A Molecular Dynamics Study to Assess the Positive Ion Distribution and the Effects of Protonation on the N-terminus Region of a Xylanase.

Kewalin Posansee[1,2], Pongsak Khunrae[3] and Thana Sutthibutpong[1,2]*

[1]Theoretical and Computational Physics Group, Department of Physics, KMUTT, Thailand. Faculty of Science, King Mongkut's University of Technology Thonburi.

[2]Theoretical and Computational Science Center (TaCS), Science Laboratory Building, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT), 126 Pracha-Uthit Road, Bang Mod, Thrung Khru, Bangkok 10140, Thailand.

[3]Department of Microbiology, Science Laboratory Building, Faculty of Science, King Mongkut's University of Technology Thonburi.

*email: thana.sut@mail.kmutt.ac.th

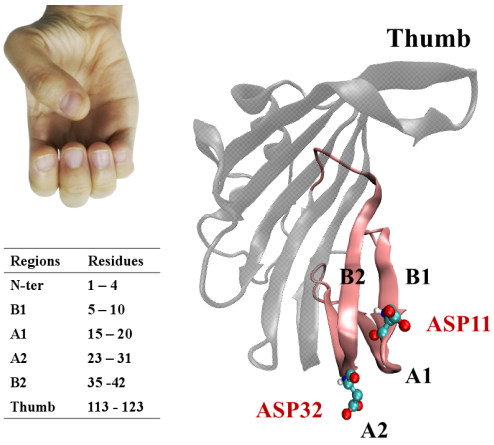
Abstract. Xylanase enzyme families play an important role in many industries due to its ability to digest the hemicellulose within plant cell walls. Developing the enzyme that can withstand a highly acidic environment in a stomach requires an extended information in molecular scale for further protein engineering. In this study, an atomistic molecular dynamics simulation was performed to assess the positive ion distribution pattern on the N-terminus region of a Xylanase enzyme molecule. Then, another series of simulations were performed for the enzyme protonated at the aspartate residues 11 and/or 32 identified by the previous calculation. Conformational analysis showed that ASP32 protonation caused a significant disruption at the N-terminus region, as the nearby arginine residue 31 (ARG31) became free from the Coulomb interaction with the negatively charged ASP32 sidechain and cause the N-terminus region to misfold.

1. Introduction

In modern animal feedstock industry, enzyme technologies are used during the processing of food ingredients [1]. One of the most popular enzyme family is Xylanase. Xylanase possess its function on hemicellulose digestion, which is an important step prior to the total cellulose digestion for the plant cell walls [2]. To implement xylanase into the feedstock, the enzyme is mixed with the other ingredients for the primary digestion before undergoing a granulation process by an extruding machine. However, to further improve the efficiency of the enzymes, secondary digestion of the enzymes within the stomachs has been proposed. To reach this goal, enzyme molecules with high resistivity to high temperature of the extruding machine and to low pH in the stomachs are needed [3]. A number of enzyme engineering attempts have been done both experimentally and computationally to improve thermostability of xylanase and other families of cellulase and hemicellulase enzymes [4]. Many studies on the molecular biology and molecular biophysics of xylanase enzymes from the GH11 family (Xyn11A) refer the three-dimensional structures of an Xyn11A molecule as a 'jellyroll' or a 'right hand' (see figure 1.) [5,6]. Atomistic molecular dynamics simulations, a physics-based computational approach, have long been used in the studies of protein structures and functions due to their abilities to provide atomistic details

of both equilibrium and non-equilibrium processes occurs within protein structures under some extreme environments. Our recent work on improving the thermostability of a xylanase enzyme of the from *Bacillus sp.* showed that introducing a pair of disulfide bond connecting the helix and the 'back-hand' beta hairpin-loop greatly improve the thermostability [7]. However, another weak-spot on the protein structure was found at the 'Pinky-finger' region at the N-terminus. Thus, one of the ongoing studies is focusing on the addition of a hydrophobic residue to stabilize the beta-strands around the 'Pinky-finger' region.

In this work, the distribution of positive ions and the stability of protonated 'Pinky-finger' region were investigated. In the first part, an atomistic molecular dynamics simulation of a wildtype Xyn11A recombinase from *Bacillus sp.* was performed to see the distribution of positive ions about each residue at the 'Pinky-finger' region (amino acid residues 1-50). High salt concentration was chosen in order to accelerate the sufficient statistics on the ion distribution. After that, two residues were chosen to create protonated enzyme structures for further series of MD simulation runs. Conformational analysis addressed the molecular mechanisms of the local structural disruptions due to the protonation, which could serve as information for further protein engineering towards the acid-stable Xyn11A.



