Research Article

Identifying USP1 Inhibitors with Allosteric Effect on Its Triple Catalytic Center through Virtual Screening

Jinhong Xu, Yongxia Li, Tao Jiang, Yi Zhu, Jinyan Zhu, Tao Xu, Longji Fang, Zujun Hong, Yuying Jia, and Fang Bao

1Navy Anqing Hospital, Anqing, China
2College of Pharmacy, Anqing Medical College, Anqing, China

Correspondence should be addressed to Fang Bao; baofangbf@126.com

Received 26 March 2023; Revised 28 May 2023; Accepted 2 June 2023; Published 10 June 2023

1. Introduction

Ubiquitination is a frequently occurring posttranslational modification that involves the addition of ubiquitin to target proteins via an isopeptide bond [1, 2]. This process is typically orchestrated by the sequential action of ubiquitin-activating enzymes (E1s), ubiquitin-conjugating enzymes (E2s), and ubiquitin ligases (E3s). Apart from its role in protein degradation, ubiquitination has many vital non-proteolytic functions, such as DNA damage repair, DNA replication, signal transduction, transcriptional regulation, membrane transport, endocytosis, and more [3, 4]. It controls the levels, localization, and activity of numerous proteins and is involved in most cellular processes, including signal transduction, cell cycle, growth, and apoptosis [5, 6]. As with other posttranslational modifications, protein ubiquitination is reversible, and deubiquitinases (DUBs) catalyze the removal of multiubiquitin chains or single ubiquitin [7].

As one of the quintessential members of DUB, USP1 plays a pivotal role in regulating DNA repair processes and cell differentiation [3]. Its absence can result in impaired DNA repair function and genomic instability, thus rendering USP1 a potential pharmaceutical target for countering drug resistance in DNA damage cancer therapy [8].

In recent years, a plethora of evidence has underscored the critical impact of ubiquitination in tumorigenesis, which can exert either suppressive or promotive effects on cancer initiation and progression [9]. Deubiquitinases mediate substrate deubiquitination by removing ubiquitin molecules, thereby preventing substrate proteins from degradation. Among these enzymes, USP1, as a member of the deubiquitinase family, is regularly upregulated in various malignant tumors and closely correlated with the tumorigenesis and progression of several cancers, including multiple myeloma, leukemia, osteosarcoma, glioma, non-small-cell lung cancer, and liver cancer [10, 11]. Therefore, inducing downregulation of USP1 by inhibitor is
2. Materials and Methods

2.1. Database Cleaning and Protein Preparation. The Enamine database, containing 1,373,646 small chemical molecules, was used as the target for virtual screening in this study. The vast number of molecules in the database covers most of the chemical space and possesses favorable physicochemical properties, providing an advantage for discovering new molecular scaffolds with high drug-likeness. Prior to virtual screening, all molecules underwent 3D structure conversion using the LigPrep module in the Schrodinger 2021-3 software [20], which involved energy minimization of all molecules under the OPLS3e force field [21], correct protonation prediction of all molecules using the Epik method, and removal of possible salt ions present in the molecule library. The cleaned molecules were used for subsequent molecular property calculations.

The crystal structure of USP1 bound to the inhibitor ML323 (PDB ID 7ZH4) [22] retrieved from the PDB database was used as the crystal structure for the calculations. The Protein Preparation Wizard module in the Schrodinger 2021-3 software was used for protein preparation, including adding missing residues and charges, removing all water molecules, adding hydrogen atoms to the protein system and optimizing their orientation, protonating amino acids at pH 7.4, and conducting structure optimization.

2.2. Pharmacophore Construction. According to the reported binding mode analysis between USP1 and the inhibitor ML323, we manually constructed a pharmacophore model using the Phase program [23]. The pharmacophore model consists of two hydrogen bond acceptors and three aromatic rings, forming the “ARRRA” hypothesis model.

2.3. Molecular Docking. In this study, molecular docking was used to predict the binding mode between the target molecules and the USP1 protein. The Glide software in the Schrodinger 2021-3 software was used for molecular docking, and ML323 was manually constructed using the Maestro module.

