J Mol Model (2009) 15:1481–1490
DOI 10.1007/s00894-009-0515-0

ORIGINAL PAPER

Covalent complexes of proteasome model with peptide aldehyde inhibitors MG132 and MG101: docking and molecular dynamics study

Siwei Zhang · Yawei Shi · Hongwei Jin · Zhenming Liu · Liangren Zhang · Lihe Zhang

Received: 3 January 2009 / Accepted: 1 April 2009 / Published online: 14 May 2009

Abstract 20S proteasome plays a critical role in the regulation of several important cellular processes and has drawn extensive interest in the field of anti-tumor research. Peptide aldehydes can inhibit the 20S proteasome activity by covalently binding to the active site of the β subunits. In this work, covalent docking in conjunction with molecular dynamics (MD) simulation was used to explore the binding mode of MG132. Two conformations with the lowest docking energy were selected as the representative binding modes. One of the conformations was confirmed as a more reasonable binding mode by molecular dynamics simulations. The binding mode analysis revealed that a space demanding aromatic group with a short linker at the P4 site of the peptide aldehyde inhibitor would form favorable hydrophobic contacts with the neighboring subunit. A bulky substituent at the P2 position would also increase the binding stability by reducing water accessibility of the covalent bond. This study contributed to our understanding of the mechanism and structure-activity relationship of the peptide aldehyde inhibitors and may provide useful information for rational drug design.

Keywords Covalent docking · Molecular dynamics · Peptide aldehydes · Proteasome · Structure-activity relationship

Electronic supplementary material The online version of this article (doi:10.1007/s00894-009-0515-0) contains supplementary material, which is available to authorized users.

S. Zhang · Y. Shi · H. Jin · Z. Liu · L. Zhang (✉) · L. Zhang
State Key Laboratory of Natural and Biomimetic Drugs,
School of Pharmaceutical Sciences, Peking University,
Beijing 100191, China
e-mail: liangren@bjmu.edu.cn

Introduction

The ubiquitin-proteasome ATP-dependent pathway, which was identified more than 20 years ago [1, 2], plays an essential role in the degradation of key components of the molecular machinery involved in many important cellular functions such as transcription, cell cycle progression, tumor suppression, and apoptosis [3–6]. The 20S proteasome, a multicatalytic complex (700 kDa), constitutes the catalytic component of the ubiquitous proteolytic machinery of the 26S proteasome. X-ray analysis showed that 20S proteasome are composed of four stacked rings, with each ring consisting of seven α- and β-type subunits. Among the seven different β subunits, β1, β2, and β5, which were originally characterized as post-glutamyl peptide hydrolytic (PGPH), trypsin-like (T-L), and chymotrypsin-like (CT-L) activity, respectively, are the catalytical sites with an N-terminal residue Thr^1 [7]. They are responsible for cleaving small chromogenic substrates and cutting polypeptides into 3–22 residues [8–10].

Proteasome inhibitors have been used extensively as research tools. The recent observation that they inhibit cell proliferation and angiogenesis and selectively induce apoptosis of tumor cells has led to their application as potential cancer therapeutics [11, 12]. Peptide aldehydes were the first discovered inhibitors of the 20S proteasome, and they are still actively investigated *in vitro* and *in vivo* for their rapid reversibility [13–15]. They inhibit the proteasome by reacting with the hydroxyl group of the N-terminal threonine of the β subunits, forming a reversible hemiacetal that has been confirmed by the crystal structure of tripeptide aldehyde MG101 (Ac-Leu-Leu-nLeu-al) (Fig. 1) in complex with 20S proteasome [16]. However, their clinical usage was greatly limited due to their undesirable selectivity and specificity. Therefore, designing more selec-

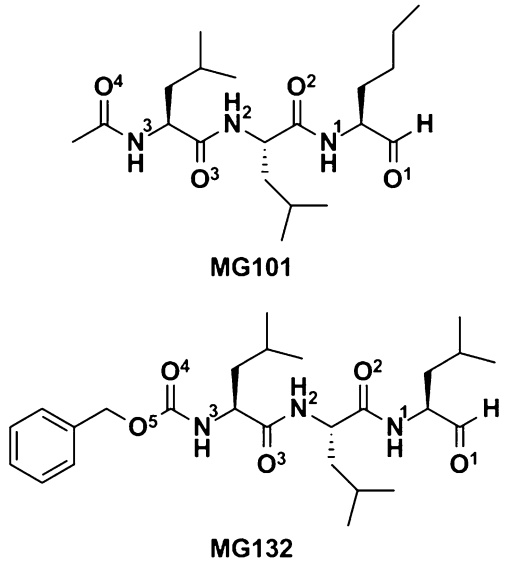

Fig. 1 The structures of MG101 and MG132. The ID number of each nitrogen and oxygen atom is shown

tive compounds with specific binding to the proteosome active sites is highly desirable.

So far, the only structural data available for the peptide aldehydes in complex with proteasome is that of MG101 [11]. MG132 (Z-Leu-Leu-Leu-al) (Fig. 1) is a leading structure of peptide aldehydes, which is not only significantly more potent against the proteasome but also much more selective than MG101 [17, 18]. It is therefore widely used for probing ubiquitin-proteasome ATP-dependent pathway. Hence, the investigation of its binding mode may shed light on the common binding behavior of peptide aldehydes and provide guides for the further improvement of the existing inhibitors.

Previous studies revealed that both S1 and S3 pockets [19] play a prominent role in the affinity and selectivity of the inhibitors, and the P1-leucine side chain of MG132 contributes to its improved selectivity comparing with MG101 [10, 11, 20]. Ligand-based QSAR approach has been applied in the investigation of peptide aldehyde proteasome inhibitors, demonstrating that medium-sized, positively charged groups at P1 along with bulky electron-rich substituents at P3 are beneficial to the activity [21]. However, P2 and P4 positions are less extensively investigated, judging from the fewer number of reports on their structure-activity relationship (SAR). And due to the lack of structural information on the receptor, ligand-based QSAR could not provide integral and precise information for the design of new proteasome inhibitors. Nevertheless, docking studies, which can take the receptor conformation into account, are also confronted with a problem in the study of proteasome inhibitors [22, 23], because covalent binding is the unique feature of proteasome inhibition and this covalent binding could not be closely imitated by the traditional non-covalent docking method.

