ORIGINAL ARTICLE

Identification of potential target endoribonuclease NSP15 inhibitors of SARS-COV-2 from natural products through high-throughput virtual screening and molecular dynamics simulation

Liang-Chang Hu[1] | **Chuan-Hua Ding**[2] | **Hong-Ying Li**[2] | **Zhen-Zhen Li**[2] | **Ying Chen**[3] | **Li-Peng Li**[4] | **Wan-Zhong Li**[3] | **Wen-Shan Liu**[2] ⓘ

[1]Department of Oncology, Affiliated Hospital of Weifang Medical University, Weifang, China

[2]Shandong Key Laboratory of Clinical Applied Pharmacology, Department of Pharmacy, Affiliated Hospital of Weifang Medical University, Weifang, China

[3]School of Pharmacy, Weifang Medical University, Weifang, China

[4]Tianjin Key Laboratory on Technologies Enabling Development of Clinical Therapeutics and Diagnostics (Theranostics), School of Pharmacy, Tianjin Medical University, Tianjin, China

Correspondence
Wen-Shan Liu, Shandong Key Laboratory of Clinical Applied Pharmacology, Department of Pharmacy, Affiliated Hospital of Weifang Medical University, Weifang 261041, Shandong Province, China.
Email: liuwenshan@wfmc.edu.cn

Wan-Zhong Li, School of Pharmacy, Weifang Medical University, Weifang 261053, Shandong Province, China.
Email: liwz@wfmc.edu.cn

Abstract

SARS-CoV-2 wreaks havoc around the world, triggering the COVID-19 pandemic. It has been confirmed that the endoribonuclease NSP15 is crucial to the viral replication, and thus identified as a potential drug target against COVID-19. The NSP15 protein was used as the target to conduct high-throughput virtual screening on 30,926 natural products from the NPASS database to identify potential NSP15 inhibitors. And 100 ns molecular dynamics simulations were performed on the NSP15 and NSP15-NPC198199 system. In all, 10 natural products with high docking scores with NSP15 protein were obtained, among which compound NPC198199 scored the highest. The analysis of the binding mode between NPC198199 and NSP15 found that NPC198199 would form H-bond interactions with multiple key residues at the catalytic site. Subsequently, a series of post-dynamics simulation analyses (including RMSD, RMSF, PCA, DCCM, RIN, binding free energy, and H-bond occupancy) were performed to further explore inhibitory mechanism of compound NPC198199 on NSP15 protein at the molecular level. The research strongly indicates that the 10 natural compounds screened can be used as potential inhibitors of NSP15, and provides valuable information for the subsequent drug discovery of anti-SARS-CoV-2.

Practical applications

Natural products play an important role in the treatment of many difficult diseases. In this study, high-throughput virtual screening technology was used to screen the natural product database to obtain potential inhibitors against endoribonuclease NSP15. The binding mechanism between natural products and NSP15 was investigated at the molecular level by molecular dynamics technology so that it is expected to become candidate drugs for the treatment of SARS-CoV-2. We hope that our research can provide new clue to combat COVID-19 and overcome the epidemic situation as soon as possible.

KEYWORDS
MD simulation, natural compounds, NSP15, SARS-CoV-2, virtual screening

---

Liang-Chang Hu and Chuan-Hua Ding contributed equally to this work.

1 | INTRODUCTION

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) causes the 2019 Coronavirus Disease (COVID-19) pandemic. The virus emerged in Wuhan, China at the end of December 2019 and spread rapidly. Soon, the World Health Organization (WHO) announced the COVID-19 a continuing epidemic. SARS-CoV-2 is highly aggressive, with higher infection and mortality rates than previous viruses (Petrosillo et al., 2020). The main symptoms of the virus infection are high fever, cough, pain, and shortness of breath. In severe cases, it can lead to body failure and death. However, asymptomatic infections carried by the virus have been found all over the world (Kumar et al., 2021; Liu et al., 2021). The virus can spread through droplets when coughing, talking, or sneezing. As of October 30, 2021, the virus has infected more than 246 million people and killed more than 5 million people, brought serious disasters to countries around the world. At present, several vaccines against the virus are on the market, but the vaccine protection rate is not 100%. Especially after the virus mutates, the protective effect of the vaccine may be further reduced. Therefore, we need to continue to develop anti-SARS-CoV-2 drugs to provide multiple protections for human health.

SARS-CoV-2 is an enveloped non-segmented positive-sense RNA virus, belonging to β-Coronavirus family, with a genome size of 29.9 kb. The virus consists of about 3000 base pairs, which are responsible for encoding important proteins, including nucleocapsid protein (N), spike protein (S), envelope protein (E), membrane protein (M), and non-Structural proteins (Nsps) (Gordon et al., 2020; Unchwaniwala & Ahlquist, 2020). SARS-CoV-2 binds to the human host cell receptor, enters the host cell, and uses the host cell system to release RNA into the cytoplasm to start the translation process. During this period, ORF1a and ORFab of RNA are translated into two kinds of polyproteins (Pp1a and Ppab), which are then processed by proteases to release non-structural viral proteins (NSPs) (Nakagawa et al., 2016; Ni et al., 2020; Yan et al., 2020). These Nsps, including NSP12, NSP13, NSP14, and NSP15, are associated with different viral functions, such as the formation of replicase transcriptase complexes (Kumar et al., 2021; Prajapat et al., 2020).

Endoribonuclease NSP15, also known as uridine-specific endorphinase, preferentially cleaves the 3′end of uridine. It is encoded by the coronavirus as an RNA-processing enzyme, and plays a key role in the viral replication. Studies have found that inhibiting NSP15 protein can slow down virus replication (Deng & Baker, 2018; Kumar et al., 2021). Therefore, NSP15 can be used as a drug target for the development of potential SARS-CoV-2 inhibitors. The catalytic site of the NSP15 protein structure contains six key residues, namely His235, His250, Lys290, Ser294, Thr341, and Tyr343, and these residues are generally conserved in SARS-CoV-2. In addition, residues His235, His250, and Lys290 will form a catalytic triad (Kim et al., 2020; Savale et al., 2021).

Natural compounds occupy an important position in drug development, and the toxicity of natural products is less than that of chemical drugs. Researchers have been actively looking for potential natural medicines for the treatment of various diseases (such as malaria and cancer) (Harvey, 2005; Ravindranath, 2010). The NPASS database contains 30,926 natural products and can be accessed free of charge. This database provides the activity values of natural products (such as IC_{50}, EC_{50}, and MIC) reported in the literature for enzyme proteins or cell experiments (Zeng et al., 2018). This research aims to perform high-throughput virtual screening, molecular docking, and molecular dynamics (MD) simulation study on the NPASS natural compounds database to obtain potential anti-COVID-19 inhibitors.

2 | MATERIALS AND METHODS

2.1 | System preparation

The crystal structure of NSP15 protein (PDB ID: 6WXC) was obtained from the protein database bank (PDB) (Kim et al., 2020). Before docking, the crystal structure was processed by the "Prepare Protein" module in Discovery Studio (DS) v4.5 software, including removing water, adding hydrogen atoms, assigning bond sequences, treating metals, treating disulfides, and supplementing missing residues. The SDF file containing 30,926 natural products was downloaded from the NPASS database. These natural products were processed by the "Prepare Ligands" module in DS v4.5, including maintaining ionization, desalting, and generating all possible three-dimensional conformations.

2.2 | High-throughput virtual screening based on molecular docking

As an effective tool, high-throughput virtual screening was used in drug development (Liu et al., 2021). In this study, DS v4.5 software was used for high-throughput virtual screening based on molecular docking. The binding site of NSP15 protein was first defined through the "Define and Edit Binding Sites" module. The "From Current Selection" module was used to construct binding pockets around the key residues His235, His250, Lys290, Ser294, Thr341, and Tyr343 at the catalytic site, which was shown as a sphere with a radius of 10.66, and its coordinates were $X = 63.38$, $Y = 69.29$, and $Z = 30.71$. Through continuous screening of LibDock module and docking optimization of CDOCKER module, natural compounds with high docking score and binding affinity would be obtained. The docking result was visualized using DS v4.5 and displayed as Libdock score and -CDOCKER_ENERGY score. The higher the docking Score, the better the combination of protein and ligand.