Regions	Residues
N-ter	1 – 4
B1	5 – 10
A1	15 – 20
A2	23 – 31
B2	35 -42
Thumb	113 - 123

Figure 1. GH11 Xylanase (Xyn11A) model represented by a 'right hand'. The 'Pinky finger' region was indicated in Pink, while the beta strands B1, A1, A2 and B2 are labelled, and residues ASP11 and ASP32 are represented by balls and sticks.

2. Methods

Atomistic structure of Xyn11A was obtained from a homology modelling and were parameterized by the AMBER03 force-field [8]. The protein was solvated in TIP3P explicit solvent [9] $(66.5\times66.5\times66.5)Å^3$ water box, with Na+ or Cl- counter-ions to neutralize the charged amino acid residues. Additional Na+ and Cl- ions were added to emulate 0.5 M salt concentration. Energy minimization using the steepest descent method and 0.5 ns MD equilibration, where each structure was gradually heated from 100K to 333K, was performed. After that, 100 ns production MD run was carried out by GROMACS 5.1.2 at 1 bar and at 333 K [10]. Berendsen and velocity rescale algorithms were employed to regulate pressure and temperature in the NPT ensemble [11]. The particle mesh Ewald (PME) method [12] and a cutoff distance of 10 Å were used. LINCS constraint algorithm [13] was applied to covalent bonds associated with hydrogen, so that the timestep of 2.0 fs was allowed. After simulations finished, all the trajectory snapshots were fitted to their starting structures by using the least square fitting method from the GROMACS module *trjconv*. Then, to observe the ion distribution and determine the sites for protonation, minimum distances between Na+ ions and Pinky-finger region were analyzed. After the protonation sites were decided, protonated starting structures were built and followed similar simulation protocol as the non-protonated structure. Root mean square deviation (RMSD) of each Xyn11A structure, excluding 'Thumb' region, relative to the starting structure was then calculated

for each 10 ps of the simulation trajectories at all protonated conditions to measure the ensemble average of local conformational changes along a 100-ns trajectory. RMSD at a given time t is defined by

$$\mathrm{RMSD}(0,t) = \left(\frac{1}{M} \sum_{i=1}^{N} m_i \|\mathbf{r}_i(0) - \mathbf{r}_i(t)\|^2 \right)^{1/2} \quad (1)$$

while m_i and r_i are the mass and the position of atom i, M is the total mass and N is the total number of atoms. Also, the Pinky-finger region (residue 1- 50) was monitored. Then, root mean square fluctuation (RMSF) per residue was calculated to measure the local fluctuation affected by amino acid protonation. RMSF of an atom i is defined by

$$\mathrm{RMSF} = \left(\frac{1}{T} \sum_{j} (\mathbf{r}_i(t_j) - \tilde{\mathbf{r}}_i)^2 \right)^{1/2} \quad (2)$$

while r_i is the position of atom i and T is the total period of time, Visual molecular dynamics (VMD) program [14] was used to visualize the molecular structures.

3. Results and Discussions

3.1. MD Simulations to Study Ion Distribution Around a Xylanase Molecule

In order to determine the positive ion binding sites within the Pinky-finger region, probability of Na+ ions with minimum distances from amino acids less than 4 Å was analyzed from the MD trajectory of Xyn11A wildtype structure at 333K. From the probability of ion binding at each residue shown in figure 2, the two most ion-populated residues were ASP11 and ASP32 (indicated by red arrows). It could be explained that aspartate sidechains had low pKa value and was negatively charged at neutral pH environment. In an acidic environment in which protons were abundant, aspartate sidechains were more likely to be protonated. Therefore, ASP11 and ASP32 were chosen as the protonated residues in the latter simulations in order to monitor the conformational changes and structural disruptions due to the amino acid protonation.