In this study, a covalent docking approach was used to investigate the binding mode of MG132 with proteasome. Further molecular dynamics simulation provided an integral and precise binding process, since the target and ligand are all considered flexible throughout the conformational search [24]. First, docking studies were done to generate the binding position of MG132 to 20S proteasome. Two conformations with the lowest docking energy were then selected for binding mode analysis including a 5 ns molecular dynamics simulation. Comparing with the crystal structures of MG101-proteasome complex and the proteasome without inhibitor (PR), a more reasonable binding mode of MG132 was finally confirmed. Through explicit analysis and comparison with current experimental data, we proposed a molecular model that provides insight into the binding mode of MG132 to proteasome. The model allows new discovery and re-examination of the correlation between the structure and activity of the inhibitors especially at the P2 and P4 sites, therefore aiding the design of more potent and promising proteasome inhibitors (Fig. 2).

Methods

Molecular docking

The crystal structure of yeast proteasome:MG101 complex was kindly provided by Prof. Groll from *Charité Institut für*

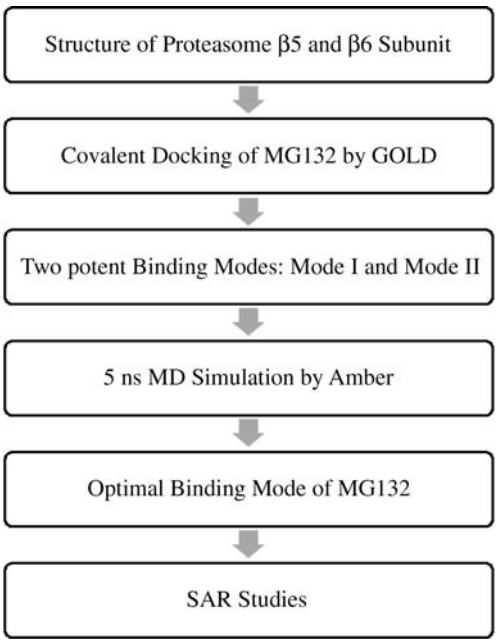

Fig. 2 Protocol of the covalent docking and dynamics simulation for an optimal binding mode of MG132 and further SAR studies

Biochemie and used for all the docking studies presented here. The protein and ligands for docking were prepared using the software Insight II [25]. The initial structure of MG101 was derived from the crystal complex coordinates, and the original structure of MG132 was constructed based upon the MG101 crystal structure conformation followed by energy minimization. MG101 and MG132 were then covalently docked to the binding pocket of β5 subunit using GOLD version 4.0 [26, 27]. A radius of 20 Å from the β5-catalytic N-terminal threonine was used to direct site location. For each of the genetic algorithm runs, a maximum number of 100,000 operations were performed on a population of 100 individuals with a selection pressure of 1.1. Operator weights for crossover, mutation, and migration were set to 95, 95, and 10, respectively, as recommended by the authors of the software. Fifty GA runs were performed in each docking experiment as done in the software validation procedure. The default GOLD fitness function was used to identify the better binding mode [28]. The distance for hydrogen bonding was set to 2.5 Å and the cut-off value for van der Waals calculation was set to 4 Å. Covalent docking was applied and the terminal carbonyl carbon of all the ligands were bonded to the hydroxyl oxygen of Thr^1. The docking process of MG101 provided results consistent with the crystal structure. Following the docking study of MG132, two conformations with the lowest docking energy, mode I and mode II, were selected as the representative binding modes.

Dynamics simulations

The docking complex of mode I and mode II of MG132, along with the crystal structures of the MG101 complex and the proteasome without inhibitor (PR), were used as the initial structures in the following 5 ns MD simulation. The Amber Molecular Dynamics Package version 8.0 was used with an Amber99 force field [29]. The partial atomic charges (AM1-BCC charges) and molecular mechanical parameters of MG132 and MG101 were obtained using Divcon and Antechamber modules (Supporting information). To make the simulation reasonable and computationally less demanding, a reduced system was used consisting of subunit β5 and β6 of the proteasome with or without the inhibitor. All the complexes were solvated in TIP3P water [30] using an octahedral box, which extended 8 Å away from any solute atoms, and six sodium ions were added into the box to obtain an electro-neutral system, resulting in about 45,000 atoms in each complex. The calculations began with 500 steps of steepest descent followed by 500 steps of conjugate gradient minimization with a large constraint of 500 kcal mol^{-1} $Å^{-2}$ on the solute atoms. Then 500 steps of steepest descent followed by 500 steps of conjugate gradient minimization without restraints on any atoms were performed. After minimization, the system was gradually heated from 0 K to 300 K in 50 ps and subsequently equilibrated for 50 ps under constant pressure (1 atm) and temperature (300 K) conditions with weak restraints of 2 kcal mol^{-1} $Å^{-2}$ on the solute atoms. The final production simulations of 5 ns for all systems were carried out at the same conditions. To avoid deformation of the structure due to the absence of neighboring subunits, the alpha carbon (Cα) atoms whose distances to the atoms of the inhibitor molecule and β5:Thr^1 in the system are>12 Å were tethered to the starting coordinates by using a harmonic potential with a force constant of 1 kcal mol^{-1} $Å^{-2}$ in the production simulation. All bonds were constrained by using the SHAKE algorithm [31] with a tolerance of 10^{-5}, allowing for the integration of Newton's equations by using a 2.0 fs time step. Periodic boundary conditions with minimum image conventions were applied to calculate the non-bonded interactions. A cutoff of 12 Å was used for the Lennard-Jones interactions. The trajectory was saved every 10 ps and then further analyzed with in-house software.

Results and discussion

Automated docking results

Knowledge of the enzymatic kinetics of the aldehyde inhibitor [7] can help direct docking solutions that would allow covalent inhibition of the proteasome. The GOLD software, which has one unique function for handling covalent docking, was adopted here in the docking study of covalent proteasome inhibitors. Because the carbonyl carbon of the ligand is bonded to the hydroxyl oxygen of Thr^1 by setting covalent parameters in GOLD, the docking results can best imitate the binding mode of the proteasome inhibitor's crystal structure compared with previously published non-covalent docking studies [22, 23].