2.3 | MD simulation

The pose with the highest docking score obtained from the virtual screening was further subjected to MD simulation. MD simulation

of NSP15 and screening inhibitor complexes was performed to evaluate their conformational movement and structural stability at the molecular level (Alffenaar et al., 2018). The MD simulations of the NSP15 system and the NSP15-NPC125597 system were carried out in the GROMACS 4.5.5 software package with the CHAMM27 force field (Pol-Fachin et al., 2009). First, the "topology" file was generated, which contained the bonding parameters and non-bonding parameters of the simulation system. Then, the simulation system was placed in a dodecahedron periodic box, the simulation system was kept at the center and with a distance of 10 Å from the edges, and simple single point charge (SPC) water molecules were filled to keep the system in balance. Before energy minimization, appropriate amounts of Na^+ and Cl^- ions were added to neutralize the system (Joung & Cheatham, 2008). Subsequently, the steepest descent method was used to eliminate conflicting and inappropriate geometric shapes and minimize the system energy to 1,000 kJ/mol/nm. Two ensembles of NVT (constant particle count, volume, and temperature) and NPT (constant particle count, pressure, and temperature) were used to further balance the system. The first round of balance involved the use of a speed adjusting thermostat for a 500 ps NVT simulation, and the second round of balance involved the use of a barostat for a 500 ps NPT simulation. In addition, the LINCS algorithm was used to restrain all the keys in the simulation system (Hess, 2008). Finally, 100 ns MD simulations were carried out, and the simulation trajectories were analyzed.

2.4 | RMSD and RMSF

After the MD simulation, the stability of the simulated trajectories should be evaluated to ensure the accuracy and reliability of the simulation result. Here, the simulated stability was assessed by calculating the root mean square deviation (RMSD) of the time-dependent structure relative to the starting structure (Dixit et al., 2006; Guilbert & James, 2008). In addition, the root-mean square fluctuation (RMSF) for the protein residues was calculated by the process of g_rms, and evaluated the flexibility of residues in the simulation system based on the RMSF value in the whole simulation time (Bao et al., 2009; Chen et al., 2018).

2.5 | Principal component analysis

Principal component analysis (PCA), as a statistical tool, was widely used to analyze the internal motion and conformational changes of biological macromolecules (Fakhar et al., 2017). The GROMACS 4.5.5 software package was applied to obtain the covariance matrix of the Cα atom position in the trajectory of the simulation system. The covariance matrix Cij of each Cα i and j was represented by the following integration:

$$C_{ij} = \left\langle \left(x_i - \langle x_i \rangle \right) \left(x_j - \langle x_j \rangle \right) \right\rangle (i, j = 1, 2, 3, \ldots, N)$$

In the formula, x_i or x_j represented the Cartesian coordinates of the ith or jth Ca atom, $\langle x_i \rangle$ or $\langle x_j \rangle$ represented the time average of all configurations selected in the simulation, and N represented the number of Cα atoms. The eigenvector of the matrix was used as the principal component (PC), which represented the projection of the trajectory on the main mode. The corresponding eigenvalue represented the magnitude in the direction of the PC. Usually, the first few principal components (PC1, PC2, and PC3) were used to describe the overall movement of the simulation system (Mesentean et al., 2006; Zhou et al., 2016). The PCA scatter plot was obtained by Bio3D software (Grant et al., 2006).

2.6 | Dynamic cross-correlation map

The cross-correlation was expressed as a three-dimensional matrix, which graphically showed time-related information about correlation motion among the domains of the biomacromolecule system (Ndagi et al., 2017; van den Berg & Hoheisel, 1990). To evaluate the correlation matrix across all Cα atoms, dynamic cross-correlation map (DCCM) analysis was carried out for simulation systems to assess the dynamic correlation among different protein domains. The correlation matrix could be denoted by the cross-correlation coefficient (C_{ij}) and defined by the following formula (Wan et al., 2013; Xu et al., 2016):

$$C_{ij} = \frac{\left\langle \Delta r_i * \Delta r_j \right\rangle}{\sqrt{\left(\left\langle \Delta r_i^2 \right\rangle * \left\langle \Delta r_j^2 \right\rangle \right)}}$$

Here, i and j represented the atom or residue of ith and jth, respectively. Δr_i and Δr_j were the displacement vector corresponding to ith and jth atom (or residue), and "$\langle\ \rangle$" indicated an ensemble average. The value of correlation matrix ranged from −1 to 0 (negative correlation) and from 0 to 1 (positive correlation). The higher the absolute value of matrix, the stronger the correlation between atoms (or residues) (Kasahara et al., 2014).

2.7 | Residue interaction network

Residue interaction network (RIN) was used as an analysis tool to explore differences in the interaction between residues in simulation system (Mehla & Ramana, 2016). The last 95 ns simulated trajectory of each system was used to construct a RIN. In RIN, the definition of residual interaction network and protein contact diagram was specified by two parameters. The node represented amino acid residue, and the edge between them represented non-covalent interactions, including Van der Waals' force (VDW), hydrogen bond (H bond), salt bridges, and pi–pi interactions. The edge represented the contact between atoms with at least one WDW interaction. Cytoscape (Shannon et al., 2003) and RINalyzer (Doncheva et al., 2011) software were applied to visualize RIN.

2.8 | Binding free energy

Free energy of binding played a pivotal role in exploring the dynamic interaction between protein and ligand. The molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) method was usually used to calculate the free energy of binding (Homeyer & Gohlke, 2012; Kumari et al., 2014). The free energy of each energy term was defined according to the following equation:

$$\Delta G_{bind} = G_{complex} - (G_{ligand} + G_{receptor})$$

$$\Delta G = \Delta E_{MM} + \Delta G_{sol} - T\Delta S$$

$$\Delta E_{MM} = \Delta E_{ele} + \Delta E_{vdw}$$

$$\Delta G_{solvation} = \Delta G_{polar} + \Delta G_{nonpolar}$$

In the equation, $G_{complex}$, $G_{receptor}$, and G_{ligand} indicated the free energies of complex, ligand, and receptor, respectively. E_{MM} indicated molecular mechanical components in the gas phase, G_{sol} indicated the free energy of solvation, and TS indicated the contribution of entropy to free energy, where T indicated temperature and S indicated entropy. E_{MM} included electrostatic interaction (E_{elec}) and Van Dehua interaction (E_{vdw}). Additionally, solvation free energy (G_{sol}) included electrostatic free energy (G_{polar}) and non-electrostatic free energy ($G_{nonpolar}$) (Fang et al., 2016; Sang et al., 2017).

3 | RESULTS AND DISCUSSION

3.1 | Analysis on molecular docking

3.1.1 | Result on high-throughput virtual screening

In this study, 30,926 natural small molecule compounds from the NPASS database were subjected to high-throughput virtual screening, aiming to obtain compounds with high docking scores and binding affinity to NSP15 protein. The virtual screening pocket was constructed based on the key residues His235, His250, Lys290, Ser294, Thr341, and Tyr343 in the catalytic site of the NSP15 protein. Before screening, the reliability of the docking model needed to be evaluated. The inhibitor Tipiracil in the crystal structure of 6WXC was docked to the crystal structure again, and it was found that the RMSD value between the docked ligand and crystallized ligand was 0.92 Å, which showed a good overlap between them, indicating that the constructed screening model was reliable. Subsequently, the high-throughput virtual screening based on the LibDock module was performed on these natural compounds, and the top 2,000 compounds were selected according to the LibDockScore values. A precise molecular docking study based on the CDOCKER module was performed on these 2,000 compounds, and 10 compounds were obtained according to the -COOCKER_ENERGY values, which acted as potential inhibitors of the NSP15 protein. The top 10 natural products with high docking scores with NSP15 protein are shown in Table 1.

