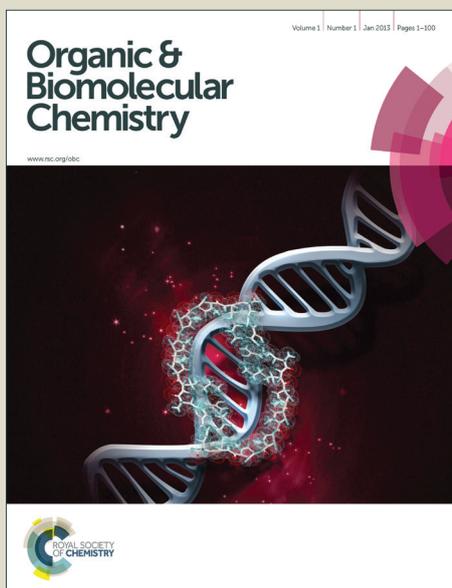


# Organic & Biomolecular Chemistry

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

# Fundamental reaction pathway and free energy profile for proteasome inhibition by syringolin A (SylA)

Donghui Wei,<sup>1,2,3</sup> Mingsheng Tang,<sup>1</sup> and Chang-Guo Zhan<sup>2,3</sup>

<sup>1</sup>Department of Chemistry, Zhengzhou University, Daxue Road, Zhengzhou, Henan 450052, China, <sup>2</sup>Molecular Modeling and Biopharmaceutical Center and <sup>3</sup>Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, 789 South Limestone Street, Lexington, KY 40536, USA

**Running title:** Reaction pathway for proteasome inhibition by SylA

## **Correspondence:**

Chang-Guo Zhan, Ph.D.

Director, *Molecular Modeling and Biopharmaceutical Center*

Endowed College of Pharmacy Professor in Pharmaceutical Sciences

Professor, Department of Pharmaceutical Sciences

College of Pharmacy

University of Kentucky

789 South Limestone Street

Lexington, KY 40536

Phone: 859-323-3943

E-mail: [zhan@uky.edu](mailto:zhan@uky.edu)

## Abstract

In this study, molecular dynamics (MD) simulations and first-principles quantum mechanical/molecular mechanical free energy (QM/MM-FE) calculations have been performed to uncover the fundamental reaction pathway of proteasome with a representative inhibitor syringolin A (SylA). The calculated results reveal that the reaction process consists of three steps. The first step is a proton transfer process, activating Thr1-O $\gamma$  directly by Thr1-N $^{\delta}$  to form a zwitterionic intermediate. The next step is nucleophilic attack on the olefin carbon of SylA by the negatively charged Thr1-O $\gamma$  atom. The last step is a proton transfer from Thr1-N $^{\delta}$  to another olefin carbon of SylA to complete the inhibition reaction process. The calculated free energy profile demonstrates that the second step should be the rate-determining step and has the highest free energy barrier of 24.6 kcal/mol, which is reasonably close to the activation free energy (~22.4 – 23.0 kcal/mol) derived from available experimental kinetic data. In addition, our computational results indicate that no water molecule can assist the rate-determining step, since the second step is not involved a proton transfer process. The obtained mechanistic insights should be valuable for understanding the inhibition process of proteasome by SylA and structurally related inhibitors at molecular level, and thus provide a solid mechanistic base and valuable clues for future rational design of novel, more potent inhibitors of proteasome.

## Introduction

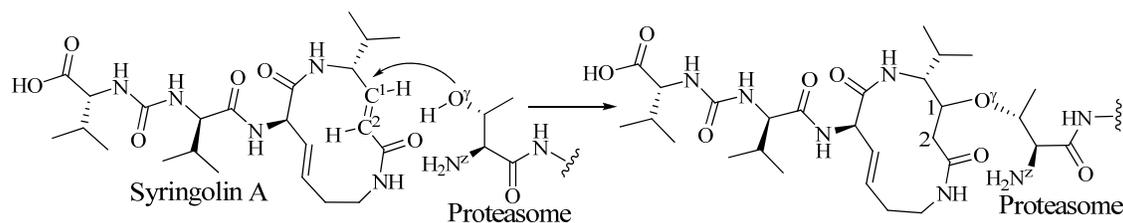
Proteasome, which contains a catalytic core particle (*i.e.* 20S proteasome) and two regulatory particles (*i.e.* 19S ‘cap’ regulatory complexes), is the major component of the nonlysosomal protein degradation pathway.<sup>1</sup> In eukaryotic and prokaryotic cells, ubiquitin can be attached to proteins and label them for destruction, then the proteins can be recognized by 19S regulatory complex and degraded by 20S proteasome.<sup>2</sup> This ubiquitin-proteasome pathway plays a primary role in the degradation of most proteins and removing the misfolded proteins in cells.<sup>3</sup> Recently, it was also found that the proteasome inhibitors have powerful anti-cancer activity, and several proteasome inhibitors designed according to the regulation mechanism of the proteasome system *in vivo* have been applied to the medical field.<sup>4-7</sup> For example, the proteasome inhibitor bortezomib has been used in clinic for the treatment of multiple myeloma.<sup>8</sup> Moreover, some of the early proteasome inhibitors have contributed to the development of new anti-cancer drugs, such as CEP-18770, Carfilzomib, and NPI-0052.<sup>4</sup> More recently, a new strategy to use HIV

protease-mediated activation of sterically capped proteasome inhibitor has been investigated for selectively killing the HIV-infected cells.<sup>9</sup> All of these facts demonstrate that proteasome inhibitors should be useful in the design of new anti-cancer tools and future therapeutics.

Due to the special anti-cancer activity, much attention has been paid to the development of proteasome inhibitors over the past decade. Thus far, there have been many kinds of proteasome inhibitors in the references, including peptide aldehydes,<sup>10, 11</sup> arecoline oxide tripeptides,<sup>12, 13</sup> retro hydrazino-azapeptoids,<sup>14</sup> proline- and arginine-rich peptides,<sup>15</sup> dipeptidyl boronates,<sup>16</sup> dipeptidyl boronic acids,<sup>17-19</sup>  $\beta$ -lactones,<sup>20-22</sup> epoxyketones,<sup>23-26</sup> vinyl sulfones,<sup>27-29</sup> substituted vinyl ketones,<sup>30</sup>  $\alpha,\beta$ -unsaturated N-acylpyrrole peptidyl derivatives,<sup>31</sup> cyclic peptides,<sup>32, 33</sup> and so on.<sup>34</sup> According to their chemical properties, the proteasome inhibitors can be mainly grouped into several types, and each type has a unique binding mode with the active sites of proteasome.<sup>1, 35-38</sup> Among various kinds of proteasome inhibitors, there are both non-covalent and covalent binding inhibitors. To the best of our knowledge, all of the current clinical inhibitors form a covalent bond with proteasome during the inhibition process. Although there have been many experimental reports on proteasome inhibitors,<sup>39, 40</sup> the detailed reaction mechanism concerning how proteasome is inhibited by a covalent inhibitor has not been understood very well, so extensive computational studies on the complicated proteasome-inhibitor reactions at molecular level are very valuable.