The crystal structure of MG101 was docked into the binding pocket of β5 subunit in order to test the credibility of the docking procedure. As shown by the docking results, the conformation that produced the highest fitness score by GOLD was the one that was closest to the experimental binding mode. The docked conformation and experimental binding mode differed by only 0.73 Å in root-mean-square deviation (RMSD). Orientation of the P1-P4 residues of the ligand in the docked conformation was similar to that observed in the crystal structure. Hydrogen bonds within β5:Thr^{21}, Gly^{47}, Ala^{49}, and β6:Asp^{114} of proteasome were also reproduced in the docked structure. The protocol thus provided a reasonable prediction of the experimental binding mode for the peptide aldehyde ligand by use of the "GOLD score" to rank the compounds.

The docking model of MG132 was built with the same method as MG101. Interestingly, the docking study of MG132 demonstrated two classes of binding modes, therefore both mode I and mode II with the lowest docking energy respectively were selected as the two representative binding modes. As shown in Fig. 3, mode I and mode II have the same backbone conformation, and they both adopt β-conformation and fill the gap between strands S2 and S4 by forming hydrogen bonds with residues Thr^{21}, Gly^{47}, and Ala^{49}, generating an antiparallel β-sheet structure. The P1-leucine side chain of the inhibitor projects into the S1 pocket and the P2-leucine side chain points outside. The difference of mode I and mode II lies in the conformation of N-terminal of MG132. For mode I, the P3-leucine side chain stretches out into the subunit-specific S3 pocket and is in close contact with residues of the adjacent β6 subunit, with its N3 atom hydrogen-bonded to the hydroxyl group of β6:Asp^{114}. On the contrary, in mode II, the side chain at the P3 site fills into the open space of the S4 pocket whereas the benzyloxy carbonyl moiety protrudes into the S3 pocket with its O4 atom hydrogen-bonded to the β6:Ser^{118}.

Preliminary structural features in dynamics simulations

Before proceeding with more detailed analysis of dynamics simulations, it is important to assess the stability of the system. For this purpose, we analyzed the RMSD of the complex atoms as a function of time for the MD trajectory. Figure 4 shows the RMSD of all three complexes with respect to the starting structure. Because most of the Cα atoms were restrained during the simulation, the RMSD leveled off rapidly after an initial increase and reached equilibrium after 1 ns and then remained stable over most of the 5 ns simulation. The average RMSD values of mode I and mode II of MG132 complexes, MG101 complex, and PR from 1 ns to 5 ns are 1.69, 1.66, 1.60, and 1.67 Å, respectively. The RMSD values of the three complexes are quite similar during the MD simulation.

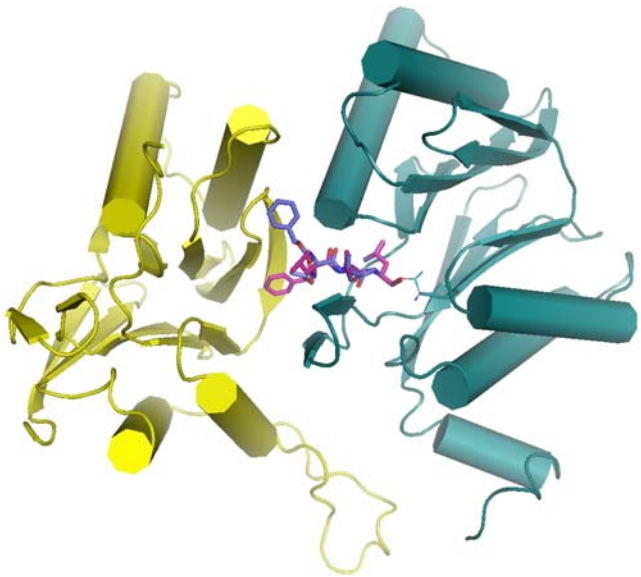

Fig. 3 An overlap of the two binding modes of MG132 extracted from the results of the docking experiment. The β5 (teal) and β6 (yellow) subunits of 20s proteasome are shown in cartoon, and the β5: Thr^{1} is shown. The ligands are represented by stick model (mode I in blue, mode II in red). The picture is rendered in Pymol 0.99 [36]

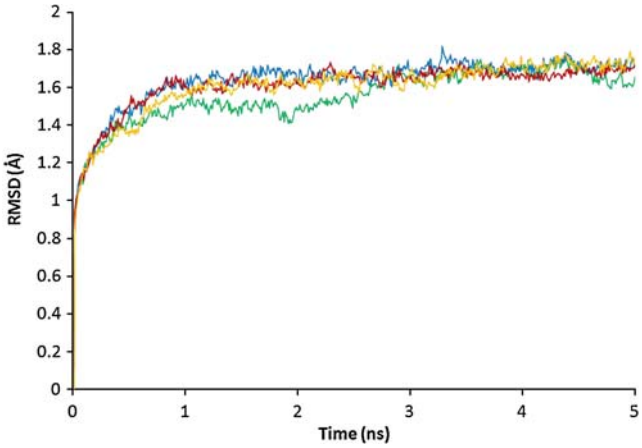

Fig. 4 RMSD of atoms in the simulation of mode I (blue), mode II (red), MG101 (green) and PR (yellow)

For further investigation on the stability of the whole system, we examined the hydrogen bonds formed at the interface between subunit β5 and β6 from 1 ns to 5 ns time scale of MD trajectories (Table 1). Six hydrogen-bonds were found in the initial crystal structure, which were formed between β5:Ala^{27} and β6:Arg^{125}, β5:Ser^{28} and β6:Glu^{120}, β5:Asp^{51} and β6:Arg^{91}, β5:Gln^{53} and β6:Tyr^{119}, β5:Gly^{94} and β6:Arg^{91}, β5:Glu^{194} and β6:Lys^{14N}. During the MD simulation, most of these H-bonds were preserved, suggesting that the simulation result is reasonable and consistent with the crystal structure. However, the inter-facial H-bond formed between β5:Ala^{27} and β6:Arg^{125} was broken during all four simulations. We believe that this collapse of H-bond is due to the absence of other