3.1.2 | Investigation on the binding pattern

The natural product NPC198199 with the highest docking score was selected for detailed binding mode research. Figure 1a shows the binding site of NSP15 protein and compound NPC198199. Figure 1b shows the binding pocket and key interactions between NSP15 protein and compound NPC198199; Figure 1c shows the two-dimensional diagram of the interactions between NSP15 protein and compound NPC198199. It was found that the compound NPC198199 formed 10 stable H-bond interactions with residues His235, Gln245, Gly248, Lys290, Val292, Ser294, and Glu340, and formed stable VDW interactions with residues Asp240, Leu246, Gly247, His250, Cys293, Thr341, Pro344, and Lys345, and formed a stable pi–pi interaction with residue Trp333. The key residues His235, His250, Lys290, Ser294, Thr341, and Tyr343 related to the catalytic activity of the NSP15 protein almost all interacted with the compound NPC198199, which helped stabilize the conformation of the NSP15 protein and weaken its activity.
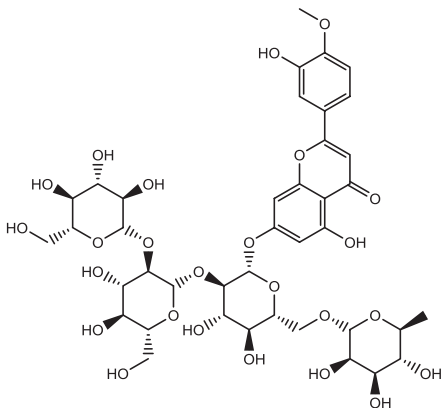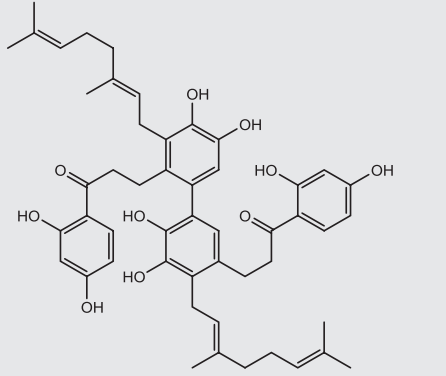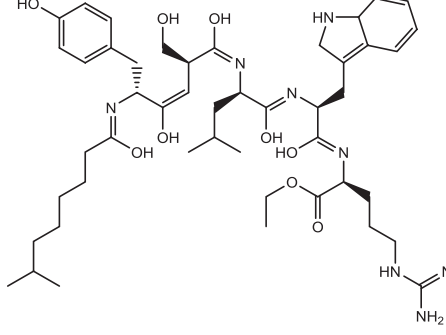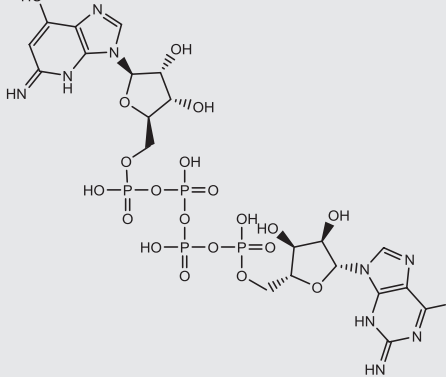
3.2 | Analysis on MD simulations

3.2.1 | Investigation on stability and flexibility of the NSP15 system and NSP15-NPC198199 complex system

In order to assess the stability of the simulation system, the RMSD values of all Cα atoms in the entire MD trajectories were obtained. Generally, the lower the RMSD value of the skeleton, the higher the stability of the system. Figure 2a shows the relationship between the RMSD of the NSP15 protein system and the NSP15-NCP198199 complex system over time. It could be found that the RMSD trajectories of the NSP15 protein system and the NSP15-NPC198199 complex system reached equilibrium around 5 ns and generated stable trajectories. The average RMSD value of the NSP15 protein system was 0.28 nm, and the average RMSD value of the NSP15-NPC198199 complex system was 0.24 nm. Therefore, the stability of the NSP15-NPC198199 complex system was higher than that of the NSP15 protein system, indicating that the compound NPC198199 played an active role in stabilizing the conformation of the NSP15 protein.

Each amino acid residue belonged to the protein system and determined the conformational characteristics of the protein. In order to obtain information on the flexibility of each amino acid residue in the protein system, RMSF of all side chain atoms of the protein was calculated. Generally, in the simulation system, the larger the RMSF value, the higher the flexibility of the residues. Conversely, the smaller the RMSF value, the lower the flexibility of the residues. Figure 2b shows the RMSF value of each residue of

TABLE 1 The top 10 natural products with high docking score with NSP15 protein obtained by virtual screening

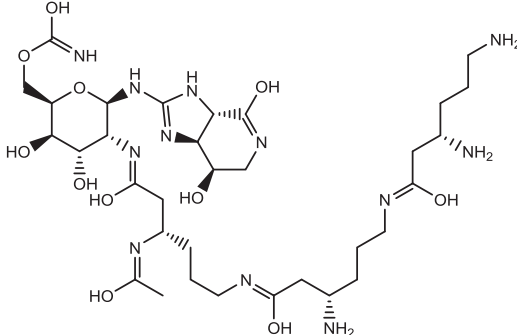| Natural products ID	Structures	LibDock score	-COOCKER_ENERGY
NPC198199		192.678	68.246
NPC35		190.278	68.011
NPC125597		190.705	67.925
NPC14590		182.909	67.316
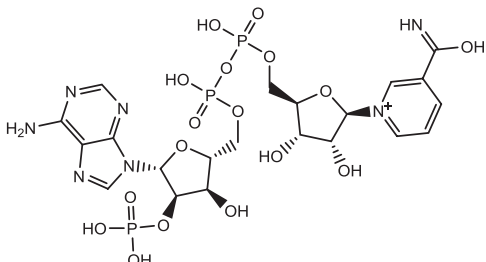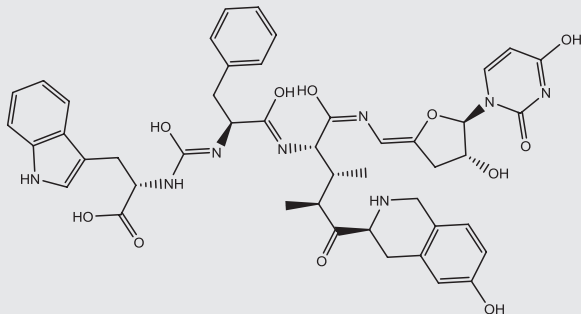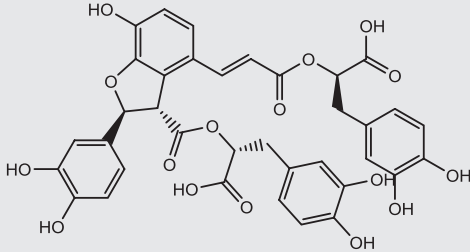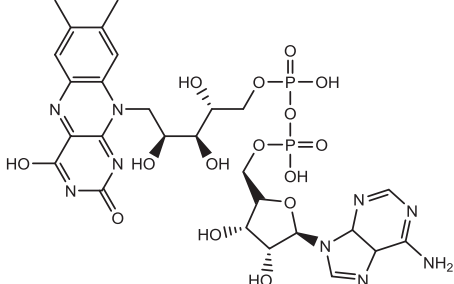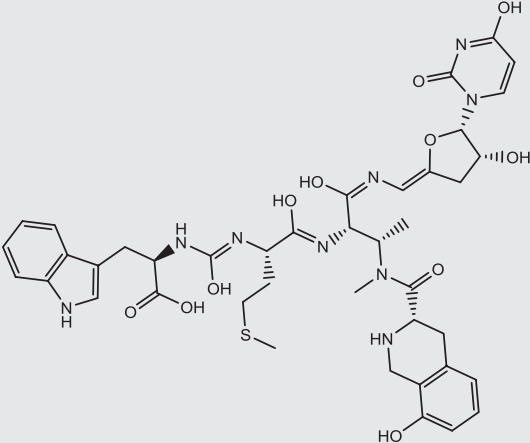
TABLE 1 (Continued)

Natural products ID	Structures	LibDock score	-COOCKER_ENERGY
NPC107160		182.938	67.018
NPC478009		179.059	66.842
NPC10897		180.382	66.229
NPC216403		179.104	65.312
NPC158055		178.926	65.147

TABLE 1 (Continued)