Before docking, ML323 was prepared using the same method as the small molecule database preparation described above. Then, the Receptor Grid Generation module was used to generate the docking box for the prepared USP1, with the center of the box set to the center of the crystal ligand and the size of the box based on the size of the crystal ligand. Molecular docking was performed using the HTVS, SP, and XP precision algorithms from the Glide module [24]. Finally, the RMSD calculation and superimposition analysis were performed on the active conformations of the docking output. 3D interaction force analysis was conducted using PyMOL 2.5.4 [25].

2.4. Virtual Screening. Prior to virtual screening, the docking box file was generated using a similar method as described above. It is noteworthy that Q97 and Q160 were additionally set as hydrogen bond donors as a restriction during docking to efficiently identify active molecules. Therefore, the docking conformation output by the compound in subsequent docking work must satisfy the requirement of forming hydrogen bonds with Q97 and Q160 on the allosteric pocket of the protein.

The virtual screening work was completed using the VSW (Virtual Screening Workflow) module in Schrodinger 2021-3 software. In the virtual screening process, the preparatory small molecule database was first filtered using pharmacophores. Then, the HTVS, SP, and XP algorithms of Glide were used to screen the pharmacophore-filtered compounds in sequence with increasing precision. The HTVS, SP, and XP algorithms gradually improve in accuracy and are able to effectively select true positive molecules. The top 10% of compounds scored by each algorithm after docking were extracted for the next round of screening. Finally, the docking output of the XP algorithm was used for prime MM/GBSA and strain energy calculations.

2.5. ADMET Parameters Prediction. The Qikprop module in Schrodinger 2021-3 was used to predict multiple pharmacological parameters, including octanol/water partition coefficient, solubility, hERG, Caco-2, MDCK permeability, human oral absorption, RuleOFFive, and RuleOFFThree.

2.6. Molecular Dynamics Simulation. Based on the above docking results, hits and USP1 complexes were used as initial structures for full-atom molecular dynamics simulations, which were conducted using AMBER 18 software [26]. Prior to simulation, the charges of hits were calculated using the Hartree–Fock (HP) SCF/6-31G* method with the Antechamber module [27] and Gaussian 09 software [28].
The small molecules and proteins were subsequently described using the GAFF2 [29] and ff14SB force fields [30], respectively. Hydrogen atoms were added to each system using the LEaP module, and an octahedral TIP3P [31] solvent box was truncated at a distance of 10 Å from the system. Na+/Cl- ions were added to balance the system’s charge. Finally, the topological and parameter files required for simulation were outputted.

In the simulation stage, solvated complexes were energy-minimized so that each system performs the 2500 steps steepest descent optimization, and then the 2500 steps conjugate gradient optimization. After minimization, a 200 ps heating process was conducted on the fixed volume and constant heating rate condition to gradually raise the temperature of the system from 0 K to 298.15 K. Under the condition of maintaining the temperature at 298.15 K, a 500 ps NVT (constant particle number, volume, and temperature) simulation was performed to further evenly distribute the solvent molecules in the solvent box. Finally, a 500 ps equilibration simulation of the entire system was carried out under NPT (constant particle number, pressure, and temperature). Subsequently, all systems were subjected to 100 ns NPT simulations under periodic boundary conditions. During the simulations, the nonbonded cut-off distance was set to 10 Å, the Particle Mesh Ewald (PME) method was used to calculate long-range electrostatic interactions [32], the SHAKE method was utilized to restrict the hydrogen bond length [33], and the Langevin algorithm [34] was applied for temperature control with a collision frequency $\gamma$ of 2 ps$^{-1}$. The system pressure was set at 1 atm, the integration time step was 2 fs, and the trajectory was saved every 10 ps. RMSD, RMSF, RoG, and SASA calculations were performed on all trajectories using the CPPTRAJ module [35].