Table 1 The occurrence rate of hydrogen bonds formed between β5 and β6 subunit during the simulations of mode I, mode II, MG101, and PR. The occurrence rate less than 20% was highlighted in bold type

H-bonds	Mode I	Mode II	MG101	PR
β5:Ala^{27} - β6:Arg^{125}	**1.00%**	**7.50%**	**15.75%**	**7.75%**
β5:Ser^{28} - β6:Glu^{120}	84.00%	**0.25%**	85.75%	100.00%
β5:Asp^{51} - β6:Arg^{91}	97.50%	97.50%	98.25%	100.00%
β5:Gln^{53} - β6:Tyr^{119}	99.00%	86.25%	99.50%	99.75%
β5:Gly^{94} - β6:Arg^{91}	56.75%	78.75%	59.00%	61.00%
β5:Glu^{194} - β6:Lys^{14N}	65.00%	76.75%	75.75%	63.50%

neighboring subunits in the MD simulation, since in the crystal structure the guanidinium group of Arg^{165} side chain in subunit β3' participates in the interface interaction near the above broken H-bonds between β5 and β6 by forming H-bonds with β5:Val^{26} at one side and with β6:Asp^{139} at the other side. Nevertheless, another H-bond that is also adjacent to β3':Arg^{165} and formed between β5:Ser^{28} and β6:Glu^{120} was preserved during all the simulations except for mode II, where β5:Lys^{32} instead of Ser^{28} formed H-bond to β6:Glu^{120} with an occurrence rate of 72.25% (data not shown).

Binding mode analysis revealed the optimal conformation of MG132 complex

Flexibility of the residues in the active site

As revealed by its X-ray conformation, MG101 adopts an antiparallel β-sheet conformation, filling the gap between strands S2 and S4. These β-sheets are stabilized by direct hydrogen bonds between the conserved residues (Thr^{21}, Gly^{47}, and Ala^{49}) of the β5 subunit and the backbone atoms of MG101. Both mode I and mode II of MG132 were in good agreement with this conformation and with these H-bonds. On the side of neighboring β6 subunit, both MG101 and MG132 in mode I interacted with the carboxyl group of Asp^{114} side chain; however, in mode II MG132 formed H-bond with the hydroxyl group of Ser^{118} side chain.

The stability and validity of the binding modes were checked during the MD process by analyzing the binding pocket of inhibitors. For all four simulations, MD trajectories from 1 to 5 ns time scale were used here for investigation. The average conformations during this simulation period were calculated and then superimposed together based on MG101 complex, and the RMSD values between the backbone atoms of mode I, mode II, PR, and MG101 complex are 0.40, 0.34, and 0.50 Å, respectively (Fig. 5). The most obviously observed effect is the difference on the average conformations of strands S2, which contains the key residue Thr^{21}.

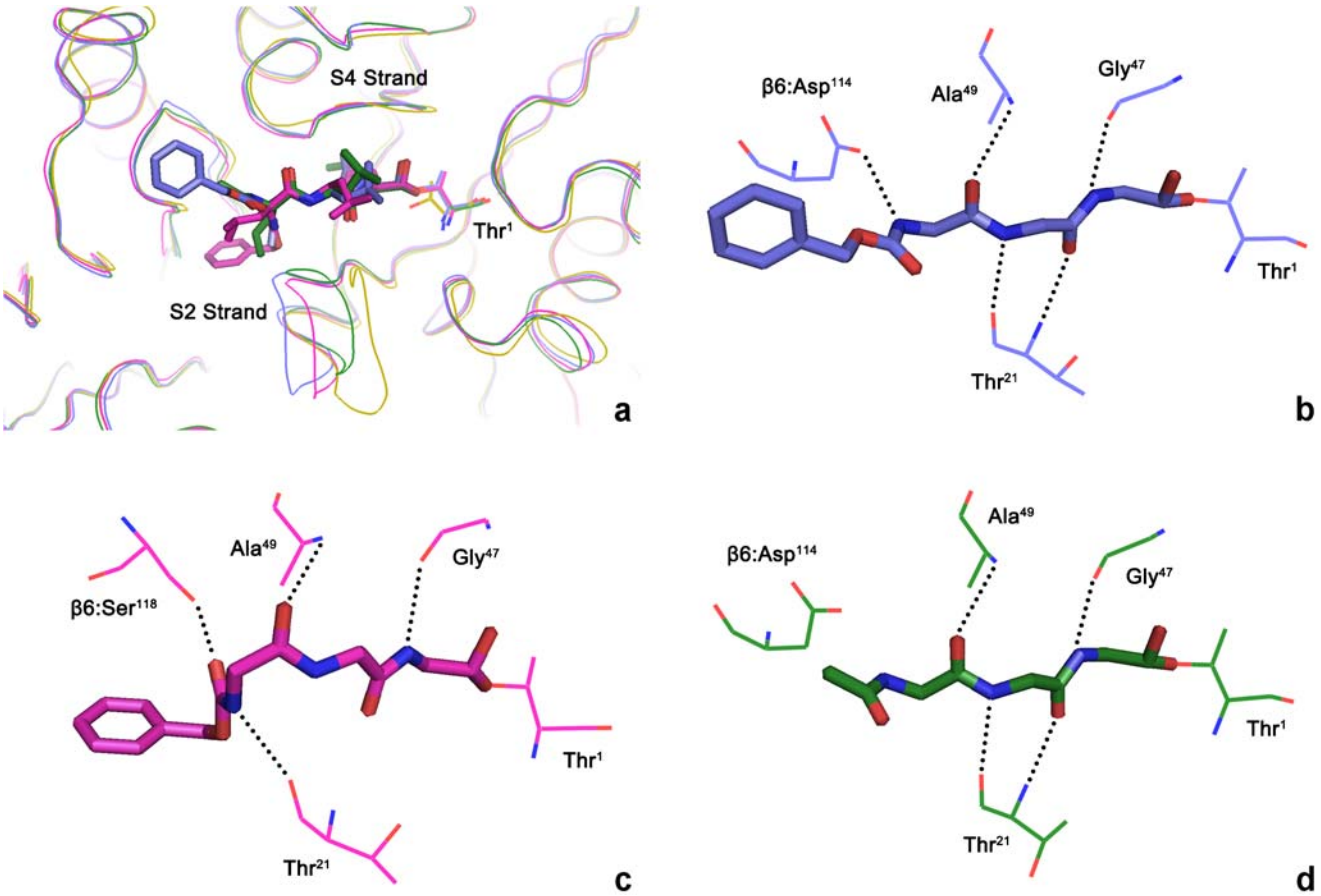

Fig. 5 The average conformations during the simulation period from 1 ns to 5 ns. **(a)** The conformations of mode I (blue), mode II (red) and PR (yellow) are superimposed on the MG101 complex (green). The protein is shown in ribbon, and the β5:Thr^{1} is shown. The ligands are represented by stick model. **(b-d)** The binding sites of mode I, mode II and MG101. Only the backbone of the ligand is shown for the purpose of clear viewing. Hydrogen bonds with an occurrence rate more than 80% are shown in dotted lines. Key residues are shown

