($>8 \text{ \AA}$), indicating that the electrostatic interaction between the carboxylic acid and Asn6841 corresponding to N13 was extremely unstable. Therefore, we removed the N13 pharmacophore characteristics. In addition, we

analyzed the exclusion volume. The flexibility of the amino acid structure around the binding pocket is poor. During the molecular dynamics simulation, the amino acid around the pocket has a micro displacement, so the amino acid that affects the compound binding effect is no longer flexible. Finally, a simplified pharmacophore model retaining the key characteristics was obtained after screening (Fig. 4).

The pharmacophore model has four characteristics, including two hydrogen bond donor D8 and D10, positive charge P14 and a hydrogen bond acceptor A2. These characteristics can reflect the stability of sinefungin and the key factor of its inhibitory activity while also, improving the possibility of screening results with inhibitory activity. At the same time, the simplified pharmacophore model also provides the possibility of diversification of molecular configuration of screening results.

3.3. Pharmacophore model screening

The simplified pharmacophore model was used to screen about 1.66 million compounds in Bionet and Chembiv databases. The minimum matching number with pharmacophore model was set to 3. The screening results were sorted according to “phase screen score” and a total of 543 compounds were retained. These candidate compounds were then screened by molecular docking and the interaction was analyzed.

3.4. Docking simulation

In total 543 compounds screened by the pharmacophore model could theoretically form the interaction similar to sinefungin. In order to examine whether these compounds fit the binding pocket and form the critical interaction mode with the active site, molecular docking method was used to further screen the pharmacophore screening results. Docking simulation contains Glide SP and XP methods. Before docking experiments, SP and XP were used for re-docking the co-crystalline compounds. As shown in Fig. 5, the RMSD values deviation between the conformation generated by the two docking methods and the original crystal conformation is not greater than 2 Å. In addition, the conformation generated by XP docking result in a more reductive crystal conformation. Because the reproducibility of XP docking is better, the

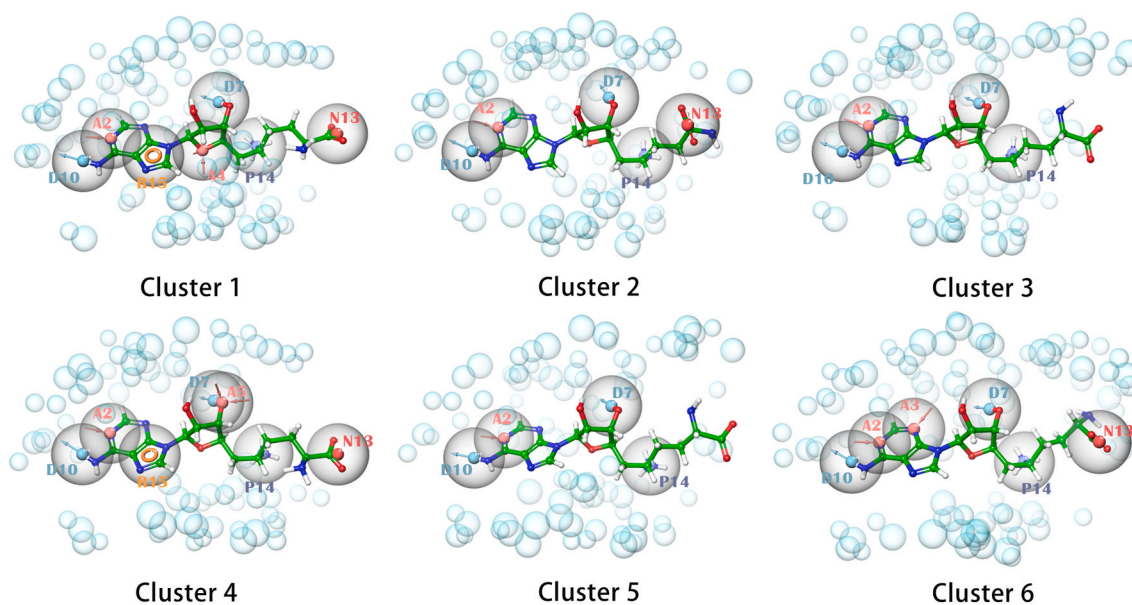


Fig. 2. Structure-based pharmacophore model derived from the representative structure of six cluster sets. Red represents hydrogen bond receptor (A), blue represents hydrogen bond donor (D), yellow represents aromatic ring (R), purple represents positive charge (P) and magenta represents negative charge (N). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

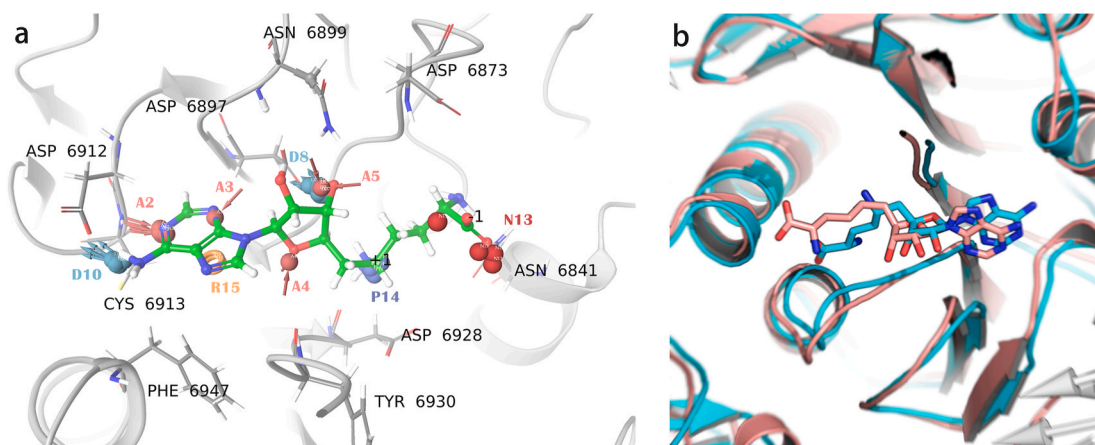


Fig. 3. (a) 6 representative pharmacophores were superimposed in the binding pocket. (b): Initial crystal structure and generate by molecular dynamics structure of sinefungin and nsp16 complex were superimposed. Blue represents the initial conformation of the complex. Pink represents the conformation of the complex in the last frame of molecular dynamics simulation. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