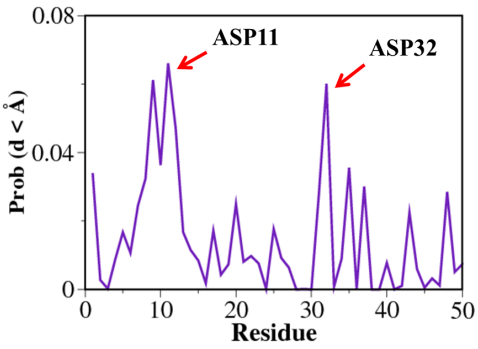
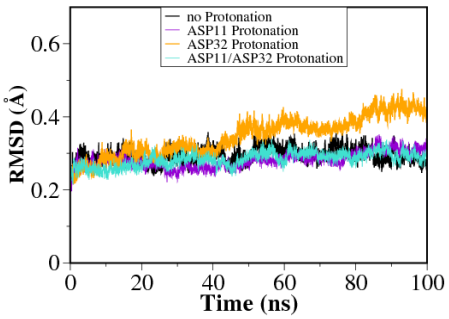


Figure 2. Probability of Na+ ions to be in close proximity (minimum distance d < 4 Å) with each amino acid. Probability at the residues ASP11 and ASP32 were indicated by arrows.

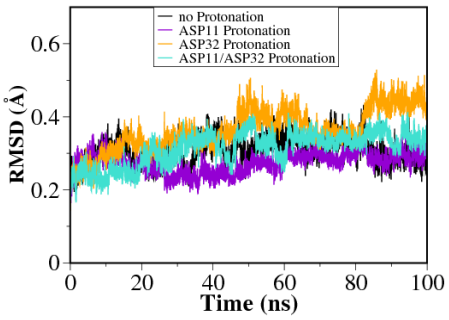
3.2. MD Simulations to Monitor Conformational Changes

Figure 3a and 3b showed the Root Mean Square Deviation (RMSD) values calculated from the MD trajectories of Xyn11A structures with different protonation states . RMSD was firstly calculated for the whole protein structure excluding the thumb region. Figure 3a showed that the RMSD at the 'nothumb' region was found the highest for the structure with ASP32 protonation after 90 ns MD simulation (4.2 ± 0.1 Å). The rest of the structures had lower RMSD values (3.0 ± 0.2 Å) than the ASP32-protonated

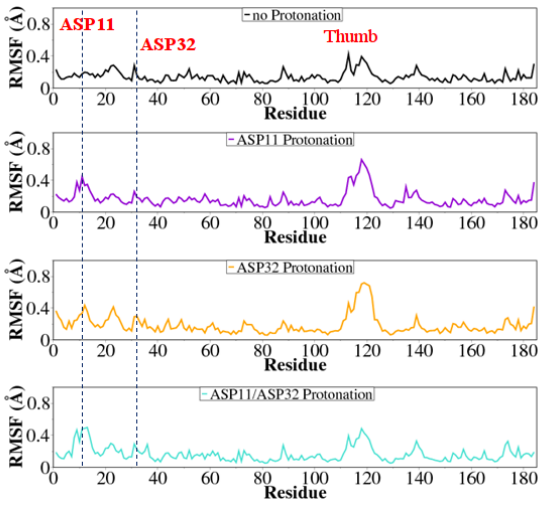
structure. Similar calculations were done for the 'Pinky-finger' region and the ASP32-protonated structure also displayed the highest RMSD (4.4 ± 0.2 Å), followed by the ASP11/ASP32-protonated structure (3.4 ± 0.2 Å). These results were surprising because the structure with a single protonation at ASP32 underwent significantly larger conformational changes than that with double protonation. To investigate the local flexibility at each amino acid residue, Root Mean Square Fluctuation (RMSF) was calculated for all four structures. The dashed lines in figure 3c indicated the positions of ASP11 and ASP32 protonation. The RMSF results showed that, for all three protonated states, there were significant increases in the fluctuations of B1/A1 loops, including the ASP11 residue, compared to the non-protonated state. However, the ASP32 residue was unaffected even when ASP32 itself was protonated.



(a) nothumb



(b) Pinky-finger



(c) RMSF

Figure 3. a) - b) Root Mean Square Deviation (RMSD) and c) Root Mean Square Fluctuation (RMSF) per residue calculated from the simulations of Xyn11A with no protonation (black), ASP11 protonation (purple), ASP32 protonation (orange) and ASP11/ASP32 protonation (cyan) for a) 'nothumb' the whole structure excluding the thumb region and b) 'Pinky-finger' region (residues 1- 50).