Comparing with mode I, the S2 strands of mode II, MG101, and the protein increasingly moved inside toward the central cavity of barrel-shaped proteasome. Moreover, in mode II and PR, the S2 strands moved downwards and away from the backbone of the inhibitor compared to mode I and MG101. The average distances between the active site flanking residues 20–22 and residues 46–48 [10] in mode I, mode II, MG101, and PR are 10.74, 10.78, 10.61, and 11.37 Å, respectively. And the average distances between Thr^{21}-O and Ala^{49}-N, which are the doorkeeper atoms of the binding site, in mode I, mode II, MG101, and PR are 7.37, 7.61, 7.24, and 9.28 Å, respectively. Obviously, the PR enlarged its binding pocket during the MD simulation because of the lack of H-bonds and the anti-parallel β-sheet conformation that stabilize the inhibitor. The remarkable difference at the binding pocket reveals the conformational change of the proteasome between its bound and unbound form. Therefore, the movement of S2 strands and consequently the larger space of binding site in mode II would generate undesirable disturbance to the H-bonds and increase water accessibility of the covalent bond between MG132 and the proteasome, as discussed below.

Hydrogen bond analysis

Hydrogen bonds were analyzed during the MD simulation. The geometric criterion for the formation of H-bonds is common with a donor-acceptor distance less than 3.5 Å and the donor-H-acceptor angle larger than 120°. In the model of MG101 complex, all four H-bonds found between MG101 and β5 subunit that formed an anti-parallel β-sheet between strands S2 and S4 were preserved with high occurrences during the last 4 ns of MD simulation. On the other side of binding pocket, comprehensive analysis of H-bond occupation and donor-acceptor distance measurement shows that β6: Asp^{114} formed H-bond alternatively on its Oδ1 or Oδ2 atom of the charged carboxyl side chain with MG101-N3 atom (Table 2a). The high stability of the H-bond network in MG101 complex indicated that the simulation is in excellent agreement with the crystal structure. And the stability of H-bonds as well as the anti-parallel β-sheet conformation account for the high binding affinity of aldehyde inhibitors.

For the inhibitor MG132, detailed analysis of the trajectories showed several distinct and interesting differences in the behavior of the two binding modes. In the MD simulation of mode I, the average conformation is similar to MG101, except for that β6:Asp^{114} forms H-bond primarily on its Oδ2 atom with MG132 with a high occupation rate (Table 2b). However, the binding site is much more flexible and unstable in mode II during the simulation. The anti-parallel β-sheet conformation between MG132 and strands S4 is stable as usual, but both H-bonds between Thr^{21} and MG132 are weakened during the simulation with lower

Table 2 The analysis of hydrogen bonds between the inhibitor and the proteasome during the simulations of mode I, mode II, and MG101

H-bond donor	H-bond acceptor	Occurrence (%)	Donor-acceptor distance (Å)	Donor-H-acceptor angle (°)
Mode I				
MG132-N1[a]	β5:Gly^{47}-O	98.00	3.078(0.17)[b]	28.70(9.50)
β5:Thr^{21}-N	MG132-O2	94.75	3.125(0.16)	20.72(8.62)
MG132-N2	β5:Thr^{21}-O	99.00	2.899(0.13)	29.06(13.22)
β5:Ala^{49}-N	MG132-O3	96.00	3.061(0.18)	26.13(11.46)
MG132-N3	β6:Asp^{114}-Oδ2	82.75	2.975(0.20)	23.16(10.65)
Mode II				
MG132-N1	β5:Gly^{47}-O	98.75	2.991(0.15)	26.55(11.56)
β5:Thr^{21}-N	MG132-O2	70.50	3.186(0.17)	17.40(8.36)
MG132-N2	β5:Thr^{21}-O	66.50	3.186(0.18)	28.55(10.79)
β5:Ala^{49}-N	MG132-O3	90.75	3.100(0.18)	26.87(10.82)
β6:Ser^{118}-Oγ	MG132-O4	93.00	2.863(0.17)	17.66(9.61)
MG132-N3	β5:Thr^{21}-O	93.25	2.982(0.17)	34.71(12.48)
MG101				
MG101-N1	β5:Gly^{47}-O	99.25	2.949(0.14)	25.63(9.40)
β5:Thr^{21}-N	MG101-O2	97.50	3.038(0.17)	19.43(8.79)
MG101-N2	β5:Thr^{21}-O	98.50	2.890(0.13)	24.82(12.30)
β5:Ala^{49}-N	MG101-O3	97.00	3.020(0.16)	25.72(11.92)
MG101-N3	β6:Asp^{114}-Oδ2	66.00	2.959(0.15)	17.88(9.83)
MG101-N3	β6:Asp^{114}-Oδ1	20.00	3.022(0.17)	19.85(12.02)

[a] Nitrogen and oxygen atoms are termed N1, N2, N3 (O1, O2, …, O5) depending on their proximity to the covalent bond, as shown in Fig. 1

[b] The standard deviation

occurrence. The average donor-acceptor distances are all near 3.2 Å, which indicates that the H-bond connection is rather weak and unstable (Table 2c). Moreover, a new H-bond formed between Thr^{21}-O and MG132-N3 during the simulation resulted in a more twisted conformation of the inhibitor, pulling Thr^{21} toward the benzyloxy carbonyl moiety of the inhibitor. On the side of β6 subunit, no H-bonds were found between MG132 and Asp^{114} even at a single snapshot during the simulation. Due to the direct interaction with β6:Asp^{114} and the stable H-bond interaction with β5:Thr^{21}, the binding conformation of MG132 in mode I is much more reasonable than in mode II.