The catalytic core particle of proteasome (*i.e.* 20S proteasome) is composed of 28 subunits arranged in a unit as four homoheptameric rings ( $\alpha_7\beta_7\beta_7\alpha_7$ ), and each homoheptameric ring contains seven different subunits.<sup>41</sup> There are three types of proteasome  $\beta$ -type subunits, *i.e.*  $\beta_1$ ,  $\beta_2$ , and  $\beta_5$  that have caspase-like (C-L), trypsin like (T-L), and chymotrypsin-like (CT-L) activities, respectively.<sup>35</sup> So, a total of six active sites of proteasome, including two  $\beta_1$  sites, two  $\beta_2$  sites, and two  $\beta_5$  sites, are functionally independent. All of them have an N-terminal threonine residue (Thr1) which can initiate a nucleophilic attack on substrate protein, small peptide or inhibitor. The X-ray crystal structures show the binding mode of proteasome with the product,<sup>10, 20, 21, 30, 32, 42-45</sup> but it cannot determine the detailed reaction process. Although the inhibition mechanism of proteasome by epoxomicin (EPX) has been investigated in our previous work,<sup>46</sup> the detailed reaction mechanisms concerning how proteasome is inhibited by other types of covalent inhibitors (such as peptide aldehydes, boronic acid inhibitors,  $\beta$ -lactones, substituted vinyl sulfones, and substituted vinyl ketones) have still been unknown so far.

As a natural product, Syringolin A (SylA, depicted in **Scheme 1**) was identified as a virulence factor in the plant–pathogen interaction, and has been found that it irreversibly inhibits all of the three catalytic activities of eukaryotic proteasomes.<sup>30</sup> SylA is an  $\alpha,\beta$ -unsaturated lactam in which the amide group cannot be hydrolyzed by proteasome. As a proteasome inhibitor, SylA was recently shown to be a powerful tool to inhibit proliferation and induce apoptosis in neuroblastoma and ovarian cancer cells.<sup>47</sup> Thus, understanding the inhibition process of proteasome by SylA should be helpful for the research community to understand the inhibitions of proteasome by many  $\alpha,\beta$ -unsaturated lactam inhibitors, and to rationally design more specific, selective and efficient new inhibitors of proteasome. This prompted us to study the detailed inhibition mechanism of proteasome by the representative  $\alpha,\beta$ -unsaturated lactam SylA.



**Scheme 1** The inhibition of proteasome by  $\alpha,\beta$ -unsaturated lactam SylA.

In this study, the inhibitor SylA, which can irreversibly react with the Thr1 residue in the active site of the  $\beta 5$  subunit in proteasome by the Michael-addition reaction (depicted in **Scheme 1**),<sup>30</sup> was chosen as the substrate of proteasome, and the possible inhibition reaction pathway has been explored by performing molecular dynamics (MD) simulations and first-principles quantum mechanical/molecular mechanical (QM/MM)-free energy (QM/MM-FE) calculations. By using the QM/MM-FE approach, first-principles QM/MM reaction-coordinate calculations were followed by free energy perturbation (FEP) simulations to account for the dynamic effects of the protein environment on the free energy profile for the reaction pathway. Our QM/MM-FE calculations were based on the pseudobond first-principles QM/MM approach<sup>48-50</sup> and the revised pseudobond QM/MM-FE implementation.<sup>51-56</sup> The computational results clearly reveal the details of the enzymatic reaction pathway and its corresponding free energy profile. Based on the calculated free energy profile for the reaction process, the rate-determining step is identified. And the roles of essential residues and substrate are discussed on the basis of the QM/MM-optimized geometries.

## Computational Methods

**MD Simulation.** The X-ray crystal structure of proteasome-inhibitor (SylA) complex (PDB ID: 2ZCY, with a resolution of 2.9 Å) was employed to construct the initial structure of the enzyme-reactant (ER) complex.<sup>30</sup> In the X-ray crystal structure, the covalent bond between Thr1 and SylA was broken, and the missing atoms of SylA were added. According to our previous work,<sup>46</sup> the terminal amino group should be neutral in the active form of the proteasome. However, it is not evident why the terminal amino group of Thr1 should be neutral because an amino group in water usually has a  $pK_a$  of  $\sim 9$ . In order to verify the choice of protonation state of the terminal amino group, the  $pK_a$  values of the active-site residues were calculated by using PROPKA method<sup>57</sup> implemented in the PROPKA3 software package.<sup>58</sup> The PROPKA method is capable of predicting  $pK_a$  shifts of active-site residues and ionizable groups in protein-ligand complexes. Using the PROPKA method, the  $pK_a$  of the terminal amino group of Thr1 was calculated to be 5.35, supporting the neutral state of the terminal amino group at pH 7. For the histidine residues, hydrogens were placed at the  $\delta$ -position for His10, His178, His196, His278, His364, and at the  $\epsilon$ -position for His237 and His259. The standard protonation states at pH 7 were used for other residues.

The restrained electrostatic potential (RESP) charges were used as the atomic charges of the substrate SylA in the MD simulation and the subsequent QM/MM calculations. The electrostatic potential used to obtain the RESP charges was calculated at the HF/6-31G\* level using the Gaussian03 program,<sup>59</sup> and then the RESP charges were determined by fitting with the standard RESP procedure implemented in the Antechamber module of the Amber11 program.<sup>60</sup> As described in our previous studies,<sup>24</sup> the initial structure of ER was also constructed by retaining only two subunits ( $\beta 5$  and  $\beta 6$ ) and inhibitor SylA. Briefly, the MD simulations were performed using the Sander module of Amber (version 11) with Amber ff03 force field. Four chloride ions were added to neutralize the ER complex. After that, the ER complex was solvated in an orthorhombic box of TIP3P water molecules<sup>61</sup> with a minimum solute-wall distance of 10 Å. A  $\sim 10$  ns MD simulation of the solvated system was performed in the same way described in our previous computational study.<sup>62</sup> Because the structure of the last snapshot in the MD simulation was close to the average structure simulated, the last snapshot of the MD-simulated structure was used as the initial structure for the first-principles QM/MM calculations.

**QM/MM Calculations and the Minimum-Energy Reaction Pathway.** Since the reaction occurs in the reaction center, the water molecules beyond 50 Å of the olefin carbon (C<sup>1</sup>, **Scheme 1**) atom of SylA were removed, leaving the QM/MM system with 7,514 water molecules and a total of 29,012 atoms. The used boundary of the QM-MM system for the whole reaction is depicted in the figures discussed below, and the QM/MM interface was treated by using a pseudobond approach.<sup>48-50</sup> Before the QM/MM geometry optimization, the initial structure of the entire reaction system was energy-minimized with the MM method by using the AMBER11 program,<sup>60</sup> and the convergence criterion for energy gradient of 0.1 kcal·mol<sup>-1</sup>·Å<sup>-1</sup> was achieved.