Mechanisms of Pinky-finger disruptions were illustrated in figure 4. Figure 4a shows the conformation of the Pinky-finger region (residue 1-50) from the wildtype, non-protonated Xyn11A. A hydrogen bonding network formed between two hairpin-loops, B1/A1 (residues 11-15) and A2/B2 (residues 31-35), was similar to that found between beta-strands, but with some mispairing due to the flexibility of the loop. Also, additional hydrogen bond were found between one of the terminal oxygen atoms of ASP11 and a secondary amine (N-H) group at the backbone of nearby residue. ASP11 protonation, shown in figure 4b, reduced the probability of h-bonding between ASP11 and the backbone, resulted in the increased flexibility of the B1/A1 loop, consistent with RMSF result. The most significant changes of RMSD were observed when Xyn11A was protonated at the ASP32 residue and the charge of the protonated-ASP32 sidechain became neutral. As a result, Coulomb interaction between ARG31 and ASP32 sidechains was removed. ARG31 then created a non-native hydrogen bonding network, which exposed the B1/A1 loop to further misfolding created by the extended hydrogen bonding network with ASP11. Meanwhile, the ASP11/ASP32 protonated structure shown in figure 4d was less disrupted, as one of the two terminal oxygen atoms became protonated and was less likely to form the extended h-bonding network seen in the ASP32 protonated structure.

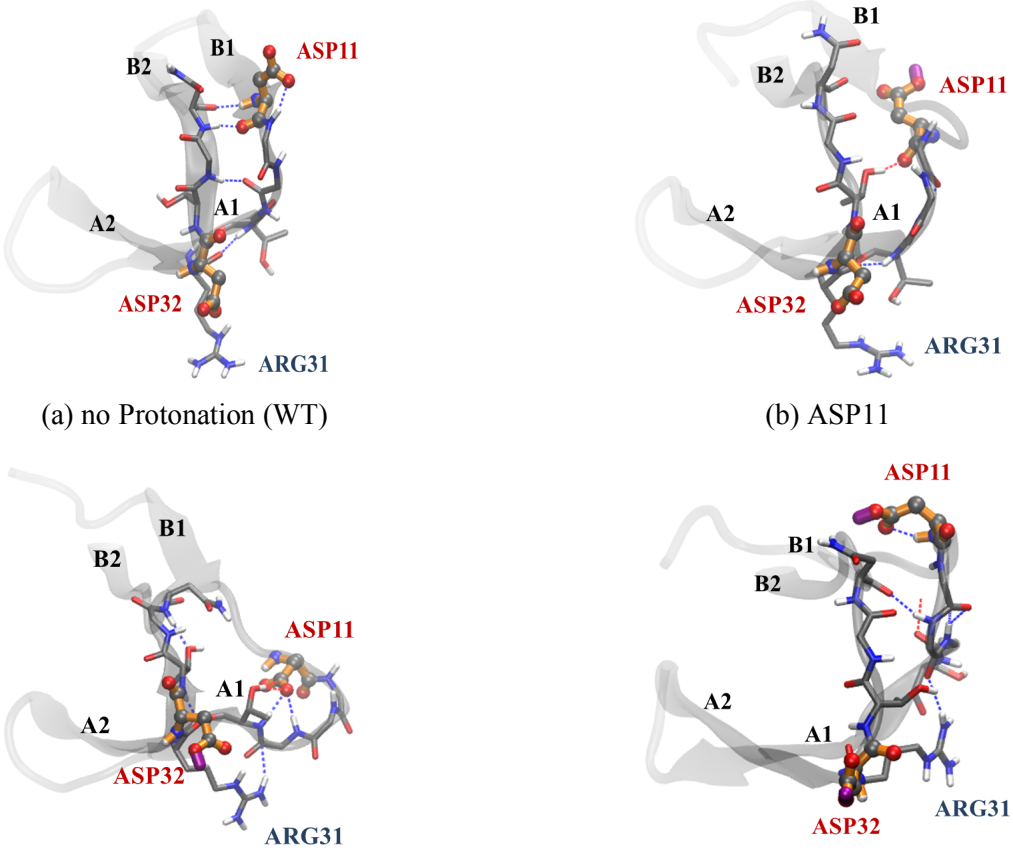


Figure 4. sample conformational snapshots of the 'Pinky finger' region (residue 1-50) taken from the simulations of Xyn11A with a) no protonation, b) ASP11, c) ASP32 and d) ASP11/ASP32 protonation. Residues 11 and 32 were represented by balls and sticks the other hairpin-loop residues (B1/A1: residues 11-15 and A2/B2: residues 31-35) were represented by licorice, in which C, H, O and N atoms were shown in grey, white, red and blue, respectively. The protonated sites were highlighted in purple and the hydrogen bonds were represented in dashed lines under the conditions that donor-acceptor distances < 3.5 Å and donor-hydrogen-acceptor angle < 45°.