Coordination of water to the covalent bond destabilizes the binding mode

Generally, aldehyde inhibitors bind to proteasome reversibly and have fast dissociation rates. They can be rapidly oxidized to inactive acids and transported out of cells by the multi-drug resistance (MDR) system. Water molecules participate actively in the dissociation process of the aldehyde inhibitors. Previous study shows that besides Thr^1, a cluster of water molecules located in the proximity of Thr^1-Oγ1, Thr^1-N, Ser^{129}-Oγ and Gly^{47}-N also plays a key role in proteolysis [16, 32]. The possible function of the water molecules is to serve as the proton shuttle between Thr^1-Oγ1 and Thr^1-N during substrate binding and to participate in the cleavage of the acyl ester intermediate, resulting in the regeneration of Thr^1-Oγ1 [16, 20]. Our study focused on the vicinity of the covalent bond formed between the inhibitors and Thr^1 of proteasome β5 subunit. An analysis of the number of collision events between water molecules and the oxygen atoms of hemiacetal (Thr^1-Oγ1) shows that the covalent bond in mode II is more accessible to water than in mode I. The average number of water molecules within 3.5 Å from Thr^1-Oγ1 during the last 4 ns of MD simulation in mode I, mode II and MG101 are 0.888, 1.22, and 1.19, respectively (Fig. 6). Moreover, according to the criterion for the formation of H-bonds described earlier, the overall occurrence of H-bonds between water molecules and Thr^1-Oγ1 is 15.5% in mode II, which is twice that in mode I (7.25%) and in MG101 (5.75%). Because of the key role that water plays in the proteolysis process, the frequent attacks of the covalent bond by solvent molecules in mode II would likely result in the destabilization of this bond between the inhibitor and the proteasome.

We also investigated why the covalent bond formed in mode I is relatively well protected from water attacks. Analysis of the trajectories shows that the water molecule accessible to the covalent bond occupies the pocket formed by β5:Thr^1, Asp^{17}, Thr^{21}, Lys^{33}, Ser^{129}, Tyr^{168} and the P2 moiety of the inhibitor (Fig. 7). Among them, Thr^{21}, Ser^{129}, Tyr^{168} and the P2 moiety are exposed to the water in the central cavity of the β ring and act as a "gateway" of the pocket. This observation indicates that a bulky substituent at the P2 site could increase the binding stability by reducing the water accessibility of the covalent bond. In the simulation of mode I, a water molecule that is 3.0-4.0 Å away from Thr^1-Oγ1 occupies this pocket and blocks the path through which other water molecules would approach and attack the covalent bond, thereby preventing the covalent bond from being attacked by free water molecules. In mode II, however, this pocket is much larger because

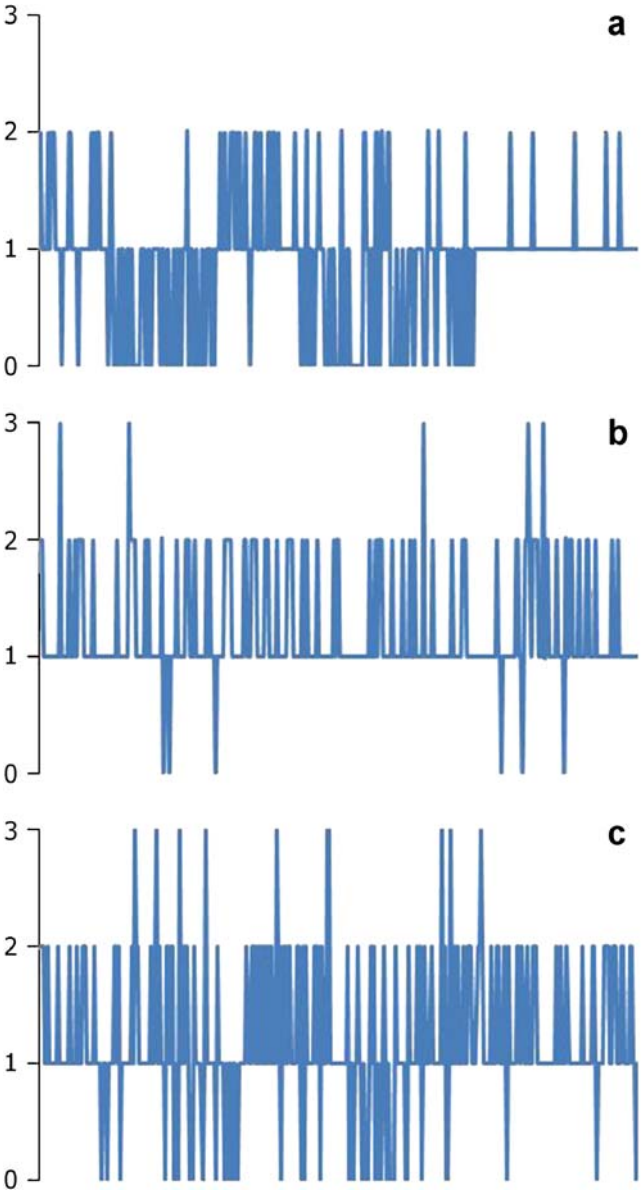

Fig. 6 Number of water contacts with the covalent bond between inhibitor and the proteasome during the last 4 ns of MD simulation. For each time step (at 10 ps intervals), the contacts were counted by identifying the number of water molecules within 3.5 Å of the covalent bond. **(a)** mode I **(b)** mode II **(c)** MG101

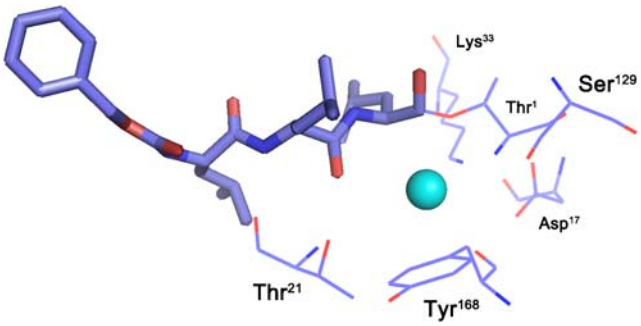

Fig. 7 The water molecule (blue sphere) accessible to the covalent bond in mode I. MG132 is represented by stick model. Residues of the pocket in which the water occupies are shown

strand S2 including Thr^{21} has moved inside toward the central cavity of the β ring and also moved downward away from the inhibitor's backbone. The larger pocket facilitated the access of free water molecules. During the last 4 ns of MD simulation, the pocket was occupied by two or more water molecules over 25 percent of the time scale. Therefore, the constrained anti-parallel β-sheet conformation with stable H-bonds between Thr^{21} and MG132 formed in mode I could be one structural reason for the lower accessibility of water to its covalent bond, leading to a more stable and reasonable binding mode.