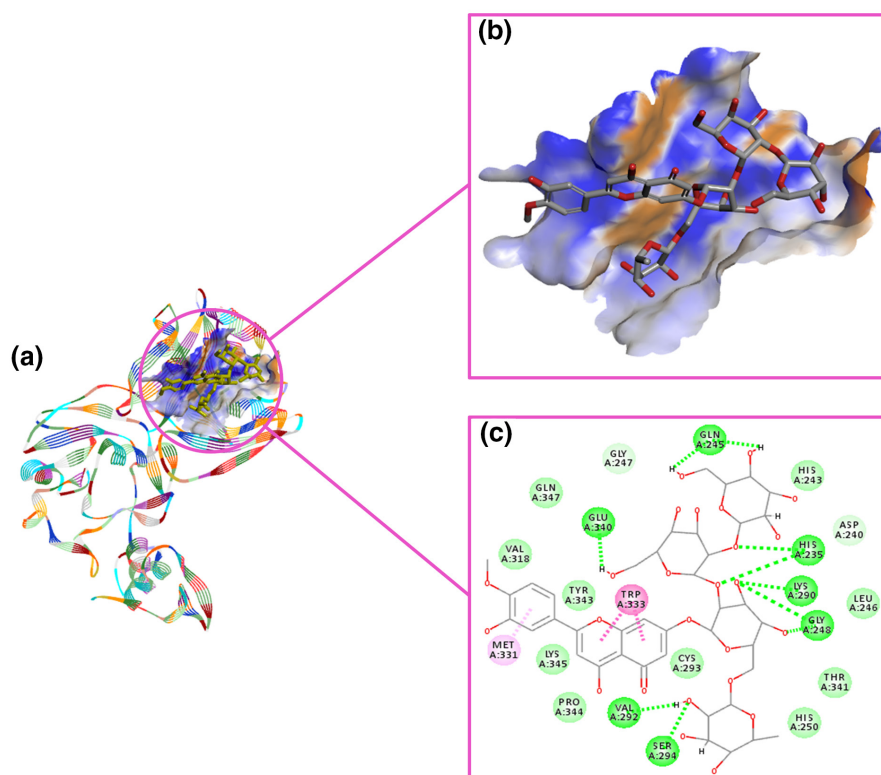| Natural products ID	Structures	LibDock score	-COOCKER_ENERGY
NPC478012		179.007	65.028

FIGURE 1 The binding mode of compound NPC198199 and NSP15 protein. (a) The docking site of the compound NPC198199 and NSP15 protein; (b) the enlarged view of the binding pocket of the compound NPC198199 and NSP15 protein; and (c) the interactions between the compound NPC198199 and NSP15 protein

the NSP15 system and the NSP15-NCP198199 complex system. It could be found that the RMSF of the NSP15 system and NSP15-NPC198199 complex system showed similar fluctuation trend. The calculation found that the average RMSF of the NSP15 system was 0.21 ns, and the average RMSF of the NSP15-NPC198199 complex system was 0.16 ns, which indicated that the NSP15-NPC198199 complex system exhibited lower fluctuation. It was worth noting

that the maximum RMSF difference between the NSP15 system and NSP15-NPC198199 complex system was more than 0.15 ns in some key regions associated with the catalytic active site (marked with the green boxes), which indicated that the binding of compound NPC198199 and NSP15 protein could significantly reduce the flexibility of residues near the catalytic site and stabilize the protein conformation.
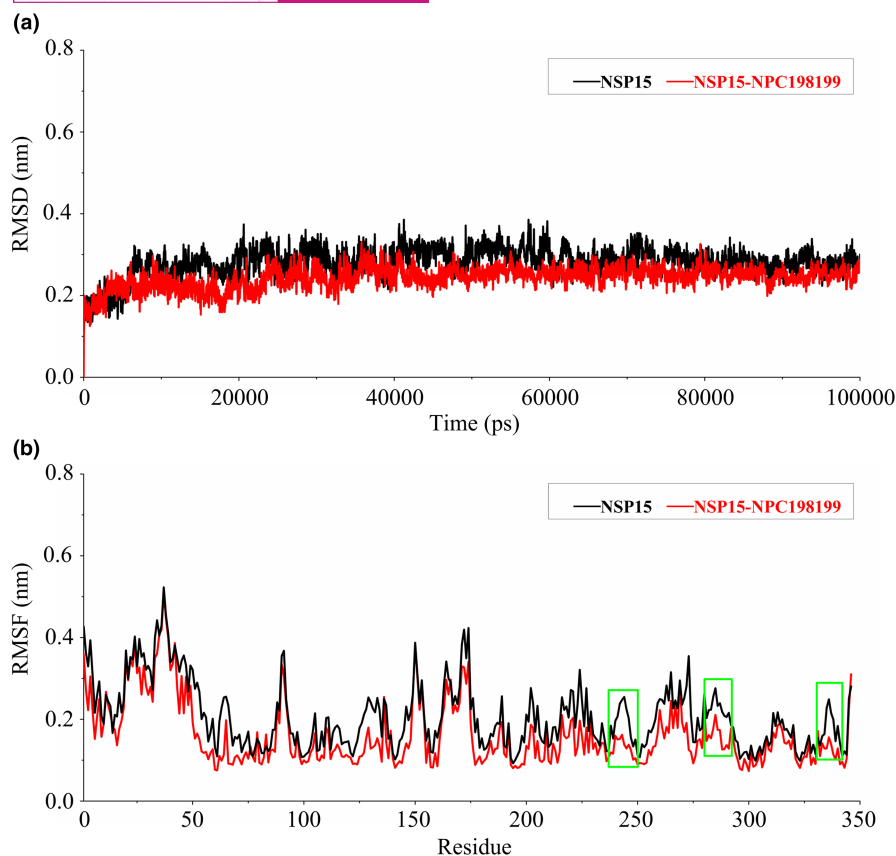
FIGURE 2 Assessment on the stability and flexibility of the systems. (a) The RMSD values of all Cα atoms for the NSP15 and NSP15-NPC198199 systems over 100 ns MD simulation; (b) the RMSF fluctuations of residues for NSP15 and NSP15-NPC198199 systems over 100 ns MD simulation. In addition, the green boxes highlighted the key areas where the two systems had large fluctuation differences

3.2.2 | Investigation on the conformational state and correlated motion of the NSP15 system and NSP15-NPC198199 complex system

PCA was performed on the intercepted 95 ns simulation trajectories of NSP15 protein system and NSP15-NPC198199 complex system to investigate the influence of compound NPC198199 on the conformational state of NSP15 protein. The eigenvector represents the overall motion of the Cα atom in the simulation system, and the percentage change of the fluctuation of the atomic position captured in each dimension is represented by the corresponding eigenvalues. According to Figure 3, the top 20 PCs of the NSP15 system and the NSP15-NPC198199 complex system accounted for 89.0% and 94.5% of the total contribution of the simulated trajectories, respectively. In the NSP15 system, the first two PCs (PC1 and PC2) contributed 59.2% and 8.5%, respectively, and the other PCs contributed less than 6.1%. In the NSP15-NPC198199 complex system, the first two PCs contributed 66.6% and 9.3%, respectively, and the other PCs contributed less than 5.8%. Therefore, the first two eigenvectors PC1 and PC2 captured most of the differences in the original distribution of the conformational space of the simulation system, indicating that the first two eigenvectors could basically reflect the changes in the conformational state. The PCA scatter plots of these two simulation systems were obtained by projecting along

the directions of the first two PC1 and PC2. The blue dot and red dot in the figure represented the unstable conformational state and the stable conformational state, respectively, and the simulation system periodically jumped between these two conformational states. Judging from the distribution of scattered points, there were significant differences in the conformational states of the two systems. In the NSP15 system, the distribution of scattered points was irregular and messy. However, in the NSP15-NPC198199 complex system, the scattered points were more organized, and the clustered red and blue points were evenly distributed parallel on both sides of the diagonal, indicating that the system occupied a smaller phase space. Therefore, the stability of NSP15-NPC198199 complex system was higher than that of NSP15 system, which stayed in sync with the analysis result of RMSD and RMSF.

