J Mol Model (2014) 20:2515
DOI 10.1007/s00894-014-2515-y

ORIGINAL PAPER

Discovery of novel covalent proteasome inhibitors through a combination of pharmacophore screening, covalent docking, and molecular dynamics simulations

Aibo Li · Haopeng Sun · Lei Du · Xiaoxin Wu · Jianqin Cao · Qidong You · Yuyan Li

Received: 18 March 2014 / Accepted: 27 October 2014 / Published online: 14 November 2014

Abstract The ubiquitin–proteasome pathway plays a pivotal role in the regulation of cellular protein processing and degradation. Proteasome inhibitors (PIs) have enormous potential to treat multiple myeloma, solid tumors, parasites, inflammation, and immune diseases, which is spurring the development of new types of PIs with enhanced efficacy, fewer side effects, and reduced drug resistance. Nevertheless, virtual screening for covalent PIs has rarely been reported because calculating the covalent binding energy is a challenging task. The aim of this study was to discover new covalent inhibitors of the 20S proteasome. The structures of PIs were manually divided into two parts: a noncovalent binding part resulting from virtual screening, and an epoxyketone group that was pre-selected as a covalent binding part. The SPECS database was screened by noncovalent docking and a pharmacophore model built with the 20S proteasome. After validating the covalent conjugation, 88 hits with epoxyketone were covalently docked into the 20S proteasome to analyze the intermolecular interactions. Four compounds were selected after multiple filtration and validations. Molecular dynamics simulations were performed to check the stability of the noncovalent and covalent docked ligand–enzyme complexes and investigate the interaction patterns of the screened inhibitors. Finally, two compounds with novel aromatic backbones, reasonable interactions, and stable covalent binding modes were retained. These compounds can serve as potential hits for further biological evaluation.

Keywords Ubiquitin–proteasome · Covalent inhibitors · Virtual screening · Pharmacophore · Docking · Molecular dynamics

Introduction

The ubiquitin–proteasome pathway (UPP) is the main quality control system for protein degradation [1]. Proteasome inhibitors (PIs) accelerate protein misfolding and induce apoptosis in cells, especially malignant plasma cells [2] due to their production of antibodies [3]. The first PI to be launched commercially, bortezomib (BTZ), was approved in May 2003 by the FDA to treat multiple myeloma (MM) [4] and mantle cell lymphoma [5]. While BTZ is still being investigated as a treatment for various hematological malignancies and solid tumors [6], some adverse neurologic and cardiovascular effects of BTZ have been reported [7]. In June 2012, carfilzomib (CFZ), a novel second-generation PI, was approved for the treatment of MM [8]. CFZ shows superior selectivity and improved adverse effect profiles, forming stable and irreversible adducts exclusively with the proteasome but not with other proteases [9, 10]. However, CFZ resistance in cell lines, likely caused by p-glycoprotein (Pgp) upregulation and multi-drug-resistance-related efflux pumps, has been observed [11]. In addition to being used to treat various blood tumors, PIs have been proposed as a treatment for solid tumors [12], parasites [13], inflammation, and immune diseases [14]. There is therefore strong interest in developing new types of PIs with enhanced efficacy, fewer side effects, and reduced drug resistance.

A. Li and H. Sun contributed to the work equally and should be regarded as co-first authors.

Electronic supplementary material The online version of this article (doi:10.1007/s00894-014-2515-y) contains supplementary material, which is available to authorized users.

A. Li · L. Du · X. Wu · J. Cao
Department of Medicinal Chemistry, China Pharmaceutical University, Nanjing 210009, China

H. Sun · Q. You · Y. Li (✉)
Jiangsu Key Laboratory of Drug Design and Optimization, China Pharmaceutical University, Nanjing 210009, China
e-mail: yuyanli@cpu.edu.cn

In UPP, proteolysis takes place in the 26S proteasome, which is composed of a 20S core particle and two 19S regulatory particles [15]. At the proteolytic site, the 20S proteasome shows three different types of catalytic activity at three distinct subunits: the β5 subunit (chymotrypsin-like), the β2 subunit (trypsin-like), and the β1 subunit (caspase-like) [16]. Among these types of catalytic activity, chymotrypsin-like activity has been shown to correspond to the rate-limiting step [17]. Therefore, the β5 subunit of the 20S proteasome has become the primary target for proteasome inhibition to treat various cancers. Recently, a variety of PIs have been designed, synthesized, and biologically evaluated. Most of these PIs are covalent inhibitors with good cell viability that selectively target the β5 subunit by forming a covalent bond with the residue Thr^1 [18], while noncovalent compounds such as PI-083 normally show low activities ($IC_{50} > 1$ μM) [19]. Thus, the study reported in the present paper, we searched for covalent PIs that target the β5 subunit of the 20S proteasome.

The design of covalent PIs requires careful optimization of both the noncovalent binding affinity and the covalent reactive warhead [20]. At present, virtual screening for covalent inhibitors of the 20S proteasome has rarely been reported since it is difficult to calculate the covalent binding energy. Furthermore, the covalent binding parts of PIs such as boronate and epoxyketone are not commonly included in silicon screening. Thus, it is quite a challenge to identify covalent PIs through virtual screening for hits with these two kinds of covalent binding parts. As shown in Fig. 1, we divided the molecule into two parts: a covalent binding part which was selected beforehand and a noncovalent binding part which was obtained by virtual screening. The epoxyketone group was selected as the covalent binding part due to the formation of stable and irreversible adducts of this group with the residue Thr^1 and because peptide epoxyketones have thus far been found to be the most selective and potent PIs [18]. The geometries of the hits were first screened using Libdock [21] to ensure that the hits could enter the pocket of the receptor. Figure 2c shows the process of building the pharmacophore model of the noncovalent binding part for further screening. In the next step, the covalent binding part was attached at the marked position for covalent docking to investigate its binding mode. In order to validate the covalent binding, the distance between the carbonyl group of the epoxyketone and the hydroxyl oxygen of the residue Thr^1 was monitored by performing molecular dynamics (MD) simulations of the noncovalent docking complexes. Based on Soliman's research, MD simulations for covalent docking complexes were performed to demonstrate the stable interaction patterns of the screened inhibitors [22]. Finally, two hits with novel backbones (Table 1), reasonable interactions, and stable binding modes (Figs. 10, 11) were found and will be validated by synthesis and biological evaluation.

Materials and methods

Pharmacophore model generation

Mouse constitutive 20S proteasome with the epoxyketone inhibitor PR-957 (PDB code: 3UNB) was applied to generate the pharmacophore model, as the human proteasome is identical to the murine one in the area of the binding pocket [23]. The software LigandScout 3.03 [24] was used to detect crucial interactions and then to automatically create an advanced pharmacophore model.

To obtain the features of the noncovalent binding part for virtual screening, we deleted the features of the covalent binding part. Meanwhile, the nitrogen of the morpholine group (shown as the blue rays in Fig. 2a) and the hydrophobic site nearby are not common features of most PIs [15] (such as BTZ and CFZ), and were also deleted manually. The resulting model used for virtual screening contained two hydrogen-bond acceptors (HBA), two hydrogen-bond donors (HBD), one hydrophobic group, and several excluded volumes (Fig. 2c). The pivotal interactions of this model (Fig. 2d) were the basic noncovalent hydrogen-bond interactions of epoxyketones [23]. The processed model was then saved in Catalyst Hypoedit Script format, which was later edited using the Hypoedit tool to enable usage in Discovery Studio 3.0 [25]. The pharmacophore model was validated with three known ligands and evaluated using ROC analysis (see sections SM-1 and SM-2 in the "Electronic supplementary material," ESM).

Virtual screening

A schematic summary of the overall procedure is presented in Fig. 3. The LibDock protocol was used for crude screening of the SPECS database (371,557 compounds) [26]. A sphere 10 Å in radius and centered on the centroid of the ligand in the 20S proteasome (3UNB) was defined, while the default parameters of other parameters were retained.

Further screening was conducted using the pharmacophore model of the noncovalent binding part. The optimal conformation of each compound in the database was obtained with the "flexible" fitting method, the "FAST" conformation generation method, and the "best mapping only" option in the Ligand Pharmacophore Mapping protocol. After pharmacophore screening, compounds were ranked according to their fit values, and the threshold value (fit value >2.5) used for filtering was set based on the results of ROC analysis.