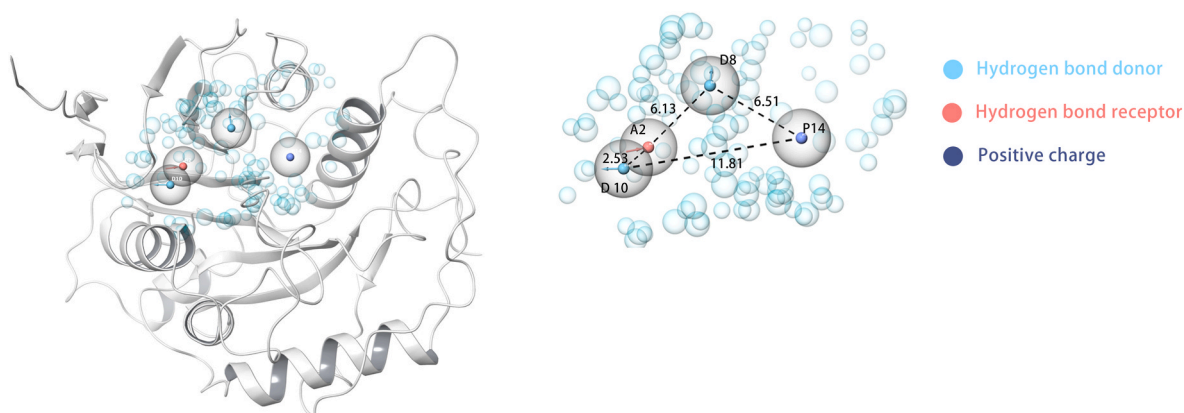


Fig. 4. Merged pharmacophore model created by 6 representative pharmacophores were superimposed to analyze the key characteristics. Black lines and numbers represent the distance between pharmacophore characteristics (Distance unit: Å).

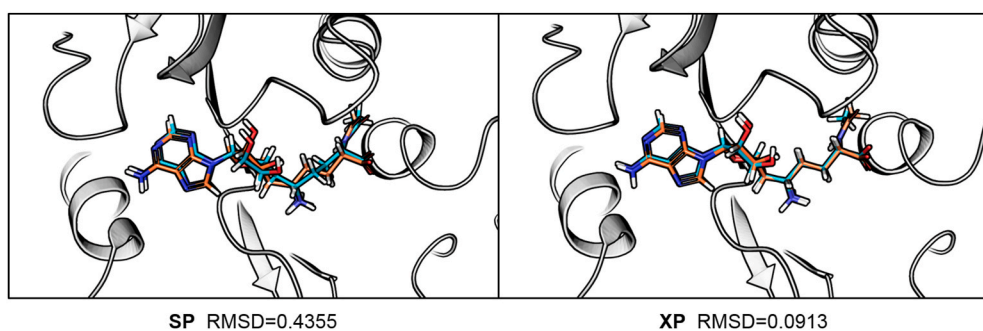


Fig. 5. Glide SP and XP docking program respectively re-docking Sinefungin crystal structure, re-docking pose and crystal structure pose compare RMSD value. Blue represents the initial pose of the crystal, and yellow represents the pose generated by the docking procedure. The RMSD values of SP and XP are 0.4355 and 0.0913 respectively. The two docking procedures have good reproducibility for the crystal structure of the ligand, and the reproducibility of XP is more accurate. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

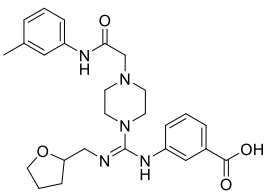
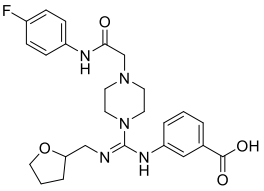
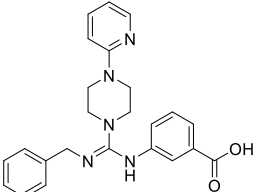
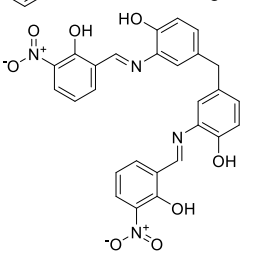
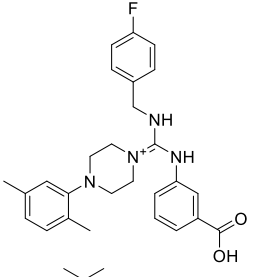
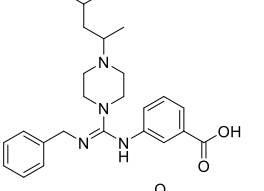
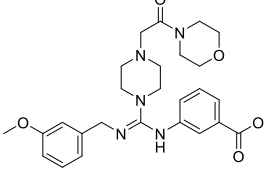
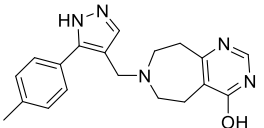
pose generated by XP docking was selected for the study of ligand-receptor interaction mode.

