

# Understanding the Autocatalytic Process of Pro-kumamolisin Activation from Molecular Dynamics and Quantum Mechanical/Molecular Mechanical (QM/MM) Free-Energy Simulations

Jianzhuang Yao,<sup>[a]</sup> Alexander Wlodawer,<sup>[b]</sup> and Hong Guo\*<sup>[a]</sup>

Proteolytic enzymes are synthesized as inactive precursors or zymogens to promote folding, prevent unwanted protein degradation, and provide a mechanism for regulating protein function through proteolytic activation. Understanding the ways by which Nature prevents unwanted activation of proteases and mechanisms for conversion of zymogens to active enzymes is of considerable interest.<sup>[1]</sup> However, a detailed understanding of the activation processes and energetics involved is still lacking. Herein, we applied molecular dynamics (MD) and quantum mechanical/molecular mechanical (QM/MM) free-energy simulations to study the process and energetics of the conversion of the kumamolisin zymogen to the active enzyme. It has been shown that the protonation of Asp164 would trigger conformational changes and generate the functional active site for autocatalysis. The mechanism of acylation for autocatalytic cleavage of prodomain is also derived from the two-dimensional QM/MM free-energy simulations. The results seem to indicate that one of the reasons for sedolisins to use an aspartate as a catalyst (e.g., Asp164 in kumamolisin) instead of asparagine (the oxyanion-hole residue in classical serine peptidase) might be due in part to the requirement for the creation of a built-in switch that delays the self-activation until secretion into acidic medium.

Kumamolisin belongs to a recently characterized family of serine-carboxyl peptidases (sedolisins),<sup>[2,3]</sup> which are present in a wide variety of organisms and are most active at low pH. The members of the family also include tripeptidyl-peptidase 1 (TPP1) for which the loss of the activity as a result of mutations in the TPP1 gene is believed to be the cause of a fatal neurodegenerative disease.<sup>[4]</sup> Similar to other proteo-

lytic enzymes, sedolisins are synthesized as inactive precursors to prevent unwanted activation and proteolysis. The inactive precursors for sedolisins include propeptides of approximately 200 amino acids in length, which may play a role in the folding of the proteins, as well as protecting them from being prematurely activated. The activation cleavage of the prodomains occurs after the release of zymogens into the acidic environment, leading to the production of the active enzymes.<sup>[3b]</sup>

The crystal structures have been determined for the inactive Ser278Ala pro-kumamolisin mutant,<sup>[5]</sup> as well as for the active kumamolisin.<sup>[6]</sup> A superposition of the structure of the Ser278Ala pro-kumamolisin mutant with that of kumamolisin showed that the catalytic domain of the mutant exhibited a virtually identical structure compared to the active enzyme.<sup>[5]</sup> The pro-kumamolisin structure has therefore proved that the catalytic domain has basically adopted a mature-like conformation already in the zymogen form. One of the key structural features of the proenzyme is the presence of a salt bridge between the P<sub>3</sub>-Arg169p linker residue (herein, the suffix "p" designates the residues in the pro-domain) and Asp164 from the catalytic domain.<sup>[5]</sup> The existence of this salt bridge seems to prevent the formation of the functional configuration for Asp164.<sup>[7]</sup> An interesting question is whether the protonation of Asp164 and subsequently breaking of this salt bridge during secretion into an acidic medium would trigger conformational changes and lead to generation of well-positioned general acid/base catalyst (Asp164) and creation of a functional active site in the preparation of the self-activation.<sup>[5]</sup>

The X-ray structure of the Ser278Ala pro-kumamolisin mutant was used to generate the model for the wild-type pro-kumamolisin through a manual change of Ala278 to Ser278. Figure 1A shows that the catalytic Ser residue, along with Glu78 and Asp82, are located at the correct position and would attack the carbonyl carbon of P<sub>1</sub>-His171p during the autocatalytic cleavage. This agrees with the earlier suggestion that the catalytic triad in the pro-kumamolisin structure has already adopted a mature-like conformation.<sup>[5]</sup> In contrast, Asp164 is located at a different position compared to that observed in the active enzyme<sup>[6]</sup> and forms a salt bridge with the P<sub>3</sub>-Arg169p (Figure 1A). Asp164 is therefore unable to participate in the stabilization of the TI during the autocatalytic cleavage of the peptide bond between P<sub>1</sub>-His171p and P<sub>1</sub>'-Phe172p.

[a] J. Yao, Prof. H. Guo  
Department of Biochemistry and Cellular and Molecular Biology  
University of Tennessee, Knoxville, TN 37996 (USA)  
Fax: (+1)865-974-6306  
E-mail: hguo1@utk.edu

[b] Dr. A. Wlodawer  
Protein Structure Section  
Macromolecular Crystallography Laboratory  
National Cancer Institute, Frederick, MD 21702 (USA)

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/chem.201301310>.