Covalent binding part attachment

Due to the restrictions of the software, it is difficult to take covalent binding parts into consideration when processing the pharmacophore model. One way of circumventing this

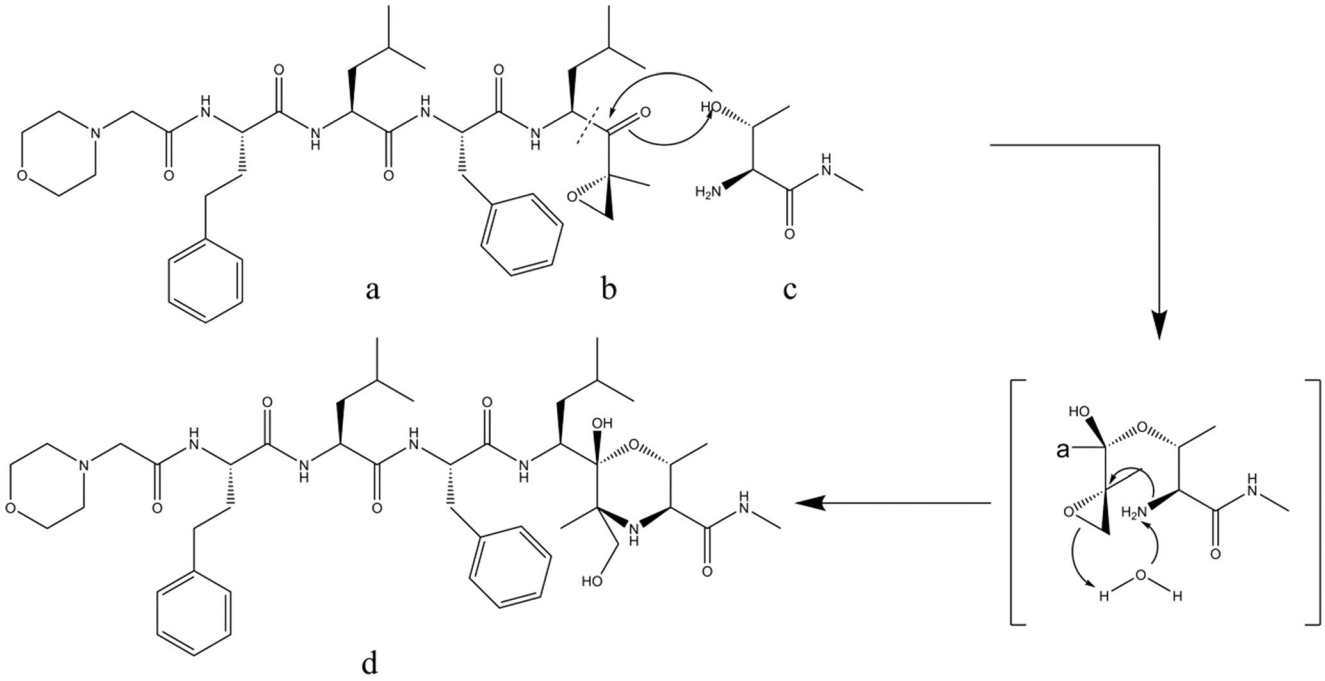


Fig. 1a–d Mechanism leading to the formation of the morpholine adduct between CFZ and the residue Thr[1]. **a** Noncovalent binding part of CFZ. **b** Covalent binding part of CFZ. **c** Residue Thr[1]. **d** Morpholine adduct

problem is to create the covalent feature by defining a series of electrophilic fragments manually [27, 28]. However, if the epoxyketone group is defined as a new feature, only a few epoxyketones (362 in the Zinc [29] database) are available, and thus no novel backbones can be screened. In this study, features of the covalent binding part were deleted and later appended with the epoxyketone group after screening to search for epoxyketones with novel backbones.

Compounds with fit values of >2.5 were selected to analyze the pharmacophore model complex derived from pharmacophore screening. The complex of PR957 (Fig. 4a) was validated when the epoxyketone group could be appended near the hydrogen-bond acceptor (HBA) feature and the hydrophobic feature. A methylene group was needed between these two features, as this allowed the epoxyketone group to be attached in the right direction without incurring any conflicts. As shown in Fig. 4b (where the excluded volume is hidden to improve visualization), the ring-opening form of the α,β-epoxyketone group was attached to the methylene group close to the HBA feature.

The candidate molecules were selected based on the following four criteria. (1) A methylene group exists between the HBA feature and the hydrophobic feature, which ensures that the epoxyketone group points in the correct direction towards residue Thr[1]. (2) Hydrogens of the methylene group should point in the opposite direction to the pharmacophore features. (3) There is no conflict between the molecules and the excluded volume. (4) The attachment operation should retain the original chirality of PR957 and cause no conflict between the molecules and the attached fragment. For compounds with the same scaffold, only the compound with the highest fit value was retained. Compounds that met the criteria were attached to the epoxyketone group for subsequent covalent docking studies.

Covalent docking

GOLD version 5.0 [30] was used to implement covalent docking studies with the same protein–ligand complex (3UNB) as described by Siwei Zhang et al. [31]. The genetic algorithm (GA) method was adopted by GOLD for conformational analysis and docking evaluation. The default GA settings of GOLD were used: population size=100; selection pressure=1.1; operations=100,000; islands=5; niche size=2; migration=10; mutation=95; crossover=95. The centroid of the co-crystallized ligand was regarded as the docking cavity. The carbonyl group of the epoxyketone (Fig. 1, part b) was set as the link atom that binds to the hydroxyl oxygen of Thr [1] (Fig. 1c) in the covalent option. In order to identify potential leads, different binding poses were ranked using the Goldscore function and the interactions were validated to select reasonable hits. The covalent docking was evaluated by docking PR957 into the same protein using GOLD (see section SM-3 in the ESM).

Molecular dynamics simulations

Molecular dynamics simulations for noncovalent docking complexes

The noncovalent binding model is a key component of investigations into the binding affinities between covalent PIs and