First, we used SP to dock 543 compounds screened by pharmacophores model, of which 527 molecules were docked to the binding pocket. The docking results were ranked by “Docking score”. We selected the top 100 docking compounds for XP simulation to reduce the false positive rate of molecular docking. Finally, the top 10 compounds generated from XP simulation were selected to analyze their hydrogen bonding interaction. The docking data and structural formulas of 10

compounds were recorded in [Table 2](#).

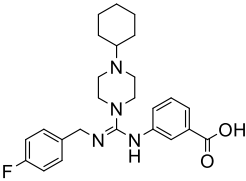
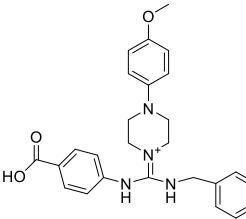
C2, C3, C4, C7, C8 and C10 can form hydrogen bonds with key amino acids Asp6928 and Asp6897. In addition, the nitrogen ions in these compounds can also form salt bridges with Asp6928. Furthermore ionic interaction can be observed in the π -cation interactions of C3, C4 and C7 with Lys6844. As show in [Table 2](#), the structural difference between C1 and C2 is only in the substituents on the benzene ring, but there is evident difference in the interaction between the C1 and C2 in the binding pocket ([Fig. S4](#) in Supplementary Materials). Moreover, the

Table 2
2D structures and docking evaluation values of 10 candidate compounds.

Compound	CAS ID	Structure	Docking score	H-bond	Distance (Å)
C1	1224032-33-0		-9.309	Met6929 Tyr6930 Gly6871	2.30 2.60 2.12
C2	1224020-56-7		-9.191	Gly6871, Asp6928	2.11 2.39
C3	1224031-61-1		-8.624	Asp6928, Asp6897, Tyr6930	2.79 1.87 2.22
C4	310458-53-8		-8.547	Asp6897, Asn6841	1.85 2.46
C5	1224023-91-9		-8.000	Tyr6930, Asn6841, Gly6869	2.39 1.75 2.47
C6	1224019-21-9		-7.970	Tyr6930	2.11
C7	1224035-30-6		-7.915	Tyr6930, Asp6928, Gly6879	2.04 2.48 1.73
C8	2096376-30-4		-7.908	Asp6928, Asn6841, Lys6933	2.45 1.86 2.50

(continued on next page)

Table 2 (continued)

Compound	CAS ID	Structure	Docking score	H-bond	Distance (Å)
C9	1224019-24-2		-7.819	Tyr6930	2.14
C10	932539-95-2		-7.780	Asp6897, Lys6933	1.94 2.49

distance parameter of the hydrogen bond between the compound and the amino acid raises questions as to whether some of the hydrogen bonds can exist the same time, as shown in Fig. 6. Hence the C1 and C2 conformations produced by the docking simulation were superimposed at the docking pocket. The amino on the double bond side of C2 forms a hydrogen bond with Asp6928, while the position of C1 in the bonding pocket is lower than C2. We observed that there is sufficient space at the end of the binding pocket, which makes it possible for C1 to shift to the position that coincides with the C2. The position change increases the possibility that the amino on the C1 double bond forms a hydrogen bond with Asp6928 (the yellow cutting head direction in Fig. 6 represents this hypothesis). In addition, the amino connected to the two segments of the double bond can undergo a certain degree of rotation, so the amino on the opposite side can be changed by rotation and shorten the distance between Asp6928 form hydrogen bond. Furthermore, the pockets where the two amino linking groups are deeper, which can eliminate the negative effect of steric hindrance when the two amino end of the double bond rotate (the red rotating cutting head in Fig. 6 represents hypothesis). The default protein structure of Glide docking simulation is rigid and insufficient sampling of the ligand would result in a docking pose that lacks accuracy. Therefore, the interaction of ligands at the binding pocket was further analyzed by MM/GBSA and molecular dynamics.

3.5. MM/GBSA calculation

The thermodynamic concept was introduced to simulate the process of ligand binding to the protein pocket, which can effectively determine whether the interaction mode between receptor and ligand is stable. Prime MM/GBSA was used to scan and analyze the receptor-ligand binding mode generated by XP docking simulation, and to determine the binding stability of the complex through calculating the free binding energy. As shown in Table 3, the free binding energy (ΔG_{bind}) of sinefungin was -51.318 kcal/mol. Among the candidate compounds, the free binding energy of C1, C2, C3 and C7 were above this value, and C2, C3 and C7 can form hydrogen bonds with key amino acid Asp6898 and Asp6928. The similar structures of C1 and C2 results in similar free binding energy values ($\Delta G_{\text{bind}}\text{C1} = -59.983$ kcal/mol; $\Delta G_{\text{bind}}\text{C2} = -70.916$ kcal/mol) which are better than those of the other two compounds. The free binding energy results of C2 show that $\Delta G_{\text{bind}}\text{Coulomb}$, $\Delta G_{\text{bind}}\text{vdW}$, $\Delta G_{\text{bind}}\text{Lipo}$ (nonpolar solvation energy) and $\Delta G_{\text{bind}}\text{Hbond}$ provide positive contributions with the main contributions from $\Delta G_{\text{bind}}\text{vdW}$ (-51.997 kcal/mol) and $\Delta G_{\text{bind}}\text{Coulomb}$ (-24.058 kcal/mol). The calculated results of C1, C3 and C7 are similar to C2 and sinefungin. The results for MM/GBSA show that the binding of C1, C2, C3 and C7 with the protein was relatively stable, and the van der Waals energy and Coulomb energy are the main contributors to the free binding energy.