A reaction-coordinate driving method and an iterative energy minimization procedure were applied to determine the reaction pathway by the pseudobond QM/MM calculations.<sup>48</sup> The step size used to scan the reaction coordinates was 0.1 Å. In the QM/MM calculations, we performed the QM calculations at the B3LYP/6-31G\* level of theory by using a modified version<sup>51</sup> of Gaussian03 program,<sup>59</sup> and carried out the MM calculation by using a modified version<sup>51</sup> of the AMBER8 program.<sup>63</sup> The reactant (ER), intermediates, transition states, and the final product of the inhibition reaction process were located and characterized by normal mode analysis. In addition, single-point energies of the QM/MM-optimized geometries were refined at the higher QM/MM(B3LYP/6-31++G\*\*):AMBER) level. We treated the boundary carbon atoms with the improved pseudobond parameters,<sup>49</sup> and used no cutoff for non-bonded interactions was throughout the QM/MM calculations. The convergence criterion applied to the QM region for the maximum force of 0.53 kcal·mol<sup>-1</sup>·Å<sup>-1</sup> (0.00045 au/Bohr), the RMS force of 0.35 kcal·mol<sup>-1</sup>·Å<sup>-1</sup> (0.00030 au/Bohr), the maximum displacement of 0.0018 au/rad, and the RMS displacement of 0.0012 au/rad were achieved. The convergence criterion for geometry optimizations of the MM subsystem was the root-mean-square deviation (rmsd) of energy gradient  $\leq 0.1$  kcal·mol<sup>-1</sup>·Å<sup>-1</sup>. In all the QM/MM calculations, the atoms within 20 Å of C<sup>1</sup> atom in the reaction center (**Scheme 1**) were allowed to move while the other atoms were kept frozen. It should also be noted that the QM and MM subsystems were energy-minimized iteratively during the QM/MM geometry optimization. For each step of the iteration, the MM subsystem was kept frozen during the QM subsystem minimizations, whereas the QM subsystem was kept frozen when the MM subsystem was energy-minimized.

**Free Energy Perturbation for the Minimum-Energy Reaction Pathway.** After the minimum-energy reaction pathway was determined by the QM/MM calculations, the FEP

method (which was implemented in our revised version of the AMBER8 program) was employed to account for the free energy changes associated with the QM-MM interactions.<sup>52</sup> In the FEP calculations, we kept the QM subsystem frozen and performed conformational sampling of the MM subsystem at each state along the reaction path. The point charges on the frozen QM atoms, which were determined by fitting the electrostatic potential (ESP) in the QM part of the QM/MM single-point energy calculations, were used in the FEP calculations. The same procedure as used in our previous work on other reaction systems was employed to calculate the total free energy difference between the transition state and the reactant.<sup>46, 51-56, 64-71</sup> The FEP calculations enabled us to determine the relative free energy changes more reasonably. Technically, by performing the QM/MM-FE calculations, there should be two terms in the final (relative) free energies, including the relative free energy change determined by the FEP calculations and the QM part of the QM/MM energy (excluding the Coulombic interaction energy between the point charges of the MM atoms and the ESP charges of the QM atoms). It should be noted that the QM/MM-FE energy did not include the finite-temperature contribution from the QM part. In the FEP calculations, the used time step was 2 fs, and all bond lengths involving a hydrogen atom were constrained. During sampling of the MM subsystem, the temperature was maintained at 298.15 K in the MD simulations. Each FEP calculation included 50 ps of equilibration and 300 ps of sampling.

A supercomputer (*i.e.* the Dell X-series Cluster with 384 nodes or 4,768 processors) at the University of Kentucky's Computer Center was employed for most of the MD and QM/MM simulations. And the SGI Fuel workstations in our own laboratory at University of Kentucky was used for the other modeling and computations.

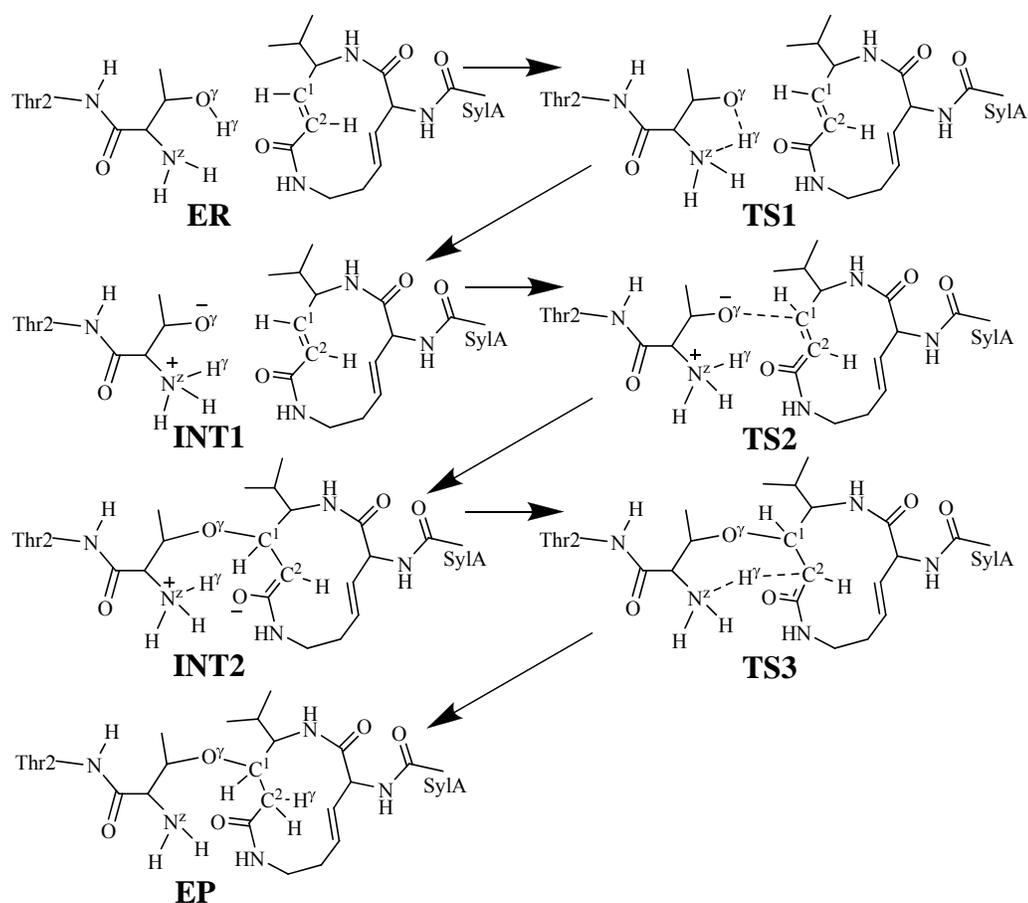
## Results and Discussion

### *MD simulation results*

Based on our previous work,<sup>46</sup> a water molecule may or may not participate in the proteasomal hydrolysis and inhibition reactions of proteasome to mediate the proton transfer process. Thus, we should consider the possibility of the water-assisted proton transfer pathway for the proton transfer step. However, the X-ray crystal structure of proteasome-inhibitor (SylA) complex (PDB ID: 2ZCY) shows that there is no any water molecule in the active site. To verify whether a water molecule could move close to the active site for the reaction, a ~10 ns MD

simulation was performed on the enzyme-reactant (ER) structure. The MD simulation results indicate that no water molecule moved close to both the Thr1-N<sup>z</sup> and Thr1-O<sup>γ</sup> atoms (<4.5 Å). In addition, as the new covalent bond C<sup>1</sup>-O<sup>γ</sup> is formed and a proton transfers from Thr1-N<sup>z</sup> to C<sup>2</sup> of inhibitor during the reaction, we tracked the changes of the O<sup>γ</sup>(Thr1)-C<sup>1</sup> distance (between the Thr1-O<sup>γ</sup> atom and the C<sup>1</sup> atom of inhibitor) and the N<sup>z</sup>(Thr1)-C<sup>2</sup> distance (between the Thr1-N<sup>z</sup> atom and the C<sup>2</sup> atom of inhibitor). The MD simulation led to a dynamically stable ER complex, and the average O<sup>γ</sup>(Thr1)-C<sup>1</sup> and N<sup>z</sup>(Thr1)-C<sup>2</sup> distances are ~3.44 and ~4.59 Å respectively, indicating that the key distances in the reaction center should be proper for the inhibition process.