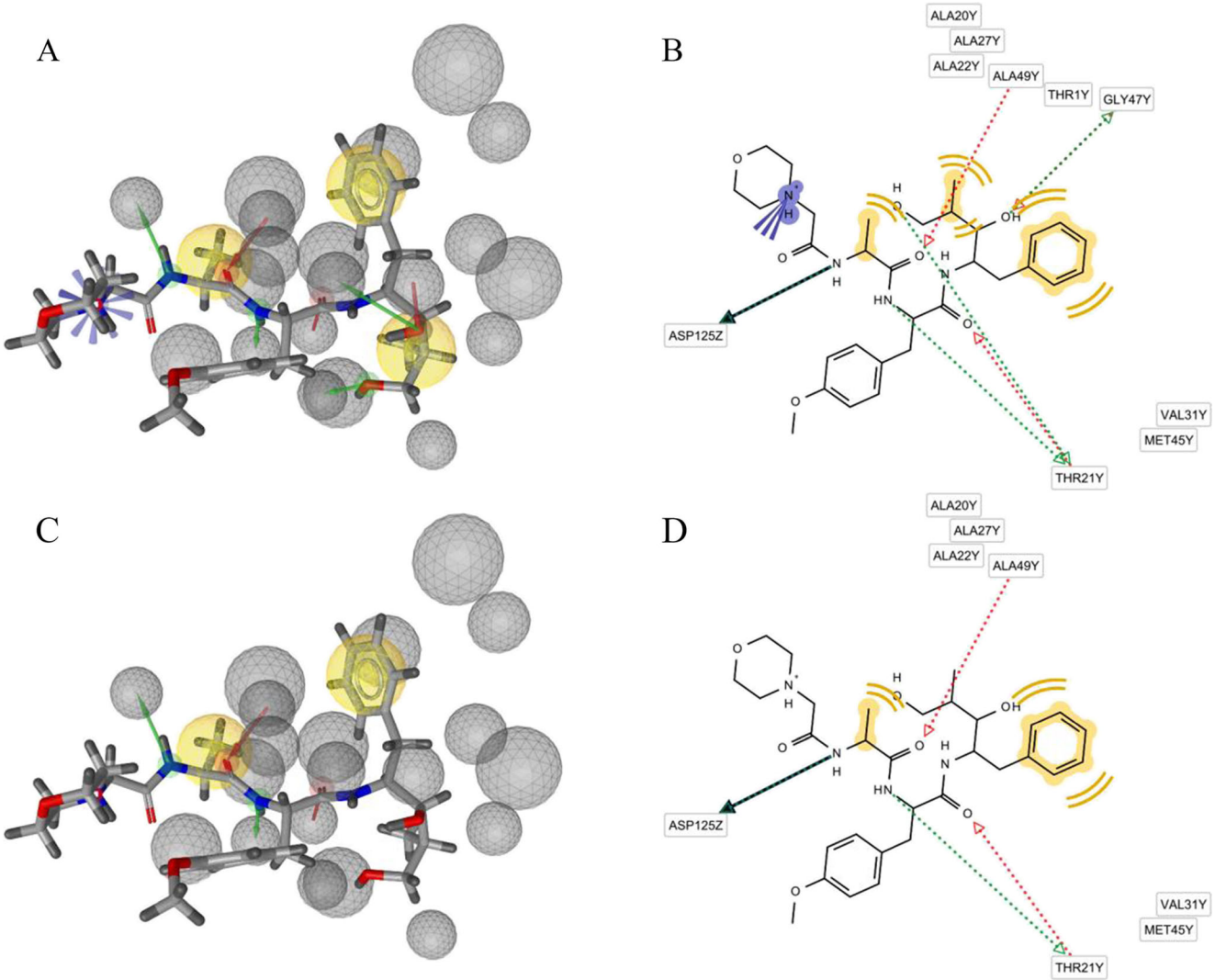


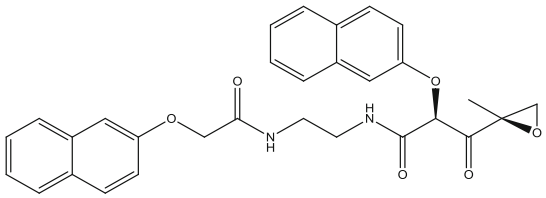
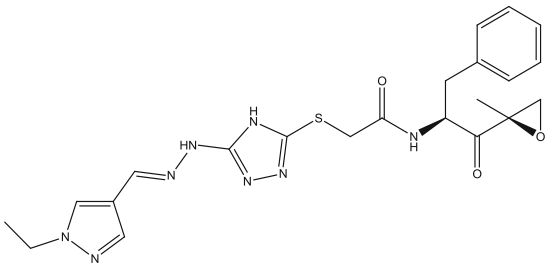
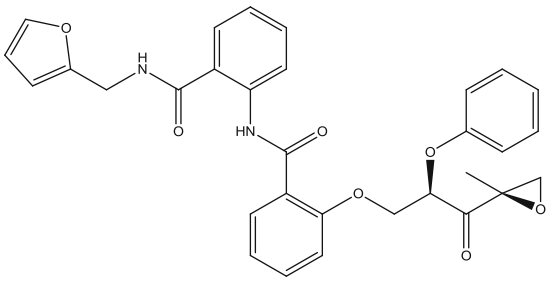
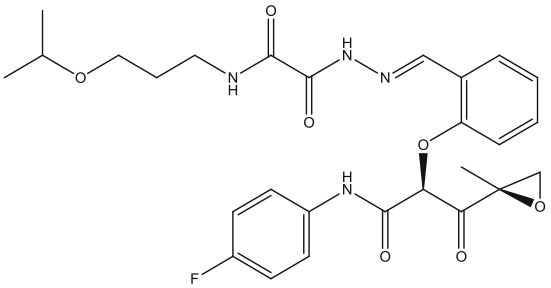
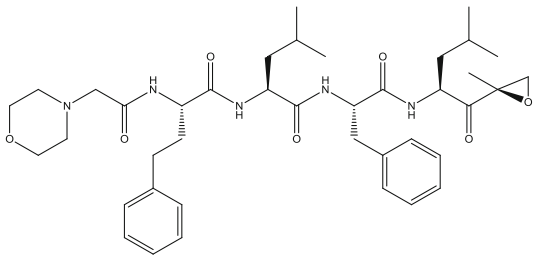
Fig. 2a–d Process of generating the pharmacophore model for screening the noncovalent binding part. **a** LigandScout pharmacophore model built from the 20S proteasome–PR957 complex (3UNB). **b** 2D interactions of PR957 in the ligand-binding pocket. **c** Processed model built from the noncovalent binding interactions. **d** Basic noncovalent hydrogen-bond interactions of PR957 in the ligand-binding pocket. *Red arrows* HBA, *green arrow* HBD, *yellow spheres* hydrophobic sites, *blue rays* positive ionizable area, *gray spheres* excluded volumes

the β5 subunit. Predicting the noncovalent binding modes of covalent PIs can aid the development of potent covalent PIs [32]. When the carbonyl group is close enough to the hydroxyl oxygen, it can be covalently bonded through a nucleophilic reaction (Fig. 1). The distance between the carbonyl group of the epoxyketone and the hydroxyl of the residue Thr^1 was monitored throughout the MD simulations. The corresponding atomic distances in CFZ and PR957 served as references.

The initial structures were acquired from noncovalent docking experiments by Glide's standard precision (SP) method [33]. MD calculations were carried out using the Desmond software [34] and the OPLS-2005 force field. TIP3P (transferable intermolecular potential 3-point) water [35] was added in a orthorhombic box (size 75 Å × 69 Å × 81 Å), which ensured that the entire surface of each complex was covered by the solvent. The systems were then neutralized by adding Cl^- ions, and salt (0.15 mol/L NaCl) was also added to construct the solvent environment. The resulting system contained approximately 37,510 atoms.

A series of predefined minimizations and MDs were executed to relax the system before the production simulation using the default relaxation protocol in Desmond. The relaxed system was simulated for 10 ns with a time step of 2 fs in the RESPA (REversible reference System Propagator Algorithm) integrator option. The NPT ensemble was performed with the Nosé–Hoover thermostat method and the Martyna–Tobias–Klein barostat method to keep the system at 300 K and a pressure of 1.01325 bar. The cutoff radius was set to 10 Å for short-range forces and the particle-mesh Ewald (PME) method was adopted to calculate long-range forces. The SHAKE

Table 1 Structures of the final four hits with epoxyketone

Hits	Structure	Distance[a](Å) ($\Delta_{distance}$[b])	ADME properties[c]	
			Stars[d]	HumanOral Absorption[e]
Compound 1		5.25(6.13)	0	3
Compound 2		5.24(3.48)	1	2
Compound 3		3.30(2.18)	4	1
Compound 4		6.86(7.07)	1	1
CFZ		4.51(2.81)	6	1

Distance[a] The average distance between carbonyl group of epoxyketone and the hydroxyl of Thr[1] from MD simuatons.

$\Delta_{distance}$[b] the difference between the maximum and the minimum distances.

ADME properties[c] : Calculated using QikProp (QikProp, version 3.1).

Stars[d] Number of property or descriptor values that fall outside the 95 % range of similar values for known drugs.

HumanOralAbsorption[e] : Predicted qualitative human oral absorption: 1, 2, or 3 for low, medium, or high.

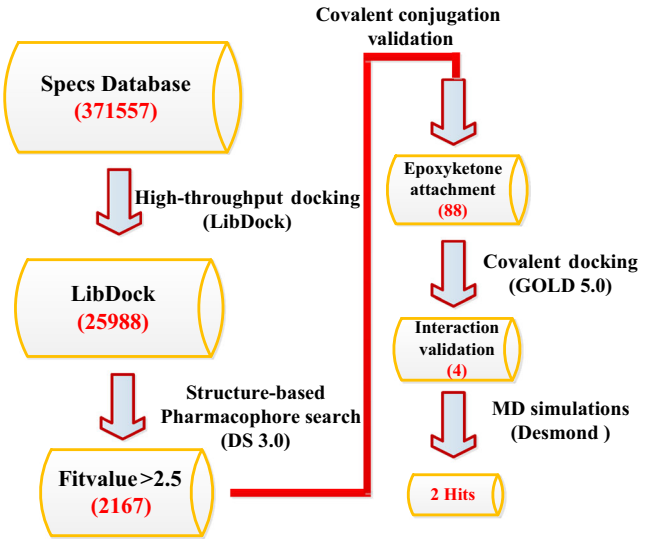


Fig. 3 Flowchart of the linear virtual screening procedure (the number of hits is shown in *parentheses*)

algorithm was used to restrict all covalent bonds involving hydrogen atoms [36]. Energies were recorded every 1.2 ps for Desmond simulation quality analysis and trajectories were collected every 4.8 ps for Desmond simulation event analysis. The RMSDs of the complexes were calculated throughout the simulations, with the first frame used as reference.