3.6. Molecular dynamics simulation

The stability of the interaction between the four compounds and nsp16 was investigated systematically by 150ns molecular dynamics simulation. We expect to use molecular dynamics simulation to investigate whether the four compounds can interact with key amino acids during the simulation, and whether these interactions can persist similar to sinefungin.

As shown in Fig. 7a, during molecular dynamics simulation of the four complexes, the structure of nsp16 showed a stable trend after 105ns, and the fluctuation of RMSD value for C2-protein was significantly weaker than other three groups of compounds. At the same time, the complex system composed of the four compounds and the protein was compared with sinefungin. The deviation amplitude of the protein structure relative to the initial structure at 50ns of molecular dynamics simulation is about the same, and the RMSD value maintained at 2.0–3.5 Å. The reasonable fluctuation range proved that the proteins in the four complexes could maintain a stable state in the simulation process.

In addition, the data of Fig. 7b shows that Stable RMSD curve fluctuations can prove that the four compounds can remain stable in the binding pocket. However, it can be evidently obvious that at the end of

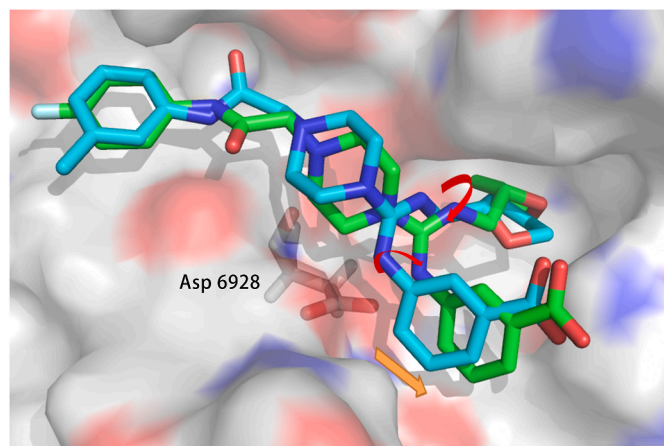


Fig. 6. C1 and C2 were superimposed at the binding pocket. The Yellow translation arrow and red rotation arrow represent the hypothesis in the text, which to explore the possibility of forming hydrogen bond between C1 and Asp6928. Blue for C1, green for C2. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Contributions of various energy components to the binding free energy (kcal/mol).

Compound	$\Delta G_{\text{bind}}^{\text{a}}$	$\Delta G_{\text{bind}}^{\text{Coulomb}}^{\text{b}}$	$\Delta G_{\text{bind}}^{\text{Covalent}}^{\text{c}}$	$\Delta G_{\text{bind}}^{\text{Hbond}}^{\text{d}}$	$\Delta G_{\text{bind}}^{\text{Lipo}}^{\text{e}}$	$\Delta G_{\text{bind}}^{\text{vdW}}^{\text{f}}$	$\Delta G_{\text{bind}}^{\text{Packing}}^{\text{g}}$	$\Delta G_{\text{bind}}^{\text{SolvGB}}^{\text{h}}$
Sinefungin	-51.318	-12.620	1.213	-5.437	-8.122	-49.469	0.000	22.924
C1	-59.983	-16.995	5.673	-3.536	-18.079	-56.289	-0.597	29.616
C2	-70.916	-24.058	-2.692	-1.911	-18.154	-51.977	-1.078	28.976
C3	-52.439	-21.808	2.741	-3.777	-14.672	-43.080	-0.365	28.379
C4	-29.571	-41.275	7.169	-3.477	-12.430	-52.678	-0.910	74.102
C5	-38.726	-16.612	5.922	-3.340	-16.133	-40.070	0.008	31.239
C6	-47.290	-11.516	2.653	-3.007	-13.941	-44.122	-0.415	23.061
C7	-54.859	-26.419	6.347	-3.657	-15.608	-45.450	-0.681	30.406
C8	-18.097	-45.129	5.969	-3.112	-10.273	-33.485	-0.013	67.861
C9	-48.864	-22.029	-0.127	-2.734	-10.839	-43.763	-0.479	30.838
C10	-48.754	-18.648	2.660	-1.612	-14.177	-38.669	0.015	21.480

Note: a free binding energy; B the contribution of Coulomb on the free binding energy; c the contribution of covalent on free binding energy; d the contribution of hydrogen bond on free binding energy; e the contribution of lipophilic on free binding energy; f the contribution of van der Waals on the free binding energy; g contribution of stacking on binding energy; h the contribution of polar solvation on free binding energy.

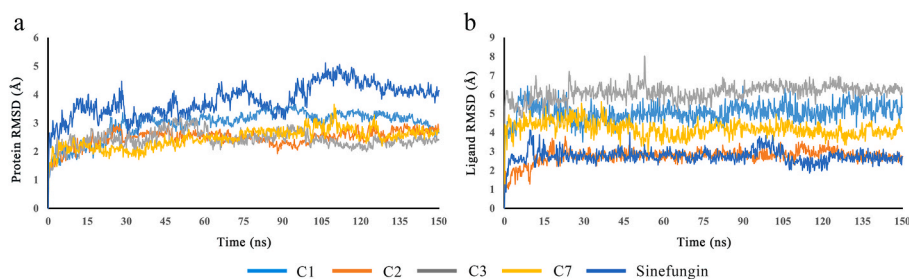


Fig. 7. The RMSD trajectory of the Complexes of nsp16 with four candidate compounds (C1, C2, C3 and C7) and sinefungin during the 150 ns simulation. (a) The stability of protein structure was analyzed during the simulation. (b) Stability analysis of ligand in binding pocket during simulation.

the simulation, the position at the binding pocket for compounds besides C2 deviate greatly from the initial docking positions (RMSD > 4 Å). However, the result of C2 is similar to sinefungin with the position at the binding pocket deviating only slightly from the initial docking position after the simulation. In conclusion, at the end of the molecular dynamics simulation, the four compounds maintained a relatively stable state in the pocket, but the interaction of C1, C3 and C7 in the binding pocket could be changed by the obvious position deviation compared with docking pose.