2.7. MM/GBSA Binding Free Energy Calculation and Decomposition. For the MM/GBSA calculations, the MMPBSA.py module [36] within the AMBER 18 software was utilized. To calculate the binding free energy for each system, snapshots were captured every 100 ps from whole MD trajectory and computed based on the equations displayed in the following:

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

(1)

$$= \Delta E_{\text{internal}} + \Delta E_{\text{VDW}} + \Delta E_{\text{elec}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}}.$$ 

(2)

In equation (2), $\Delta E_{\text{internal}}$ represents the internal energy, $\Delta E_{\text{VDW}}$ stands for van der Waals energy, and $\Delta E_{\text{elec}}$ is the electrostatic energy. The internal energy includes bond energy, angle energy, and torsional energy. The solvent free energy terms, $\Delta G_{\text{GB}}$ and $\Delta G_{\text{SA}}$, are collectively referred to as solvation free energy. Herein, $\Delta G_{\text{GB}}$ denotes the polar solvation free energy, while $\Delta G_{\text{SA}}$ refers to the nonpolar solvation free energy. For $\Delta G_{\text{GB}}$, we employed the GB model developed by Nguyen et al. (igb = 8) for calculation [37]. The nonpolar solvation free energy ($\Delta G_{\text{SA}}$) was calculated based on the product of the surface tension ($\gamma$) and solvent accessible surface area (SASA): $\Delta G_{\text{SA}} = 0.0072 \times \text{SASA}$ [38].

3. Results and Discussion

3.1. Structure-Activity Relationship (SAR) Analysis and Construction of Pharmacophore Model. There exists a plethora of approaches to inhibit the USPs enzyme. Formerly, Lange et al. collated multiple inhibitory methods, including binding within the active site, covalent binding within the active site, and allosteric binding in the distal ubiquitin site [39]. These modalities competitively associate with ubiquitin either directly or indirectly to impede the completion of the deubiquitination process by the USPs enzyme. Recently, Rennie et al. employed cryo-electron microscopy to elucidate the co-crystalline structure between ML323 and USP1, uncovering a novel USPs pocket [22]. This discovery constitutes the sole binding modality between small molecule inhibitors and USP1 thus far disclosed. As depicted by Figure 1(a), ML323 binds just above the catalytic triad at a conformational site. As illustrated by Figure 1(b), under normal circumstances, the catalytic triad (C90/H593/D751) is mutually coupled, H593 and D751 undergo hydrogen bonding, H593’s N terminus accepts the proton yielded by C90 (mutated SER in 7ZHH crystals), thereby promoting the catalytic reaction between C90 and G76. However, ML323 binding results in the downward shifting of the loop containing D751. Consequently, H593 interacts with D752, disallowing the acceptance of the proton transferred by C90 and hindering the catalytic process. Notably, the conformational shift resulting from the interaction between ML323 and USP1 precisely obliterates the function of the catalytic triad.

By analyzing the crystal structure reported by Rennie et al. [22], we can observe the detailed interaction between ML323 and USP1 (Figure 1(c)). Multiple aromatic rings on the ML323 molecule can undergo Pi-Pi conjugation with F750. Additionally, ML323 can also undergo hydrogen bonding with Q97 and Q160. These interactions form the basis for the stable binding of ML323 to the USP1 protein and the occurrence of conformational changes leading to the loss of function of the catalytic triad.

In order to extract information conducive to the discovery of new hit compounds for further research, we generated pharmacophores based on USP1/ML323 complexes using the auto (e-pharmacophore) method in the Maestro suite. The result is shown in Figure S2(a), which consists of three aromatic ring centers and excluded volumes. The automatically generated pharmacophores are relatively simple, which may not fully explain the activity of the molecule. Based on our reading of the ML323 modification articles published by Dexheimer et al. [16], we found that the hydrogen bond acceptors on the pyrimidine ring and the triazole of the ML323 molecule are necessary to enhance the activity. Therefore, we manually added these two pharmacophores (hydrogen bond receptor) based on auto pharmacophore, namely the pharmacophores described in Figures 1(d) or S2(b). Aromatic ring center contributes to Pi-Pi stacking interactions, hydrogen bond acceptors can undergo crucial hydrogen bonding interactions, and excluded volume can restrict the binding...
space of molecules. This pharmacophore model will be used for subsequent virtual screening.