Molecular dynamics simulations for covalent docking complexes

MD simulations for covalent docking complexes were carried out in order to check the stability of the docked ligand–enzyme complexes and to provide insight into the binding affinities and interaction patterns of the screened inhibitors.

The initial structures were acquired from the covalent docking experiments by GOLD. The method used to dock covalently bound ligands in GOLD was to force the link atom in the ligand to fit onto the link atom in the protein instead of covalent bonding [30]. Thus, a slight modification was utilized for MD simulations of covalent docking complexes [37]. The inhibitors were linked via a covalent bond between the hydroxyl oxygen of Thr^1 and the carbon atom of the inhibitor while removing a redundant link atom from the structure. During this operation, the positions of the atoms remained unchanged. The ligand topologies of both noncovalent complexes (ligands only) and covalent complexes (ligands with Thr^1), including atom types and partial charges, are summarized in section SM-6 of the ESM. The modified structures were used to build MD systems and run MD simulations utilizing the same procedure as that described in "Molecular dynamics simulations for noncovalent docking complexes."

Results and discussion

Hit compound analysis

By using complementary docking-based and pharmacophore-based screening, we searched the SPECS database via a linear virtual screening strategy (Fig. 3). After noncovalent docking screening, 25,988 compounds were filtered by pharmacophore-based screening. 2167 compounds meeting the requirement (fit value >2.5) were examined with their pharmacophore model complexes to check whether the epoxyketone group could be appended in place. Only 88 compounds passed this screen, and these compounds were covalently docked into the active site of the 20S proteasome. Four compounds with reasonable interactions were finally chosen for MD simulations and ADME investigations (Table 1). Compounds 2 and 3 exhibited stable interactions, and compound 2 exhibited good drug-like characteristics according to ADME predictions.

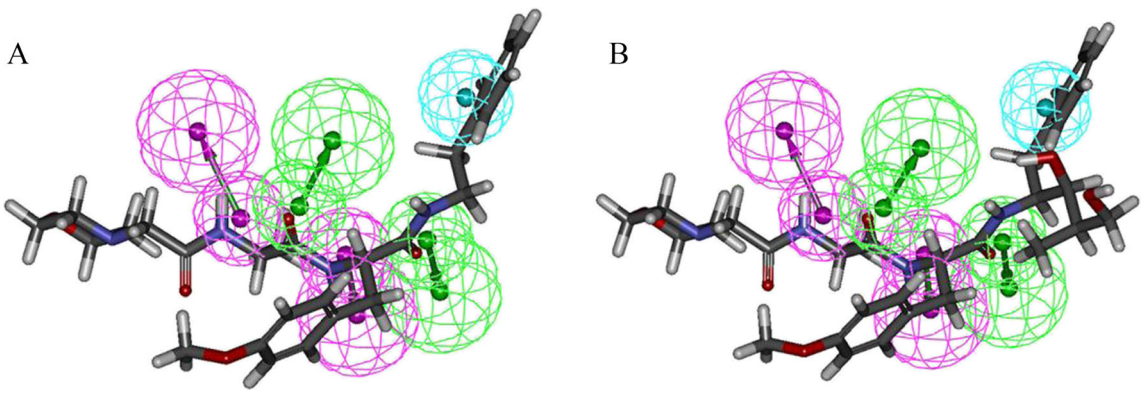


Fig. 4 **a** PR957 without epoxyketone mapped to the pharmacophore model. **b** PR957 remapped to the pharmacophore model after attaching the ring-opening form of the α,β-epoxyketone group. For simplicity, the excluded volume is not shown

Pharmacophore mapping

The four hits without epoxyketone were mapped to the pharmacophore model to conveniently examine the spatial structure (Fig. 5). In compound 1, the naphthyl group occupies the hydrophobic feature. The carbonyl group and the oxygen of the ether serve as two HBA features, while two amino groups serve as two HBD features. In compound 2, the hydrophobic feature is occupied by the phenyl group. The triazole group serves as an HBA feature and an HBD feature, and the sulfur atom maps the HBA feature. In compound 3, the phenyl group also occupies the hydrophobic feature, and other features mimic the patterns of compound 1. The features of compound 4 are similar to those of compound 3, except for one HBA feature. Comparing the mapping results with those of PR957 (Fig. 4b), hit performance can be ranked as follows: compound 2, compound 3, compound 4, and compound 1.

Covalent docking

Covalent docking was performed to visualize the favorable interactions between the four hits and the active site of the 20S proteasome. The best docking poses of the four hits were analyzed by monitoring crucial residues. In the binding mode of compound 1, two acid amide groups form four hydrogen bonds with the residues Gly^{47}, Ala^{49}, and Thr^{21}, respectively. A hydrogen-bond interaction is also observed between the terminal oxygen and the carbonyl group of the residue Asp^{125} (Fig. 6a). This binding mode favors all hydrogen bonds between the co-crystallized ligand (PR957) and the