In light of the results obtained from the X-ray crystal structure and MD simulation, and our previously study on the inhibition mechanism of proteasome by EPX,<sup>46</sup> we may reasonably assume that Thr1-O<sup>γ</sup> should be activated directly by its N-terminal amino group (Thr1-N<sup>z</sup>H<sub>2</sub>). As shown in **Scheme 2**, we have suggested the possible reaction pathway of the inhibition reaction of proteasome by SylA. The possible reaction pathway depicted in **Scheme 2** has been confirmed by our QM/MM reaction-coordinate calculations discussed below. In order to demonstrate the reaction mechanism for the inhibition of proteasome by SylA, the QM/MM reaction-coordinate calculations were performed in this study. The uncovered reaction pathway and the obtained free energy profile were described in detail as follows.

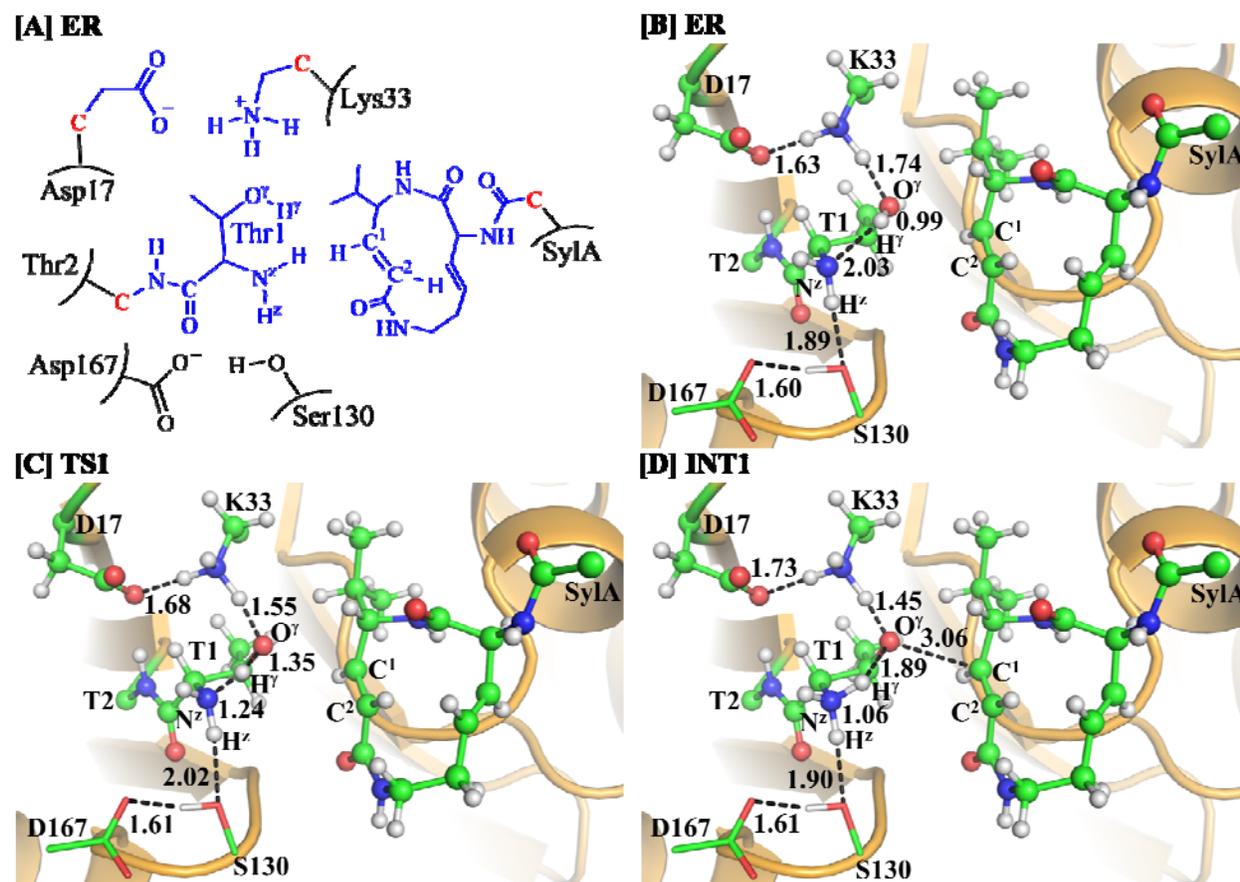


**Scheme 2** Possible pathway for the reaction of proteasome (catalytic subunit  $\beta 5$ ) with SylA.

### *Fundamental reaction pathway of the inhibition*

The QM/MM optimized ER structure of the enzyme-SylA reaction system was employed as the starting point of the QM/MM reaction-coordinate calculations, which were performed at the B3LYP/6-31G\*:AMBER level to determine the minimum-energy reaction pathway. As depicted in **Scheme 2**, the results obtained from the QM/MM calculations revealed that the inhibition of proteasome by SylA indeed consists of three reaction steps. The first reaction step is a direct proton ( $H^\gamma$ ) transfer from the Thr1- $O^\gamma$  atom to the Thr1- $N^z$  atom, forming a zwitterionic intermediate INT1 *via* transition state TS1. The second reaction step is the nucleophilic attack on the SylA- $C^1$  atom by the activated Thr1- $O^\gamma$ , coupled with the transformation of  $C^1=C^2$  double bond to a single bond *via* transition state TS2. The third reaction step is the proton ( $H^\gamma$ ) transfer from Thr1- $N^z$  to SylA- $C^2$ , resulting in the formation of product EP *via* transition state TS3. **Figs. 1 and 2** depict all the QM/MM-optimized geometries of the reactant ER, intermediates, transition

states, and product EP existing in the inhibition reaction process. In **Fig. 1A**, the atoms with blue color were treated by QM method, the boundary carbon atoms with red color were treated with the improved pseudobond parameters, and the other atoms of the entire reaction system were considered as MM subsystem during the QM/MM calculations.