3.2. Docking Accuracy Verification. Molecular docking is an important screening method in this study. Prior to virtual screening, we employed a re-docking strategy to verify the conformational search and scoring abilities of the docking algorithms. As shown in Figure 2, all three algorithms generated docked poses that align with the active pose (crystal pose), with HTVS, SP, and XP almost completely overlaying the active conformation.

The RMSD values were 0.4053 Å, 0.2804 Å, and 0.2551 Å, respectively (Table 1), and the docking scores were −10.197 kcal/mol, −10.424 kcal/mol, and −10.310 kcal/mol, respectively. This indicates that HTVS, SP, and XP algorithms can accurately search for the correct conformation of USP1 inhibitors. Therefore, they were chosen for docking-based virtual screening.

3.3. Virtual Screening. Based on the above results, a sequential virtual screening strategy will be employed utilizing both pharmacophore-based and molecular docking-based virtual screening (Figure 3(a)). In this study, the Enamine database, which includes 1,373,646 compounds with diverse chemical scaffolds, was utilized for virtual screening. The distribution of molecular weight (MW), log P, TPSA, rotatable bond count, Fsp3, similarity, and hydrogen bond acceptor/donor count in the database molecules met the requirements of drug-likeness, indicating that the compounds in this database may have good drug-like properties (Figure S1). The “ARRRA” pharmacophore model was applied to screen the database, requiring any four or more of the five pharmacophore features to match in order for a molecule to be considered a potential hit. Moreover, in our research, we collected 31 molecules with biochemical activity within 1 μM and fit them with pharmacophores [16]. Fitness score results are shown in Table S1, with the worst three fitness values being 1.44, 1.482, and 1.539. Therefore, to better match active molecules, we set the fitness score screening threshold to 1.5. Based on the filtering limitations described above, 207,226 compounds were obtained through pharmacophore-based virtual screening.

In the docking-based virtual screening, a total of 207,226 compounds were first docked into the allosteric pocket of the USP1 protein using the HTVS algorithm, resulting in 16,578 successful docking hits that met the hydrogen bonding criteria. The top 10% of the docking scores were then retained for further docking using the SP algorithm, resulting in 1,657 compounds. Subsequently, the XP algorithm was employed to dock these compounds, yielding 166 hits. It is noteworthy that our docking was performed under restrictive conditions, requiring the hits to form hydrogen bonds with the critical amino acids Q97 and Q167, as emphasized above. After docking screening, the prime MM/GBSA method was used to calculate the high-precision binding energies and strain energies of these compounds. Based on the selection criteria of prime MM/GBSA binding energy greater than 60.00 kcal/mol (reference compound prime MM/GBSA binding energy was −60.84 kcal/mol) and
strain energy lower than 5 kcal/mol, 124 compounds were selected as potential hits.

3.4. Hits Identification. The core idea behind the screening of the molecules mentioned above is to achieve stable binding with the USP1 protein. However, for a drug to be effective, it must exhibit strong interactions with the protein while also possessing good ADMET properties. Hence, we computed the ADMET properties for 124 molecules (as shown in Table S2). Through manual screening, we identified four hit molecules: Z28551960, Z423064076, Z1037335610, and Z1129923146, which also showed good binding affinity to USP1 protein (as shown in Table 2). The structures of these four molecules, along with the reference molecule ML323, are shown in Figure 3(b). We selected
Table 2: Docking score, strain energy, and prime MM/GBSA binding energies for hits.