4. Conclusions

A series of atomistic molecular dynamics simulations were performed to predict the positive ion distribution on the 'Pinky-finger' region of a GH11 xylanase enzyme, and to investigate the conformational changes due to the protonation at the predicted sites. Conformational analysis showed that ASP32 protonation caused ARG31 to form a dynamic hydrogen bonding networks with the B1/A1 hairpin loop (residues 11-15). This non-native local stabilization caused the 'Pinky-finger' region to be misfolded and became the nucleation point for total destabilization. Simulation results provided information for further protein engineering to increase the enzyme acido-tolerability, as the removal of the misfolding source - residue ARG31 - was suggested.

Acknowledgements.

K.P. was supported by the Theoretical and Computational Science (TaCS) center. T.S. was supported by Thailand Center of Excellence in Physics (ThEP).

References

[1] Birhanu M Y, Girma A and Puskur R 2017 Determinants of success and intensity of livestock feed technologies use in Ethiopia: Evidence from a positive deviance perspective *Technol. Forecast. Soc.* **115** 15–25

[2] Polizeli M L T M, Rizzatti A C S, Monti R, Terenzi H F, Jorge J A and Amorim D S 2005 Xylanases from fungi: properties and industrial applications *Appl. Microbiol. Biotechnol.* **67** 577–91

[3] Nasrollahi S M, Zali A, Ghorbani G R, Moradi Shahrbabak M and Heydari Soltan Abadi M 2017 Variability in susceptibility to acidosis among high producing mid-lactation dairy cows is associated with rumen pH, fermentation, feed intake, sorting activity, and milk fat percentage *Anim. Feed Sci. Technol.* **228** 72–82

[4] Childers M C and Daggett V 2017 Insights from molecular dynamics simulations for computational protein design *Mol. Syst. Des. Eng.* **2** 9–33

[5] Törrönen A, Harkki A and Rouvinen J 1994 Three-dimensional structure of endo-1,4-beta-xylanase II from Trichoderma reesei: two conformational states in the active site *EMBO J.* **13** 2493–501

[6] Purmonen M, Valjakka J, Takkinen K, Laitinen T and Rouvinen J 2007 Molecular dynamics studies on the thermostability of family 11 xylanases *Protein. Eng. Des. Sel.* **20** 551–9

[7] Sutthibutpong T, Rattanarojpong T and Khunrae P 2017 Effects of helix and fingertip mutations on the thermostability of xyn11A investigated by molecular dynamics simulations and enzyme activity assays *J. Biomol. Struct. Dyn.* 1–15

[8] Cornell W D, Cieplak P, Bayly C I, Gould I R, Merz K M, Ferguson D M, Spellmeyer D C, Fox T, Caldwell J W and Kollman P A 1995 A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules *J. Am. Chem. Soc.* **117** 5179–97

[9] Price D J and Brooks C L 2004 A modified TIP3P water potential for simulation with Ewald summation The *J. Chem. Phys.* **121** 10096–103

[10] Abraham M J, Murtola T, Schulz R, Páll S, Smith J C, Hess B and Lindahl E 2015 GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers *SoftwareX* **1–2** 19–25

[11] Berendsen H J C, Postma J P M, van Gunsteren W F, DiNola A and Haak J R 1984 Molecular dynamics with coupling to an external bath *J. Chem. Phys.* **81** 3684–90

[12] Darden T, York D and Pedersen L 1993 Particle mesh Ewald: An N · log(N) method for Ewald sums in large systems *J. Chem. Phys.* **98** 10089–92

[13] Hess B, Bekker H, Berendsen H J C and Fraaije J G E M 1997 LINCS: A linear constraint solver for molecular simulations *J. Comput. Chem.* **18** 1463–72

[14] Humphrey W, Dalke A and Schulten K 1996 VMD: Visual molecular dynamics *J. Mol. Graph. Model.* **14** 33–8