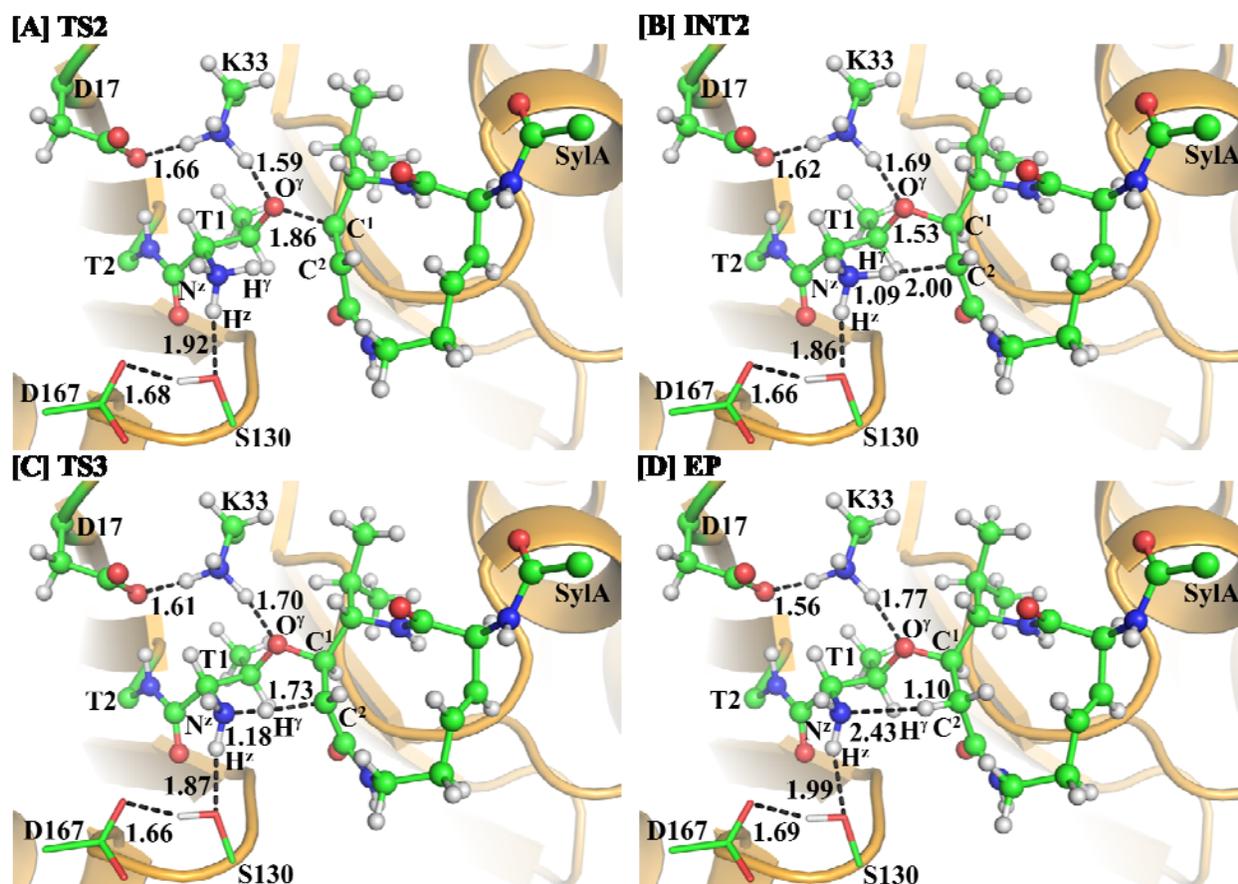


**Fig. 1** (A) Division of the QM/MM system. Atoms in blue color were treated as QM part. The boundary carbon atoms colored in red were treated with the improved pseudobond parameters. All of the other atoms were considered as the MM subsystem. (B-D) Optimized geometries for the key states ER, TS1, and INT1 optimized at the QM/MM(B3LYP/6-31G\*:AMBER) level. The key distances in the figures are given in Å. Carbon, hydrogen, nitrogen, and oxygen atoms are colored in green, white, blue, and red, respectively. The backbone of the protein is rendered as ribbon and colored in orange. The QM atoms are represented as balls and sticks and the surrounding residues are rendered as sticks.

**Step 1: Proton H<sup>γ</sup> Transfers Directly from Thr1-O<sup>γ</sup> to Thr1-N<sup>z</sup>.** Fig. 1B shows the QM/MM-optimized ER structure, in which both the hydrogen bond distance between the Thr1-O<sup>γ</sup> atom and the closest hydrogen atom of the Lys33-NH<sub>3</sub><sup>+</sup> group and the hydrogen bond distance between the Thr1-H<sup>z</sup> and Ser130-O atoms are shorter than 1.90 Å. The hydrogen bond distance between the Lys33-NH<sub>3</sub><sup>+</sup> hydrogen and Asp17-CO<sub>2</sub><sup>-</sup> oxygen atoms is 1.63 Å, while the hydrogen bond distance between the Ser130-H (hydroxyl hydrogen) and the Asp167-CO<sub>2</sub><sup>-</sup> oxygen is 1.60 Å, showing a hydrogen-bond network in the reaction center.

As shown in **Scheme 2**, the reaction step 1 involves the formation of the H<sup>γ</sup>-N<sup>z</sup> bond and the breaking of the H<sup>γ</sup>-O<sup>γ</sup> bond. So, the changes of the H<sup>γ</sup>-N<sup>z</sup> distance (R<sub>H<sup>γ</sup>-N<sup>z</sup></sub>) and H<sup>γ</sup>-O<sup>γ</sup> distance (R<sub>H<sup>γ</sup>-O<sup>γ</sup></sub>) can reflect the nature of reaction step 1. Therefore, R<sub>H<sup>γ</sup>-O<sup>γ</sup></sub>-R<sub>H<sup>γ</sup>-N<sup>z</sup></sub> was set as the reaction coordinate for reaction step 1. As shown in the QM/MM-optimized geometries (**Fig. 1B** to **2D**), R<sub>H<sup>γ</sup>-O<sup>γ</sup></sub> gradually elongates from 0.99 Å (ER, **Fig. 1B**) to 1.35 Å (TS1, **Fig. 1C**), while R<sub>H<sup>γ</sup>-N<sup>z</sup></sub> gradually shortens from 2.03 Å (ER, **Fig. 1B**) to 1.24 Å (TS1, **Fig. 1C**) and then to 1.06 Å (INT1, **Fig. 1D**). It should be noted that the intermediate INT1 is a very active zwitterion, which can react with the olefin carbon of SylA readily. During this reaction step, the hydrogen bond between the Thr1-O<sup>γ</sup> atom and Lys33-NH<sub>3</sub><sup>+</sup> group is strengthened, which should help to stabilize the zwitterionic intermediate INT1.

**Step 2: Nucleophilic Attack on SylA-C<sup>1</sup> by the Activated Thr1-O<sup>γ</sup>.** The negatively charged Thr1-O<sup>γ</sup> atom initiates the nucleophilic attack on SylA-C<sup>1</sup> atom. The nature of such process can be represented by the change of the O<sup>γ</sup>-C<sup>1</sup> distance (R<sub>O<sup>γ</sup>-C<sup>1</sup></sub>). Thus the reaction coordinate for this step was chosen as R<sub>O<sup>γ</sup>-C<sup>1</sup></sub>. In this process, the distance R<sub>O<sup>γ</sup>-C<sup>1</sup></sub> is shortened from 3.06 Å (INT1, **Fig. 1D**) to 1.86 Å (TS2, **Fig. 2A**) and then to 1.53 Å (INT2, **Fig. 2B**), respectively. As shown in **Scheme 2**, the C<sup>1</sup>=C<sup>2</sup> double bond changes to a single bond in this reaction step, and the negatively charged SylA-C<sup>2</sup> should be very active in the next reaction step.