<table>
<thead>
<tr>
<th>Compound ID</th>
<th>Docking score</th>
<th>Strain energy</th>
<th>Prime MM/GBSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML323</td>
<td>−10.197</td>
<td>5.511</td>
<td>−60.84</td>
</tr>
<tr>
<td>Z28551960</td>
<td>−10.203</td>
<td>1.272</td>
<td>−73.9</td>
</tr>
<tr>
<td>Z423064076</td>
<td>−10.529</td>
<td>1.136</td>
<td>−72.42</td>
</tr>
<tr>
<td>Z1037335610</td>
<td>−11.116</td>
<td>3.332</td>
<td>−82.44</td>
</tr>
<tr>
<td>Z1129923146</td>
<td>−10.475</td>
<td>0.75</td>
<td>−74.8</td>
</tr>
</tbody>
</table>

these core molecules based on their superior solubility, bioavailability, and drug-likeness compared to ML323 (as presented in Table 3); prime MM/GBSA binding energies should be above −70 kcal/mol; they should possess a novel chemical scaffold distinct from ML323; and exhibit an “L” shape similar to the active conformation of ML323. These four molecules have a high potential for activity against USP1, and subsequently, the binding mode of these molecules with the protein will be analyzed. Molecular dynamics simulations will also be employed to investigate the dynamic properties of the interactions between these molecules and USP1.

3.5. The Binding Mode between Hit and USP1 Protein. These small molecules bind to the allosteric pocket located above the catalytic center of ubiquitin. They can form hydrogen bonds with Q97 and Q160 on the protein, as well as Pi-Pi conjugation with F570 (Figure 4). The overall interaction pattern is consistent with that of ML323, suggesting that these molecules have the potential to strongly bind to USP1. Subsequently, molecular dynamics simulations will be employed to further investigate the binding kinetics between small molecules and the protein.

3.6. The Dynamic Characteristics of the Binding between Hit and USP1 Protein. The root mean square deviation (RMSD) in molecular dynamics simulations reflects the movement process of the system [40]. The larger and more fluctuating the RMSD, the more intense the movement is; conversely, the movement is smoother. As shown in Figure 5(a), the RMSD of all systems can reach convergence after a certain simulation time, indicating no anomalous behavior in the simulations. It is noteworthy that the RMSD of USP1 without the bound inhibitor (black line) shows a smoother trend, indicating that the binding of the inhibitor can cause more conformational changes in the protein. Remarkably, the RMSD of the USP1 that binding with Z423064076 is significantly higher than that of the other systems. We analyzed the conformation of this protein before and after the simulation, and found that the loop (aa 462–481) on the protein could be transformed into a helix (Figure S3). This conformational change is the main reason for the higher RMSD, which means that the protein loop region undergoes further folding under the binding of Z423064076. In addition, we also calculated the RMSD of the ligands. From Figure 5(a), we can see that, except for Z28551960, other molecules are within 2 Å, indicating that the small molecules remained stably bound in the protein pocket during the 100 ns simulation. In the later stage of the simulation, the RMSD of the Z28551960 molecule remained between 1.5 and 2.5 Å, indicating that its stability was relatively weak. Similarly, as shown in Figure 5(b), the root mean square fluctuation (RMSF) of the protein in the apo state is higher than that of USP1 in the state with the bound inhibitor in most regions, indicating that the inhibitor can increase the flexibility of the protein in the extrasequence, which is advantageous for the protein’s conformational changes.

The radius of gyration (RoG) reflects the compactness of the system and can indicate the degree of compactness of the system. As shown in Figure 5(c), the RoG of USP1 without the bound inhibitor is the lowest compared to that with the bound inhibitor. This indicates that the inhibitor can reduce the compactness of the protein system, which is advantageous for the conformational changes of the protein, such as the rupture of the catalytic triad. In addition, from Figure 5(d), we can also observe a similar phenomenon, that is, the solvent accessible surface area (SASA) of USP1 is higher after binding of the inhibitor, indicating that the binding of the inhibitor makes the USP1 protein looser, with more structures exposed to solvent. This may be detrimental to the normal functioning of the USP1 in deubiquitination. Due to the randomness in MD simulations, we conducted repeated simulations. We found that the difference in complex RMSD between the three simulations was small (Figure S4), indicating that the simulation results mentioned above were nonrandom.

3.7. Binding Affinity Calculated by MM/GBSA. Based on the whole trajectory of molecular dynamics simulations, we again used the MMGBSA method to calculate the binding energy, which can more accurately reflect the binding effect between small molecules and the target protein [41, 42].