By analyzing the histogram of the interaction in the simulation process (Fig. 8c and d), it evident that the critical hydrogen bonds (interactions fraction was 0.995 and 0.901 respectively) and ionic interactions (0.606 and 0.887 respectively) between C2 and Asp6897, Asp6928 are the main interaction. The RMSD value of C2 at the binding pocket was maintained between 2-3 Å, so the critical hydrogen bond with the key amino acid Asp6928 in the docking simulation was retained during the molecular dynamics simulation. In addition, new main hydrogen bonds were formed between the fluorobenzene linked imide group and Asp6897, Tyr6930 (0.754). The change of position in the binding pocket brought a positive effect on C1, which forms critical hydrogen bonds with key amino acids Asp6928 (0.645) and 6897 (0.709). Main hydrogen bonds formed by the carboxylic acid on the C1 benzene ring with Lys6874 (0.577) and the carbonyl oxygen with Asn6899 (0.907) did not appear in the docking simulation (Fig. 8a and b). However, for C3 and C7, obvious displacement change brings negative effects. In the docking simulation, both C3 and C7 can form a hydrogen bond with Asp6928 (Table 2), but this hydrogen bond was not stable in the molecular dynamics simulation. Upon stable binding, C3 and C7 deviate from the initial docking position and lose the dominant interaction mode with key amino acids (Fig. 8e, f, g and h). The carboxylic acid at C3 terminal forms strong hydrogen bonds with Leu6855 (0.399) and Thr6854 (1.428). The nitrogen cation on C7 piperazine forms a strong water bridge with Lys7051 (0.904), and the amino groups that linked to two benzene form a strong hydrogen bond with Thr6854

(0.970) and Thr6856 (0.974).

In summary, molecular dynamics simulation verified our doubts about the docking simulation results. In the molecular dynamics simulation, the candidate compounds with superior free binding energy were stable in the binding pocket. The results showed that C1 and C2 could form continuous and stable hydrogen bonds with Asp6897 and Asp6928, while C3 and C7 could not reproduce the hydrogen bonds formed by the docking simulation results with Asp6928. Among the four compounds, the change of RMSD values of protein and ligand during both the molecular dynamics simulation and the interaction mode analysis of C2 was similar to sinefungin C2 also obtained the optimal value in MM/GBSA calculation.

3.7. MM/PBSA energy decomposition

The results of Desmond MD simulation showed that the candidate compounds C1 and C2 could form continuous and stable interactions with the key amino acids of nsp16. In order to further explore the energy contribution of these key amino acids in the two candidate compounds and sinefungin, MM/PBSA calculations were performed on the complex system composed of the three compounds and nsp16 after Gromacs MD simulation. The RMSD values of the C α atoms were monitored along the whole MD trajectory, as shown in Fig. 9 All the trajectories indicated that the MD simulations arrived at equilibrium state. MM/PBSA binding free energy calculation results show that the candidate compound and sinefungin can stably bind to the nsp16 binding pocket (Fig. S5). Fig. 10 shows the results of amino acid energy decomposition. Amino acids whose energy value is lower than -0.5 kcal/mol were used for analysis. During the binding of sinefungin to nsp16, the polar interaction ($\Delta G_{\text{electrostatic}}$ and $\Delta G_{\text{Polar Solvation}}$) affects the binding stability to the greatest extent. The histogram shows that the key amino acids (Asp6912, Cys6913, Asp6897 and Asp6928) all have favorable energy contribution, and the electrostatic potential energy contributions are more prominent than other amino acids. This suggests that there are