To further explore the influence of the compound NPC198199 on the conformation of NSP15 protein, the application domain cross-correlation map (DCCM) analyses of all Cα atoms were performed by intercepting the last 95 ns simulation trajectories of the NSP15 system and the NSP15-NPC198199 complex system. DCCM displays a correlation diagram of the overall simulation system, which can be used to identify differences in relative motion between residues. As shown in Figure 4, the two-dimensional images of the DCCM of the NSP15 system and the NSP15-NPC198199 complex system are shown. The positively correlated
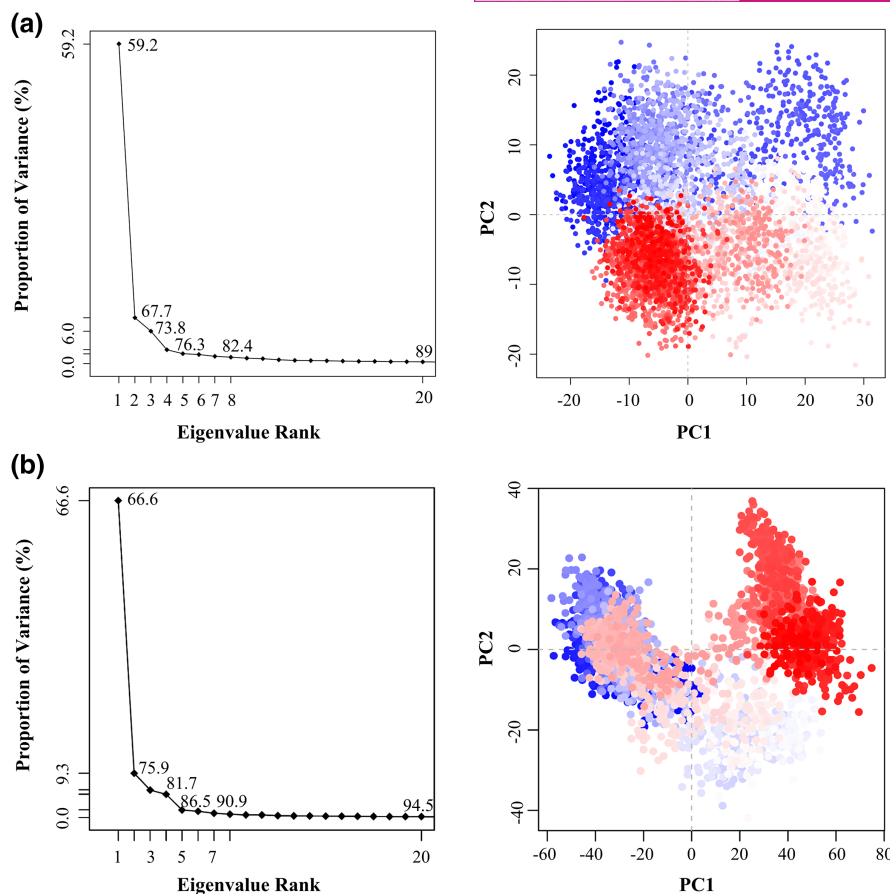
FIGURE 3 The variance contribution of the principal components and PCA scatter plots projected along the direction of the principal components PC1 and PC2 for NSP15 system (a) and NSP15-NPC198199 complex system (b)

motion is shown in cyan and the range is from 0 to 1, meaning that the residues move in the same direction and negatively correlated motion is shown in pink and range from −1 to 0, meaning that the residues move in the opposite direction. The darker the displayed color, the stronger the correlation. It could be seen from the figure that the correlated motion of the two systems showed similar motion trend as a whole, but it was obvious that the correlated motion of the two systems had changed. Especially in key areas related to catalytic site (marked as the black boxes), the two systems showed significant differences in correlated motion. Compared with the NSP15 system, in the NSP15-NPC198199 system, the negatively correlated motions of region (residues Val70-Thr99) and region (residues Met210-Phe303), region (residues Leu120-Gln160), region (residues Ile212-Gly247), region (residues Val166-Leu190), and region (residues Met210-His250) were significantly reduced. It was generally believed that the reduction of negatively correlated motion would reduce the flexibility of the protein. Therefore, the binding of the compound NPC198199 and NSP15 protein would reduce the negatively correlated motion of the complex system, which might make the catalytic site in the NSP15 protein difficult to expose to the substrate, and help stabilize its conformation, thereby achieving the purpose of inhibiting protein activity.

3.2.3 | The exploration on the differences of residue interaction between NSP15 system and NSP15-NPC198199 complex system

In order to explore the differences in the interaction of key residues inside the protein after NSP15 binding with the compound NPC198199, RINs were generated by intercepting the NSP15 system and NSP15-NPC198199 complex system at the last 95 ns simulation trajectories. Using NetworkAnalyzer and RINalyzer plugin of Cytoscape, the network differences between residues near catalytic site in NSP15 and NSP15-NPC198199 systems are shown (Figure 5). The three connecting lines of different colors between residues in the figure reflected the covalent and non-covalent interactions between residues. The red dotted line meant that the interaction existed only in NSP15-NPC198199 system, the green dotted line meant that the interaction existed only in NSP15 system, and the black solid line meant that the interaction existed in both systems. It was found that the interactions between residues in the NSP15-NPC198199 system were significantly increased and more new interactions were formed compared with the NSP15 system. Specifically, in the NSP15-NPC198199 complex system, residue His250 formed an additional H-bond interaction with Pro344 and an additional VDW interaction with residue Val295. Residue
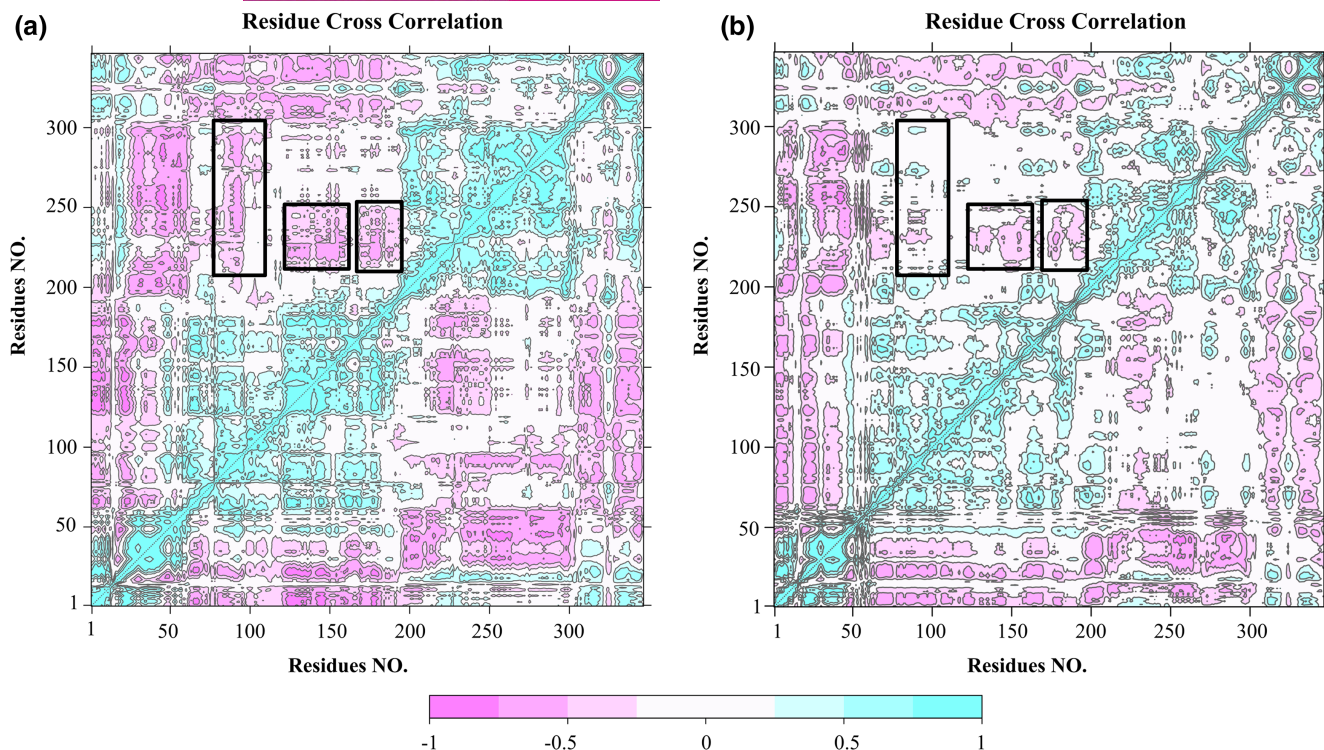
(a)