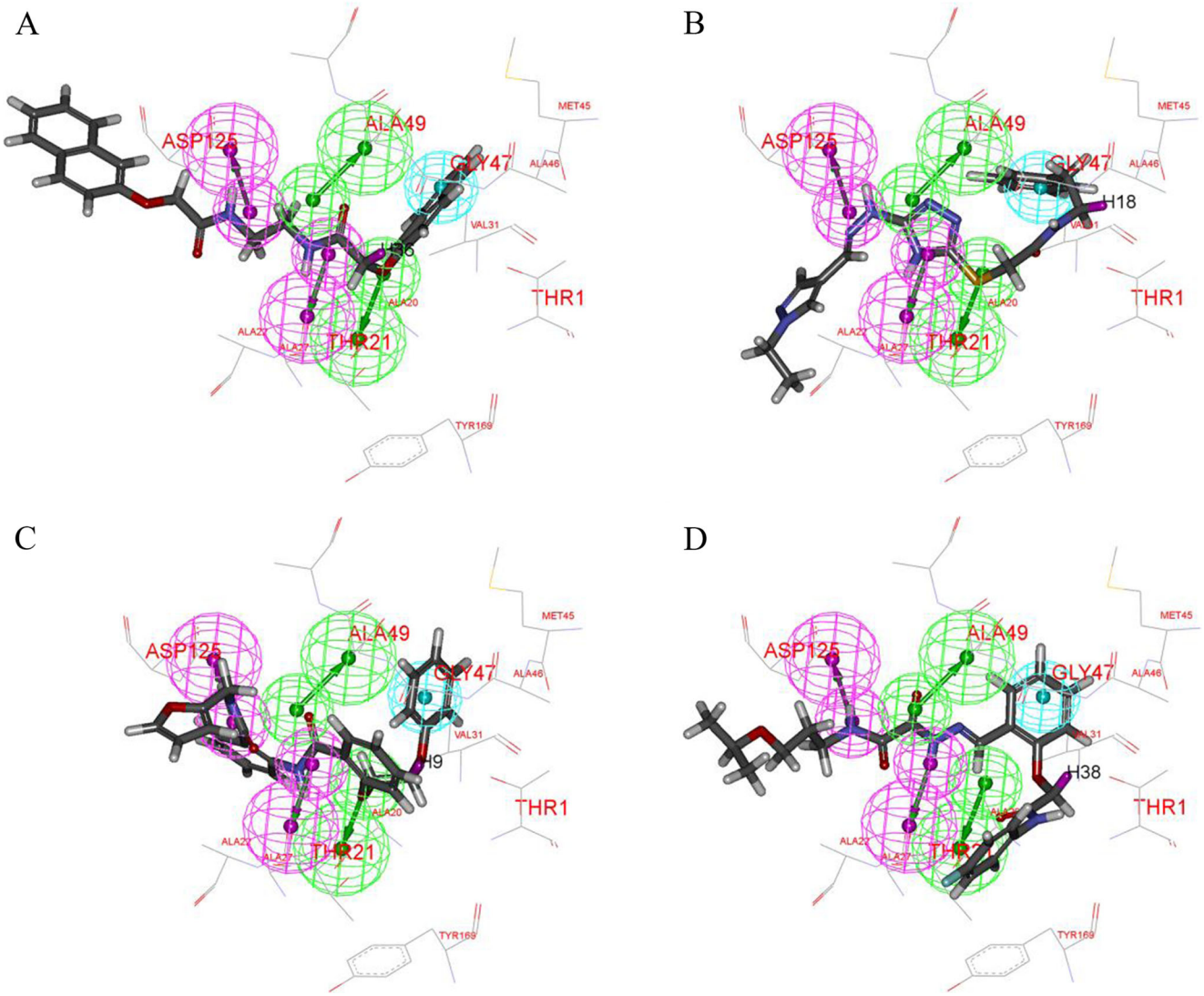


Fig. 5a–d Four hits without epoxyketone were mapped to the pharmacophore model. **a** Compound 1. **b** Compound 2. **c** Compound 3. **d** Compound 4. In the diagrams, the excluded volume is replaced with protein residues to improve visualization. The hydrogen through which the epoxyketone group is attached is labeled and is colored *purple*

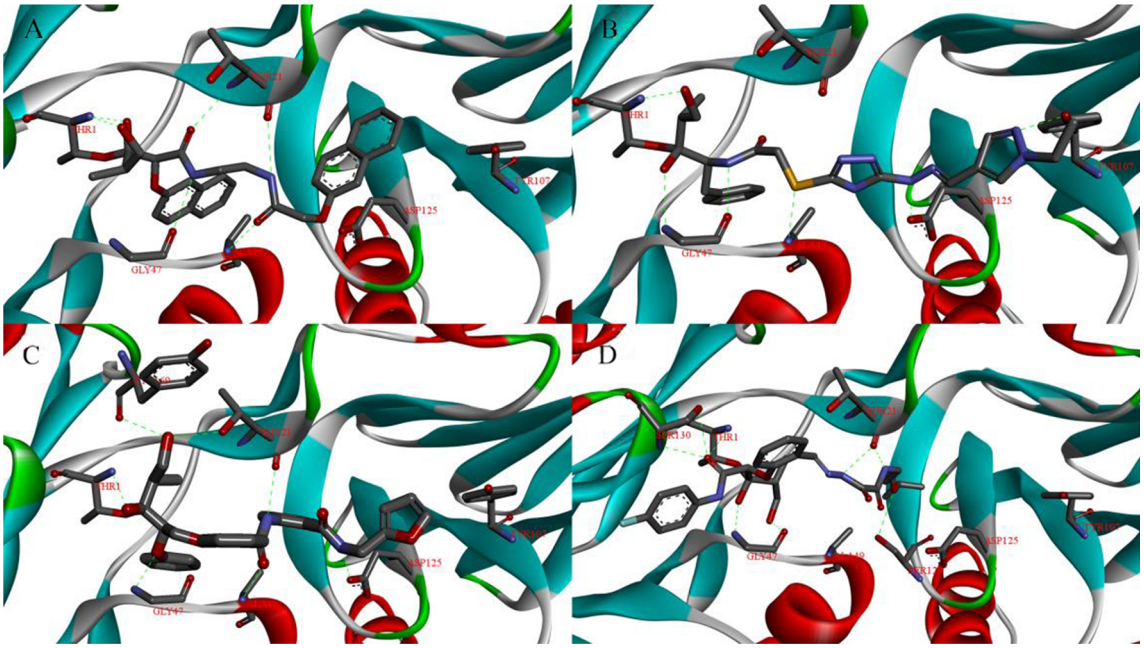


Fig. 6a–d Best binding poses of the four hits from the covalent docking results. **a** Compound 1. **b** Compound 2. **c** Compound 3. **d** Compound 4

active pocket of the β5 subunit in 3UNB (see Fig. S3a in the ESM), as well as that of compound 2 (Fig. 6c). The binding mode of compound 3 is special in that no hydrogen bonds are formed with residue Thr^{21} (Fig. 6c). Unlike other hits, the phenyl group of compound 4 does not occupy the hydrophobic pocket below residue Thr^{1}, and no hydrogen bond is observed with residue Asp^{125} (Fig. 6d), which is known to be a significant residue in inhibitory activities [38].

Molecular dynamics results

Molecular dynamics simulations for noncovalent docking complexes

MD simulations of noncovalent docking complexes were performed to detect whether covalent bonds can form between the four hits and the residue Thr^{1}. The atomic distance between the carbonyl group of the epoxyketone and the hydroxyl oxygen of the residue Thr^{1} was monitored during MD simulations. The corresponding atomic distances in CFZ and PR957 were used as references because the two compounds are both active 20S PIs.

Throughout the 10-ns MD simulations, the RMSD values for all backbones and compounds remained within the limits of 5 Å and 3 Å, respectively, and all RMSD variations were stable after 2.5 ns of simulation. All of the variations in the total potential energy were less than 900 kcal mol^{-1} (see Figs. S4–S7 in the ESM). When the carbonyl group is close enough to the hydroxyl oxygen, a covalent bond can form. For CFZ, the atomic distance was around 3.09–4.63 Å, and the average distance was 4.51 Å. The difference between the maximum and minimum distances ($\Delta_{distance}$) was 2.81 Å,

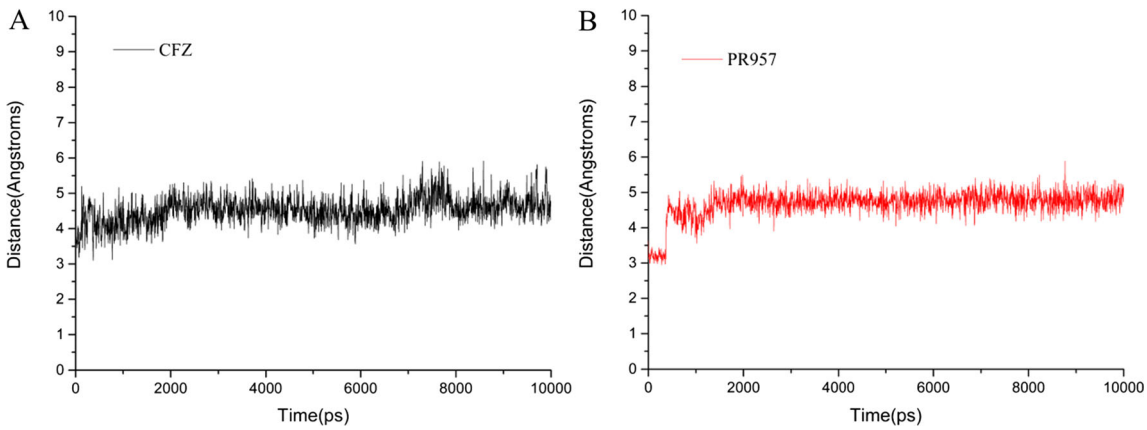


Fig. 7a–b Time evolution of the distance between the carbonyl group of the epoxyketone and the hydroxyl oxygen of the residue Thr^{1} in **a** carfilzomib (CFZ) and **b** PR957