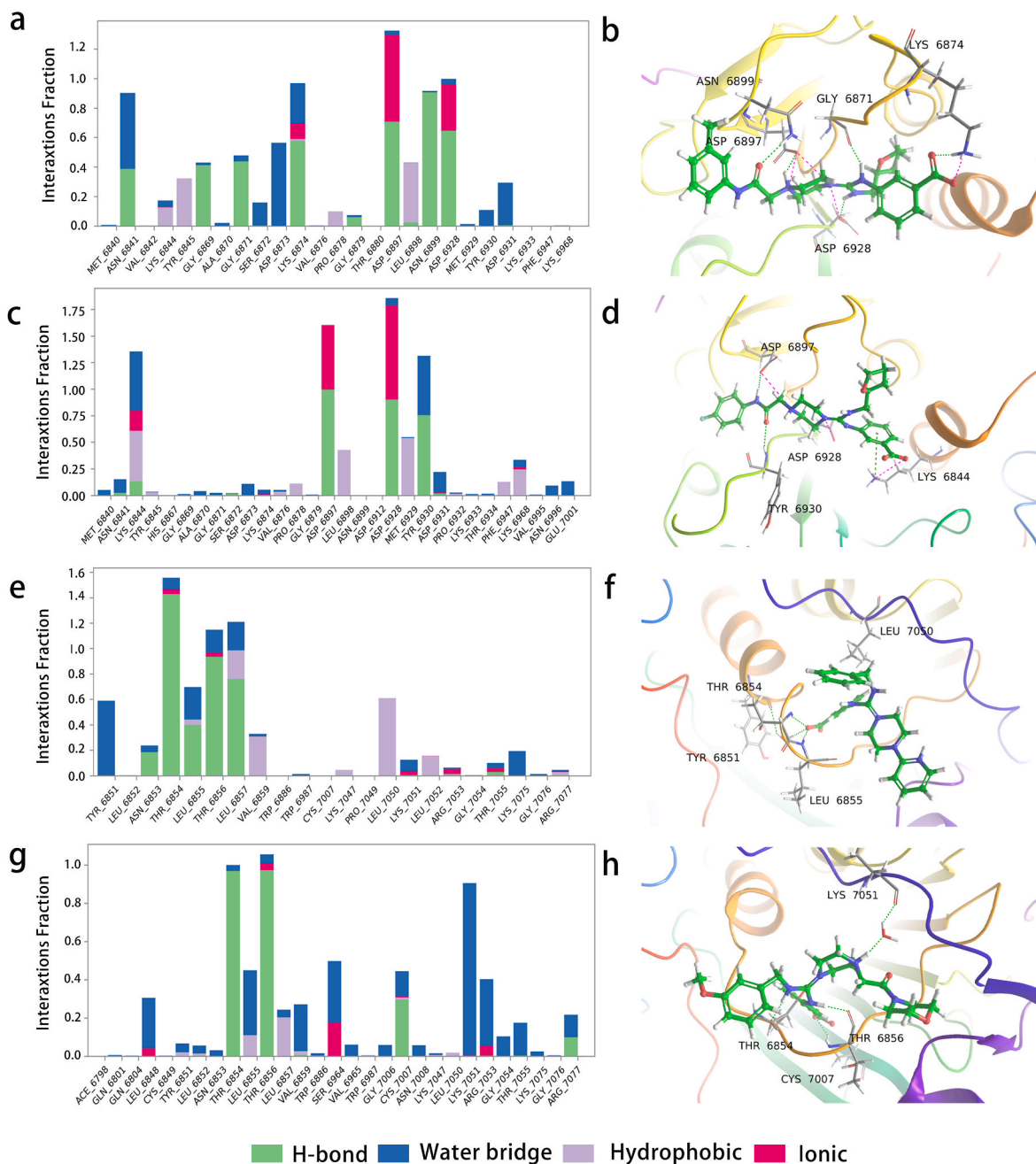


Fig. 8. (a, c, e and g) The histogram of protein-ligand contact over the course of the trajectory. (b, d, f and h) Stable protein ligand complex system generated by molecular dynamics simulation of the 3D interaction mode of the. The green, pink, and dark green lines represent hydrogen bonds, salt bridges, and π -cation interaction respectively. a and b represent C1, c and d represent C2, e and f represent C3, moreover g and h represent C7. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

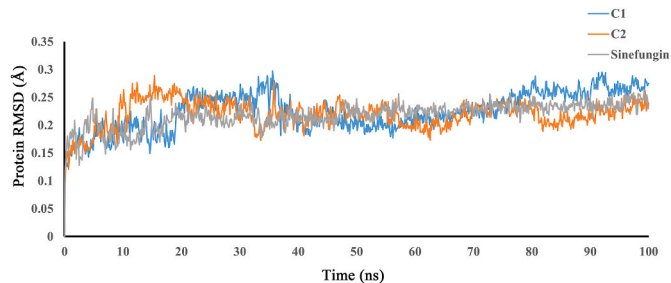


Fig. 9. RMSD values of Gromacs MD simulation.

strong hydrogen bonds between sinefungin and these amino acids. Asp6897 and Asp6928 also have crucial electrostatic potential energy contributions in the binding process of C1 and C2, and the total energy contribution is greater than other amino acids. Similar to sinefungin, Asp6897 produced the ultimate contribution value. The results of amino acid energy decomposition further verified the rationality involved in the determination of key amino acids. Candidate compounds C1 and C2 and sinefungin can stably bind to the docking pocket, and both form strong hydrogen bonds with the key amino acids Asp6897 and Asp6928.