Residue Cross Correlation

(b)

Residue Cross Correlation

FIGURE 4 The analyses of domain cross-correlation map (DCCM) for NSP15 system (a) and NSP15-NPC198199 system (b). Moreover, the black boxes highlighted the key areas where the correlated motions of the two systems were significantly different

Ile281 formed additional H-bond interactions with Lys290 and Leu255, and an additional VDW interaction with residue Phe264. Residues His235 formed an additional H-bond interaction with Gly247. Residues Phe233 formed an additional H-bond interaction with Ile236. Residues Leu255 formed additional VDW interactions with Phe264 and Leu249. Residue Val276 formed additional VDW interactions with residues Ile296 and Pro344. Therefore, due to the binding of NSP15 and NPC198199, the interaction between residues in the complex system was obviously enhanced, which helped to stabilize the system conformation and inactivate the protein.

Furthermore, to explain the behavior of RIN, the topological parameters of each amino acid residue were defined, including the shortest path betweenness and closeness centrality. Residues with high shortest path betweenness and closeness centrality played key role in stabilizing the protein conformation (Liu et al., 2019; Xue et al., 2014). Both the shortest path betweenness and closeness centrality of each residue is between 0 and 1. When the shortest path betweenness is greater than 0.1, the residue will be regarded as a prominent residue, which means that this residue has strong interaction with other residues in the network. Tables S1and S2 showed the shortest path betweenness and closeness centrality values of all residues in the NSP15 system and the NSP15-NPC19899 complex system. According to the analysis results, it was found that there were 14 prominent residues in NSP15 system, which were residues Phe16, Leu73, Ile80, Asn83, Ile100, Val102, Phe123, Phe124, Ile144, Tyr194, Glu211, Phe303, Ile306, and Tyr325, respectively. But there were 29 prominent residues in NSP15-NPC198199 complex system,

which were residues Phe16, Lys61, Lys71, Ile72, Leu73, Asn75, Leu76, Ile80, Asn83, Ile100, Val102, Phe123, Phe124, Gln131, Phe135, Ile144, Tyr179, Thr193, Tyr194, Ile236, Leu252, Leu255, Ile296, Phe303, Ile306, Thr322, Ile323, Tyr325, and Thr326, respectively. In addition, it was found that the average closeness centrality value of NSP15 system was 0.45, while that of NSP15-NPC198199 complex system was 0.50. Obviously, the SNP15-NPC198199 complex system had more significant residues and a higher closeness centrality value, which meant that the interaction between residues in the NSP15-NPC198199 complex system was stronger and the system was more stable.

In conclusion, according to the analysis results of RIN, it was found that compound NPC198199 would help to stabilize the protein conformation, which also revealed the underlying reason for the inactivation of NSP15 protein after binding with NPC198199.

3.2.4 | The calculation of binding free energy

To assess the binding degree and binding affinity of the NSP15 protein and compound NPC198199, the binding free energy was calculated for the final 95 ns MD simulation trajectories by the MM-PBSA method. The binding free energy consists of four characteristic items, including VDW energy, electrostatic energy, polar solvation energy, and non-polar solvation energy. Generally, the value of the binding free energy is negatively correlated with the affinity between the ligand and the residues. The smaller the value

of the binding free energy, the stronger the binding affinity. VDW energy, electrostatic energy, and non-polar solvation energy are usually beneficial to the total binding free energy, while the polar solvation energy is detrimental to the total binding free energy, meaning that the VDW energy, electrostatic energy, and non-polar solvation energy help to enhance the stability of the complex. As shown in Table 2, the binding free energy of NSP15 protein and compound NPC198199 was −263.640 kJ/mol, indicating that they had a solid binding affinity. By decomposing the binding free energy, it was found that the VDW energy was −237.874 kJ/mol, the electrostatic energy was −91.418 kJ/mol, the non-polar solvation energy was −25.964 kJ/mol, and the polar solvation energy was 91.616 kJ/mol. In addition, the binding energy decomposition method was used to calculate the binding energy value of each residue to further clarify the contribution of each residue in the protein to the binding energy of the complex system. It was found that residues His235, Gln245, Gly248, His250, Lys290, Val292, Ser294, Tpr333, Glu340, Thr341, Tyr343, and Pro344 were key residues, and their binding energies were less than −10 kcal/mol. Obviously, these key residues were mainly distributed around the

catalytic active site. Therefore, the interactions between these key residues and NPC198199 played an important role in stabilizing the conformation of the complex.

3.2.5 | The analysis on the H-bond occupancy

The H-bond interaction formed between ligand and protein was the most important factor affecting protein conformation. In order to determine whether the H-bond interaction formed between the compound NPC198199 and the NSP15 protein could remain stable during the MD simulation process, the occupancy of H-bond interactions in the NSP15-NPC198199 complex system was calculated. When the H-bond occupancy exceeds 50%, it means that the H-bond could remain stable. Table 3 shows the calculation result of the H-bond occupancy in the NSP15-NPC198199 complex system. According to the calculation result, the H-bond occupancy of residues Glu340, Ser294, Val292, Lys290, Gly248, Gln245, and His235 with ligand NPC198199 were 91.7%, 87.9%, 91.8%, 85.9%, 61.8%, 73.4%, and 86.6% respectively, indicating that these formed
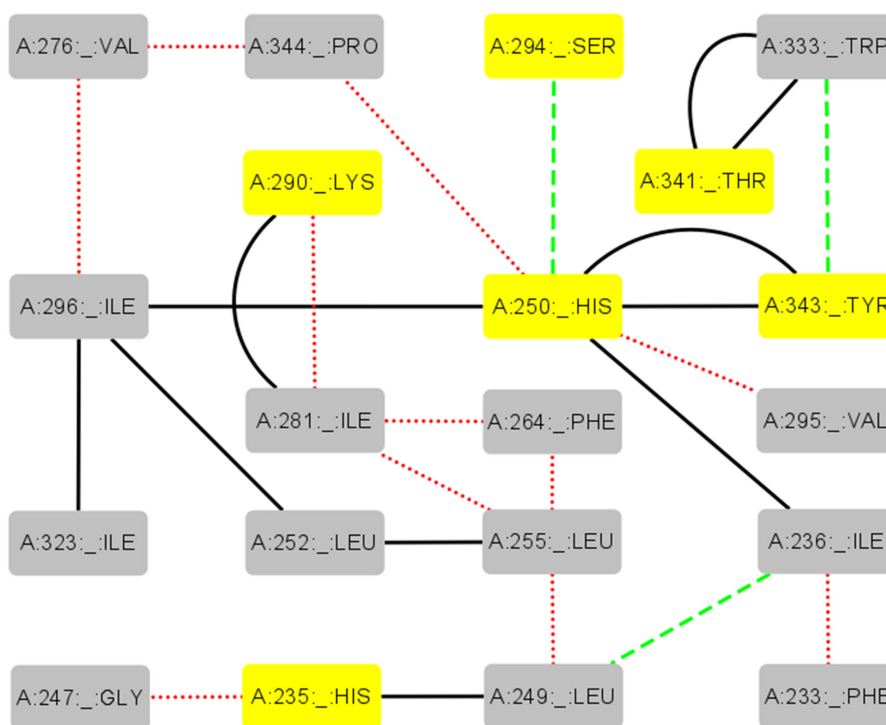

FIGURE 5 Residue interaction network (RIN) map of NSP15 system and NSP15-NPC198199 complex system. The interactions represented by the edge styles were presented in two systems (black solid line), NSP15 system only (green dotted line) and NSP15-NPC198199 complex system (red dotted line)