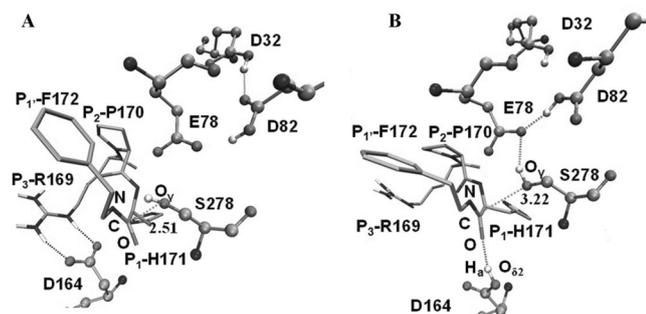


Figure 1. A) Active site of pro-kumamolisin with Ala278 changed to Ser (wild type) before the protonation of Asp164. The residues from the catalytic domain are shown in balls and sticks, and those from the prodomain are in sticks. For clarity, most of hydrogen atoms were not added. B) Average structure after the protonation of  $O_{\delta 2}$  and 20 ns MD simulations. A similar result was obtained with the protonation of the other oxygen of Asp164.

The average active-site structure for pro-kumamolisin after the protonation of Asp164 and 20 ns MD simulations is shown in Figure 1B. It is evident that the salt bridge between  $P_3$ -Arg169p and Asp164 has already been broken; Arg169p moved away from its position in pro-kumamolisin and interacts instead with solvent molecules. Interestingly, after the MD simulations, the active site has changed to the functional configuration with Asp164 well aligned to act as a general acid/base catalyst. This was confirmed with five independent 20 ns simulations with different initial velocities. As was mentioned earlier, the catalytic triad in the pro-kumamolisin structure has already adopted a mature-like conformation. Figure 1B shows that this catalytic machinery has not been significantly altered by the MD simulations and remains functional. Thus, our simulations have confirmed the hypothesis that the protonation of Asp164 would trigger conformational changes to generate the well-positioned general acid/base catalyst (Asp164) and a functional active site for the next step of the autocatalytic pro-kumamolisin activation.

The two-dimensional free-energy map for the acylation step of the autocatalytic cleavage is presented in Figure 2A. The free-energy curve along the minimum free-energy path in Figure 2A is shown in Figure 2B. A comparison of the curve in Figure 2B with the one for the acylation reaction involving kumamolisin As published earlier<sup>[7d]</sup> shows that the two curves are rather similar, even though the barriers in Figure 2B are slightly lower (by about 2 kcal mol<sup>-1</sup>). We also determined the free-energy curve without the protonation of Asp164, and the free-energy barrier was found to be considerably higher (by more than 10 kcal mol<sup>-1</sup>). This result suggests that the protonation of Asp164 and breaking of the salt bridge are the necessary steps for autocatalysis, consistent with experimental observations. Some average active-site structures obtained from the QM/MM simulations are shown in Figure 3A–C. As is evident from Figure 3, Ser278 is the nucleophile that attacks the carbonyl carbon atom C-