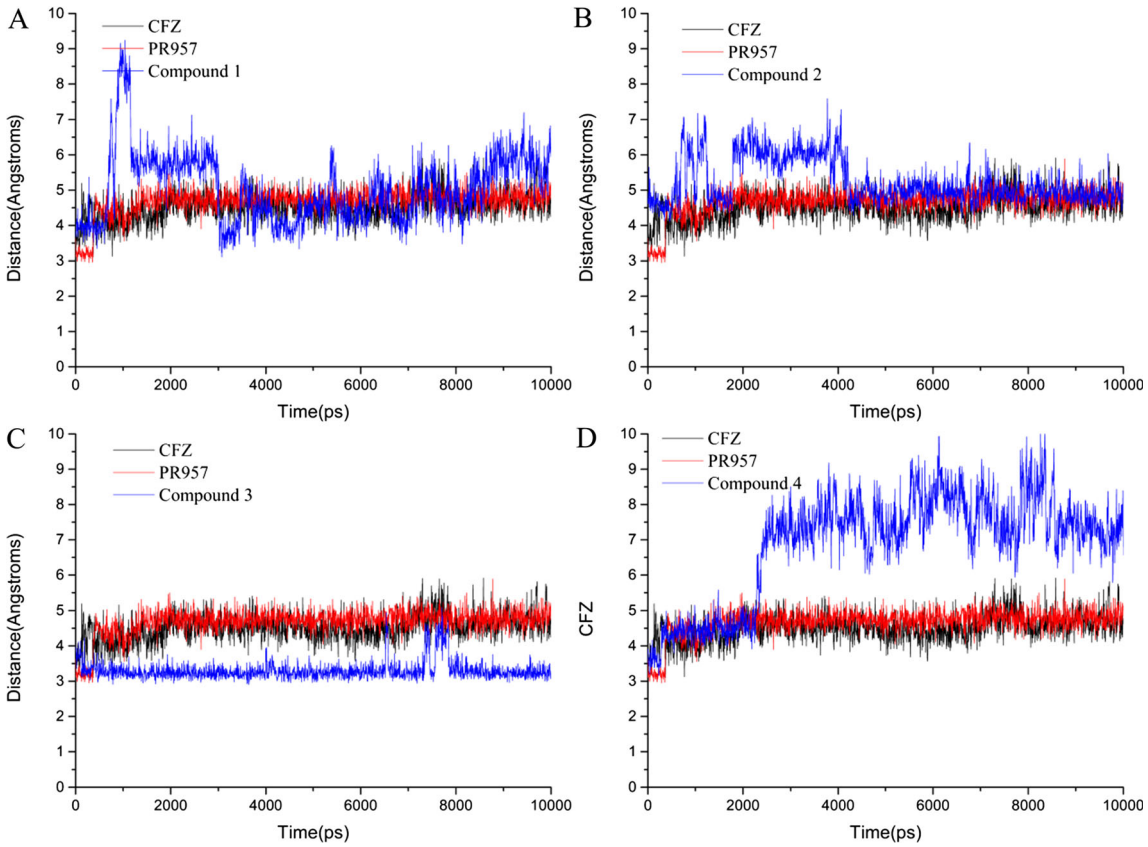


Fig. 8a–d Time evolution of the distance between the carbonyl group of the epoxyketone and the hydroxyl oxygen of the residue Thr[1] in **a** compound 1, **b** compound 2, **c** compound 3, **d** compound 4 (the corresponding distances in CFZ and PR957 are also plotted for reference purposes)

which indicated that a stable binding mode can form (Fig. 7a). For PR957, the atomic distance was around 2.94–5.88 Å, and the average distance was 4.67 Å. $\Delta_{distance}$ was 2.94 Å (Fig. 7b). The similar performances of CFZ and PR957 indicated a stable binding model environment for the formation of the covalent bond. Thus, these distances can be used as references for the hits.

The distance between the carbonyl group and the hydroxyl oxygen was also monitored during the 10-ns MD simulations of the hits.

Due to its rigid structure, the atomic distance (2.90–5.08 Å), the average distance (3.30 Å), and $\Delta_{distance}$ (2.18 Å) of compound 3 (Fig. 8c) were all shorter than those of CFZ and PR957, indicating a relatively stable binding mode. For compound 2, the atomic distance was around 4.11–7.60 Å and the average distance was 5.24 Å (Fig. 8b). The maximum distance and $\Delta_{distance}$ (3.48 Å) appear to be a bit longer than those in CFZ and PR957. However, the stable average distance was 4.94 Å, with a small $\Delta_{distance}$ (2.14 Å) observed in the last 5.7 ns, similar to the values observed for

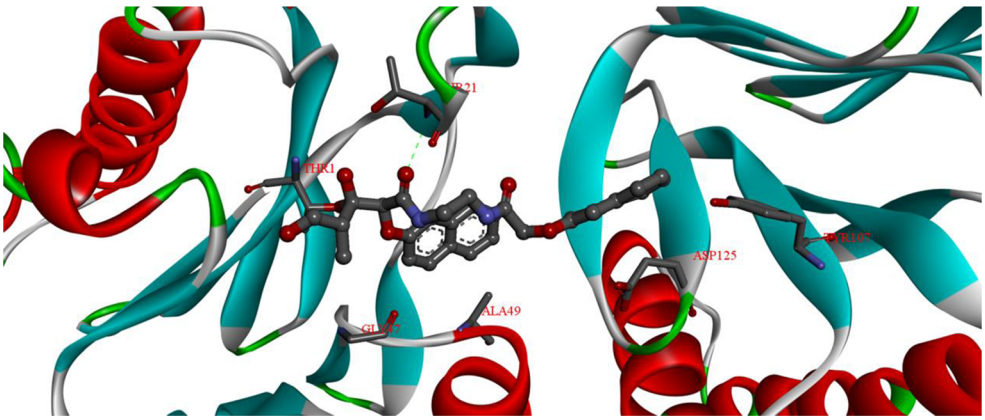


Fig. 9 Average structure of compound 1 covalently bonded to the residue Thr[1] of the 20S proteasome obtained from 10-ns MD

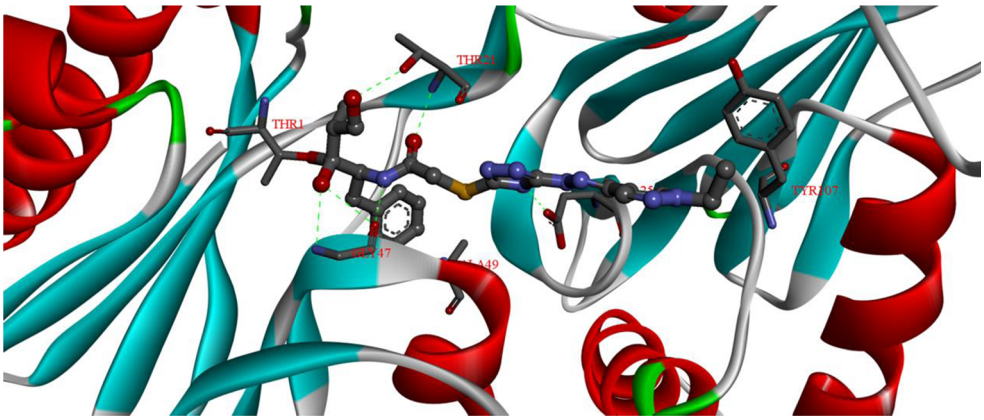


Fig. 10 Average structure of compound 2 covalently bonded to the residue Thr[1] of the 20S proteasome obtained from 10-ns MD

CFZ and PR957. The atomic distances of compound 1 (3.10–9.23 Å, Fig. 8a) and compound 4 (3.16–10.24 Å, Fig. 8d) fluctuated greatly. Both of these compounds are highly flexible and thus more likely to flip, indicating a reduced probability of covalently bonding to the residue Thr[1].

The MD simulations for noncovalent docking complexes reveal that the docking complexes of compound 2 and compound 3 maintain high conformational stability as well as short atomic distances for covalent binding. Moreover, MD simulations show that highly flexible compounds such as compound 1 and compound 4 are easy to flip, causing energy loss and reducing the chance of covalent binding.

Molecular dynamics simulations for covalent docking complexes

Ten-nanosecond simulations were performed for compounds covalently bonded to the 20S proteasome to provide insight into their binding affinities and interaction patterns. The RMSD values for all backbones and compounds remained below 4 Å and 3 Å, respectively, and all RMSD variations were stable after 2.5-ns simulations. All of the variations in the total potential energy were less than 900 kcal mol^{-1} (see Figs. S8–S9 in the ESM). These findings verify that all four compounds remain stable across the MD simulations.