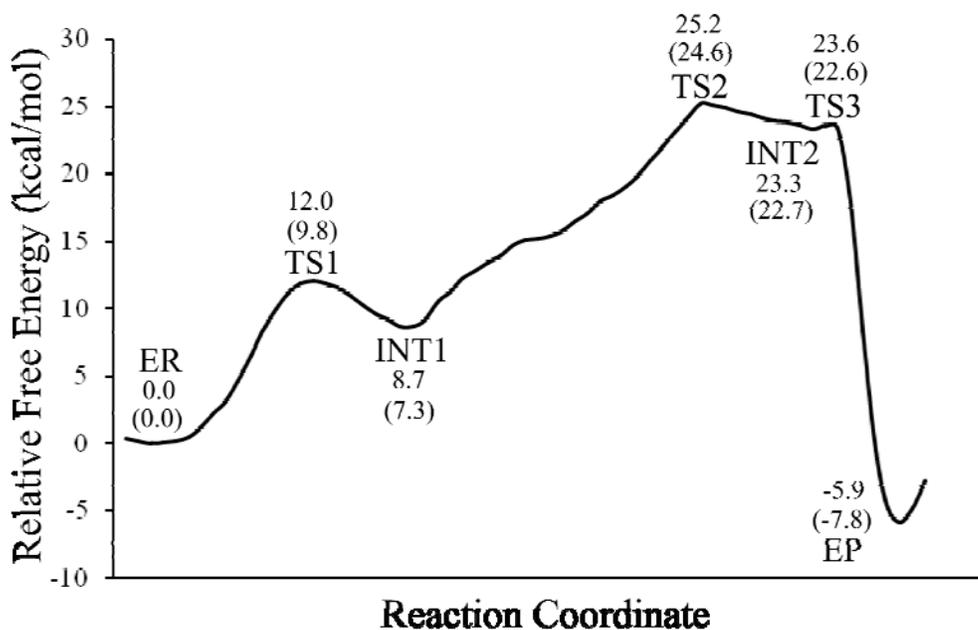


**Fig. 2** Optimized geometries for the key states TS2, INT2, TS3, and EP optimized at QM/MM(B3LYP/6-31G\*:AMBER) level. **Fig. 2** is represented using the same method with **Fig. 1**.

**Step 3: Proton  $H^\gamma$  Transfers Directly from Thr1- $N^z$  to SylA- $C^2$ .** The proton ( $H^\gamma$ ) will transfer from Thr1- $N^z$  to the negatively charged SylA- $C^2$ , which occurs simultaneously with the transformation of the  $C^1=C^2$  double bond to a single bond. Such a process involves the breaking of the  $N^z-H^\gamma$  bond and the formation of the  $H^\gamma-C^2$  bond (**Scheme 2**). Hence, the distances  $R_{N^z-H^\gamma}$  and  $R_{H^\gamma-C^2}$  were chosen to represent the reaction coordinate as  $R_{N^z-H^\gamma}-R_{H^\gamma-C^2}$  for the current reaction step. In TS3 (**Fig. 2C**), the distances  $R_{N^z-H^\gamma}$  and  $R_{H^\gamma-C^2}$  are 1.18 and 1.73 Å, respectively. In this step, the product EP is generated *via* transition state TS3, indicating that the Michael-type addition on the olefin is completed.

*Free energy profile for the fundamental inhibition reaction pathway*

As described above, the minimum-energy reaction pathway for the inhibition process calculated at the B3LYP/6-31G\*:AMBER level contains three reaction steps. To determine the corresponding free energy profile of the reaction, we performed QM/MM single-point energy calculations at a higher level (B3LYP/6-31++G\*\*.:AMBER) for each QM/MM optimized geometry along the obtained minimum-energy path. For each geometrical structure along the reaction path, the ESP charges determined in the QM part of the QM/MM single-point energy calculation were used in the subsequent FEP simulations to estimate the free energy changes. Depicted in **Fig. 3** is the free energy profile for this reaction process, which is determined by the QM/MM-FE calculations first without the zero-point and thermal corrections for the QM subsystem, and then with the zero-point and thermal corrections for the QM subsystem (values given in parentheses).



**Fig. 3** Free energy profile determined by the QM/MM-FE calculations for the inhibition reaction process. The relative energies are the QM/MM-FE energies without the zero-point and thermal corrections for the QM subsystem, and the values given in parentheses are corrected with the zero-point and thermal corrections for the QM subsystem.

As shown in **Fig. 3**, without the zero-point and thermal corrections for the QM subsystem, the free energy barriers calculated for the first to third reaction steps of the reaction pathway are 12.0, 16.5, and 0.3 kcal/mol, respectively. With the zero-point and thermal corrections for the

QM subsystem, the free energy barriers calculated for the first and second reaction steps are 9.8 and 17.3 kcal/mol, respectively. Notably, the free energy change from INT2 to TS3 is a negative value (-0.1 kcal/mol) after the zero-point and thermal corrections were included. The free energy profile suggests that the zwitterionic intermediate INT2 is very unstable and does not really exist during the reaction process. As indicated in Table S1 of ESI, the absolute QM/MM energy barriers of the three steps are 12.7, 14.2, and 0.4 kcal/mol, respectively, showing that the FEP corrections and the zero-point and thermal corrections for the QM subsystem do not significantly alter the energy profile. Noteworthy, the energy barrier of the entire inhibition process should be the energy difference between the lowest-energy state ER and the highest-energy state TS2, which is 24.6 kcal/mol with the zero-point and thermal corrections for the QM subsystem. Further, the free energy of EP is 7.8 kcal/mol below that of ER, and the free energy difference between EP and TS2 is 32.4 kcal/mol, indicating that the reverse reaction should be extremely slow and, thus, the nucleophilic attack of Thr1-O<sup>γ</sup> on the olefin should be irreversible at the room temperature. This phenomenon is in agreement with the known experimental observation that SylA is an irreversible proteasome inhibitor.

Based on the free energy profile depicted in **Fig. 3**, the free energy barrier for the entire reaction process is determined by that (24.6 kcal/mol) for the second reaction step (the rate-determining step). We wanted to know whether the calculated free energy barrier of 24.6 kcal/mol is reasonable in comparison with the available experimental reaction rate constant ( $k_{\text{obs}}$ ) or not. According to the reported experimental data,  $k_{\text{obs}}/[I] = 863 \pm 106 \text{ M}^{-1}\text{s}^{-1}$  when  $[I] = 100 - 200 \text{ nM}$  in which  $[I]$  represents the concentration of inhibitor SylA.<sup>30</sup> Thus,  $k_{\text{obs}} = 0.7 - 1.9 \times 10^{-4} \text{ s}^{-1}$ , which is associated with an free energy barrier of  $\sim 22.4 - 23.0 \text{ kcal/mol}$  at room temperature (25°C) according to the conventional transition state theory.<sup>72</sup> Our calculated free energy barrier of 24.6 kcal/mol is reasonably close to the experimentally derived free energy barrier of  $\sim 22.4 - 23.0 \text{ kcal/mol}$ , suggesting that the mechanistic insights obtained from this computational study are reasonable.

## Conclusion

In this study, the detailed mechanism for the inhibition reaction of proteasome (catalytic subunit  $\beta 5$ ) with SylA has been investigated by carrying out the first-principles QM/MM-FE calculations. Based on the results from the QM/MM calculations, the reaction process consists of

three steps. The reaction is initiated by a direct proton transfer from the Thr1-O $\gamma$  atom to the Thr1-N $^z$  atom to activate the Thr1-O $\gamma$ . Then, the negatively charged Thr1-O $\gamma$  atom initiates the nucleophilic attack on the olefin carbon of SylA. The third step is also a proton transfer process, *i.e.* the proton (H $\gamma$ ) transfer from the Thr1-N $^z$  to the negatively charged C $^2$  atom of SylA.