Structure-activity relationship derived from binding mode I

We obtained two binding modes of MG132 using a computational method based on the inhibitor bound proteasome structure. Through in-depth MD simulations and analysis, mode I was shown to be more stable and reasonable than mode II. The results are in good agreement with the known MG101 crystal structure and can explain the structure-activity relationship (SAR) of the aldehyde inhibitors.

Previous studies revealed that hydrophobic contacts could enhance the inhibitory activity against the chymotrypsin-like activity of the 20S proteasome [33]. A critical difference between MG101 and MG132 is that the P4 site of MG132 carries a benzyloxy carbonyl moiety, which is capable of forming hydrophobic, CH/π or π/π interactions with the proteasome. For this reason, we examined the hydrophobic contacts between inhibitor and the proteasome in the average conformation of MG101 and mode I. As shown in Fig. 8, four residues from β6 subunit are involved in the binding process of MG132. β6:Pro^{94}, Tyr^{96}, and Pro^{115} may form π/π interactions, and Val^{116} would possibly form CH/π interactions with the benzyloxy carbonyl moiety of MG132. This result is consistent with our energy decomposition analysis of the β6 subunit during the MD simulation, which showed that MG132 in mode I has higher binding affinity with β6:Pro^{94}, Tyr^{96}, Pro^{115}, and Val^{116} than MG101. Similarly, Stein *et al*. synthesized and evaluated the biological activities of a series of peptide aldehydes, and they demonstrated that the activity decreased by 15-fold when the benzyloxy carbonyl substituent at P4 site of MG132 was replaced with acetyl group [34]. Therefore, an aromatic substituent at the P4 site may form hydrophobic, CH/π or π/π interactions with β6 subunit, which could be the critical reason for the higher activity of MG132, compared with that of MG101.

In the average conformation of mode I, the P4 moiety of MG132 lies at the N-terminal of β5:inhibitor complex. The benzyloxy carbonyl substituent protrudes into the center of the S4 pocket in β6 subunit, and it is kept at the place by two rings of predominantly aromatic residues (Fig. 8). Side chains of β6:Pro^{94}, Tyr^{96}, Pro^{115} and Val^{116} as discussed above, form the inner ring and contact the aromatic group directly. Residues β6:Tyr^{-5}, Phe^{48}, Tyr^{88}, Phe^{92} and Phe^{93} in the outer ring contact the inner residues, and they also provide a hydrophobic interface for residues flanking the P4 moiety of MG132 or other peptide aldehyde proteasome inhibitors. Distance analysis revealed that these two rings of residues in the S4 pocket are suitable for the binding of a bulky aromatic group with a linker of 2–4 atoms long connecting with the N3 atom of the peptide aldehyde inhibitor. A shorter linker would possibly result in a farther distance between the P4 moiety and the residues of inner ring, leading to a decreased activity. Stein *et al*. confirmed that (pyridin-4-yl)methyl carbonyl group at the P4 site has a similar activity as MG132, whereas the P4 (pyridin-4-yl) carbonyl analogue results in a 50–fold loss of activity [34]. This linker could also be replaced by an aromatic group, such as (naphthalen-1-yl)methyl carbonyl substituent at the P4 site, which shows an excellent activity and selectivity

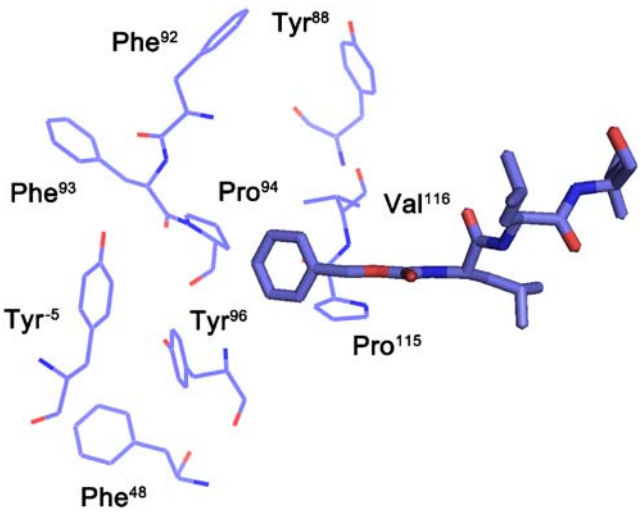

Fig. 8 Two rings of predominantly aromatic residues in the S4 pocket of mode I. MG132 is represented by stick model

[33]. Thus, rational design of more potent inhibitors for the 20 S proteasome could be aimed at this site.

Although the P2-leucine side chain is not in contact with the protein, the MD simulation revealed that this position also has important roles in the SAR. Rydzewski *et al*. reported that in a series of synthesized peptide-like vinyl sulfones, the P2-Phe analogue is 30-fold more potent than the P2-Gly analogue, and the replacement of P2-Phe with Ala results in a 3-fold loss of proteasome inhibitory activity. These results suggest that there is an increase in the activity for the compounds having bulky and sterically favored substituents at the P2 position [35]. Similarly, Stein *et al*. showed that the activity increases when the methyl substituent at P2 site is replaced with cyclohexyl methyl group and even higher with 1-naphthyl methyl group [34]. Our result is consistent with their reports. Although the P2 side chain of MG132 is not in contact with the protein and points into open space, substitution effects at the P2 position also play a definite role in compound activity, possibly because of conformational effect. As shown by our analysis, a bulky substituent at the P2 site would prevent water molecules from attacking the covalent bond. Therefore, our structural information on the flexibility of residues and the space restrictions of the protein ligand binding pockets could be used for the structure-based design of inhibitor in order to increase their affinity for specific active sites.