The average complex structures are shown in Figs. 9–12, and they are slightly different from the covalent docking complexes. Hydrogen bonds were observed between compound 2 and the residues Thr[21], Gly[47], Ala[49], and Asp[125] (Fig. 10). This binding mode favors all hydrogen bonds between the co-crystallized ligand (PR957) and the active pocket of the β5 subunit in 3UNB (see Fig. S3a in the ESM), indicating a similar covalent binding model. Compound 3 forms a hydrogen bond with the residue Tyr[107] instead of Gly[47] and Ala[49] (Fig. 11). However, its rigid structure enhances its stability, leading to low RMSD variations during the 10-ns MD simulation (see Fig. S8 in the ESM). The results suggest that compounds 2 and 3 may exhibit high binding affinities towards the 20S proteasome due to favorable interactions.

Compound 1 has only one hydrogen bond with the binding pocket, and compound 4 interacts with the 20S proteasome by reacting with Met[45] and Gln[131]. Both of these compounds are highly flexible and less stable than compounds 2 and 3 according to the RMSD variations (see Fig. S8 in the ESM). This result corresponds well to the findings obtained from MD

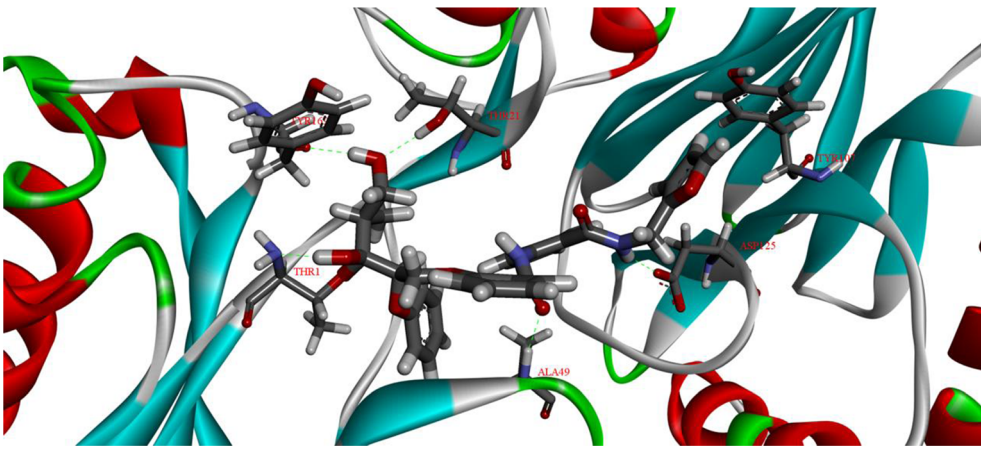


Fig. 11 Average structure of compound 3 covalently bonded to the residue Thr[1] of the 20S proteasome obtained from 10-ns MD

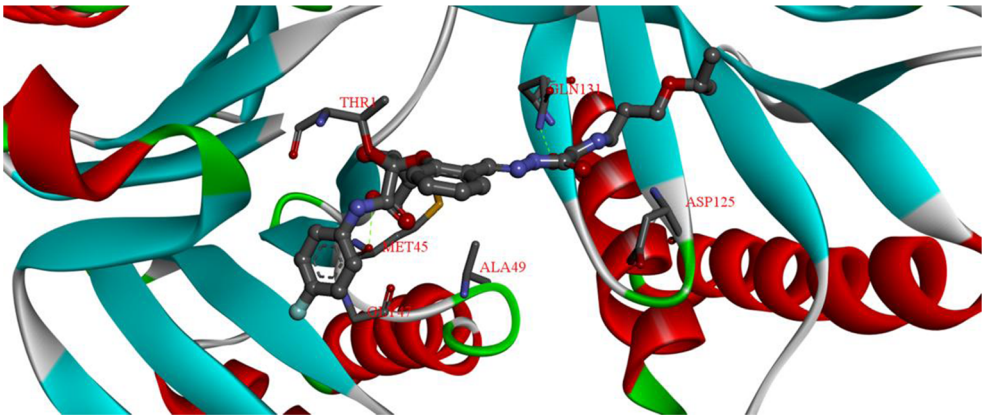


Fig. 12 Average structure of compound 4 covalently bonded to the residue Thr^1 of the 20S proteasome, obtained from 10-ns MD

simulations of the noncovalent complex: compounds 2 and 3 showed constant distances from the residue Thr^1, similar to the control PIs. Meanwhile, compounds 1 and 4, which showed fluctuating distances from the residue Thr^1, may find it difficult to form covalent bonds with the residue Thr^1 of the proteasome. Futhermore, the results of MD simulations for the noncovalent complex showed that the monitored distances were stable for the control PIs CFZ and PR957, and the average distances were <5 Å, which is close enough to allow covalent binding.

Further validation can be achieved by combining the results of the MD simulations of the noncovalent and covalent complexes to check the stability of the covalent ligands. We can thus conclude that compounds 2 and 3 may exhibit high binding affinities and compounds 1 and 4 may exhibit low binding affinities for the 20S proteasome.

ADME prediction

In addition, we investigated the ADME profiles of the compounds using Qikprop [33]. The results are summarized in Table 1. The predictions indicate that compound 2 possesses good drug-like characteristics: good oral absorption, few descriptor values that fall outside the 95 % range of similar values for known drugs, and no effects on the central nervous system; on the other hand, CFZ has low oral absorption and violates the accepted ranges of six descriptor values.

Conclusions

Covalent PIs have proven to be clinically effective, with two such compounds currently approved by the FDA. Covalent PIs with novel backbones are a promising approach to reducing drug resistance and improving ADME properties. As virtual screening for covalent inhibitors remains a challenge, we fixed the covalent binding part and screened the noncovalent binding part to indirectly obtain covalent PIs. In this study, a linear strategy which combined pharmacophore-based and docking-based methods was used to screen the novel noncovalent binding part. Then the resulting compounds were attached to the pre-selected epoxyketone group to create covalent PIs which were validated by covalent docking and MD simulation of the noncovalent and covalent complex. In the end, four potential covalent PIs were found with new scaffolds and reasonable interactions with the substrate. Among them, compounds 2 and 3 showed constant distances from the residue Thr^1 before covalent binding and exhibited stable interactions with the β5 subunit after covalent binding. Compound 2 exhibited good drug-like characteristics according to ADME predictions and will be further investigated. This procedure was designed to identify compounds with a valid covalent warhead which are rarely synthesized, such as 20S PIs, but studies attempting to find other covalent compounds with a specific covalent warhead may also benefit from using this method.

Acknowledgments This work was supported by the National Natural Science Foundation of China (grant nos. 81102317, 81230078) and the Fundamental Research Funds for the Central Universities (JKP2011001). We gratefully thank Inte: Ligand for kindly supplying the scientific trial version of Ligandscout 3.0, and would also like to thank Dr. Li Dong for proofreading this manuscript.

References

1. Ciechanover A (1998) The ubiquitin-proteasome pathway: on protein death and cell life. EMBO J 17(24):7151–7160
2. Kisselev AF, Callard A, Goldberg AL (2006) Importance of the different proteolytic sites of the proteasome and the efficacy of inhibitors varies with the protein substrate. J Biol Chem 281(13): 8582–8590
3. Obeng EA, Carlson LM, Gutman DM, Harrington WJ Jr, Lee KP, Boise LH (2006) Proteasome inhibitors induce a terminal unfolded protein response in multiple myeloma cells. Blood 107(12):4907–4916