The calculated free energy profile of the reaction process indicates that the free energy barriers for the first and second reaction steps are 9.8 and 17.3 kcal/mol, respectively, whereas the third reaction step is a barrierless process. The free energy barrier of the entire inhibition reaction process should be the free energy difference between the lowest-energy state ER and the highest-energy state TS2 (24.6 kcal/mol), indicating that the second step should be the rate-determining step of the inhibition reaction process. Our calculated results also indicate that no water molecule can assist the rate-determining step (*i.e.* the second step), since this reaction step does not involve a proton transfer. The calculated free energy barrier of 24.6 kcal/mol for the rate-determining step is reasonably close to the experimentally derived free energy barrier of ~22.4 – 23.0 kcal/mol, suggesting that our calculated results are reasonable. Further, the computational results demonstrate that the reverse reaction should be extremely slow and, thus, the nucleophilic attack of Thr1-O $\gamma$  on the olefin should be irreversible, which is qualitatively consistent with the experimental observation. The computational insights obtained should be valuable not only for understanding the detailed inhibition reaction mechanism of proteasome by  $\alpha,\beta$ -unsaturated lactam SylA, but also for the future rational design of novel, more potent inhibitors of proteasome.

**Acknowledgments.** This work was supported in part by the National Science Foundation (NSF grant CHE-1111761 to Zhan), the National Institutes of Health (NIH grants R01 DA035552, R01 DA032910, R01 DA013930, and R01 DA025100 to Zhan), the National Natural Science Foundation of China (NSFC grant No. 21303167 to Wei) and China Postdoctoral Science Foundation (grant No. 2013M530340 to Wei). The authors also acknowledge the Computer Center at University of Kentucky for supercomputing time on a Dell X-series Cluster with 384 nodes or 4,768 processors.

**References**

1. A. J. Marques, R. Palanimurugan, A. C. Matias, P. C. Ramos and R. J. Dohmen, *Chem. Rev.*, 2009, 109, 1509-1536.
2. A. Ciechanover, *Angew. Chem. Int. Edit.*, 2005, 44, 5944-5967.
3. D. Voges, P. Zwickl and W. Baumeister, *Annu. Rev. Biochem.*, 1999, 68, 1015-1068.
4. E. Genin, M. Reboud-Ravaux and J. Vidal, *Curr. Top. Med. Chem.*, 2010, 10, 232-256.
5. J. Adams, *Nat. Rev. Cancer*, 2004, 4, 349-360.
6. A. L. Goldberg and K. Rock, *Nat. Med.*, 2002, 8, 338-340.
7. P. G. Richardson, C. Mitsiades, T. Hideshima and K. C. Anderson, *Annu. Rev. Med.*, 2006, 57, 33-47.
8. P. G. Richardson, P. Sonneveld, M. W. Schuster, D. Irwin, E. A. Stadtmauer, T. Facon, J. L. Housseau, D. Ben-Yehuda, S. Lonial, H. Goldschmidt, D. Reece, J. F. San-Miguel, J. Blade, M. Boccadoro, J. Cavenagh, W. S. Dalton, A. L. Boral, D. L. Esseltine, J. B. Porter, D. Schenkein, K. C. Anderson and A. Investigators, *New Engl. J. Med.*, 2005, 352, 2487-2498.
9. D. L. Buckley, T. W. Corson, N. Aberle and C. M. Crews, *J. Am. Chem. Soc.*, 2011, 133, 698-700.
10. J. Lowe, D. Stock, B. Jap, P. Zwickl, W. Baumeister and R. Huber, *Science*, 1995, 268, 533-539.
11. M. Vivier, M. Rapp, J. Papon, P. Labarre, M. J. Galmier, J. Sauziere and J. C. Madelmont, *J. Med. Chem.*, 2008, 51, 1043-1047.
12. M. Marastoni, J. McDonald, A. Baldisserotto, A. Canella, C. De Risi, G. P. Pollini and R. Tomatis, *Bioorg. Med. Chem. Lett.*, 2004, 14, 1965-1968.
13. M. Marastoni, A. Baldisserotto, A. Canella, R. Gavioli, C. De Risi, G. P. Pollini and R. Tomatis, *J. Med. Chem.*, 2004, 47, 1587-1590.
14. S. Aubin, B. Martin, J. G. Delcros, Y. Arlot-Bonnemains and M. Baudy-Floc'h, *J. Med. Chem.*, 2005, 48, 330-334.
15. M. Gaczynska, P. A. Osmulski, Y. Gao, M. J. Post and M. Simons, *Biochemistry*, 2003, 42, 8663-8670.
16. Y. Q. Zhu, X. Zhao, X. R. Zhu, G. Wu, Y. J. Li, Y. H. Ma, Y. X. Yuan, J. Yang, Y. Hu, L. Ai and Q. Z. Gao, *J. Med. Chem.*, 2009, 52, 4192-4199.

17. C. R. Berkers, M. Verdoes, E. Lichtman, E. Fiebigler, B. M. Kessler, K. C. Anderson, H. L. Ploegh, H. Ovaa and P. J. Galardy, *Nat. Methods*, 2005, 2, 357-362.
18. Y. Q. Zhu, G. Wu, X. R. Zhu, Y. H. Ma, X. Zhao, Y. J. Li, Y. X. Yuan, J. Yang, S. Yu, F. Shao and M. Lei, *J. Med. Chem.*, 2010, 53, 8619-8626.
19. L. J. Milo, J. H. Lai, W. G. Wu, Y. X. Liu, H. Maw, Y. H. Li, Z. P. Jin, Y. Shu, S. E. Poplawski, Y. Wu, D. G. Sanford, J. L. Sudmeier and W. W. Bachovchin, *J. Med. Chem.*, 2011, 54, 4365-4377.
20. M. Groll, E. P. Balskus and E. N. Jacobsen, *J. Am. Chem. Soc.*, 2008, 130, 14981-+.
21. M. Groll, R. Huber and B. C. Potts, *J. Am. Chem. Soc.*, 2006, 128, 5136-5141.
22. M. Nett, T. A. Gulder, A. J. Kale, C. C. Hughes and B. S. Moore, *J. Med. Chem.*, 2009, 52, 6163-6167.
23. L. H. Meng, R. Mohan, B. H. B. Kwok, M. Elofsson, N. Sin and C. M. Crews, *Proc. Natl. Acad. Sci. USA*, 1999, 96, 10403-10408.
24. L. Meng, B. H. Kwok, N. Sin and C. M. Crews, *Cancer Res.*, 1999, 59, 2798-2801.
25. S. D. Demo, C. J. Kirk, M. A. Aujay, T. J. Buchholz, M. Dajee, M. N. Ho, J. Jiang, G. J. Laidig, E. R. Lewis, F. Parlati, K. D. Shenk, M. S. Smyth, C. M. Sun, M. K. Vallone, T. M. Woo, C. J. Molineaux and M. K. Bennett, *Cancer Res.*, 2007, 67, 6383-6391.
26. D. J. Kuhn, Q. Chen, P. M. Voorhees, J. S. Strader, K. D. Shenk, C. M. Sun, S. D. Demo, M. K. Bennett, F. W. van Leeuwen, A. A. Chanan-Khan and R. Z. Orlowski, *Blood*, 2007, 110, 3281-3290.
27. M. Bogyo, J. S. McMaster, M. Gaczynska, D. Tortorella, A. L. Goldberg and H. Ploegh, *Proc. Natl. Acad. Sci. USA*, 1997, 94, 6629-6634.
28. T. Nazif and M. Bogyo, *Proc. Natl. Acad. Sci. USA*, 2001, 98, 2967-2972.
29. R. M. Rydzewski, L. Burrill, R. Mendonca, J. T. Palmer, M. Rice, R. Tahilramani, K. E. Bass, L. Leung, E. Gjerstad, J. W. Janc and L. Pan, *J. Med. Chem.*, 2006, 49, 2953-2968.
30. M. Groll, B. Schellenberg, A. S. Bachmann, C. R. Archer, R. Huber, T. K. Powell, S. Lindow, M. Kaiser and R. Dudler, *Nature*, 2008, 452, 755-758.
31. A. Baldisserotto, V. Ferretti, F. Destro, C. Franceschini, M. Marastoni, R. Gavioli and R. Tomatis, *J. Med. Chem.*, 2010, 53, 6511-6515.
32. M. Groll, Y. Koguchi, R. Huber and J. Kohno, *J. Mol. Biol.*, 2001, 311, 543-548.