TABLE 2 The binding free energy between NSP15 protein and compound NPC198199, and the energy values of each component

Complex	Binding free energy (kJ/Mol)	VDW (kJ/Mol)	Electrostatic (kJ/Mol)	Polar solvation (kJ/Mol)	Non-polar solvation (kJ/Mol)
Nsp15-NPC198199	−263.640	−237.874	−91.418	91.616	−25.964

Pair ID	Interaction type	Residue	Ligand	Occupancy (%)
1	H-bond	Thr341(HG1)	NPC198199	19.9
2	H-bond	Glu340(O)	NPC198199	91.7
3	H-bond	Trp333(NE1)	NPC198199	16.6
4	H-bond	Leu332(H)	NPC198199	17.9
5	H-bond	Lys 317(O)	NPC198199	13.5
6	H-bond	Ser316(HG)	NPC198199	3.4
7	H-bond	Ser294(H)	NPC198199	87.9
8	H-bond	Val292(O)	NPC198199	91.8
9	H-bond	Lys290(HZ3)	NPC198199	85.9
10	H-bond	His250(NE2)	NPC198199	3.3
11	H-bond	Gly248(H)	NPC198199	61.8
12	H-bond	Leu246(O)	NPC198199	2.4
13	H-bond	Gln245(O)	NPC198199	73.4
14	H-bond	Asp240(OD1)	NPC198199	6.2
15	H-bond	His235(HE2)	NPC198199	86.6

H-bonds were stable and further indicating that the docking result was reliable.

4 | CONCLUSION

SARS-CoV-2 has become the deadliest virus of this century, causing the COVID-19 pandemic and bringing a heavy burden to the world. Although a variety of new coronavirus vaccines have been developed, the protective effect of the vaccine is not 100%, and a variety of virus variants have been found. Therefore, there is an urgent need to develop specific drugs to provide multiple protections for human health. It had been confirmed that inhibition of endoribonuclease NSP15 protein could slow down virus replication; therefore, NSP15 could be used as a potential drug target for anti-COVID-19. In this study, we used the NSP15 protein as the target and applied computer-aided drug design to perform high-throughput virtual screening and docking studies on 35,032 compounds from the NPASS data. According to the docking score, 10 potential compounds that could inhibit SARS-CoV-2 were identified, and the compound NPC198199 had the highest docking score. Through the investigation on the binding mode of the compound NPC198199 and NSP15 protein, it was found that NPC198199 could be well docked into the NSP15 protein and form the stable H-bond interactions with multiple key residues His235, Gln245, Gly248, Lys290, Val292, Ser294, and Glu340 at the catalytic site. Subsequently, in order to explore the internal changes of NSP15 protein after binding with the compound NPC198199, 100 ns MD simulations were conducted on the NSP15 system and NSP15-NPC198199 complex system, respectively, and the post-dynamic analyses were conducted. The analysis results of RMSD and RMSF found that the simulated trajectories of the two systems reached equilibrium after 5 ns, and the NSP15-NPC198199

complex system had higher stability. The analysis results of PCA and DCCM found that compared with the NSP15 system, the NSP15-NPC198199 complex system occupies a smaller phase space, and the negatively correlated motions between the residues were obviously reduced. The analysis results of RIN found that compared with the NSP15 system, the interactions between the residues in the NSP15-NPC198199 complex system were significantly increased, which also revealed the underlying reason for the enhanced stability of the complex system. The analysis results of binding free energy showed that the binding free energy of NSP15 and compound NPC198199 was −263.640 kJ/mol, indicating a solid binding affinity. The analysis results of the H-bond occupancy found that the occupancy of all H-bond interactions formed between the NSP15 and compound NPC198199 exceeded 50%, meaning that the H-bond interactions formed were very stable. In conclusion, through a series of analyses, it was proved that NPC198199 and other screened compounds were identified as potential NSP15 inhibitors. The research also provides valuable clues for subsequent drug development against SARS-CoV-2.

AUTHOR CONTRIBUTIONS

Wen-Shan Liu: Conceptualization; software; supervision. **Liang-Chang Hu:** Methodology; supervision. **Chuan-Hua Ding:** Formal analysis; funding acquisition. **Hong-Ying Li:** Data curation; validation. **Zhen-Zhen Li:** Data curation; validation. **Ying Chen:** Methodology. **Li-Peng Li:** Software. **Wan-Zhong Li:** Project administration; writing – review and editing.

REFERENCES

Alffenaar, J. W. C., Peloquin, C. A., & Migliori, G. B. (2018). Making optimal use of available anti-tuberculosis drugs: First steps to investigate terizidone. *International Journal of Tuberculosis and Lung Disease*, *22*(1), 2. https://doi.org/10.5588/ijtld.17.0632

Bao, J., Dong, X. Y., Zhang, J. Z., & Arora, P. S. (2009). Dynamical binding of hydrogen-bond surrogate derived Bak helices to antiapoptotic protein Bcl-xL. *The Journal of Physical Chemistry B*, *113*(11), 3565–3571. https://doi.org/10.1021/jp809810z

Chen, C., He, Z., Xie, D., Zheng, L., Zhao, T., Zhang, X., & Cheng, D. (2018). Molecular mechanism behind the resistance of the G1202R-mutated anaplastic lymphoma kinase to the approved drug Ceritinib. *The Journal of Physical Chemistry B*, *122*(17), 4680–4692. https://doi.org/10.1021/acs.jpcb.8b02040

Deng, X., & Baker, S. C. (2018). An "old" protein with a new story: Coronavirus endoribonuclease is important for evading host antiviral defenses. *Virology*, *517*, 157–163. https://doi.org/10.1016/j.virol.2017.12.024

Dixit, S. B., Ponomarev, S. Y., & Beveridge, D. L. (2006). Root mean square deviation probability analysis of molecular dynamics trajectories on DNA. *Journal of Chemical Information and Modeling*, *46*(3), 1084–1093. https://doi.org/10.1021/ci0504925

Doncheva, N. T., Klein, K., Domingues, F. S., & Albrecht, M. (2011). Analyzing and visualizing residue networks of protein structures. *Trends in Biochemical Sciences*, *36*(4), 179–182. https://doi.org/10.1016/j.tibs.2011.01.002

Fakhar, Z., Govender, T., Maguire, G. E. M., Lamichhane, G., Walker, R. C., Kruger, H. G., & Honarparvar, B. (2017). Differential flap dynamics in l,d-transpeptidase2 from mycobacterium tuberculosis revealed by molecular dynamics. *Molecular BioSystems*, *13*(6), 1223–1234. https://doi.org/10.1039/c7mb00110j

Fang, L., Wang, X., Xi, M., Liu, T., & Yin, D. (2016). Assessing the ligand selectivity of sphingosine kinases using molecular dynamics and MM-PBSA binding free energy calculations. *Molecular BioSystems*, *12*(4), 1174–1182. https://doi.org/10.1039/c6mb00067c

Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., O'Meara, M. J., Guo, J. Z., Swaney, D. L., Tummino, T. A., Hüttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., Foussard, H., Batra, J., Haas, K., Modak, M., Kim, M., Haas, P., ... Krogan, N. J. (2020). A SARS-CoV-2-human protein-protein interaction map reveals drug targets and potential drug-repurposing. *bioRxiv*. https://doi.org/10.1101/2020.03.22.002386

Grant, B. J., Rodrigues, A. P., ElSawy, K. M., McCammon, J. A., & Caves, L. S. (2006). Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics*, *22*(21), 2695–2696. https://doi.org/10.1093/bioinformatics/btl461

Guilbert, C., & James, T. L. (2008). Docking to RNA via root-mean-square-deviation-driven energy minimization with flexible ligands and flexible targets. *Journal of Chemical Information and Modeling*, *48*(6), 1257–1268. https://doi.org/10.1021/ci8000327

Harvey, A. (2005). Natural Products in Drug Discovery and Development. 27-28 June 2005, London, UK. *IDrugs*, *8*(9), 719–721.