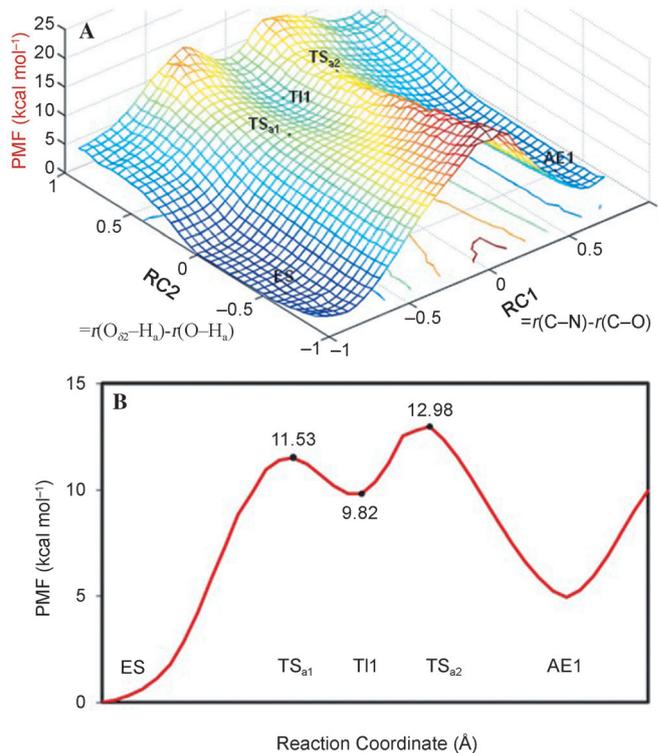


Figure 2. A) 2D free-energy (PMF) map for the acylation step of self-activation obtained from the QM/MM simulations. B) Free-energy curve along the minimum free-energy path in A. Herein, ES is used for the initial state of pro-kumamolisin before acylation (i.e., after the conformational changes to generate the active configuration), TSa1 is the first TS, TI1 is the tetrahedral intermediate for acylation, TSa2 is the second TS, and AE1 is the state corresponding to the acylzyme in the enzyme-catalyzed reaction.

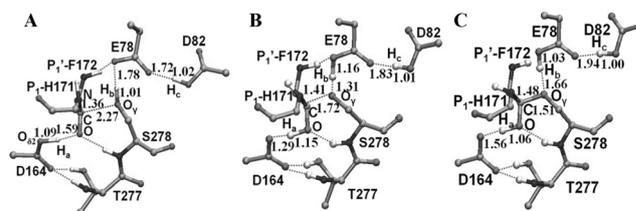


Figure 3. Certain active-site structures along the reactant path (see Figure 2B). A) Initial structure before acylation. B) TSa1. C) TI1.

(His171p), whereas Glu78 and Asp164 act as the general base and acid catalysts, respectively.

In conclusion, we were able to confirm the hypothesis that the protonation of Asp164 would trigger conformational changes to position this residue as a general acid/base catalyst, thus reconstructing the functional active site. We have also studied the mechanism of the acylation reaction for the autocatalytic cleavage of the prodomain and obtained the free-energy map and catalytic mechanism for the acylation reaction.

## Experimental Section

**Computational studies:** The initial coordinates for pro-kumamolisin were based on the X-ray structure (protein data bank (PDB) code: 1T1E) of the S278 A mutant;<sup>[5]</sup> Ala287 was converted back to Ser278 manually. The structure was constructed under acidic (pH 3.0) condition based on calculated  $pK_a$  values. The whole system was solvated in a TIP3P<sup>[11]</sup> water box of  $85 \times 106 \times 93 \text{ \AA}$  by using the NAMD 2.9 program.<sup>[12]</sup> Fifty three chloride ions ( $\text{Cl}^-$ ) were added to neutralize the charged system. Five independent MD simulations of 20 ns each with different initial velocities and protonation sites (i.e.,  $\text{O}_{\delta 2}$  or  $\text{O}_{\delta 1}$ ) were performed on the TACC Ranger at the University of Texas at Austin by using 128 processes, periodic boundary conditions and a time step of 0.002 ps. The CHARMM force field (PARAM 27)<sup>[13]</sup> was used along with the SHAKE algorithm.<sup>[14]</sup> The system was gradually heated to 300 K from 100 K within a time period of 100 ps, followed by 10 ns equilibration and 10 ns product run (NPT) with temperature controlled by Langevin thermostat.<sup>[15]</sup> A typical reactant structure (close to the average structure) was selected to perform the QM/MM MD simulation.

The stochastic boundary molecular dynamics<sup>[16]</sup> was used for the QM/MM MD and free-energy simulations. The reaction region was a sphere with radius  $r$  of  $20 \text{ \AA}$ , and the buffer region extended over  $20 \leq r \leq 22 \text{ \AA}$ . The reference center for partitioning the system was chosen to be the carbonyl carbon atom of P1-His171p. The QM part includes the side chains of Glu78, Asp82, Asp164, and Ser287, as well as a part of the linker peptide containing the carbonyl of P2-Pro170p, the backbone of P1-His171p, and the  $\text{C}_\alpha$  and amide group of P1'-Phe172p. The atoms in the QM region were simulated with the self-consistent charge density functional tight binding (SCC-DFTB) method implemented in the CHARMM program.<sup>[9]</sup> The rest of the system was treated with the MM method by using the all-hydrogen potential function (PARAM 27). The link-atom approach was used to separate the QM and MM regions.<sup>[17]</sup>