4. Conclusion

In this work, we constructed a pharmacophore model that can

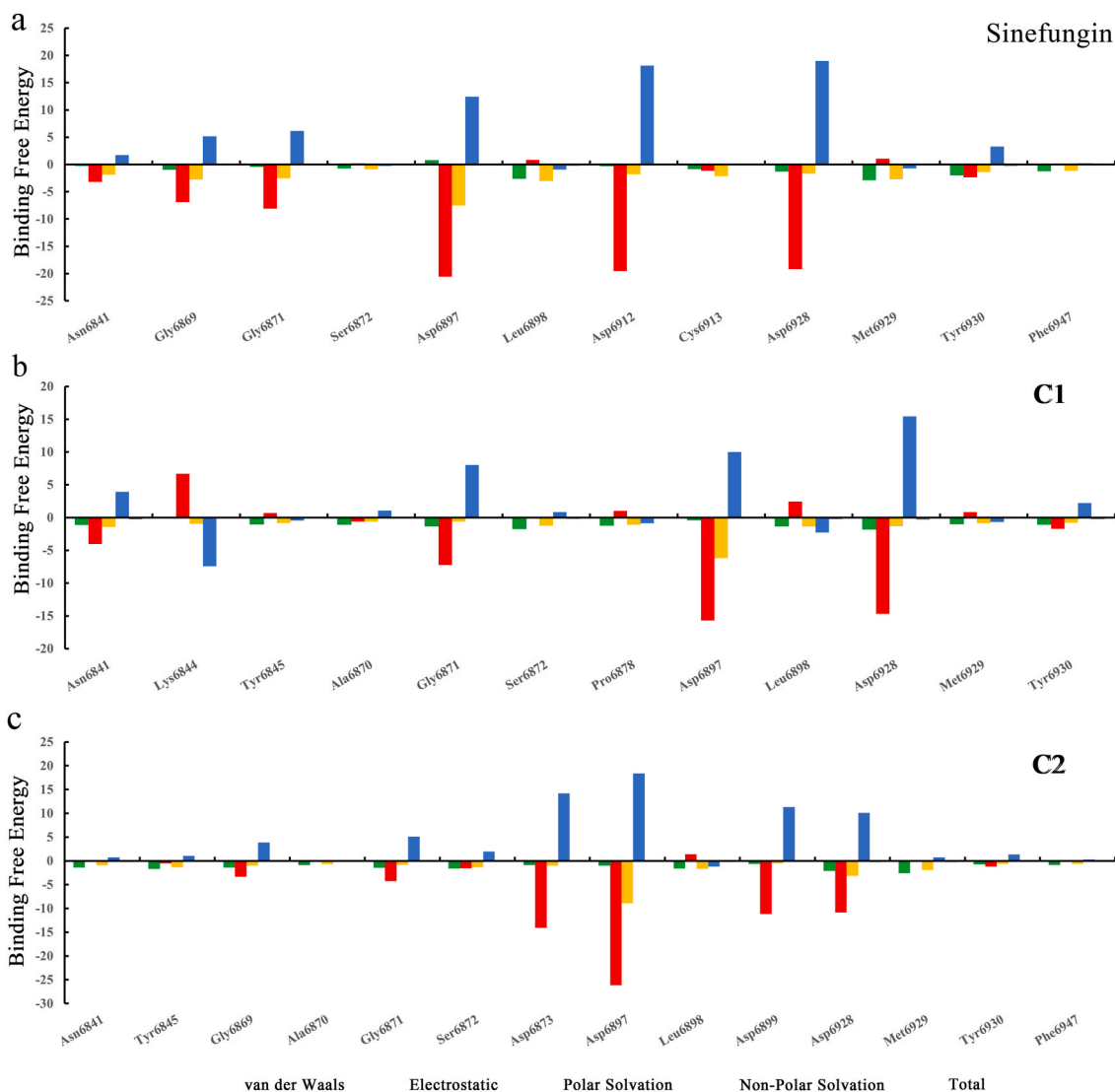


Fig. 10. Decomposition of the MM/PBSA binding energy on a per-residues in the binding pocket of each complex. (a) MM/PBSA energy decomposition scheme of nsp16/sinefungin complex. (b) MM/PBSA energy decomposition scheme of nsp16/compound 1 complex. (c) MM/PBSA energy decomposition scheme of nsp16/compound 2 complex.

accurately reflect the crucial interaction mode between sinefungin and nsp16. Molecular dynamic simulation was integrated into the structure-based pharmacophore construction, and the pharmacophore model was used to screen Bionet and Chembiv databases in order to find novel and potent nsp16 inhibitors. Through docking simulation, MM/GBSA, molecular dynamics RMSD stability analysis and main interaction analysis, it is concluded that C1 and C2 can stably bind to the binding pocket and form strong and continuous hydrogen bonds with the key amino acids Asp6897/Asp6928. MM/PBSA energy decomposition of sinefungin showed that Asp6897 and Asp6928 contributed major energy values during C1 and C2 binding, which was similar to sinefungin. Therefore C1 and C2 can be used as nsp16 inhibitors for further design and development.

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

No data was used for the research described in the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmgn.2022.108306>.

References

- [1] S. Yamayoshi, Y. Sakai-Tagawa, M. Koga, et al., Comparison of Rapid Antigen Tests for COVID-19 Viruses, *12*, 2020, p. 1420, 12.
- [2] K. Habas, C. Nganwuchu, F. Shahzad, et al., Resolution of coronavirus disease 2019 (COVID-19), *Expert Rev. Anti Infect. Ther.* 18 (12) (2020) 1201–1211.
- [3] M. Mohamadian, H. Chiti, A. Shoghli, et al., COVID-19: virology, biology and novel laboratory diagnosis, *J. Gene Med.* 23 (2) (2021) 3303.
- [4] M. Wang, R. Cao, L. Zhang, et al., Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro, *Cell Res.* 30 (3) (2020) 269–271.
- [5] S. Hussain, J. Pan, Y. Chen, et al., Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus, *J. Virol.* 79 (9) (2005) 5288–5295.