33. A. Desvergne, E. Genin, X. Maréchal, N. Gallastegui, D. L., N. Richy, M. Groll, J. Vidal and M. Reboud-Ravaux, *J. Med. Chem.*, 2013, 56, 3367-3378.
34. N. Basse, M. Montes, X. Marechal, L. Qin, M. Bouvier-Durand, E. Genin, J. Vidal, B. O. Villoutreix and M. Reboud-Ravaux, *J. Med. Chem.*, 2010, 53, 509-513.
35. L. Borissenko and M. Groll, *Chem. Rev.*, 2007, 107, 687-717.
36. P. M. Voorhees and R. Z. Orłowski, *Annu. Rev. Pharmacol. Toxicol.*, 2006, 46, 189-213.
37. R. Z. Orłowski and D. J. Kuhn, *Clin. Cancer Res.*, 2008, 14, 1649-1657.
38. A. F. Kisselev, *Chem. Biol.*, 2008, 15, 419-421.
39. V. Kasam, N. R. Lee, K. B. Kim and C.-G. Zhan, *Bioorg. Med. Chem. Lett.*, 2014, 24, 3614-3617.
40. Z. Miller, K. S. Kim, D. M. Lee, V. Kasam, S. E. Baek, K. H. Lee, Y.-Y. Zhang, L. Ao, K. Carmony, N. R. Lee, S. Zhou, Q.-Q. Zhao, Y. J. Jang, H. Y. Jeong, C.-G. Zhan, W. Lee, D. E. Kim and K. B. Kim, *J. Med. Chem.*, 2015, 58, 2036-2041.
41. M. Unno, T. Mizushima, Y. Morimoto, Y. Tomisugi, K. Tanaka, N. Yasuoka and T. Tsukihara, *Structure*, 2002, 10, 609-618.
42. M. Groll, K. B. Kim, N. Kairies, R. Huber and C. M. Crews, *J. Am. Chem. Soc.*, 2000, 122, 1237-1238.
43. M. Groll, C. R. Berkers, H. L. Ploegh and H. Ovaas, *Structure*, 2006, 14, 451-456.
44. J. Hines, M. Groll, M. Fahnestock and C. M. Crews, *Chem. Biol.*, 2008, 15, 501-512.
45. M. Groll, T. Nazif, R. Huber and M. Bogoy, *Chem. Biol.*, 2002, 9, 655-662.
46. D.-H. Wei, B.-L. Lei, M.-S. Tang and C.-G. Zhan, *J. Am. Chem. Soc.*, 2012, 134, 10436-10450.
47. C. S. Coleman, J. P. Rocetes, D. J. Park, C. J. Wallick, B. J. Warn-Cramer, K. Michel, R. Dudler and A. S. Bachmann, *Cell Proliferat.*, 2006, 39, 599-609.
48. Y. K. Zhang, H. Y. Liu and W. T. Yang, *J Chem Phys*, 2000, 112, 3483-3492.
49. Y. K. Zhang, *J Chem Phys*, 2005, 122, 024114.
50. Y. K. Zhang, *Theor Chem Acc*, 2006, 116, 43-50.
51. F. Zheng, W.-C. Yang, M. C. Ko, J.-J. Liu, H. Cho, D.-Q. Gao, M. Tong, H. H. Tai, J. H. Woods and C.-G. Zhan, *J. Am. Chem. Soc.*, 2008, 130, 12148-12155.
52. J.-J. Liu, A. Hamza and C.-G. Zhan, *J. Am. Chem. Soc.*, 2009, 131, 11964-11975.
53. J.-J. Liu, Y.-K. Zhang and C.-G. Zhan, *J. Phys. Chem. B*, 2009, 113, 16226-16236.

54. J.-J. Liu, X.-Y. Zhao, W.-C. Yang and C.-G. Zhan, *J. Phys. Chem. B*, 2011, 115, 5017-5025.
55. J.-J. Liu and C.-G. Zhan, *J. Chem. Theory Comput.*, 2012, 8, 1426-1435.
56. D.-M. Li, X.-Q. Huang, K.-L. Han and C.-G. Zhan, *J. Am. Chem. Soc.*, 2011, 133, 7416-7427.
57. C. R. Sondergaard, M. H. M. Olsson, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, 7, 2284-2295.
58. M. H. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2010, 7, 525-537.
59. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. Montgomery, J. A., T. Vreven, K. N.udin, J. C. Burant, J. M. Millam, S. S. yengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. shida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian 03*, Gaussian, Inc., Wallingford, CT, Version C.02 edn., 2004.
60. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross and P. A. Kollman, *AMBER11*, University of California, San Francisco 2010.
61. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, 79, 926-935.
62. B. Lei, M. D. Abdul Hameed, A. Hamza, M. Wehenkel, J. L. Muzyka, X. J. Yao, K. B. Kim and C.-G. Zhan, *J. Phys. Chem. B*, 2010, 114, 12333-12339.

63. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross and P. A. Kollman, AMBER8, University of California, San Francisco 2004.
64. X. Chen, L. Fang, J. Liu and C.-G. Zhan, *J. Phys. Chem. B*, 2011, 115, 1315-1322.
65. X. Chen, L. Fang, J.-J. Liu and C.-G. Zhan, *Biochemistry*, 2012, 51, 1297-1305.
66. Y. Yao, J.-J. Liu and C.-G. Zhan, *Biochemistry*, 2012, 51, 8980-8992.
67. D.-H. Wei, L. Fang, M.-S. Tang and C.-G. Zhan, *J. Phys. Chem. B*, 2013, 117, 13418-13434.
68. M. Yao, W. Tu, X. Chen and C.-G. Zhan, *Org. Biomol. Chem.*, 2013, 11, 7595-7605.
69. Y. Qiao, K. Han and C.-G. Zhan, *Biochemistry*, 2013, 52, 6467-6479.
70. D.-H. Wei, X.-Q. Huang, J.-J. Liu, M.-S. Tang and C.-G. Zhan, *Biochemistry*, 2013, 52, 5145-5154.
71. D.-M. Li, X.-Q. Huang, J. Lin and C.-G. Zhan, *Dalton Trans.*, 2013, 42, 3812-3820.
72. J. R. Alvarez-Idaboy, A. Galano, G. Bravo-Perez and M. E. Ruiz, *J. Am. Chem. Soc.*, 2001, 123, 8387-8395.