4. Wolf J, Richardson PG, Schuster M, LeBlanc A, Walters IB, Battleman DS (2008) Utility of bortezomib retreatment in relapsed or refractory multiple myeloma patients: a multicenter case series. Clin Adv Hematol Oncol 6(10):755–760
5. O'Connor OA, Wright J, Moskowitz C, Muzzy J, MacGregor-Cortelli B, Stubblefield M, Straus D, Portlock C, Hamlin P, Choi E, Dumetrescu O, Esseltine D, Trehu E, Adams J, Schenkein D, Zelenetz AD (2005) Phase II clinical experience with the novel proteasome inhibitor Bortezomib in patients with indolent non-Hodgkin's lymphoma and mantle cell lymphoma. J Clin Oncol 23(4):676–684
6. Harvey RD, Owonikoko TK, Lewis CM, Akintayo A, Chen Z, Tighiouart M, Ramalingam SS, Fanucchi MP, Nadella P, Rogatko A, Shin DM, El-Rayes B, Khuri FR, Kauh JS (2013) A phase 1 Bayesian dose selection study of Bortezomib and Sunitinib in patients with refractory solid tumor malignancies. Br J Cancer 108(4):762–765
7. Richardson PG, Briemberg H, Jagannath S, Wen PY, Barlogie B, Berenson J, Singhal S, Siegel DS, Irwin D, Schuster M, Srkalovic G, Alexanian R, Rajkumar SV, Limentani S, Alsina M, Orlowski RZ, Najarian K, Esseltine D, Anderson KC, Amato AA (2006) Frequency, characteristics, and reversibility of peripheral neuropathy during treatment of advanced multiple myeloma with Bortezomib. J Clin Oncol 24(19):3113–3120
8. Kortuem KM, Stewart AK (2013) Carfilzomib. Blood 121(6):893–897
9. Demo SD, Kirk CJ, Aujay MA, Buchholz TJ, Dajee M, Ho MN, Jiang J, Laidig GJ, Lewis ER, Parlati F, Shenk KD, Smyth MS, Sun CM, Vallone MK, Woo TM, Molineaux CJ, Bennett MK (2007) Antitumor activity of PR-171, a novel irreversible inhibitor of the proteasome. Cancer Res 67(13):6383–6391
10. Meng L, Mohan R, Kwok BHB, Elofsson M, Sin N, Crews CM (1999) Epoxomicin, a potent and selective proteasome inhibitor, exhibits in vivo antiinflammatory activity. Proc Natl Acad Sci USA 96(18):10403–10408
11. Verbrugge SE, Assaraf YG, Dijkmans BA, Scheffer GL, Al M, den Uyl D, Oerlemans R, Chan ET, Kirk CJ, Peters GJ, van der Heijden JW, de Gruijl TD, Scheper RJ, Jansen G (2012) Inactivating PSMB5 mutations and P-glycoprotein (multidrug resistance-associated protein/ATP-binding cassette B1) mediate resistance to proteasome inhibitors: ex vivo efficacy of (immuno) proteasome inhibitors in mononuclear blood cells from patients with rheumatoid arthritis. J Pharmacol Exp Ther 341(1):174–182
12. Papandreou CN, Daliani DD, Nix D, Yang H, Madden T, Wang X, Pien CS, Millikan RE, Tu S-M, Pagliaro L, Kim J, Adams J, Elliott P, Esseltine D, Petrusich A, Dieringer P, Perez C, Logothetis CJ (2004) Phase I trial of the proteasome inhibitor Bortezomib in patients with advanced solid tumors with observations in androgen-independent prostate cancer. J Clin Oncol 22(11):2108–2121
13. Lindenthal C, Weich N, Y-S CHIA, Heussler V, klinkert M-Q (2005) The proteasome inhibitor MLN-273 blocks exoerythrocytic and erythrocytic development of plasmodium parasites. Parasitology 131(01):37–44
14. Nencioni A, Grunebach F, Patrone F, Ballestrero A, Brossart P (2006) Proteasome inhibitors: antitumor effects and beyond. Leukemia 21(1):30–36
15. Kisselev AF, Goldberg AL (2001) Proteasome inhibitors: from research tools to drug candidates. Chem Biol 8(8):739–758
16. Unno M, Mizushima T, Morimoto Y, Tomisugi Y, Tanaka K, Yasuoka N, Tsukihara T (2002) The structure of the mammalian 20S proteasome at 2.75 Å resolution. Structure 10(5):609–618
17. Rock KL, Gramm C, Rothstein L, Clark K, Stein R, Dick L, Hwang D, Goldberg AL (1994) Inhibitors of the proteasome block the degradation of most cell-proteins and the generation of peptides presented on MHC class-I molecules. Cell 78(5):761–771
18. Kisselev AF, van der Linden WA, Overkleeft HS (2012) Proteasome inhibitors: an expanding army attacking a unique target. Chem Biol 19(1):99–115
19. Xu K, Xiao Z, Tang YB, Huang L, Chen CH, Ohkoshi E, Lee KH (2012) Design and synthesis of naphthoquinone derivatives as anti-proliferative agents and 20S proteasome inhibitors. Bioorg Med Chem Lett 22(8):2772–2774
20. Singh J, Petter RC, Baillie TA, Whitty A (2011) The resurgence of covalent drugs. Nat Rev Drug Discov 10(4):307–317
21. Diller DJ, Merz KM (2001) High throughput docking for library design and library prioritization. Proteins Struct Funct Bioinforma 43(2):113–124
22. Blake L, Soliman MES (2013) Identification of irreversible protein splicing inhibitors as potential anti-TB drugs: insight from hybrid non-covalent/covalent docking virtual screening and molecular dynamics simulations. Med Chem Res 23(5): 2312–2323
23. Huber EM, Basler M, Schwab R, Heinemeyer W, Kirk CJ, Groettrup M, Groll M (2012) Immuno- and constitutive proteasome crystal structures reveal differences in substrate and inhibitor specificity. Cell 148(4):727–738
24. Wolber G, Langer T (2004) LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model 45(1):160–169
25. Accelrys, Inc. (2010) Studio D 3.0: advanced software package for life science researchers. Accelrys, Inc., San Diego
26. Janca M, Dobrovolny P (2009) Methods for preparing Bortezomib and intermediates used in its manufacture. Patent WO2009004350 A1
27. Steindl T, Laggner C, Langer T (2005) Human rhinovirus 3C protease: generation of pharmacophore models for peptidic and nonpeptidic inhibitors and their application in virtual screening. J Chem Inf Model 45(3):716–724
28. Markt P, McGoohan C, Walker B, Kirchmair J, Feldmann C, Martino GD, Spitzer G, Distinto S, Schuster D, Wolber G, Laggner C, Langer T (2008) Discovery of novel cathepsin S inhibitors by pharmacophore-based virtual high-throughput screening. J Chem Inf Model 48(8):1693–1705
29. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG (2012) ZINC: a free tool to discover chemistry for biology. J Chem Inf Model 52(7):1757–1768
30. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267(3):727–748
31. Zhang S, Shi Y, Jin H, Liu Z, Zhang L, Zhang L (2009) Covalent complexes of proteasome model with peptide aldehyde inhibitors MG132 and MG101: docking and molecular dynamics study. J Mol Model 15(12):1481–1490
32. Kawamura S, Unno Y, Tanaka M, Sasaki T, Yamano A, Hirokawa T, Kameda T, Asai A, Arisawa M, Shuto S (2013) Investigation of the noncovalent binding mode of covalent proteasome inhibitors around the transition state by combined use of cyclopropylic strain-based conformational restriction and computational modeling. J Med Chem 56(14):5829–5842
33. Schrödinger, LLC (2008) Schrödinger Suite. Schrödinger, LLC, New York
34. Bowers KJ, Chow E, Huageng X, Dror RO, Eastwood MP, Gregersen BA, Klepeis JL, Kolossvary I, Moraes MA, Sacerdoti FD, Salmon JK, Yibing S, Shaw DE (2006) Scalable algorithms for molecular dynamics simulations on commodity clusters. In: Proc ACM/IEEE Conf on Supercomputing, 11–17 Nov 2006, pp 43–43
35. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926
36. Ryckaert J-P, Ciccotti G, Berendsen HJ (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J Comput Phys 23(3):327–341

37. Kojtari A, Shah V, Babinec JS, Yang C, Ji HF (2014) Structure-based drug design of diphenyl alpha-aminoalkylphosphonates as prostate-specific antigen antagonists. J Chem Inf Model 54(10):2967–2979
38. Geurink PP, van der Linden WA, Mirabella AC, Gallastegui N, de Bruin G, Blom AEM, Voges MJ, Mock ED, Florea BI, van der Marel GA, Driessen C, van der Stelt M, Groll M, Overkleeft HS, Kisselev AF (2013) Incorporation of non-natural amino acids improves cell permeability and potency of specific inhibitors of proteasome trypsin-like sites. J Med Chem 56(3):1262–1275