The 2D free-energy map (potential of mean force, PMF) from the substrate complex to the acyl enzyme was determined with the umbrella-sampling method<sup>[18]</sup> and two-dimensional weighted-histogram analysis method.<sup>[19]</sup> The time step for QM/MM MD simulation was 0.001 ps. The first reaction coordinate,  $\xi_1 = r(\text{C}-\text{N}) - r(\text{C} \cdots \text{O}_\gamma)$ , is the distance difference between the scissile peptide bond  $\text{R}[\text{C}(\text{His171p})-\text{N}(\text{Phe172p})]$  and nucleophilic attack distance  $[\text{C}(\text{His171p}) \cdots \text{O}_\gamma(\text{Ser278})]$ . The second reaction coordinate is  $\xi_2 = r(\text{O}_{\delta 2}-\text{H}_a) - r(\text{O}-\text{H}_a)$ . The reason for choosing the reaction coordinate involving Asp164 instead of Glu78 is that dynamic substrate-assisted catalysis involving Asp164 has been observed in the earlier study,<sup>[7c]</sup> whereas the energetic effect of the proton transfer involving Glu78 could be reflected in the 1D PMF profile.<sup>[7b]</sup> 1296 windows were used in the 2D PMF calculations, and free-energy mesh-contour plot was created with a bin size of  $0.05 \times 0.05 \text{ \AA}^2$ . For each window, 50 ps heating, 50 ps equilibration 50 ps production

run with a force constant of  $1000 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$  were performed. It is of interest to notice that previous computational studies on the members of sedolisin family led to rather similar barriers for the acylation reaction based on QM-(SCC-DFTB)/MM PMF simulations<sup>[7c]</sup> and high-level ab initio potential energy calculations.<sup>[20]</sup>

## Acknowledgements

We thank Dr. Toru Nakayama and Dr. Qin Xu for helpful discussions and Professor Martin Karplus for a gift of the CHARMM program. This work was supported by the National Science Foundation (NSF) (Grant number: 0817940 to H.G.) and the NSF TeraGrid resources provided by the University of Texas at Austin, as well as by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

**Keywords:** acid/base catalysis • autocatalysis • enzymes • molecular dynamics • zymogens

- [1] A. R. Khan, M. N. G. James, *Protein Sci.* **2008**, *7*, 815–836.
- [2] A. Wlodawer, M. Li, A. Gustchina, H. Oyama, B. M. Dunn, K. Oda, *Acta Biochim. Pol.* **2003**, *50*, 81–102.
- [3] a) K. Oda, M. Sugitani, K. Fukuhara, S. Murao, *Biochimica Et Biophysica Acta* **1987**, *923*, 463–469; b) K. Oda, *J. Biochem.* **2012**, *151*, 13–25.
- [4] D. E. Sleat, R. J. Donnelly, H. Lackland, C. G. Liu, I. Sohar, R. K. Pullarkat, P. Lobel, *Science* **1997**, *277*, 1802–1805.
- [5] M. Comellas-Bigler, K. Maskos, R. Huber, H. Oyama, K. Oda, W. Bode, *Structure* **2004**, *12*, 1313–1323.
- [6] M. Comellas-Bigler, P. Fuentes-Prior, K. Maskos, R. Huber, H. Oyama, K. Uchida, B. M. Dunn, K. Oda, W. Bode, *Structure* **2002**, *10*, 865–876.
- [7] a) H. B. Guo, A. Wlodawer, H. Guo, *J. Am. Chem. Soc.* **2005**, *127*, 15662–15663; b) H. B. Guo, A. Wlodawer, T. Nakayama, Q. Xu, H. Guo, *Biochemistry* **2006**, *45*, 9129–9137; c) Q. Xu, H. Guo, A. Wlodawer, *J. Am. Chem. Soc.* **2006**, *128*, 5994–5995; d) Q. Xu, H. B. Guo, A. Wlodawer, T. Nakayama, H. Guo, *Biochemistry* **2007**, *46*, 3784–3792.
- [8] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187–217.
- [9] Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, M. Karplus, *J. Phys. Chem. B* **2001**, *105*, 569–585.
- [10] Q. Xu, L. Li, H. Guo, *J. Phys. Chem. B* **2010**, *114*, 10594–10600.
- [11] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [12] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- [13] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wierkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- [14] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341.
- [15] S. Adelman, J. Doll, *J. Comput. Phys.* **1976**, *64*, 2375.
- [16] C. L. Brooks, A. Brunger, M. Karplus, *Biopolymers* **1985**, *24*, 843–865.
- [17] M. J. Field, P. A. Bash, M. Karplus, *J. Comput. Chem.* **1990**, *11*, 700–733.

- [18] G. M. Torrie, J. P. Valleau, *Chem. Phys. Lett.* **1974**, 28, 578–581.  
[19] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, J. M. Rosenberg, *J. Comput. Chem.* **1992**, 13, 1011–1021.

- [20] K. Bravaya, A. Bochenkova, B. Grigorenko, L. Topol, S. Burt, A. Nemukhin, *J. Chem. Theory Comput.* **2006**, 2, 1168–1175.

Received: April 7, 2013  
Revised: May 30, 2013  
Published online: July 2, 2013