As shown in Table 3, the binding energies of the USP1/ML323, USP1/Z28551960, USP1/Z423064076, USP1/Z1037335610, and USP1/Z1129923146 complexes are −39.35 ± 3.89 kcal/mol, −35.43 ± 2.81 kcal/mol, −28.86 ± 3.12 kcal/mol, −39.68 ± 3.55 kcal/mol, and −43.95 ± 3.18 kcal/mol, respectively. The negative values indicate that the molecules have binding affinity with the target protein, and the lower the value, the stronger the binding. Obviously, our calculations show that these molecules and their corresponding proteins all have a strong binding affinity. Compared with the control molecule ML323, we found that Z1037335610 and Z1129923146 have stronger binding effects, especially Z1129923146, which has a significantly higher binding strength than ML323. In addition, our energy decomposition shows that the binding energies of these complexes are mainly contributed by van der Waals energy, followed by electrostatic energy. These high-precision binding energy calculations once again validate the high affinity effect of the four hits and USP1. Table 4
### Table 3: ADME and pharmacological parameters prediction for hits using QikProp.

<table>
<thead>
<tr>
<th>Compound ID</th>
<th>QPlogPo/w</th>
<th>QPlogSp</th>
<th>QPlogHERG</th>
<th>QPPCaco</th>
<th>QPPMDCK</th>
<th>PercentHumanOralAbsorption</th>
<th>RuleOfFive</th>
<th>RuleOfThree</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML323</td>
<td>5.413</td>
<td>-6.929</td>
<td>-6.616</td>
<td>1834.018</td>
<td>797.101</td>
<td>100</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Z28551960</td>
<td>3.411</td>
<td>-4.947</td>
<td>-6.167</td>
<td>1137.4</td>
<td>968.245</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Z423064076</td>
<td>3.08</td>
<td>-4.325</td>
<td>-5.428</td>
<td>2128.1</td>
<td>1119.12</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Z1037335610</td>
<td>4.299</td>
<td>-5.456</td>
<td>-6.25</td>
<td>990.54</td>
<td>1022.704</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Z1129923146</td>
<td>2.707</td>
<td>-4.315</td>
<td>-4.405</td>
<td>294.97</td>
<td>326.516</td>
<td>87.002</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

1Predicted octanol/water partition co-efficient log \( p \) (acceptable range: –2.0 to 6.5). 2Predicted aqueous solubility; \( S \) in mol/L (acceptable range: –6.5 to 0.5). 3Predicted IC50 value for blockage of HERG \( K^+ \) channels (concern below –7). 4Predicted Caco-2 cell permeability in nm/s (acceptable range: <25 is poor and >500 is great). 5Predicted apparent MDCK cell permeability in nm/sec. (<25 is poor and >500 is great great). 6Percentage of human oral absorption (<25% is poor and >80% is high). 7Number of violations of Lipinski’s rule of five (maximum is 4). 8Number of violations of Jorgensen’s rule of three. Compounds with fewer (and preferably no) violations of these rules are more likely to be orally available.

**Figure 4:** The binding modes of Z28551960_USP1 (a), Z423064076_USP1 (b), Z1037335610_USP1 (c), and Z1129923146_USP1 (d) to the USP1 protein. The white cartoon represents the USP1 protein, the yellow cartoon represents ubiquitin, and the sticks represent small molecules. The yellow dashed lines indicate hydrogen bonds, and the green dashed lines indicate Pi-Pi conjugation.
Figure 5: The dynamic characteristics of proteins after the combination of different small molecules: (a) protein and ligand RMSD changes over time; (b) protein RMSF in different complexes; (c) protein SASA changes over time; (d) protein RoG changes over time.

Table 4: Binding free energies and energy components predicted by MM/GBSA (kcal/mol).