- [6] Y. Chen, C. Su, M. Ke, et al., Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex, *PLoS Pathog.* 7 (10) (2011), e1002294.
- [7] E. Decroly, C. Debarnot, F. Ferron, et al., Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex, *PLoS Pathog.* 7 (5) (2011), e1002059.
- [8] N. Vithani, M.D. Ward, M.I. Zimmerman, et al., SARS-CoV-2 Nsp16 Activation Mechanism and a Cryptic Pocket with Pan-Coronavirus Antiviral Potential *bioRxiv*, 2020.
- [9] G.K. Azad, Identification of novel mutations in the methyltransferase complex (Nsp10-Nsp16) of SARS-CoV-2, *Biochem. Biophys. Rep.* 24 (2020), 100833.
- [10] D. Kim, J.Y. Lee, J.S. Yang, et al., The Architecture of SARS-CoV-2 Transcriptome *Cell*, 181, 2020, pp. 914–921, 4.
- [11] M. Sedova, L. Jaroszewski, A. Alisoltani, et al., Coronavirus3D: 3D Structural Visualization of COVID-19 Genomic Divergence *Bioinformatics*, 36, 2020, pp. 4360–4362, 15.
- [12] I. Sola, F. Almazán, S. Zúñiga, et al., Continuous and discontinuous RNA synthesis in coronaviruses, *Ann. Rev. Virol.* 2 (1) (2015) 265–288.
- [13] T. Viswanathan, S. Arya, S.H. Chan, et al., Structural basis of RNA cap modification by SARS-CoV-2, *Nat. Commun.* 11 (1) (2020) 3718.
- [14] A.K. Maurya, N. Mishra, In silico validation of coumarin derivatives as potential inhibitors against Main Protease, NSP10/NSP16-Methyltransferase, Phosphatase and Endoribonuclease of SARS CoV-2, *J. Biomol. Struct. Dyn.* (2020) 1–16.
- [15] Y. Furuichi, A.J. Shatkin, Viral and cellular mRNA capping: past and prospects, *Adv. Virus Res.* 55 (2000) 135–184.
- [16] S. Daffis, K.J. Szretter, J. Schriewer, et al., 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members, *Nature* 468 (7322) (2010) 452–456.
- [17] R. Züst, L. Cervantes-Barragan, M. Habjan, et al., Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor, *Mda5 Nat. Immunol.* 12 (2) (2011) 137–143.
- [18] Y. Chen, D. Guo, Molecular mechanisms of coronavirus RNA capping and methylation, *Virol. Sin.* 31 (1) (2016) 3–11.
- [19] P. Krafcikova, J. Silhan, R. Nencka, et al., Structural analysis of the SARS-CoV-2 methyltransferase complex involved in RNA cap creation bound to sinefungin, *Nat. Commun.* 11 (1) (2020) 3717.
- [20] M.A. El Hassab, T.M. Ibrahim, S.T. Al-Rashood, et al., In silico identification of novel SARS-COV-2 2'-O-methyltransferase (nsp16) inhibitors: structure-based virtual screening, molecular dynamics simulation and MM-PBSA approaches, *J. Enzym. Inhib. Med. Chem.* 36 (1) (2021) 727–736.
- [21] B.J.M. Aldahham, K. Al-Khafaji, M.Y. Saleh, et al., Identification of naphthyridine and quinoline derivatives as potential Nsp16-Nsp10 inhibitors: a pharmacoinformatics study, *J. Biomol. Struct. Dyn.* (2020) 1–8.
- [22] A. Bhardwaj, S. Sharma, S.K. Singh, Molecular docking studies to identify promising natural inhibitors targeting SARS-CoV-2 nsp10-nsp16 protein complex *Turk. J. Pharmacol. Sci.* 19 (1) (2022) 93–100.
- [23] T.H. Fatoki, O. Ibraheem, I.O. Ogunyemi, et al., Network analysis, sequence and structure dynamics of key proteins of coronavirus and human host, and molecular docking of selected phytochemicals of nine medicinal plants, *J. Biomol. Struct. Dyn.* 39 (16) (2021) 6195–6217.
- [24] A. Malik, M. Kohli, N.A. Jacob, et al., In silico screening of phytochemical compounds and FDA drugs as potential inhibitors for NSP16/10 5' methyl transferase activity, *J. Biomol. Struct. Dyn.* (2021) 1–13.
- [25] M. Rosas-Lemus, G. Minasov, L. Shuvalova, et al., High-resolution Structures of the SARS-CoV-2 2'-O-Methyltransferase Reveal Strategies for Structure-Based Inhibitor Design *Sci Signal*, 13, 2020, 651.
- [26] G.M. Sastry, M. Adzhigirey, T. Day, et al., Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments, *J. Comput. Aided Mol. Des.* 27 (3) (2013) 221–234.
- [27] Y. Fu, T. Ye, Y.X. Liu, et al., Based on the virtual screening of multiple pharmacophores, docking and molecular dynamics simulation approaches toward the discovery of novel HPPD inhibitors, *Int. J. Mol. Sci.* 21 (15) (2020).
- [28] E.B. Lenselink, J. Louvel, A.F. Forti, et al., Predicting binding affinities for GPCR ligands using free-energy perturbation, *ACS Omega* 1 (2) (2016) 293–304.
- [29] D.J. Rhoades, D.H. Kinder, T.M. Mahfouz, A comprehensive ligand based mapping of the σ_2 receptor binding pocket, *Med. Chem.* 10 (1) (2014) 98–121.
- [30] J. Shen, C. Tan, Y. Zhang, et al., Discovery of potent ligands for estrogen receptor beta by structure-based virtual screening, *J. Med. Chem.* 53 (14) (2010) 5361–5365.
- [31] R.A. Friesner, R.B. Murphy, M.P. Repasky, et al., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, *J. Med. Chem.* 49 (21) (2006) 6177–6196.
- [32] K. Roos, C. Wu, W. Damm, et al., OPLS3e: extending Force Field Coverage for Drug-Like Small Molecules, *J. Chem. Theor. Comput.* 15 (3) (2019) 1863–1874.
- [33] M.J. Robertson, G.C.P. van Zundert, K. Borrelli, et al., GemSpot: A Pipeline for Robust Modeling of Ligands into Cryo-EM Maps Structure, 28, 2020, pp. 707–716, 6, e703.
- [34] J.L. Velázquez-Libera, F. Durán-Verdugo, A. Valdés-Jiménez, et al., LigRMSD: a Web Server for Automatic Structure Matching and RMSD Calculations Among Identical and Similar Compounds in Protein-Ligand Docking *Bioinformatics*, 36, 2020, pp. 2912–2914, 9.
- [35] Z. Liu, F. Guo, Y. Wang, et al., BATMAN-TCM: a bioinformatics analysis Tool for molecular mechANism of traditional Chinese medicine, *Sci. Rep.* 6 (2016), 21146.