Summary

We used automated covalent docking combined with molecular dynamics simulation to elucidate the three-dimensional structures of proteasome inhibitor MG132 bound to the active sites. The MD simulation revealed an optimal conformation of MG132-proteasome complex (mode I), in which the inhibitor interacts tightly with the binding pocket by forming several stable H-bonds and maintains a stable anti-parallel β-sheet structure. Moreover, the MD simulation showed that this binding mode could prevent the covalent bond from being attacked by free water molecules and thereby stabilize the covalent bond between inhibitor and the proteasome. We believe that these attributes are important for the binding affinity of peptide aldehydes or other covalent inhibitors. We also found that a space demanding aromatic group with a linker of 2–4 atoms long at the P4 site of the peptide aldehyde inhibitor would form favorable hydrophobic interactions with the neighboring subunit of proteasome, which could be a critical reason for the higher activity of MG132 than that of MG101. A bulky substituent at the P2 position could also increase the binding stability by reducing the water accessibility of the covalent bond. Structural data obtained with computer-assisted molecular modeling could be valuable in the structure-based design of more potent and selective peptide aldehyde inhibitors of proteasome.

Acknowledgments We thank Prof. Groll for kindly providing us with the crystal structure of yeast proteasome:MG101 complex. We thank Dr. Jundong Zhang, Prof. Haiyan Liu and Prof. Shuqin Yu for their help in revising the manuscript. This study was supported by the National Natural Science Foundation of China (Grant No. 20672010).

References

1. Etlinger JD, Goldberg AL (1977) Proc Natl Acad Sci USA 74:54–58
2. Hershko A (1996) Trends Biochem Sci 21:445–449
3. Almond JB, Cohen GM (2002) Leukemia 16:433–443
4. Dou QP, Li B (1999) Drug Resist Updat 2:215–223
5. Drexler HC, Risau W, Konerding MA (2000) FASEB J 14:65–77
6. Oikawa T, Sasaki T, Nakamura M, Shimamura M, Tanahashi N, Omura S, Tanaka K (1998) Biochem Biophys Res Commun 246:243–248
7. Orlowski M, Wilk S (1981) Biochem Biophys Res Commun 101:814–822
8. Kisselev AF, Akopian TN, Woo KM, Goldberg AL (1999) J Biol Chem 274:3363–3371
9. Nussbaum AK, Dick TP, Keilholz W, Schirle M, Stevanovic S, Dietz K, Heinemeyer W, Groll M, Wolf DH, Huber R, Rammensee HG, Schild H (1998) Proc Natl Acad Sci USA 95:12504–12509
10. Lowe J, Stock D, Jap B, Zwickl P, Baumeister W, Huber R (1995) Science 268:533–539
11. Borissenko L, Groll M (2007) Chem Rev 107:687–717
12. Adams J, Behnke M, Chen S, Cruickshank AA, Dick LR, Grenier L, Klunder JM, Ma YT, Plamondon L, Stein RL (1998) Bioorg Med Chem Lett 8:333–338
13. Vinitsky A, Michaud C, Powers JC, Orlowski M (1992) Biochemistry 31:9421–9428
14. Lee DH, Goldberg AL (1996) J Biol Chem 271:27280–27284
15. Kisselev AF, Goldberg AL (2001) Chem Biol 8:739–758
16. Groll M, Ditzel L, Lowe J, Stock D, Bochtler M, Bartunik HD, Huber R (1997) Nature 386:463–471
17. Tsubuki S, Saito Y, Tomioka M, Ito H, Kawashima S (1996) J Biochem 119:572–576
18. Palombella VJ, Rando OJ, Goldberg AL, Maniatis T (1994) Cell 78:773–785
19. The Schechter-Berger nomenclature is used to name the substrate binding pockets and the residues of peptide aldehydes, Schechter I, Berger A (1967) Biochem Biophys Res Commun 27:157–162
20. Groll M, Huber R, Potts BC (2006) J Am Chem Soc 128:5136–5141
21. Zhu YQ, Pei JF, Liu ZM, Lai LH, Cui JR, Li RT (2006) Bioorg Med Chem 14:1483–1496
22. Fu Y, Xu B, Zou X, Ma C, Yang X, Mou K, Fu G, Lu Y, Xu P (2007) Bioorg Med Chem Lett 17:1102–1106
23. Smith DM, Daniel KG, Wang Z, Guida WC, Chan TH, Dou QP (2004) Proteins 54:58–70
24. Cundari TR, Dinescu A, Zhu DM, Hua L (2007) J Mol Model 13:685–690
25 Accelrys Software Inc (2001) Insight II, Accelrys, San Diego, CA, USA
26. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) J Mol Biol 267:727–748

27. The Cambridge Crystallographic Data Centre (2008) GOLD 4.0, CCDC, Cambridge, UK
28. Zeng J, Liu GX, Tang Y, Jiang HL (2007) J Mol Model 13:993–1000
29. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, DeBolt S, Ferguson D, Seibel G, Kollman P (1995) Comput Phys Commun 91:1–41
30. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926–935
31. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) J Comput Phys 23:327–341
32. Unno M, Mizushima T, Morimoto Y, Tomisugi Y, Tanaka K, Yasuoka N, Tsukihara T (2002) Structure 10:609–618
33. Momose I, Umezawa Y, Hirosawa S, Iinuma H, Ikeda D (2005) Bioorg Med Chem Lett 15:1867–1871
34. Stein RL (1997) U.S. Patent 5,693,617; Stein RL (1997) Chem Abstr 128:48498
35. Rydzewski RM, Burrill L, Mendonca R, Palmer JT, Rice M, Tahilramani R, Bass KE, Leung L, Gjerstad E, Janc JW, Pan L (2006) J Med Chem 49:2953–2968
36. DeLano WL (2006) The PyMOL molecular graphics system 0.99, DeLano Scientific, San Carlos, CA, USA