<table>
<thead>
<tr>
<th>Compound ID</th>
<th>$\Delta E_{vdw}$</th>
<th>$\Delta E_{elec}$</th>
<th>$\Delta G_{GB}$</th>
<th>$\Delta G_{SA}$</th>
<th>$\Delta G_{bind}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML323</td>
<td>$-52.31 \pm 4.74$</td>
<td>$-16.86 \pm 4.36$</td>
<td>$36.47 \pm 4.33$</td>
<td>$-6.64 \pm 0.54$</td>
<td>$-39.35 \pm 3.89$</td>
</tr>
<tr>
<td>Z28551960</td>
<td>$-50.54 \pm 2.84$</td>
<td>$-9.22 \pm 3.50$</td>
<td>$30.94 \pm 3.09$</td>
<td>$-6.61 \pm 0.25$</td>
<td>$-35.43 \pm 2.81$</td>
</tr>
<tr>
<td>Z423064076</td>
<td>$-44.04 \pm 4.50$</td>
<td>$-20.17 \pm 4.76$</td>
<td>$41.20 \pm 4.06$</td>
<td>$-5.84 \pm 0.37$</td>
<td>$-28.86 \pm 3.12$</td>
</tr>
<tr>
<td>Z1037335610</td>
<td>$-51.20 \pm 3.6$</td>
<td>$-4.05 \pm 4.81$</td>
<td>$21.62 \pm 4.23$</td>
<td>$-6.04 \pm 0.31$</td>
<td>$-39.68 \pm 3.55$</td>
</tr>
<tr>
<td>Z1129923146</td>
<td>$-53.36 \pm 3.11$</td>
<td>$-12.87 \pm 5.71$</td>
<td>$28.90 \pm 3.72$</td>
<td>$-6.61 \pm 0.23$</td>
<td>$-43.95 \pm 3.18$</td>
</tr>
</tbody>
</table>

$^1$Van der Waals energy. $^2$Electrostatic energy. $^3$Electrostatic contribution to solvation. $^4$Nonpolar contribution to solvation. $^5$Binding free energy.
4. Conclusions
This study explored the conformational mechanism of ML323 on USP1 and analyzed the binding mode of ML323, based on which a pharmacophore model was constructed. Subsequently, virtual screening was performed on the Enamine database (1,373,646 compounds), which has a wide coverage of the chemical space. First, we used pharmacophore-based filtering to obtain 207,226 compounds with high matching, and then further screened 124 small molecules through molecular docking, binding energy, and strain energy. These molecules have good physico-chemical properties, can match the USP1 inhibitor pharmacophore model, and can bind to the USP1 protein with high affinity. Then, after ADMET filtering and manual selection, four potential hits were identified: Z28551960, Z423064076, Z1037335610, and Z1129923146. Finally, molecular dynamics simulations confirmed that these four molecules have good binding characteristics with the USP1 protein and can lead to different degrees of improvement in the RMSD, RMSF, RoG, and SASA dynamic properties, showing a trend of affecting the stability, flexibility, compactness, and extensibility of the protein, forming the basis for the protein’s conformational effects. In summary, our study provides important reference value for the development and structural optimization of USP1 inhibitors in the future.

Data Availability
The original contributions presented in the study are included in the article/Supplementary Materials and further inquiries are available from the corresponding author upon reasonable request.

Conflicts of Interest
The authors declare that there are no conflicts of interest.

Authors’ Contributions
JX and FB conceptualized the manuscript. JX, YL, TJ, YZ, JZ, TX, LF, ZH, and YJ collected the literature, wrote the manuscript, and made the figures. JX and FB edited and made significant revisions to the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments
This research was funded by the Anqing City 2022 Medical and Health Self-Raised Funds Science and Technology Project (No. 2022Z4009).

Supplementary Materials
Figure S1: Database of physical and chemical property distribution; Figure S2: Pharmacophore model; Figure S3: The alignment of USP1/Z423064076 in simulation before and after; Figure S4: The RMSD of the complex changes over time in three dynamic simulation processes; Table S1: Fitness score of 31 molecules with biochemical activity within 1 μM; Table S2: Prime MMGBSA, strain energy, ADMET, and pharmacological parameters prediction for ML323 and 124 compounds. (Supplementary Materials)

References