Hess, B. (2008). P-LINCS: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation*, *4*(1), 116–122. https://doi.org/10.1021/ct700200b

Homeyer, N., & Gohlke, H. (2012). Free energy calculations by the molecular mechanics Poisson-Boltzmann surface area method. *Molecular Informatics*, *31*(2), 114–122. https://doi.org/10.1002/minf.201100135

Joung, I. S., & Cheatham, T. E. (2008). Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *Journal of Physical Chemistry B*, *112*(30), 9020–9041. https://doi.org/10.1021/jp8001614

Kasahara, K., Fukuda, I., & Nakamura, H. (2014). A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer-DNA complex. *PLoS One*, *9*(11), e112419. https://doi.org/10.1371/journal.pone.0112419

Kim, Y., Jedrzejczak, R., Maltseva, N. I., Wilamowski, M., Endres, M., Godzik, A., Michalska, K., & Joachimiak, A. (2020). Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. *Protein Science*, *29*(7), 1596–1605. https://doi.org/10.1002/pro.3873

Kumar, S., Kashyap, P., Chowdhury, S., Kumar, S., Panwar, A., & Kumar, A. (2021). Identification of phytochemicals as potential therapeutic agents that binds to Nsp15 protein target of coronavirus (SARS-CoV-2) that are capable of inhibiting virus replication. *Phytomedicine*, *85*, 153317. https://doi.org/10.1016/j.phymed.2020.153317

Kumari, R., Kumar, R., Open Source Drug Discovery, Consortium, & Lynn, A. (2014). g_mmpbsa—A GROMACS tool for high-throughput MM-PBSA calculations. *Journal of Chemical Information and Modeling*, *54*(7), 1951–1962. https://doi.org/10.1021/ci500020m

Liu, W. S., Li, H. G., Ding, C. H., Zhang, H. X., Wang, R. R., & Li, J. Q. (2021). Screening potential FDA-approved inhibitors of the SARS-CoV-2 major protease 3CL(pro) through high-throughput virtual screening and molecular dynamics simulation. *Aging (Albany NY)*, *13*(5), 6258–6272. https://doi.org/10.18632/aging.202703

Liu, W. S., Wang, R. R., Sun, Y. Z., Li, W. Y., Li, H. L., Liu, C. L., Ma, Y., & Wang, R. L. (2019). Exploring the effect of inhibitor AKB-9778 on VE-PTP by molecular docking and molecular dynamics simulation. *Journal of Cellular Biochemistry*, *120*(10), 17015–17029. https://doi.org/10.1002/jcb.28963

Mehla, K., & Ramana, J. (2016). Travelers' diarrhea-associated Enterotoxigenic Escherichia coli gyrA mutants and quinolone antibiotic affinity: A molecular dynamics simulation and residue interaction network analysis. *OMICS*, *20*(11), 635–644. https://doi.org/10.1089/omi.2016.0104

Mesentean, S., Fischer, S., & Smith, J. C. (2006). Analyzing large-scale structural change in proteins: Comparison of principal component projection and Sammon mapping. *Proteins*, *64*(1), 210–218. https://doi.org/10.1002/prot.20981

Nakagawa, K., Lokugamage, K. G., & Makino, S. (2016). Viral and cellular mRNA translation in coronavirus-infected cells. *Advances in Virus Research*, *96*, 165–192. https://doi.org/10.1016/bs.aivir.2016.08.001

Ndagi, U., Mhlongo, N. N., & Soliman, M. E. (2017). The impact of Thr91 mutation on c-Src resistance to UM-164: Molecular dynamics study revealed a new opportunity for drug design. *Molecular BioSystems*, *13*(6), 1157–1171. https://doi.org/10.1039/c6mb00848h

Ni, W., Yang, X., Yang, D., Bao, J., Li, R., Xiao, Y., Hou, C., Wang, H., Liu, J., Yang, D., Xu, Y., Cao, Z., & Gao, Z. (2020). Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. *Critical Care*, *24*(1), 422. https://doi.org/10.1186/s13054-020-03120-0

Petrosillo, N., Viceconte, G., Ergonul, O., Ippolito, G., & Petersen, E. (2020). COVID-19, SARS and MERS: Are they closely related? *Clinical Microbiology and Infection*, *26*(6), 729–734. https://doi.org/10.1016/j.cmi.2020.03.026

Pol-Fachin, L., Fernandes, C. L., & Verli, H. (2009). GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations. *Carbohydrate Research*, *344*(4), 491–500. https://doi.org/10.1016/j.carres.2008.12.025

Prajapat, M., Sarma, P., Shekhar, N., Avti, P., Sinha, S., Kaur, H., Kumar, S., Bhattacharyya, A., Kumar, H., Bansal, S., & Medhi, B. (2020). Drug targets for corona virus: A systematic review. *Indian Journal of Pharmacology*, *52*(1), 56–65. https://doi.org/10.4103/ijp.IJP_115_20

Ravindranath, V. (2010). Natural products as a resource for treatment of Alzeimer's disease. *Journal of Neurochemistry*, *115*, 6.

Sang, P., Xie, Y. H., Li, L. H., Ye, Y. J., Hu, W., Wang, J., Wan, W., Li, R., Li, L. J., Ma, L. L., Li, Z., Liu, S. Q., & Meng, Z. H. (2017). Effect of the R119G mutation on human P5CR structure and its interactions with NAD: Insights derived from molecular dynamics simulation and free energy analysis. *Computational Biology and Chemistry*, *67*, 141–149. https://doi.org/10.1016/j.compbiolchem.2016.12.015

Savale, R. U., Bhowmick, S., Osman, S. M., Alasmary, F. A., Almutairi, T. M., Abdullah, D. S., Patil, P. C., & Islam, M. A. (2021). Pharmacoinformatics approach based identification of potential Nsp15 endoribonuclease modulators for SARS-CoV-2 inhibition. *Archives of Biochemistry and Biophysics*, *700*, 108771. https://doi.org/10.1016/j.abb.2021.108771

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Unchwaniwala, N., & Ahlquist, P. (2020). Coronavirus dons a new crown. *Science*, *369*(6509), 1306–1307. https://doi.org/10.1126/science.abe0322

van den Berg, H. P., & Hoheisel, C. (1990). Dynamic cross correlation in isotopic two-component liquids: Molecular-dynamics calculation results compared with predictions of kinetic theory. *Physical Review A*, *42*(6), 3368–3373.

Wan, H., Hu, J. P., Tian, X. H., & Chang, S. (2013). Molecular dynamics simulations of wild type and mutants of human complement receptor 2 complexed with C3d. *Physical Chemistry Chemical Physics*, *15*(4), 1241–1251. https://doi.org/10.1039/c2cp41388d

Xu, L., Kong, R., Zhu, J., Sun, H., & Chang, S. (2016). Unraveling the conformational determinants of LARP7 and 7SK small nuclear RNA by theoretical approaches. *Molecular BioSystems*, *12*(8), 2613–2621. https://doi.org/10.1039/c6mb00252h

Xue, W., Ban, Y., Liu, H., & Yao, X. (2014). Computational study on the drug resistance mechanism against HCV NS3/4A protease inhibitors vaniprevir and MK-5172 by the combination use of molecular dynamics simulation, residue interaction network, and substrate envelope analysis. *Journal of Chemical Information and Modeling*, *54*(2), 621–633. https://doi.org/10.1021/ci400060j

Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., & Zhou, Q. (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, *367*(6485), 1444–1448. https://doi.org/10.1126/science.abb2762

Zeng, X., Zhang, P., He, W., Qin, C., Chen, S., Tao, L., Wang, Y., Tan, Y., Gao, D., Wang, B., Chen, Z., Chen, W., Jiang, Y. Y., & Chen, Y. Z. (2018). NPASS: Natural product activity and species source database for natural product research, discovery and tool development. *Nucleic Acids Research*, *46*(D1), D1217–D1222. https://doi.org/10.1093/nar/gkx1026

Zhou, Y., Zhang, N., Chen, W., Zhao, L., & Zhong, R. (2016). Underlying mechanisms of cyclic peptide inhibitors interrupting the interaction of CK2alpha/CK2beta: Comparative molecular dynamics simulation studies. *Physical Chemistry Chemical Physics*, *18*(13), 9202–9210. https://doi.org/10.1039/c5cp06276d

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.